Essays on the Impact of Internal Migration and Intergenerational Dynamics on Human Capital: Evidence from Indonesia and China

by

RUYI SHI

Thesis submitted in fulfilment
of the requirements for the degree of

Doctor of Philosophy

June, 2025

Department of Economics
School of Economics and Public Policy
The University of Adelaide
Australia

Contents

List of Figures

List of Tables

Abstract

This thesis includes three independent chapters that explore how internal migration, education, and family structure affect human capital outcomes and intergenerational support in Indonesia and China.

The first chapter examines the causal impact of internal migration on formal employment in Indonesia using longitudinal household survey data from the Indonesia Family Life Survey (IFLS). We use rainfall as an instrumental variable for migration to address the endogeneity caused by reverse causality. Our empirical results show that, in general, internal migration is expected to increase the chance of having a formal job by approximately 16 percentage points. This impact is heterogeneous across gender and levels of education. The magnitude of this impact rises to around 25 percentage points for males but drops to only 7 percentage points for females. We also find migration only has a significant impact on individuals with senior high school education or higher, with the estimated impact being roughly 20 percentage points.

The second chapter examines the heterogeneous effect of adult children's education on elderly parents' mental health in China. We employ a quantile regression model to examine how the impact of children's education on parents' mental health varies across different levels of mental health status. To address the endogeneity of education, we use children's exposure to the compulsory schooling law as an instrument. Our results from the instrumental variable quantile estimation reveal that children's education has no effect on parents with mild to moderate mental health issues (below the 0.7 quantile) but has a positive and significant impact on those with severe issues (at the 0.7 quantile and above). Our split-sample and comparative analyses indicate that the impact varies according to children's gender, parental residency type, children's educational ranking within the household, and children's birth order.

The third chapter provides rich new evidence on China's evolving birth order effects on education. We use family-by-wave fixed effects to explore within-family variations in birth order across siblings. A positive birth order effect on education is observed

among older generations (born before 1967), while the birth order effect reverses to negative among younger generations (born after 1978). For the positive effect, school dropout among older siblings is the primary driver, whereas child labour laws and resource dilution are the key factors for the negative effect. Brothers have consistently attained higher levels of education than sisters, but this gap has narrowed over time. In addition, differences in old-age support levels are negligible both among siblings and between brothers and sisters. This implies that children who are expected to provide more old-age support to their parents in return for the favours they received during childhood—as suggested by the birth order effect—do not, in fact, offer more support. The altruistic motive emerges as a potential reason why siblings, as well as brothers and sisters, provide equal levels of old-age support to their parents. Furthermore, we find no evidence that in-kind support and monetary support offset or reinforce each other, suggesting that they function neither as substitutes nor complements. Our results offer a comprehensive understanding of the relationship between birth order, human capital investment, and intergenerational support in the face of China's rapid economic transition.

The contribution of this thesis is to provide micro-level evidence on enhancing human capital development. Specifically, the first chapter fills a gap in the literature by examining how internal migration affects individual labour market outcomes in Indonesia. The second chapter is the first to explore the heterogeneous effects of children's education on parents' mental health in China. The third chapter offers new and detailed evidence on how birth order effects have evolved during China's period of rapid economic transformation. The policy implications of the thesis include easing regional mobility restrictions to improve employment opportunities, as well as increasing educational investment and reducing within-family disparities to support parental mental health and promote children's long-term development.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Acknowledgements

First and foremost, I would like to express my deepest appreciation to my principal supervisor, Dr. Nadezhda V. Baryshnikova, for her continuous support throughout my PhD journey. Her insightful ideas and thoughtful feedback have been a tremendous help to my thesis. More importantly, her patience, respect, encouragement and optimism helped me overcome the most challenging moments. I am grateful to my co-supervisor, Associate Professor Stephane Mahuteau, for his supervision and support. This thesis would not have been possible without their help.

I would also like to thank the postgraduate coordinators, Associate Professor Stephanie McWhinnie, Dr. Raul Barreto, and Associate Professor Alexandra Peralta, for their support. I am grateful to Dr. Emiliano Carlevaro, Professor Firmin Tchatoka, Dr. Qazi Haque, Dr. Akwasi Ampofo, Dr. Florian Ploeckl, and Dr. Jacob Wong for their valuable help and insightful comments on my thesis. I thank all the seminar speakers for their engaging discussions and thoughtful comments. I also thank all the participants of all the conferences and workshops where I presented my work.

Finally, I would like to thank my late grandfather, who supported me throughout my studies since primary school. I am grateful to my family and lovely girlfriend for their companionship, encouragement and support.

Dedication

To my beloved parents, Shi Yiming and Shan Bing.

For their love and support, and for inspiring me to pursue a PhD.

Introduction

The trajectory of human capital development in the developing world differs significantly from that in the developed world. Relatively lower levels of fairness and social justice make it harder for disadvantaged groups to accumulate human capital in most developing countries. One typical example reflecting social inequality is formal employment, which offers higher salaries and various social security benefits to its employees. Over the past three decades, the proportion of the labour force working in the formal sector has remained low at approximately 30% in developing economies, whereas this proportion is high at 86.5% for advanced economies (Elgin et al., 2021). In Indonesia, this trend is especially evident, where only about 20% of the labour force is employed in the formal sector (Elgin et al., 2021). This raises an important question: does internal migration help Indonesians access formal employment and build stronger pathways for human capital development? This question is particularly relevant for Indonesia and many other developing countries, where internal migration is subject to fewer regulatory barriers (Sukamdi & Mujahid, 2015).

Beyond labour market outcomes, human capital development also involves intergenerational linkages. In many Asian societies, children's education affects not only their own well-being but also their parents' mental health—an important aspect of human capital. Investigating the relationship between education and mental health leads to a better understanding of human capital development. In the context of the developing world, especially in East and Southeast Asian countries, it is more interesting to extend this relationship to an intergenerational perspective. This is because of two unique features of these countries. Firstly, the tighter family kinships in these

countries enhance the interdependent relationship between children and parents (Xu & Xia, 2014). Secondly, due to Confucian traditions that emphasise filial responsibility, Chinese parents may derive emotional satisfaction or social standing from their children's educational success (C. K. C. Lee & Morrish, 2012). In this case, better-educated children may have a greater impact on their parents' mental health in China or Vietnam than in the US or Australia.

Birth order is a key factor affecting intergenerational relationships. On the one hand, exploring birth order effects provides insight into disparities in parental investment among siblings. Such within-family inequalities tend to hinder children's human capital development more than inequalities between families. The literature on birth order effects suggests conflicting outcomes, with positive effects in developing countries (Ejrnæs & Pörtner, 2004) and negative effects in developed countries (Black et al., 2005). The conflicting results offer a great opportunity to investigate birth order effects in fast-growing economies such as China, as it reveals how within-family inequality changes during the rapid economic transition. On the other hand, children's perceptions of unequal parental investment during childhood could potentially influence the level of old-age support they offer to their parents now. Such family support is critical for elders in most developing countries including China, as it is their main source of income due to the underdeveloped social security system (Rao & Zhang, 2024).

This thesis investigates the human capital outcomes of internal migration and the intergenerational dynamics of education, mental health, and family support in the developing world, with a particular focus on Indonesia and China. We specifically focus on using econometric tools, such as the fixed-effects and instrumental variable approaches, to conduct causal inferences. Chapters 1 to 3 are written as stand-alone papers.

In the first chapter, we investigate the impact of internal migration on formal employment in Indonesia. Two major threats to the identification are the selection bias and reverse causality. The selection bias mainly arises from systematic differences be-

tween migrants and nonmigrants in characteristics such as risk aversion and internal locus of control (Piracha & Vadean, 2010; Caliendo et al., 2019). These unobserved personal traits are typically correlated with both the migration decision and formal employment. To address this issue, we use a longitudinal household survey to apply the panel fixed-effects estimator. In this case, all unobserved personal traits are captured by the individual fixed effects. Furthermore, while migration affects the labour market outcome, the expected chance of having a formal job also affects the migration decision. To address this reverse causality issue, we employ rainfall as an instrumental variable for internal migration. The main idea of this approach is that rainfall only affects migration decisions but does not directly affect chances of obtaining formal jobs, after controlling for multiple labour market indicators, such as nighttime lights and agricultural GDP.

While the literature mainly focuses on exploring the macro-level relationship between migration and employment (Munshi, 2003; Kleemans & Magruder, 2018; Bharati et al., 2024), the micro-level perspective has received less attention. To the best of our knowledge, chapter one is the first paper investigating the individual-level impact of internal migration on formal employment. We find internal migration increases the chance of having a formal job by 16.5 percentage points. This impact varies across genders, with an increase of around 25 percentage points for male migrants and 7 percentage points for females.

In the second chapter, we explore the impact of children's education on parents' mental health in China. We use a quantile approach to further investigate whether this impact is heterogeneous across the distribution of parental mental health. Two main threats to the identification are reverse causality (mentally healthier parents being better equipped to invest in their children's education) and the omitted variable bias (unobserved genetic factors that affect both parents and children). To address the endogeneity issue, we use children's exposure to the compulsory schooling law (CSL) as an instrumental variable for their education. Our identification is based on the assumption that exposure to CSL exogenously increases educational attainment for

the later-born cohort but has no effect on their parents' mental health.

One key motivation for using a quantile approach is that, on average, we find an insignificant effect of children's education on parents' mental health, which aligns with findings in the literature (Ma, 2019; Cornelissen & Dang, 2022). Given the insignificant effect on average, it is more meaningful to explore whether the effect varies across different parts of the parental mental health distribution. To our understanding, chapter two is the first paper using a quantile approach to investigate the effect of children's education on parental mental health. We find children's education has no effect on parents with mild to moderate mental health issues (below the 0.7 quantile) but has a positive and significant impact on those with severe issues (at the 0.7 quantile and above). In our comparative analysis, we show that parents' mental well-being benefits more from the education of their least educated children when they are not depressed, but shifts to benefit more from their most educated children when they are depressed. Moreover, the education of first-born children has a greater influence on parental mental health compared to that of second-born children.

In the third chapter, I examine how birth order and gender composition affect years of education, and how birth order affects the old-age support siblings provide to their parents, in the context of China. By using the family fixed effects, I am able to evaluate within-family variations in birth order. I find that later-born children tend to outperform their earlier-born siblings in education for older generations of Chinese, while the later-born siblings are less educated for younger generations of Chinese. This result indicates that birth order effects on education in China have evolved over its rapid economic transition in recent decades. I also find brothers have consistently attained higher levels of education than sisters, but this gap has narrowed over time.

While the literature primarily examines the impact of gender composition and the number of children on old-age support (Oliveira, 2016; Ho, 2019), chapter three is the first paper to examine the impact of birth order on old-age support. I find insignificant differences in old-age support levels both among siblings and between brothers and sisters. This implies that children who are expected to provide more old-age support

to their parents in return for the favours they received during childhood do not, in fact, offer more support now. The altruistic motive emerges as a key reason why siblings, as well as brothers and sisters, provide equal levels of old-age support to their parents. Furthermore, I find no evidence that in-kind support and monetary support function as substitutes or complements.

The roadmap for this thesis is as follows. Chapter 1 estimates the impact of internal migration on formal employment using longitudinal data from Indonesia. Chapter 2 uses variation in children's exposure to China's compulsory schooling law to study the effect of education on parents' mental health. Chapter 3 investigates the relationship between birth order, education, and old-age support within Chinese families. Together, these studies contribute to a broader understanding of how mobility, education, and family dynamics shape human capital and intergenerational outcomes in developing countries.

Chapter 1

Internal migration and formal employment: evidence from Indonesia

Ruyi Shi[†], Nadezhda V. Baryshnikova[† ‡], Stephane Mahuteau[†]

[†] School of Economics and Public Policy, The University of Adelaide.

[‡] The Graduate School of Business and Commerce, Keio University.

Statement of Authorship

Title of Paper	Internal migration and formal employment: evidence from Indonesia
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☒ Unpublished and Unsubmitted work written in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Ruyi Shi
Contribution to the Paper	Conceptualization, Data curation, Methodology, Software, Formal analysis and investigation, Visualization, Writing - original draft preparation, Writing - review and editing
Overall percentage (%)	70%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 12/03/2025

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Nadezhda V. Baryshnikova
Contribution to the Paper	Conceptualization, Methodology, Writing - review and editing, Supervision
Signature	Date 12/3/25

Name of Co-Author	Stephane Mahuteau
Contribution to the Paper	C are, Supervision
Signature	Date 12/03/25

Please cut and paste additional co-author p

Abstract

This paper examines the causal impact of internal migration on formal employment in Indonesia using longitudinal household survey data from the Indonesia Family Life Survey (IFLS). We use rainfall as an instrumental variable for migration to address the endogeneity caused by reverse causality. Our IV estimate suggests that internal migration increases the chance of having a formal job by approximately 16 percentage points. This impact is heterogeneous across gender and levels of education. The magnitude of this impact rises to around 25 percentage points for males but drops to only 7 percentage points for females. We also find migration only has a significant impact on individuals with senior high school education or higher, with the estimated impact being roughly 20 percentage points.

keywords: Internal migration, Formal employment, Rainfall, Instrumental variable, Indonesia

JEL Classification: O15, O17, R23

1.1 Introduction

Formal employment is favored by people for the higher salary and the presence of social security benefits, such as minimum wage and pension (Katungi et al., 2006). This preference may become stronger when the welfare gained from formal jobs significantly outweighs that of informal ones. This is particularly the case for Indonesia, where nearly 80% of the labour force worked in the informal sector, yet contributed only around 20% to the country's GDP in 2018 (Elgin et al., 2021). This indicates a pronounced welfare disadvantage associated with informal jobs in Indonesia when compared to formal jobs. This disadvantage is further amplified when compared to other developing countries. On the one hand, for countries with the same 20% GDP share of informal sector such as Chile, Argentina, and Mongolia, their respective informal sector proportions are only 64%, 44% and 50%, which are lower than Indonesia's 80%. On the other hand, for countries with the same 80% informal sector proportion such as Bolivia and Nepal, their respective GDP shares of informal sector are 63% and 35%, which are higher than Indoneisa's 20%. Consequently, it can be inferred that Indonesians should have a stronger desire towards seeking formal employment.

Internal migration is one of the most popular ways for Indonesians to improve their employment opportunities. Early in the 19th century, around 11.5% of the country's population migrated between the Dutch-established plantations and mining plants for better job opportunities (Van Lottum & Marks, 2012). Since independence, the Indonesian government has implemented the transmigration project, which improves job opportunities for more than 1.3 million landless households by relocating them from overpopulated islands to outer islands (Hardjono, 1988). In recent decades, with less stringent regulations on internal migration, Indonesia's migrant population has remained high, reaching approximately 10 million in 2010, which accounted for 4.1% of the country's population (Sukamdi & Mujahid, 2015). Given the high tendency of Indonesians to improve their employment opportunities through migration and the substantial advantages of the formal sector over the informal sector in Indonesia, it is interesting to study whether migration actually improves formal employment

opportunities.

This paper aims to examine the individual-level causal impact of internal migration on formal employment in Indonesia. By using 17 years of panel data from the Indonesia Family Life Survey (IFLS2 to IFLS5), we are able to apply the panel fixed-effects estimator, which addresses the selection bias. This bias mainly arises from systematic differences between migrants and nonmigrants in characteristics such as risk aversion, ability endowments, and internal locus of control[1]. These unobserved personal traits are typically correlated with both the migration decision and formal employment. Since risk aversion, ability endowments, and internal locus of control generally remain stable over time, they are captured by the individual fixed effects. Furthermore, while migration affects the labour market outcome, the expected chance of having a formal job also affects the migration decision. To address this reverse causality issue, we employ rainfall, a popular IV for migration in agriculture-based countries, as an instrumental variable for internal migration. This approach is complemented by controls for a series of labour market indicators, such as nighttime lights and agricultural GDP. Both rainfall and nighttime lights data are acquired from satellite imagery with high-resolution rates of 5 and 1.1 kilometres, respectively.

Our paper contributes to the literature in the following ways. Firstly, among the limited literature on the impact of migration on formal employment, the majority of studies examine the macro-level impact. For example, Kleemans & Magruder (2018) and Bharati et al. (2024) explore the effect of changes in migration stock on the city-level average formal employment rate in Indonesia[2]. Munshi (2003) focuses on the impact of past migration stock on the average employment rate in the Mexico-US context. While macro-level analysis captures the overall impact of migration on em-

[1]Locus of control refers to a broad belief or expectation about the connection between an individual's actions and the outcomes they experience (Rotter, 1966). According to Caliendo et al. (2019), individuals with an internal locus of control believe that their achievements result from their own efforts, whereas those with an external locus of control attribute outcomes to outside forces such as luck or chance.

[2]Although both papers use household survey data, their analyses are conducted at the city level and focus on how the migration stock influences the employment rate, rather than analysing effects at the individual level.

ployment, micro-level evidence provides a clearer and more precise understanding by focusing on the specific experiences of individual migrants. The precision of micro-level analysis arises from its ability to control for individual characteristics such as age, family size, education, and poverty indices, while also accounting for unobserved personal traits through the use of a fixed-effects estimator. In contrast, the generalisation of outcomes across the entire migrant group in macro-level analysis reduces the precision of the estimation. To the best of our knowledge, this is the first paper to provide individual-level evidence on the relationship between migration and formal employment opportunities.

Secondly, papers that investigate migration at the individual level focus on whether migration increases household income or expenditure (Beegle et al., 2011; Bryan et al., 2014; Lagakos et al., 2020; Hamory et al., 2021), rather than on whether migration improves formal employment opportunities. While income growth is one important benefit of migration, securing a formal job not only ensures a higher income in general but also provides social protection (Nguyen et al., 2015). This paper fills the literature gap regarding the impact of migration on formal employment.

The primary finding of the paper is that internal migration increases the likelihood of migrants obtaining a formal job by 16.5 percentage points. A second finding is that the impact of migration varies across genders, with an increase of around 25 percentage points for male migrants and 7 percentage points for females.

The remainder of this paper is structured as follows. The next section introduces the background of internal migration and formal employment in Indonesia. Section 1.4 describes the data source and variable definitions. Section 1.5 discusses the identification and estimation strategies. Section 1.6 presents our empirical results, followed by the robustness check in Section 1.7. We conclude this paper in section 1.8.

1.2 Background

1.2.1 Internal migration in Indonesia

Indonesia has a long history of internal migration. Tracing back to the Dutch colonial period of the 1930s, internal migrants accounted for 11.5% of the country's indigenous population due to the high labour demand of the Dutch-established plantations and mining operations in various parts of the archipelago (Van Lottum & Marks, 2012). Since gaining independence in 1949, the Indonesian government introduced the transmigration policy to improve the economic situation for low-income farmers by relocating them from overpopulated areas like Java and Bali to outer islands such as Kalimantan, Maluku, Nusa Tenggara, Papua, Sumatra, and Sulawesi (Bazzi et al., 2016). After migrating to previously unsettled areas, these farmers could gain access to houses and lands that provide enough food, helping to avoid hunger (Kebschull, 1986). During the peak period of the program, 535,474 families were resettled from 1979 to 1983 and 750,000 families were targeted to move between 1984 to 1988 (Hardjono, 1988). A total of $6.6 billion was spent during this period, which works out to roughly $3,330 per person (Bazzi et al., 2016).

Although the transmigration project aimed to alleviate poverty among landless farmers, it has been criticized for the environmental damage and the ignorance of customs and property rights of local residents. This policy also exacerbated ethnic conflicts due to the resentment among natives who felt that migrants often received privileged treatment (Tirtosudarmo, 1997). Combined with the oil price collapse in 1986, government-led migration was gradually replaced by independent migration in the 1990s (Bharati et al., 2024). The number of internal migrants in Indonesia was 14.8 million in 1990 and has remained high, with 10.7 million in 2000 and 9.8 million in 2010 (S. Alatas, 1993; Sukamdi & Mujahid, 2015). Internal migrants also made up around 5 to 6% of the country's population in the last 30 years, of which half moved within the province. For those who migrate between provinces, more than 30% choose West Java or Jakarta as their top destinations (Sukamdi & Mujahid, 2015).

1.2.2 Formal and informal employment in Indonesia

Since the 1970s, Indonesia has experienced rapid economic growth led by structural transformation and urbanization. The fast-growing manufacturing and service sectors in urban areas have created a huge demand for formal employment. As a result, the percentage of workers with formal employment increased from nearly 35% in 1990 to 44% in 1997 (V. Alatas & Newhouse, 2010). However, this expansion was halted in 1997 due to the significant contraction in Indonesia's economy caused by the Asian financial crisis. The previous formal workers then started to enter the informal sector in the 2000s. In the recent decade, it is estimated that nearly 80% of the labour force has been working in the informal sector in Indonesia, however, the informal sector only contributed to around 20%[3] of the country's GDP (Elgin et al., 2021).

Given the substantial size of the informal sector in Indonesia, public opinions toward it are quite diverse. Supporters argue that the informal sector's low human capital requirements and fewer entry barriers provide a safety net for less-educated or self-employed individuals to make a living, especially during periods of economic recession (Loayza & Rigolini, 2011; Bharati et al., 2024). In contrast, skeptics cast their doubt from two perspectives. From a social standpoint, the lack of market and law regulations for the informal sector not only allows informal firms to pay less or no tax to the government, but also creates unfair competition for formal firms. This, in turn, limits the government's capacity to offer quality public goods (Levy, 2010), and harms economic growth by reducing the efficiency of formal firms (Rothenberg et al., 2016). From the perspective of an individual, the informal job provides less financial security such as insurance and employment protection, leaving informal workers in a vulnerable position, both in terms of their ability to endure adverse economic shocks and their bargaining power in the job market (Perry, 2007; D. Andrews et al., 2011).

[3]This number is the Multiple indicators multiple causes model-based (MIMIC) estimates of informal output (% of official GDP). The percentage estimated by the dynamic general equilibrium model is between 15 to 20%.

1.3 Literature review

There is a large literature studying the macro-level impact of migration on formal employment. Kleemans & Magruder (2018) using the Indonesia Family Life Survey to examine the impact of growth in migration stock on local formal employment for the whole population, including both natives and migrants. They find that an increase in the migration stock reduces the average formal employment rate in the destination by 33 percentage points. The authors argue that this drop is driven by the increased labour supply due to the immigration shock, which in turn crowded out the previous formal workers. Bharati et al. (2024) extend the work of Kleemans & Magruder (2018) by splitting the sample into migrants and natives. Using the country's nationally representative survey data, they find that a one percent rise in the population share of new migrants increases the formal employment rate for migrants by around 22 percentage points, whereas it reduces the formal employment rate for natives by 17 percentage points. This result implies that the drop in formal employment observed in Kleemans & Magruder (2018) for the overall population may be solely driven by the native workers.

In addition, Munshi (2003) find positive evidence of past migration stock on employment for current migrants in the context of the Mexico-US migration. The author finds that a one percent increase in the past migration stock leads to a more than 40% rise in the average employment rate for new migrants in their destination. In terms of identification, all three papers apply the instrumental variable approach to address possible endogeneity issues. Bharati et al. (2024) use historical migration patterns as the IV for current migration, while the instruments used in Kleemans & Magruder (2018) and Munshi (2003) are rainfall, which provides insights into our identification strategies later.

The migration literature at the individual level primarily examines how migration affects household income or consumption, and consistently finds that the impact is positive. Beegle et al. (2011) use thirteen-year panel data in Tanzania to find that migration increases household consumption by 36%. Lagakos et al. (2020) suggest

a similar result in household income in Indonesia, though with a smaller magnitude of around 11%. They also find the magnitude of this impact changes considerably after accounting for the individual fixed effect, with the magnitude dropping from 54% to 11%. This finding sheds light on the importance of considering individual fixed effects in our paper. Though migration can bring significant income growth to the household, a large number of migrants are still working informally and without social security benefits (Nguyen et al., 2015). However, studies exploring the impact of migration on formal employment are relatively rare, and this paper aims to fill this gap.

1.4 Data

The data used in this paper is drawn from the Indonesia Family Life Survey (IFLS), a longitudinal household survey conducted over five waves and including individuals above 15 years old. The 5 waves were surveyed respectively in 1993 (IFLS1), 1997 (IFLS2), 2000 (IFLS3), 2007 (IFLS4) and 2014 (IFLS5). Although IFLS tracks the same households for more than 20 years, the attrition rate between waves is remarkably low at 5% with 87% of households participating in all 5 waves (Strauss et al., 2016). Except for the low attrition rate, being highly representative is another good feature of the data. IFLS1 conducted the survey in 13 out of a total of 27 provinces, which together contain 83% of the country's population. To ensure a more representative survey, the enumeration areas were randomly selected using a sample frame from the Indonesian Bureau of Statistics (BPS) that represents the entire country (Strauss et al., 2016).

This paper uses the later 4 waves of IFLS because the questionnaire format used in IFLS1 differs from that used in the subsequent waves. We also restrict our sample to people who migrate before 30 years old[4]. One rationale behind this decision is that young people are the main force among migrants. In 2010, roughly 76% of migrants aged 15-64 in Indonesia were under the age of 35 (Sukamdi & Mujahid, 2015). Another

[4]The age at migration for the nonmigrant is the age in the survey year.

reason for restricting migration age under 30 is our belief that the impact of migration is more pronounced among younger people. With fewer professional networks and family obligations, they face lower opportunity costs compared to older individuals. This is further supported by the fact that when the sample is split by age groups, the significant results are driven solely by the under-30 cohort[5]. We also exclude individuals who are currently attending school because the dependent variable here is related to labour market outcomes. In total, we have 9,947 observations and 4,667 individuals.

Following Kleemans & Magruder (2018), we define the independent variable of primary interest, migration, as an indicator variable that takes the value of 1 if an individual moves across city borders and stays in the destination for more than six months. We define nonmigrants as those who do not move or move only across village borders. We also compute the migration distance manually by using ArcGIS software. For migrants, the mean and median for migration distance are 192 and 69 kilometres, respectively. The mean for those within the 10th to the 90th percentile drops to 126 kilometres, while the median remains consistent at roughly 70. The average migration distance for nonmigrants is not 0, as this group also includes individuals who move across village borders. For people who migrate multiple times between two waves, we ignore transit destinations and only consider locations in the last wave and the current wave. In the full sample, as illustrated in Figure 1.1, migrants account for nearly 17% of our sample. Figure 1.2 reveals that urban-to-urban migration is the most prevalent type, making up 42% of all migrations in our sample, followed by rural-to-urban migration, which accounts for 28.2%. This result also reflects the urbanization process of Indonesia in the recent few decades.

The outcome variable is binary and is defined by whether the individual holds a formal job. There is no universal definition in the literature for formal employment due to frequent changes in government regulations toward formal sectors (Rothenberg et al.,

[5]We find insignificant results for subsamples using other age cohorts. Results can be provided upon request.

2016; Harriss-White, 2010). Some studies define formal employment as wage workers (Kleemans & Magruder, 2018; Comola & De Mello, 2011); others consider both the occupation and the firm size of where the individual works (Mondragón-Vélez et al., 2010); while Pérez (2020) uses the pension criterion. Based on the above literature, we define a formal worker as an individual working in government or government-owned companies, or working in private companies with more than 10 people, or being self-employed and hiring more than 10 people, or having a pension after retirement[6]. As depicted in Figure 1.3, migrants are more likely to have a formal job than nonmigrants among all groups, especially for male workers. In the full sample, the proportion of having a formal job among migrants is more than 30%, compared to 20% of nonmigrants. This gap becomes wilder for the male group, reaching almost 20 percentage points.

We control for socioeconomic characteristics, including age, residency type, education, marital status, household size, and poverty indices. As shown in Table 1.1, migrants tend to be younger and are more likely to reside in urban areas with smaller household sizes compared to nonmigrants. Additionally, both migrants and their parents exhibit higher levels of education. In terms of poverty indices, migrants possess a higher level of total assets as well as a higher income. Following Jayawardana et al. (2021), we use household expenditure per capita as a proxy for income because the income variable in IFLS is poorly constructed. We also observe statistically significant mean differences in the above variables between migrants and nonmigrants. The detailed variable description is shown in Table A1 in the Appendix.

[6]We don't use pension as the only criteria because the second and the third wave didn't include the pension question.

1.5 Methodology

1.5.1 Model

Our model equation is as follows:

$$EMP_{it} = C + \beta MIG_{it} + X_{it}^{'}\eta + \alpha_i + \gamma_t + \epsilon_{it} \tag{1.1}$$

For each individual i at wave t ($t = 2, 3, 4, 5$), where EMP_{it} is the formal employment indicator at the destination[7], C is the constant term, MIG_{it} equals one if this person is a migrant and zero otherwise, X_{it} is a column vector of socioeconomic control variables including age in 5-year intervals, household size, farmland, urban, region of household location[8], education, parents' education, marital status, poverty status such as poor sanitation, total assets and income. α_i is the unobserved time-invariant individual heterogeneity, which could be fixed effects or random effects. γ_t is the wave dummy, while ϵ_{it} denotes the idiosyncratic errors.

1.5.2 Identification Issues

To identify the causal effect of migration on employment, we need to exclude the correlation between migration and unobserved variables that affect formal employment. Two issues that potentially threaten the model identification are self-selection and reverse causality.

The self-selection problem is mainly caused by the existence of systematic differences between migrants and nonmigrants, such as different levels of risk aversion (Piracha & Vadean, 2010; Bauernschuster et al., 2014; Mueller et al., 2020), internal locus of control (Caliendo et al., 2019), ability endowments (Beegle et al., 2011). All factors play a critical role in the migration decision. For instance, migrants are typically less risk-averse than nonmigrants, therefore, they tend to self-select for risky activities such as

[7]It takes the value of one if the individual has a formal job and zero otherwise. For nonmigrants, the destination and origin are the same.

[8]For the sake of parsimony, we decided not to include dummy variables for the individual provinces. Instead, we divide the 13 sampled provinces into three major regions, namely Sumatra, Java, and others, and add dummy variables for the three regions.

migration rather than being randomly selected. Similarly, individuals with an internal locus of control or stronger ability endowments tend to exhibit a greater willingness to migrate in search of better opportunities. Meanwhile, the above three factors are also associated with one's labour market performance such as obtaining a formal job. In this case, because the above three factors correlate with both the migration decision and the probability of having a formal job, it is challenging to disentangle the effect of migration from their influence on formal employment. To address this issue, we apply the panel fixed-effects estimator to eliminate the unobserved individual heterogeneity α_i in equation 1.1, which contains time-invariant variables related to self-selection (Munshi, 2003).

The second threat to model identification is reverse causality. While migration affects the labour market outcome, the expected chance of having a formal job also affects the migration decision. Specifically, if individuals have higher expectations of obtaining a formal job in the destination based on their previous working experiences or educational attainment, they are more likely to migrate. In this case, migration and formal employment are jointly determined and the model suffers from issues of reverse causality. To address this issue, we use rainfall in both origin and destination as the instrumental variables for migration, which has been widely used in the migration literature, such as Kleemans & Magruder (2018); Adams Jr & Cuecuecha (2010) for Indonesia and Munshi (2003) for Mexico. We use rainfall rather than other weather variables because it is the most important factor contributing to weather variation in Indonesia (Maccini & Yang, 2009).

Next, we will explain the validity of using rainfall as an IV for migration from the perspective of relevance, independence, and the exclusion restriction. Rainfall serves as a relevant instrument for migration when it has a significant impact on economic outcomes, which is the case in Indonesia in the sense that a large number of Indonesians rely on rainfed agriculture (Kleemans & Magruder, 2018). The World Bank (2010) data also indicate that the agricultural sector in Indonesia contributed 13.9% to the national GDP and provided jobs to approximately 45 million Indonesians (39% of the

labour force) in 2010, ranking Indonesia among the top one-third (around 60th) of approximately 190 countries, sorted from highest to lowest percentage. If the drought season lasts for a long period and leads to a shrinkage in the agriculture sector, the affected people are likely to migrate to other areas with more favourable weather conditions. Thus, precipitation is highly relevant to migration activities in Indonesia.

To satisfy the independence condition, rainfall must be uncorrelated with any other factors that might affect formal employment. However, since equation 1.1 includes only individual-level variables, macro-level factors such as the destination GDP and performances of the agriculture sector, which correlate with both destination employment and destination rainfall, would be omitted and included in the error term ϵ_{it}. Our rainfall instrument would be invalid in this case. To address this issue, we control for potentially omitted macro-level variables, such as destination GDP and agriculture at the city level. Since city-level GDP data was not available in the 1990s or early 2000s, we use the National Oceanic and Atmospheric Administration (NOAA) nighttime lights data as a proxy. Nighttime luminosity has been widely used as a proxy for economic development in the literature (Henderson et al., 2012; Castelló-Climent et al., 2018; Mamo et al., 2019; Kocornik-Mina et al., 2020). NOAA data measures the brightness of lights at night detected by satellites with the digital number ranging from 0 to 63, where higher numbers indicate greater brightness. We use the Global DMSP-OLS Nighttime Lights Time Series 1992 - 2013 (Version 4). For nighttime lights data in 2014, we use VIIRS Nighttime Light. In addition, the resolution of this data is as high as 1.1 kilometers, which enables us to specify the household location at the suburb level. Detailed data is presented in Figure A1 in the Appendix.

City-level agriculture data in Indonesia was not available either and we could only access province-level agriculture data[9], which is too broad and not informative in this study because most migrants in our sample move only across cities but remain in the same province. Due to this reason, adding the province-level agriculture variable is

[9]The measurement we use here is the logarithm of the provincial GDP in the agricultural sector at constant prices. The data source is the Indonesian Bureau of Statistics (BPS).

likely to cause the maximum likelihood estimator (MLE) to fail to converge, as there is little variation across individuals for this variable. Therefore, we choose not to add the agricultural variable to our main specification and only to use it as a robustness check[10].

To satisfy the exclusion restriction, rainfall should only affect the probability of obtaining formal employment in the destination through migration. However, due to Indonesia's heavy dependence on its agricultural sector, recent rainfall may directly affect destination employment. To address this issue, we use distant-past rainfall as the IVs and control for the recent-past rainfall[11]. We resort to the following example to illustrate the idea more clearly. The employment of a person in year t is emp_t, and the migration decision of this person in year t is affected by rainfall in the past 6 six years ($rain_t$ to $rain_{t-6}$) in both origin and destination. We define recent-past rainfall as $rain_t$ and $rain_{t-1}$, and distant-past rainfall as $rain_{t-2}$ to $rain_{t-6}$. In this case, through omitted macro-level variables other than nighttime lights or agriculture, recent-past rainfall is still likely to be correlated with emp_t, while this correlation for distant-past rainfall is expected to be much weaker[12]. Therefore, if we control for recent-past rainfall ($rain_t$ and $rain_{t-1}$), distant-past rainfall ($rain_{t-2}$ to $rain_{t-6}$) should not be correlated with emp_t anymore. In this case, our instrumental variables

[10]We acknowledge that our identification for the main specification might be problematic without controlling for agriculture variables. However, this may not be a big concern in this study because macro-level variables here are less important than individual-level variables. Since the outcome variable (formal employment) is at the individual level, after controlling for a series of personal-level variables, macro-level variables should have little effect on formal employment. In other words, even if we have access to the city-level agriculture variable and include it as a control, its coefficient is not significant after controlling for individual-level variables. This assumption is supported by the insignificant coefficients of the other macro-level variables such as city-level nighttime lights in Table 1.5

[11]The distant-recent-past rainfall strategy is introduced by Munshi (2003), while he uses $rain_{t-3}$ to $rain_{t-6}$ as the distant-past rainfall. To make our instruments more informative, we also add $rain_{t-2}$ to our specification.

[12]To provide statistical evidence, we first obtain the residual of regressing formal employment on macro-level variables. Then, we test the correlation between the residual and recent-past rainfall and distant-past rainfall, we find that the correlation in the former is twice as large as that in the latter.

distant-past rainfall are exogenous[13].

We acquire rainfall data from Climate Hazards Center InfraRed Precipitation with Station data (CHIRPS), which is a satellite dataset with a high-resolution rate of 5 by 5 kilometers[14](Funk et al., 2014). In this paper, rainfall is measured by the monthly average precipitation for the year in both the current and the last wave location, as the decision to migrate is based on the weather in both locations. To better capture the growing seasons in Indonesia, we measure each year's rainfall from July of the previous year to June of the current year in our main specification, which is in line with Kleemans & Magruder (2018). In addition, we also apply the z-score of precipitation in our main specification, which is calculated by subtracting the mean of precipitation in the last 35 years[15] and then dividing the corresponding standard deviation. The reason for using the z-score instead of levels of precipitation is that it captures deviations from the mean, which is more suitable for countries like Indonesia where there are significant differences in precipitation levels across multiple regions (Mueller et al., 2020).

1.5.3 Estimation strategies

To estimate equation 1.1, a general estimation strategy is the fixed-effects two-stage least squares (FE-2SLS) approach. We use rainfall as the instrumental variables of migration. The first-stage regression model is shown in equation 1.2.

$$MIG_{it} = C + Z_{it}^{'}\delta + X_{it}^{'}\eta + \alpha_i + \gamma_t + \epsilon_{it} \tag{1.2}$$

[13]To further test whether our IV approach satisfies the independence and exclusion restriction conditions, we compare regression results from two models: one that includes only rainfall instruments, and another that includes rainfall instruments along with recent-past rainfall, nighttime lights, and destination agriculture GDP. We find that the joint F-test for the rainfall IVs is significant in the former model but insignificant in the latter, suggesting that our IVs have no direct impact on formal employment after controlling for recent-past rainfall and macro-level variables.

[14]Detailed data is presented in Figure A2 in the Appendix.

[15]The Chirps data is available from 1981 to 2016, in total 35 years. Although 2016 is beyond our sample, we still include it to obtain a more general and accurate average and standard deviation of rainfall in the country.

Where MIG_{it} is the dummy variable of migration. $Z_{it}^{'}$ represents the 10 instrumental variables used for migration, which include distant-past rainfall ($rain_{t-2}$, $rain_{t-3}$, $rain_{t-4}$, $rain_{t-5}$, and $rain_{t-6}$) in both the origin and destination regions. $X_{it}^{'}$ contains all socioeconomic controls in equation 1.1, recent-past rainfall in the destination, and the city-level nighttime lights in the destination. α_i is the unobserved individual fixed effects, γ_t is the wave dummy, and ϵ_{it} denotes the idiosyncratic errors.

One big advantage of the FE-2SLS method is its ability to address the self-selection bias by eliminating the unobserved individual heterogeneity α_i. However, this method requires continuity in both the dependent variable and the endogenous variable, which is not the case in this study, as both formal employment and migration are binary. In light of the binary nature of this study, we further apply a binary choice model in an instrumental variable framework and compare the results with those obtained by the FE-2SLS method. We use the panel random-effects probit regression model with endogenous covariates[16], which is shown in equation 1.3 and 1.4.

$$EMP_{it}^{*} = C_1 + \beta MIG_{it} + X_{it}^{'}\eta_1 + \alpha_i + \gamma_t + \epsilon_{1,it} \tag{1.3}$$

$$MIG_{it}^{*} = C_2 + Z_{it}^{'}\delta + X_{it}^{'}\eta_2 + \alpha_i + \gamma_t + \epsilon_{2,it} \tag{1.4}$$

Where, EMP_{it}^{*} is a latent variable of EMP_{it}, and $EMP_{it} = 1$ if $EMP_{it}^{*} > 0$, $EMP_{it} = 0$ if $EMP_{it}^{*} < 0$. MIG_{it}^{*} is a latent variable of MIG_{it}, and $MIG_{it} = 1$ if $MIG_{it}^{*} > 0$, $MIG_{it} = 0$ if $MIG_{it}^{*} < 0$. $Z_{it}^{'}$ represents the 10 instrumental variables used for migration, which include distant-past rainfall ($rain_{t-2}$, $rain_{t-3}$, $rain_{t-4}$, $rain_{t-5}$, and $rain_{t-6}$) in both the origin and destination regions. $X_{it}^{'}$ is a vector of control variables, identical to $X_{it}^{'}$ in equation 1.2. Different from the linear model, α_i in the binary model is treated as a random effect. γ_t is the wave dummy. The error terms of equation 1.3 and 1.4 follow the bivariate normal distribution: $(\epsilon_{1,it}, \epsilon_{2,it}) \sim bvn(0, 0, 1, 1, \rho)$, where ρ is the correlation between the unobserved factors affecting both migration and formal employment. If ρ is significantly different from zero, MIG_{it} is endogenous.

[16] We use the STATA command of xteprobit.

The above model is estimated by the maximum likelihood method. Applying 2SLS to the above model yields inconsistent estimates due to the nonlinearity of equation 1.4. In this case, equation 1.4 no longer generates the predicted $\hat{MIG}_{it}$ as the dependent variable is a latent variable. Substituting MIG_{it} by its predicted value is no longer feasible and would lead to the forbidden regression, as discussed in Wooldridge (2010).

1.6 Empirical results

We will first present the result of the migration equation (equation 1.4) for the random-effects probit model and the first stage for the FE-2SLS, which shows the relevance of the IV for migration. Reduced rainfall in a region often compels residents to migrate to areas with more favourable precipitation levels; thus, lower rainfall in the origin is associated with a higher probability of migration. In other words, the sign of the rainfall shock is expected to be negative at the place of origin and positive at the destination. Results in Table 1.2 follow this logic, where origin rainfall has negative coefficients and destination rainfall has positive coefficients. We also notice that both models exhibit similar results. Since the weak IV test for nonlinear models is difficult to compute, we test the IV strength in the FE-2SLS (Han & Lee, 2019). We use the Olea and Pflueger effective F statistic (F^{eff})[17], which is proper for heteroscedastic errors (Olea & Pflueger, 2013; I. Andrews et al., 2018, 2019). The F^{eff} are greater than the 10% critical value of 14, indicating that rainfall based on the growing season serves as a strong instrument.

Next, we show estimated results of the effect of migration on formal employment. The values shown in Table 1.3 are the average marginal effects, the results of estimated coefficients are shown in Appendix Table A6. Column (1) shows the marginal effect estimates of equation 1.3 for the full sample. The correlation between the error terms

[17]Because the Olea & Pflueger F statistic is applicable only in the cross-sectional setting, we use the demeaned model instead of the panel fixed effects model to satisfy this requirement. To adjust for standard errors, we cluster them at the city level. By doing so, the demeaned model yields the same coefficients and standard errors as the standard panel fixed effects model.

(ρ) is significant, which means the endogeneity problem exists, and unobserved factors between migration and formal employment are negatively correlated[18]. Overall, migration significantly increases the chance of obtaining a formal job by 16.5 percentage points. In particular, if a non-migrant would have migrated, their probability of acquiring a formal job would increase by 16.5 percentage points. This increment is substantial given that there are only 20% nonmigrants in our sample having a formal job, as shown in Table 1.1. The corresponding FE-2SLS results are similar, although with a slightly smaller effect. Additionally, as shown in Appendix Table A7, the magnitude of the impact estimated by the fixed-effects model without IV is smaller than that of the model with IV. This result suggests that endogeneity such as omitted variables bias and reverse causality leads to an underestimation of the impact of migration on formal employment.

We also apply the correlated random effects estimator proposed by Wooldridge (2019). This estimator produces the same results as the fixed effects estimator but allows for the estimated coefficients for time-invariant variables, such as the variable *male* in our case. The result is shown in Table A8 in the Appendix. In addition, gender inequality is readily apparent in the job market, as males enjoy a 9% privilege in the chance of obtaining formal jobs compared to females. Given the considerable gender gap in the labour market, it is meaningful to investigate whether this gap persists in the marginal effect of migration on formal employment. One possible way is to interact migration with gender. However, as both migration and the interaction term are endogenous, it is complicated to construct the instrument for the interaction term. Instead, we split our sample by gender and conduct 2 separate regression analyses. The results are shown in columns (3) to (6).

Columns (3) and (4) show results for males, and columns (5) and (6) show results for females. For both gender groups, migration significantly increases their probability of having a formal job, while this increment for males is almost four times that of females.

[18]The negative correlation contradicts our risk aversion example. However, ρ measures the overall correlation between unobserved factors, while risk aversion is only one of them.

On average, if a male nonmigrant would have migrated, his chance of getting a formal job would increase by almost 25 percentage points, while for a female nonmigrant, this number is only 7 percentage points. This result suggests our assumption in the last paragraph, namely the gender gap persists in the marginal effect of migration on formal employment. We notice that for split samples, FE-2SLS coefficients for migration are weakly significant for males and not significant for females, and differences in coefficient magnitudes are also bigger between the random-effects probit regression model and FE-2SLS. Such discrepancies may indicate that our results for split samples are not as robust as the results for the full sample. One possible reason might be the much smaller sample sizes[19].

The impacts of socioeconomic controls are also worth noting, as they may provide insights for the potential robustness check. For example, educational attainment plays an important role in the labour market, and it provides more help for males, with the effect of education on formal employment for males being twice that for females. Interestingly, marriage has opposite effects on employment for different genders. While getting married mildly increases men's chances of having a formal job, it lowers women's chances dramatically by more than 14 percentage points. The above interesting observations are taken into consideration during our robustness check.

1.7 Robustness check

In this section, we check the robustness of our results in three ways. Firstly, we compare the results of our main IV specification with three other types of measurements: levels of rainfall based on growing seasons, rainfall z-score and levels of rainfall based on the calendar year. The results are shown in Table 1.4 columns (2) to (4), respectively. To make the comparison easier, we copy the previous results of our main specification from Table 1.3 to column (1) in Table 1.4. In general, all IV measurements suggest that migration significantly increases the probability to have a formal

[19]The reduced sample size also leads to a lower Olea & Pflueger F^{eff}, but the value is still greater than the 20% critical value and we may not consider it as a weak IV.

job, and the effect size is larger and more significant for males than for females. We notice that the differences in coefficient magnitudes among the four IV measurements are not substantial. The estimated coefficient for the calendar year measurement is slightly larger than that for the growing season measurement. However, the former measurement is likely to be a weak IV as its Olea & Pflueger F^{eff} is even lower than the 30% critical value. So, we only need to compare the discrepancy between the z-score and levels of rainfall based on the growing season. As the differences among all three panels between columns (1) and (2) are minimal, we are confident in the robustness of our results across various rainfall measurements

Secondly, we add destination agriculture controls that are discussed in section 1.5.2 to our model, and compare the results with our main specification. Results with agriculture controls are presented in Table 1.5 columns (2), (4) and (6). As shown in Table 1.5, adding agriculture controls has little impact on estimated coefficients of migration for all three samples. We also find that macro-level controls are not significant when we control for individual-level variables. This may be because formal employment as an individual-level variable, is affected more by individual-level factors than macro-level ones. This further justifies the validity of our identification that omitted macro-level variables are less likely to bias our estimation.

Thirdly, we examine whether our results are robust across individuals with different educational attainments. Before conducting the empirical analysis, we first check the percentage of individuals with formal employment in both high and low education groups within our sample. As Figure A3 illustrates, there is roughly a 10% difference in the formal employment rate between migrants and nonmigrants for the educated group, while this difference reduces to just 1% for the less educated group. This sharp contrast observed in our sample suggests that migration is expected to improve formal employment opportunities primarily for educated individuals. Our empirical result shown in Table 1.6 is in line with this summary statistics. For migrants with senior high school education or higher, migration significantly increases their chances of having a formal job by more than 20 percentage points, while this number for the

full sample is only around 16 percentage points as indicated in Table 1.3. In contrast, the impact of migration on formal employment is negligible for the less educated group. One possible explanation is that formal jobs typically require employers to have higher levels of education and skill compared to informal jobs.

This result also provides a hint regarding policy implications. For example, the Indonesian government could provide less educated young people with more free training courses, such as Program Kartu Prakerja, a nationwide training program launched in 2019 to enhance the competencies of less-educated young people with limited skills (Halimatussadiah et al., 2020). By enhancing their education and skill levels, individuals are more likely to seek formal employment opportunities through migration.

In addition, we find the pronounced gender gap in the labour market is even wilder among the educated group. Educated males are 13.4% more likely to find a formal job than their female counterparts, while this number is 9% in the full sample and is only 5.8% in the less educated group. Due to the limited sample size, we do not split the sample into males and females as above.

1.8 Conclusion

According to World Bank data, in 2018, Indonesia had nearly 80% of its labour force working informally while only contributing 20% to its GDP (Elgin et al., 2021). Given the low salaries and inadequate social welfare benefits associated with informal jobs, people in Indonesia may migrate to other cities to seek better opportunities for formal employment to improve their standard of living. In this paper, we find a significant and positive causal impact of internal migration on formal employment in Indonesia. To conduct the causal analysis, we solve the self-selection issue by using 17 years of panel data (IFLS) to account for the unobserved individual heterogeneity. We then deal with the reverse causality problem by using rainfall as the instrument for migration.

Our primary finding suggests that moving internally in Indonesia increases the chance for migrants to secure a formal job by approximately 16 percentage points. This

result indicates that internal migration serves as a potentially effective way for less-advantaged groups in Indonesia to improve their human capital development. We also find that the impact of migration is heterogeneous across genders and levels of education. The impact on males rises to around 25 percentage points, but it drops to only 7 percentage points for females. For migrants with senior high school education or higher, we find migration increases their formal employment by around 20 percentage points, while the effect is not significant for the less educated group. Besides, our results are robust across different rainfall measurements.

In addition to the findings presented, it is important to acknowledge the limitations of this study. Firstly, the random-effects probit model with endogenous covariates used in this study treats the unobserved individual heterogeneity as a random effect. As discussed earlier, unobserved factors such as risk aversion are likely to be correlated with migration. Thus, the random-effects estimator could potentially be biased, though it provides results similar to those of the FE-2SLS. Another shortcoming is the identification strategy of the rainfall instrument. Although we control for macro-level factors such as nighttime lights and province-level agriculture performance, some time-varying unobserved factors such as changing job networks or the unemployment rate may remain in the error term and lead to inconsistent estimation.

Figures for Chapter 1

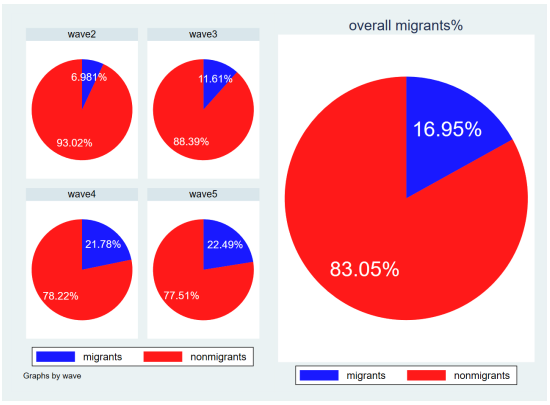


Figure 1.1: Migrants% by wave

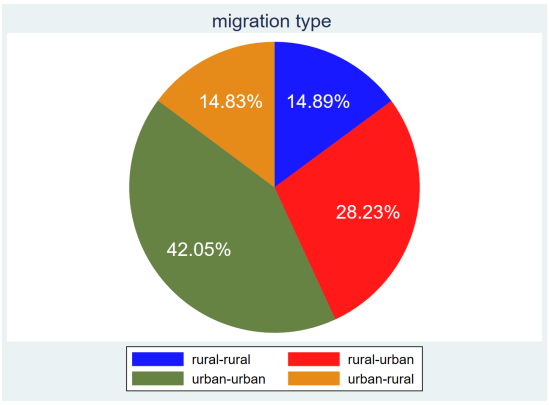


Figure 1.2: Migration type

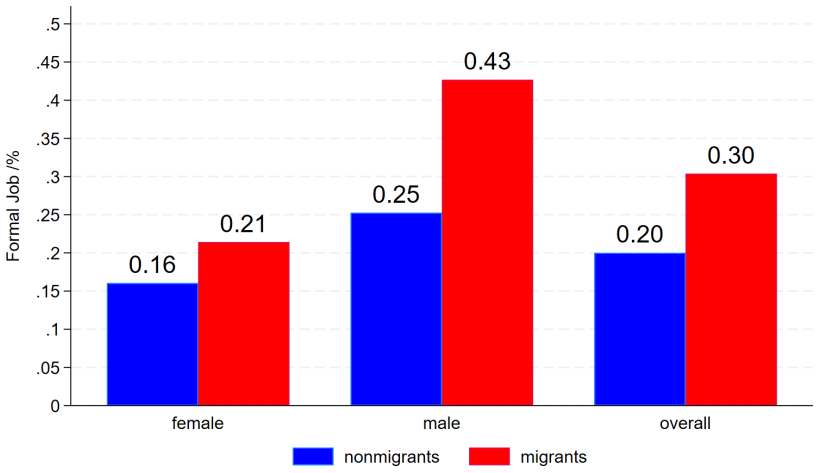


Figure 1.3: Percentage of having a formal job

Tables for Chapter 1

Table 1.1: Descriptive Statistics

	Full		Migrants		Nonmigrants		Mean
	mean	sd	mean	sd	mean	sd	Difference
formal job	0.22	0.41	0.30	0.46	0.20	0.40	0.10***
migration	0.17	0.38	1.00	0.00	0.00	0.00	1.00
migration distance	187.04	287.44	191.77	292.09	0.58	7.91	191.18***
male	0.43	0.49	0.42	0.49	0.43	0.50	-0.01
age	24.14	4.68	25.95	4.82	23.78	4.56	2.17***
urban	0.54	0.50	0.70	0.46	0.51	0.50	0.20***
senior high school edu	0.42	0.49	0.59	0.49	0.39	0.49	0.20***
parents senior edu	0.16	0.36	0.24	0.43	0.14	0.35	0.10***
married	0.57	0.49	0.69	0.46	0.55	0.50	0.14***
HH size	6.00	3.15	4.24	2.85	6.35	3.09	-2.11***
farmland	0.28	0.45	0.14	0.34	0.31	0.46	-0.17***
poor sanitation	0.21	0.41	0.14	0.35	0.22	0.42	-0.08***
lntotassetpc	15.06	1.73	15.38	1.83	15.00	1.70	0.38***
lnexpc	12.18	1.20	12.97	1.16	12.02	1.15	0.95***
Obervations	9947		1686		8261		

notes: The mean difference represents the disparity in means between migrants and nonmigrants. Descriptive statistics by wave are shown in Appendix Table A2 to A5. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Table 1.2: Migrants' responses to rainfall shocks (growing season)

	xteprobit migration equation	FE-2SLS 1st stage
origin rain t-2	-0.063***	-0.100***
	(0.012)	(0.024)
origin rain t-3	-0.000	0.006
	(0.013)	(0.026)
origin rain t-4	-0.052***	-0.047**
	(0.012)	(0.022)
origin rain t-5	0.015	0.013
	(0.012)	(0.026)
origin rain t-6	-0.013	-0.036
	(0.012)	(0.025)
destination rain t-2	0.019	0.055**
	(0.013)	(0.025)
destination rain t-3	0.037***	0.039
	(0.013)	(0.026)
destination rain t-4	0.024**	0.025
	(0.012)	(0.022)
destination rain t-5	0.008	0.013
	(0.012)	(0.026)
destination rain t-6	0.010	0.041*
	(0.012)	(0.025)
F statistic of joint significance	-	22.30
Olea & Pflueger F^{eff}	-	18.19
unobserved individual heterogeneity	RE	FE
wave & regional dummies	✓	✓
socioeconomic controls	✓	✓
observations	9,905	9,905

notes: The dependent variable is migration. The values shown are marginal effects. Destination nighttime lights and destination recent-past rainfall are both included in the regression. IVs are distant-past rainfall in both origin and destination. Rainfall is measured by the precipitation z-scores based on the growing season for both columns. t is the time of migration, which could be earlier than the time of the interview. The Olea & Pflueger F^{eff} are obtained from the FE-2SLS. Critical values for Olea-Pflueger weak IV test are 25 for 5%, 14 for 10%, 9 for 20%, and 7 for 30%. Standard errors are clustered at the city level. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Table 1.3: The effect of migration on formal employment (average marginal effects)

	Full sample		Male		Female	
	xteprobit ME	FE-2SLS	xteprobit ME	FE-2SLS	xteprobit ME	FE-2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
migration	0.165***	0.157**	0.247***	0.181*	0.070*	0.089
	(0.036)	(0.065)	(0.063)	(0.108)	(0.039)	(0.077)
male	0.090***					
	(0.011)					
married	-0.068***	-0.048***	0.038***	0.044*	-0.144***	-0.129***
	(0.010)	(0.015)	(0.014)	(0.023)	(0.012)	(0.021)
senior high school edu	0.085***	0.041	0.113***	0.053	0.062***	0.015
	(0.011)	(0.030)	(0.019)	(0.044)	(0.014)	(0.028)
corr(ϵ_1, ϵ_2)	-0.258***	-	-0.304***	-	-0.167*	-
	(0.064)		(0.084)		(0.091)	
Olea & Pflueger F^{eff}	-	18.189	-	9.235	-	11.631
P value for Hansen-J	-	0.732	-	0.107	-	0.531
UIH	RE	FE	RE	FE	RE	FE
wave dummies	✓	✓	✓	✓	✓	✓
regional dummies	✓	✓	✓	✓	✓	✓
socioeconomic controls	✓	✓	✓	✓	✓	✓
observations	9,905	9,905	4,241	4,241	5,664	5,664
number of individuals	4,646	4,646	2,000	2,000	2,667	2,667

notes: The dependent variable is formal employment. Destination nighttime lights and destination recent-past rainfall z-score are both included in the regression. IVs are distant-past rainfall in both origin and destination. Rainfall is measured by the precipitation z-scores based on the growing season for all 6 columns. Results for other IV measurements are shown in the robustness check part later. The Olea & Pflueger F^{eff} are obtained from the FE-2SLS. Critical values for Olea-Pflueger weak IV test are 25 for 5%, 14 for 10%, 9 for 20%, and 7 for 30%. UIH represents unobserved individual heterogeneity. Standard errors are clustered at the city level. *** means p < 0.01, ** means p < 0.05, * means p < 0.1.

Table 1.4: The effect of migration on formal employment using 4 different IV measurements

	Growing season		Calendar year	
	Rainfall z-score	levels of rainfall	Rainfall z-score	levels of rainfall
	(1)	(2)	(3)	(4)
Panel (a): full sample				
migration	0.165***	0.167***	0.200***	0.198***
	(0.036)	(0.036)	(0.047)	(0.049)
male	0.090***	0.090***	0.090***	0.090***
	(0.011)	(0.011)	(0.011)	(0.011)
Olea & Pflueger F^{eff}	18.189	13.518	5.142	3.396
Observations	9,905	9,905	9,905	9,905
Panel (b): male sample				
migration	0.247***	0.253***	0.261***	0.274***
	(0.063)	(0.068)	(0.080)	(0.077)
Olea & Pflueger F^{eff}	9.235	7.879	2.058	1.522
Observations	4,241	4,241	4,241	4,241
Panel (c): female sample				
migration	0.070*	0.068	0.106**	0.093
	(0.039)	(0.044)	(0.054)	(0.058)
Olea & Pflueger F^{eff}	11.631	8.975	5.256	3.437
Observations	5,664	5,664	5,664	5,664

notes: The dependent variable is formal employment. The results shown are the marginal effect of xteprobit. Individual fixed effects are not included. Control variables added in the regressions are the same as in previous tables, including wave & regional dummies, all socioeconomic controls, destination nighttime lights and destination recent-past rainfall. The Olea & Pflueger F^{eff} are obtained from the FE-2SLS. Critical values for Olea-Pflueger weak IV test are 25 for 5%, 14 for 10%, 9 for 20%, and 7 for 30%. Standard errors are clustered at the city level. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Table 1.5: The effect of migration on formal employment adding destination agriculture controls

	Full sample		Male		Female	
	(1)	(2)	(3)	(4)	(5)	(6)
migration	0.165***	0.161***	0.247***	0.238***	0.070*	0.072*
	(0.036)	(0.037)	(0.063)	(0.065)	(0.039)	(0.040)
destination nighttime lights	0.001	0.001	-0.001	-0.001	0.002	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
destination nighttime lights t-1	0.000	-0.000	0.001	0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
des agriculture		0.017		-0.104		0.043
		(0.122)		(0.153)		(0.144)
des agriculture t-1		-0.017		0.104		-0.044
		(0.121)		(0.153)		(0.144)
Olea & Pflueger F^{eff}	18.189	18.274	9.235	9.396	11.631	11.409
Observations	9,905	9,892	4,241	4,236	5,664	5,656

notes: The dependent variable is formal employment. The results shown are the marginal effect of xteprobit. The unobserved individual heterogeneity is considered a random effect in all columns. Control variables include wave & regional dummies, all socioeconomic controls, destination nighttime lights and destination recent-past rainfall. IVs are distant-past rainfall in both origin and destination. Rainfall is measured by the precipitation z-scores based on the growing season for all 6 columns. The Olea & Pflueger F^{eff} are obtained from the FE-2SLS. Critical values for Olea-Pflueger weak IV test are 25 for 5%, 14 for 10%, 9 for 20%, and 7 for 30%. Standard errors are clustered at the city level. * means p < 0.01, ** means p < 0.05, * means p < 0.1.

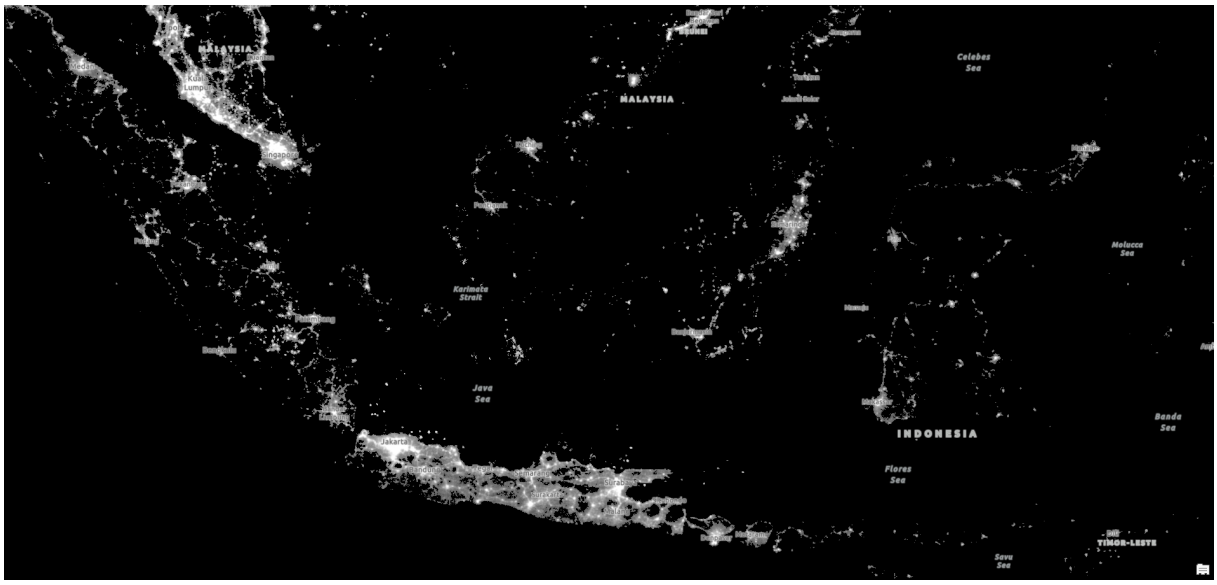
Table 1.6: The effect of migration on formal employment by education

	senior high school edu=1		senior high school edu=0	
	xteprobit ME	FE-2SLS	xteprobit ME	FE-2SLS
migration	0.204***	0.244**	0.058	0.072
	(0.050)	(0.110)	(0.053)	(0.091)
male	0.134***		0.058***	
	(0.017)		(0.015)	
unobserved individual heterogeneity	RE	FE	RE	FE
observations	3,962	3,962	5,493	5,493
number of individuals	1,919	1,919	2,544	2,544

notes: The dependent variable is formal employment. Control variables include wave & regional dummies, all socioeconomic controls, destination nighttime lights and destination recent-past rainfall. IVs are distant-past rainfall in both origin and destination. Rainfall is measured by the precipitation z-scores based on the growing season for all 4 columns. Standard errors are clustered at the city level. * means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

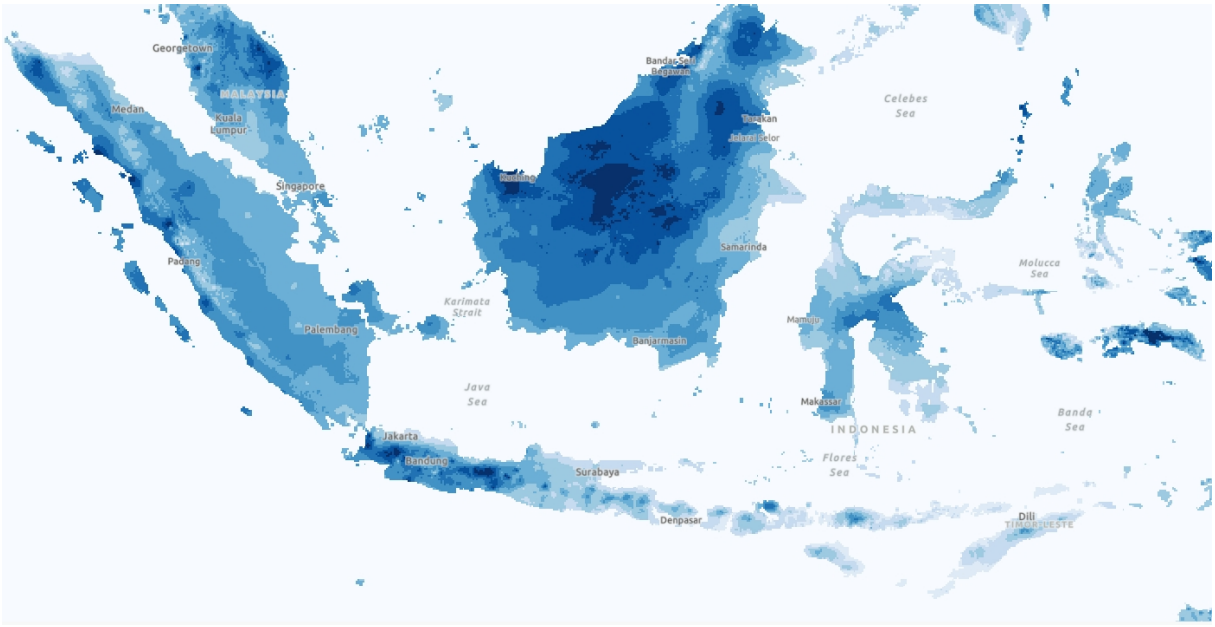
Appendix A

Figure A1: Nighttime lights data



Data source: National Oceanic and Atmospheric Administration (NOAA). This is one example showing nighttime lights in 2013. The data is processed using ArcGIS software.

Figure A2: Rainfall data



Data source: Climate Hazards Center InfraRed Precipitation with Station data (CHIRPS). This figure illustrates the average rainfall from 1981 to 2015, with darker blue indicating higher levels of precipitation. The data is processed using ArcGIS software.

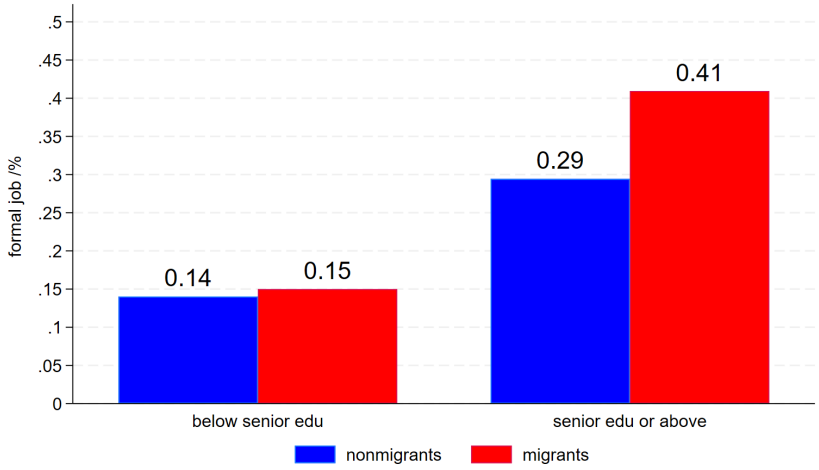


Figure A3: Percentage of having a formal job by education

Table A1: Variable Description

Variables	Description
formal job	=1 if has a formal job when taking the interview
migration	=1 if moves across city borders and stays > 6 months
migration	=0 if does not move or moves only across village borders
migration distance	distance between the last and the current location, in KMs
age	age in the year of the interview
urban	=1 if lives in the urban area when taking the interview
senior high school edu	=1 if has had senior high school education or higher
parents senior high school edu	=1 if mother or father has had senior high school education or higher
married	=1 if ever married
HH size	number of household members
farmland	=1 if own a farmland
poor sanitation	=1 if live in a poor sanitation condition
lntotassetpc	log of total household assets (in Indonesian Rupiah) per person
lnexpc	log of total household expenditure (in Indonesian Rupiah) per person

Table A2: Descriptive Statistics (wave 2)

	Full		Migrants		Nonmigrants		Mean
	mean	sd	mean	sd	mean	sd	Difference
formal job	0.19	0.39	0.22	0.42	0.19	0.39	0.04
migration	0.07	0.25	1.00	0.00	0.00	0.00	1.00
migration distance	14.71	90.42	188.80	287.32	1.65	14.15	187.15***
male	0.45	0.50	0.37	0.49	0.45	0.50	-0.08
age	21.30	3.72	22.31	3.80	21.23	3.70	1.08**
urban	0.47	0.50	0.66	0.48	0.45	0.50	0.20***
senior high school edu	0.29	0.45	0.47	0.50	0.28	0.45	0.19***
parents senior edu	0.11	0.31	0.17	0.37	0.10	0.30	0.06
married	0.37	0.48	0.63	0.49	0.35	0.48	0.28***
HH size	6.34	2.58	5.46	3.02	6.40	2.53	-0.94**
farmland	0.34	0.47	0.15	0.36	0.35	0.48	-0.20***
poor sanitation	0.22	0.42	0.15	0.36	0.23	0.42	8/1/00
lntotassetpc	13.84	1.62	14.40	1.65	13.80	1.61	0.60***
lnexpc	10.74	0.76	11.35	0.85	10.69	0.73	0.66***
Obervations	1547		108		1439		1547

notes: The mean difference represents the disparity in means between migrants and nonmigrants. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Table A3: Descriptive Statistics (wave 3)

	Full		Migrants		Nonmigrants		Mean
	mean	sd	mean	sd	mean	sd	Difference
formal job	0.17	0.38	0.21	0.41	0.16	0.37	0.05*
migration	0.12	0.32	1	0	0	0	1
migration distance	19.38	101.38	165.53	253.35	0.17	6.41	165.37***
male	0.44	0.5	0.36	0.48	0.45	0.5	-0.09**
age	22.28	4.03	22.13	3.88	22.3	4.05	-0.17
urban	0.49	0.5	0.66	0.47	0.47	0.5	0.19***
senior high school edu	0.36	0.48	0.54	0.5	0.34	0.47	0.20***
parents senior edu	0.14	0.34	0.22	0.41	0.13	0.33	0.09***
married	0.47	0.5	0.53	0.5	0.46	0.5	0.07*
HH size	6.16	2.92	4.16	2.75	6.42	2.84	-2.26***
farmland	0.34	0.47	0.16	0.37	0.36	0.48	-0.19***
poor sanitation	0.2	0.4	0.14	0.35	0.21	0.41	-0.07**
lntotassetpc	14.55	1.52	14.59	1.88	14.54	1.47	0.05
lnexpc	11.56	0.73	12.16	0.76	11.48	0.69	0.67***
Obervations	2626		305		2321		2626

notes: The mean difference represents the disparity in means between migrants and nonmigrants. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Table A4: Descriptive Statistics (wave 4)

	Full		Migrants		Nonmigrants		Mean
	mean	sd	mean	sd	mean	sd	Difference
formal job	0.23	0.42	0.34	0.47	0.2	0.4	0.14***
migration	0.22	0.41	1	0	0	0	1
migration distance	42.43	158.91	194.18	296.1	0.72	6.77	193.46***
male	0.42	0.49	0.42	0.49	0.42	0.49	0.01
age	24.22	4.5	25.56	4.27	23.84	4.49	1.72***
urban	0.56	0.5	0.72	0.45	0.51	0.5	0.20***
senior high school edu	0.47	0.5	0.6	0.49	0.43	0.5	0.17***
parents senior edu	0.17	0.38	0.23	0.42	0.16	0.36	0.07***
married	0.59	0.49	0.66	0.48	0.57	0.49	0.08***
HH size	5.78	3.26	3.94	2.76	6.29	3.2	-2.35***
farmland	0.24	0.43	0.11	0.31	0.27	0.45	-0.16***
poor sanitation	0.21	0.41	0.13	0.34	0.23	0.42	-0.10***
lntotassetpc	15.29	1.51	15.2	1.68	15.32	1.46	-0.12
lnexpc	12.59	0.93	13.1	1.09	12.45	0.83	0.64***
Obervations	3600		784		2816		3600

notes: The mean difference represents the disparity in means between migrants and nonmigrants. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Table A5: Descriptive Statistics (wave 5)

	Full		Migrants		Nonmigrants		Mean
	mean	sd	mean	sd	mean	sd	Difference
formal job	0.27	0.45	0.32	0.47	0.26	0.44	0.06**
migration	0.22	0.42	1	0	0	0	1
migration distance	45.52	168.5	205.21	308.92	0.03	1.12	205.19***
male	0.42	0.49	0.46	0.5	0.41	0.49	0.06*
age	28.3	3.14	29.75	3.39	27.88	2.94	1.88***
urban	0.61	0.49	0.71	0.45	0.58	0.49	0.13***
senior high school edu	0.51	0.5	0.64	0.48	0.47	0.5	0.17***
parents senior edu	0.19	0.39	0.28	0.45	0.17	0.37	0.11***
married	0.82	0.38	0.86	0.35	0.81	0.39	0.05**
HH size	5.91	3.55	4.51	2.91	6.32	3.62	-1.81***
farmland	0.24	0.43	0.16	0.37	0.27	0.44	-0.11***
poor sanitation	0.2	0.4	0.15	0.35	0.21	0.41	-0.06***
lntotassetpc	16.19	1.6	16.38	1.61	16.13	1.59	0.25**
lnexpc	13.28	0.82	13.63	0.87	13.17	0.78	0.45***
Obervations	2174		489		1685		2174

notes: The mean difference represents the disparity in means between migrants and nonmigrants. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Table A6: The effect of migration on formal employment (coefficients)

	Full sample	Male	Female
migration	0.739***	0.965***	0.397**
	(0.143)	(0.230)	(0.202)
male	0.449***		
	(0.059)		
married	-0.343***	0.164***	-0.881***
	(0.048)	(0.062)	(0.076)
senior high school edu	0.427***	0.486***	0.379***
	(0.059)	(0.083)	(0.090)
observations	9,905	4,241	5,664
number of individuals	4,646	1,988	2,658

notes: The dependent variable is formal employment. The results shown are the estimated coefficients of the random-effects probit model. This table shows the estimated coefficients of Table 1.3. Standard errors are clustered at the city level. *** means p < 0.01, ** means p < 0.05, * means p < 0.1.

Table A7: The effect of migration on formal employment (OLS-FE)

	Full sample	Male	Female
migration	0.038**	0.066**	0.010
	(0.018)	(0.029)	(0.022)
married	-0.040***	0.048**	-0.122***
	(0.015)	(0.023)	(0.021)
senior edu	0.042	0.053	0.015
	(0.029)	(0.043)	(0.027)
observations	9,905	4,241	5,664
number of individuals	4,646	1,988	2,658

notes: The dependent variable is formal employment. Standard errors are clustered at the city level. *** means p < 0.01, ** means p < 0.05, * means p < 0.1.

Table A8: The effect of migration on formal employment (FE-2SLS & CRE-2SLS)

	FE-2SLS	CRE-2SLS
migration	0.157**	0.157**
	(0.065)	(0.066)
male		0.086***
		(0.012)
married	-0.048***	-0.048***
	(0.015)	(0.015)
senior edu	0.041	0.041
	(0.030)	(0.030)
observations	9,905	9,905
number of individuals	4,667	4,667

notes: The dependent variable is formal employment. The FE-2SLS column is the same as column (2) in Table 1.3. Destination nighttime lights and destination recent-past rainfall z-score are both included in the regression. IVs are distant-past rainfall in both origin and destination. Rainfall is measured by the precipitation z-scores based on the growing season for both columns. Standard errors are clustered at the city level. CRE-2SLS represents for the two-stage correlated random effects least squares model. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Chapter 2

Adult children's education and parents' mental health in China: a quantile approach

Ruyi Shi[†], Nadezhda V. Baryshnikova[† ‡]

[†] School of Economics and Public Policy, The University of Adelaide.

[‡] The Graduate School of Business and Commerce, Keio University.

Statement of Authorship

Title of Paper	Adult children's education and parents' mental health in China: a quantile approach
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☒ Unpublished and Unsubmitted work written in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Ruyi Shi
Contribution to the Paper	Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing - original draft, Writing - review and editing
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 12/03/2025

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Nadezhda V. Baryshnikova
Contribution to the Paper	Conceptualization, Methodology, Supervision, Writing - review and editing
Signature	Date: 12/3/25

Name of Co-Author	
Contribution to the Paper	
Signature	Date:

Please cut and paste additional co-author panels here as required.

Abstract

This study examines the heterogeneous effect of adult children's education on elderly parents' mental health in China. To address endogeneity, we use children's exposure to the compulsory schooling law, which generated variation in years of schooling across birth cohorts and provinces, as an instrument for children's education. Our results from the instrumental variable quantile estimation reveal that the impact of children's education is heterogeneous across the distribution of parental mental health. Specifically, children's education has no effect on parents with mild to moderate mental health issues (below the 0.7 quantile) but has a positive and significant impact on those with severe issues (at the 0.7 quantile and above). Our split-sample analysis indicates that a son's education has a greater positive impact on parental mental health than a daughter's. We also find the positive effect is greater for parents living in rural regions than those in urban regions. In the comparative analysis, we show that when parents are not depressed, the schooling of the least educated children improves parents' mental well-being more. In contrast, when parents have moderate to severe mental health issues, the schooling of the most educated children has a greater impact. Additionally, the educational impact of first-born children is more substantial compared to that of second-born children.

keywords: Education, Mental health, Instrumental variable, Quantile regression

JEL Classification: I12, I21

2.1 Introduction

In recent decades, there has been a growing concern about mental health. Poor mental health can cause a series of problems to individuals, including low workplace productivity and adverse physical health effects such as diabetes, heart disease, and stroke. Severe mental disorders may even lead to fatal outcomes, such as death by suicide (Brådvik, 2018). In 2010, around 1 billion people worldwide suffered from a mental disorder with an estimated economic cost of 2.5 trillion US dollars, and this number is expected to increase to 6 trillion US dollars by 2030 (The Lancet Global Health, 2020). The mental health issue is particularly prevalent among middle-aged and elderly people. According to the Global Burden of Disease (GBD) 2019 estimates, worldwide depressive disorders prevalence for people aged 40 and above varies from 5.14% to 5.96%, which is much higher than the rate for younger groups ranging from 0.52% to 4.81% (Vos et al., 2020). This data also indicates a slight increasing trend in depressive disorders prevalence among individuals aged 55 to 70 over the past three decades, contrasting with the decreasing trend observed in younger groups.

One of the primary factors affecting the mental health of middle-aged and elderly people is their worry about children's well-being. When children are not doing well, the emotional connection makes parents feel a deep sense of responsibility for their children, resulting in more pronounced parental worry. In this circumstance, parents also tend to give certain financial support to children from their limited retirement budget (M. Guo et al., 2013). In contrast, children with a good job, stable incomes, or higher educational attainment can not only achieve independence from their parents, alleviating parental concerns, but also provide support when parents are facing difficulties by offering an upward transfer of material resources, such as social and financial support, and healthcare services (Torssander, 2013; Frase et al., 2023). Education stands out as one of the most important determinants of one's well-being. There is extensive literature indicating that education has a positive impact on one's earnings, socioeconomic status, health, and other aspects of well-being (Angrist & Krueger, 1991; Card, 2001; Duflo, 2001; Cutler & Lleras-Muney, 2006; Conti et al.,

2010; Mirowsky, 2017). As a result, the education of children emerges as a crucial factor in affecting the mental health of elderly parents.

This is particularly the case in China. First, the mental health issue for middle-aged and elderly people is more pronounced in China. The Global Burden of Disease (GBD) data shows that from 1990 to 2019, people aged 50 and above in China have experienced a more significant increase in the prevalence of depressive disorders than the global average. The prevalence percentage in Chinese elders aged 60 and above increased from approximately 5% to 6%, making it the second-highest increment among all 204 countries in the dataset. In addition, the pattern of rising prevalence of depressive disorders among the elderly and declining prevalence among younger age groups observed over the past 30 years also appears in China, which aligns with global trends mentioned earlier.

Second, the connection between children's education and parents' mental health is potentially stronger in China. On the one hand, Chinese parents greatly value children's education due to the influence of traditional culture, such as Confucianism (C. K. C. Lee & Morrish, 2012). The implementation of education policies such as the 1986 compulsory schooling law and the 1999 higher education expansion also reflects the importance of education in China. On the other hand, compared to Western countries, Chinese families often exhibit tighter family bonds and a deeper emphasis on family kinship (Xu & Xia, 2014). In turn, children's education as a major predictor of their well-being potentially has a more substantial impact on parental mental health in China.

In this study, we examine the causal effect of children's education on parents' mental health in China. We further explore the heterogeneity in the effect of children's education along the distribution of parental mental health by using the quantile approach. Our data is sourced from the China Family Panel Studies (CFPS), recognized as the first and most authoritative large-scale panel household survey in China (Y. Xie et al., 2017). We use the 8-item Center for Epidemiologic Studies Depression (CES-D) scale to measure parental mental health. Deriving a causal interpretation in this question

presents two main challenges: (1) reverse causality, as parents with better mental health conditions are more equipped to invest in their children's education; and (2) potential omitted variable bias due to unobserved genetic factors that affect both parents and children. To address the endogeneity problem and examine the causation, we follow the approach of Ma (2019) and Everding (2019) by employing natural experiments—specifically, the implementation of the compulsory schooling law (CSL) in China—as an instrumental variable for children's education.

Our paper makes a key methodological contribution to the literature on parental mental well-being. Unlike the studies of Ma (2019) and Everding (2019), which primarily focus on using linear models to explore the average effect of children's education on parental mental health, we employ a quantile regression approach to examine its distributional impact. This shift is important for three reasons. First, it allows us to explore whether the relationship between children's education and parental mental health differs across the mental health distribution. For instance, children's education may have a significant positive effect when parents experience depression, whereas its impact may be negligible or different under other circumstances. Focusing solely on the average effect overlooks these potential variations. Second, the literature consistently finds that, on average, children's education has no significant impact on parental mental health (Ma, 2019; Cornelissen & Dang, 2022), a result we also observe. However, an insignificant average effect does not imply the absence of meaningful effects at different points in the distribution. By adopting a quantile approach, we uncover potential heterogeneity that remains hidden in mean-based analyses, offering a more comprehensive understanding of the relationship. Third, focusing on the entire distribution rather than just the average helps policymakers identify which groups benefit the most, allowing for more targeted and effective interventions.

To estimate the quantile model, we use the Generalized Quantile Regression (GQR) approach proposed by Powell (2020). The main advantage of the GQR method lies in its ability to generate unconditional quantile treatment effects, making both computation and interpretation easier as including covariates no longer alters the quantile

function within this framework. In addition, by using numerical methods instead of demeaning, the GQR approach allows for the estimated coefficients for time-invariant variables while accounting for individual fixed effects. To our understanding, this is the first paper using a quantile approach in an instrumental variable framework to investigate the impact of children's education on parental mental health.

The second contribution of this study is our comparison of how education from different children affects parents' mental health. Unlike previous studies, which focus only on a single dimension of the impact of children's education—such as the average education of all children (C. Lee et al., 2017), the percentage of children with a bachelor's degree (Yahirun et al., 2017, 2020), the eldest child (Torssander, 2013), or the most educated child (Ma, 2019)—this study compares the impact of schooling between the least and most educated children, as well as between first-born and second-born children on parents' mental health. By shedding light on which children have a greater impact, such comparative analyses provide a more comprehensive understanding of the relationship between children's education and parents' mental health.

The main finding of this paper is that there is an insignificant average effect but a significant and positive distributional effect. Specifically, children's education has no effect on parents with mild to moderate mental health issues (below the 0.7 quantile) but has a positive and significant impact on those with severe issues (at the 0.7 quantile and above). Additionally, our results indicate that the impact varies according to children's gender, parental residency type, children's educational ranking within the household (highest vs. lowest), and children's birth order.

The remainder of this paper is organised as follows. The next section reviews related literature and introduces the background of the compulsory schooling law in China. Section 2.3 describes the data source and shows the summary statistics. Section 2.4 discusses the identification and estimation strategies, followed by the results in Section 2.5. Comparative analyses are presented in Section 2.6, followed by the mechanisms in Section 2.7. We conclude this paper in Section 2.8.

2.2 Background

2.2.1 Children's education and parental mental health

There is extensive literature in health and sociology that has found a positive correlation between children's education and parental mental health in both developed and developing countries. Yahirun et al. (2020) and Peng et al. (2019) suggest that children's college education reduces parental depressive symptoms in the US. Sabater & Graham (2016) propose that children's completion of upper secondary or tertiary education has a positive impact on mitigating parental depression in Europe. C. Lee et al. (2017) find that in comparison with the education of other family members, children's education has a stronger positive correlation with parental mental health in Taiwan. Pei et al. (2020) show that adult children's education alleviates the mental health of parents living in rural China. Mustafa et al. (2023) indicate that parents with well-educated children experienced considerably reduced chances of depression than their counterparts in India.

Unlike the health and sociology literature mentioned above, economic literature in this realm aims to explore causal relationships. Ma (2019) uses the compulsory schooling law (CSL) in China as an instrumental variable to investigate the causal impact of children's education on parents' physical and mental health. While the findings reveal a significant and positive impact on physical health, including improved cognitive functions, greater survival expectations, and enhanced lung function, the results indicate an insignificant average effect on mental health after endogeneity is taken into account. Cornelissen & Dang (2022) use a difference-in-differences approach to identify the multigenerational impacts of the primary school expansion program in Vietnam. Similar to Ma (2019), they also fail to find a significant causal effect of children's education on parental depression, and only observe significance in parents' self-reported health status.

While our paper is closely related to the work of Ma (2019), there are three main differences. First and foremost, we employ the quantile regression to investigate the

distributional effect, while Ma (2019) uses the two-stage least squares to examine the effect on average. As previously discussed, the distributional effect is superior to the average effect in this context as children's education is likely to have different impacts on mentally healthy parents compared to unhealthy ones[20]. Secondly, our paper conducts several comparative analyses, such as most educated versus least educated children and first-born versus second-born children, while Ma (2019) only explores the impact of the most educated children. Thirdly, this paper employs the CFPS data, while Ma (2019) uses the China Health and Retirement Longitudinal Study (CHARLS). With more comprehensive questions on family demographics, the CFPS is better suited for topics related to intergenerational interaction, whereas the CHARLS primarily emphasizes health-related questions.

To the best of our knowledge, the only study suggesting a significant causal result in this topic is a working paper by Everding (2019), which shares great similarities with Ma (2019). By using compulsory schooling reforms in Europe as an IV, the author argues that an extra year of education reduces the chance of parental depression by 5.2 percentage points. The author further splits the sample and suggests that this effect is 8.2 percentage points for the father sample and 4 percentage points for the mother sample, both of which are statistically significant.

2.2.2 The compulsory schooling law (CSL) in China

Since 1986, the Chinese government has implemented a nationwide compulsory schooling law to improve the basic education level of the entire population. The law stipulates that all children aged 6 and above must enroll in school and receive compulsory education until reaching the age of 16. Those aged 16 and above were not affected by the law. Because a typical child who is fully exposed to the law is required to attend six years of primary school and three years of junior high school, the law is also called

[20]Several mental health studies indicate that individuals with depression tend to be more vulnerable to the adverse effects of negative events such as natural disasters (Baryshnikova & Pham, 2019) and child labour (Jayawardana et al., 2023). Thus, a plausible assumption is that depressive parents, due to their greater vulnerability, could potentially benefit more from receiving support from their children.

the nine-year compulsory schooling law. Due to disparities in economic and resource status between provinces, the central government allowed provinces to have different law effective dates. The detailed law effective years for each province are listed in Table 2.1.

To ensure the successful implementation of the law, all students in principle should be offered a tuition fee waiver from schools and local authorities. However, due to the limited funding of local governments, students still had to pay part of the tuition fee until 2008 (Xiao et al., 2017). In this case, children of parents unwilling to pay the tuition fee tended to drop out of school early or complete only part of the nine-year compulsory schooling. Their parents in turn would receive verbal warnings and criticisms from the local supervisory staff (S. Xie & Mo, 2014). Though verbal punishments were not severe penalties, they still motivated parents who were not in extreme poverty to send their children to school. In addition, the Chinese government also enacted a series of policies related to child labour since 1986. For example, The *Temporary provisions on State Businesses Employing Staff* established a minimum age of 16 for working in the public sector in 1986. This requirement was further extended to the private sector in 1988 by the *Provisional Regulations of the People's Republic of China on Private Enterprises* (Cui et al., 2019). The enactment of legislation related to child labour has further enhanced the effectiveness of the CSL, as it is illegal for dropout children to work.

The compulsory schooling reform achieved an early stage completion in major cities by 1990, followed by cities of medium size and certain economically developed counties by the mid-1990s. In the early 2000s, around 85% of China's population was covered by the CSL (Cui et al., 2019). According to the 2004 China Statistical Yearbook, the enrollment rate for the junior high school increased roughly from 70% in 1985 to 98% in 2003.

2.3 Data

The data is acquired from CFPS, a 10-year household longitudinal survey conducted by the Peking University Institute of Social Science Survey. The first wave was launched in 2010, and subsequent 5 waves were collected every 2 years. The CFPS covers 25 out of all 33 administrative regions in China, and these surveyed provinces account for 95% of the country's population (Y. Xie et al., 2017). This data is also considered the first and most authoritative large-scale panel household survey in China (Y. Xie et al., 2017). Since the 2010 and 2014 survey ask different mental health questions, this paper uses a sample from the remaining four waves (2012, 2016, 2018, 2020) of the CFPS.

As this paper focuses on adult children who have completed their basic education, we restrict the age of the children to be above 22. While Ma (2019) proposes a similar age restriction of 20 and above, we prefer a threshold of 22 as most Chinese students graduate from university at this age or older. We also exclude children who hold a master's or PhD degree, as their completion age is normally older than 22. This group accounts for 2.1% of the sample. In addition, we exclude adult children whose educational attainment changes across waves, constituting approximately 9% of the sample size, to eliminate potential confounding effects of adult education or continuing education on parents' mental health. We also restrict parents' age to be above 40 to exclude those parents who have experienced early marriage and early childbearing. This is because early marriage or early childbearing may simultaneously affect parents' mental health and children's education, and then may potentially bias our estimation (Jayawardana, 2022; Wahhaj, 2022). After removing missing values, the full sample contains 18,244 parents and 10,788 parent-wave observations.

Our dependent variable is parents' mental health, measured by the 8-item Center for Epidemiologic Studies Depression (CES-D) scale. The 8-item CES-D scale[21] is a self-reported test comprising 8 questions designed to measure depression symptoms for the

[21]For simplicity reasons, we refer to the 8-item CES-D scale as CES-D 8 for the remainder of this paper.

general public. Adapted from the original 20-question CES-D developed by Radloff (1977), this shortened version maintains equivalent reliability to the full 20-question scale (Van de Velde et al., 2009; Missinne et al., 2014; Karim et al., 2015). The full list of questions is presented in Appendix Table B2. Answers for each question range from 0 to 3 based on the frequency with which the respondents experienced depression during the past week. The CES-D score is produced by summing the scores of all questions, where positively phrased questions are reversely coded (Hsieh & Qin, 2018). The range of the CES-D 8 score is from 0 (no depression) to 24 (deep depression). Higher scores indicate more severe depression symptoms.

The independent variable of main interest is the education of adult children, measured by their years of schooling. According to the CFPS user's manual, the years of schooling are as follows: 0 for illiterate or semi-illiterate individuals, 6 for primary school, 9 for middle school (junior high school), 12 for high school (senior high school), 15 for 2 or 3-year college, and 16 for a bachelor's degree. The main specification of this study considers only the most educated child when there is more than one child in the household, as those with the highest education are likely to have the greatest impact on parents' health (Zimmer et al., 2007). We then pick the eldest child if there are multiple most educated children in the household, which is in line with Ma (2019). In the comparative analyses, we examine the impact of education of the least and most educated children on parents' mental health. Additionally, we explore the effect of the education of first-born and second-born children on parents' mental health.

The empirical density function of parents' mental health score by their children's education shown in Figure 2.1 reflects a positive correlation between children's education and parents' mental health. Initially, when the mental health score is low (indicating no depression), the solid blue line is above the red dashed line, as the mental health score increases (indicating more depression), the latter gradually surpasses the former. Figure 2.1 also indicates that the distribution of mental health scores is right-skewed.

Following Ma (2019), Du et al. (2021) and Liang & Yu (2022), we define children's exposure to the compulsory schooling law (CSL) as a linear function of children's

birth year. Because the CSL has been enacted since 1986 and is only effective for children aged between 6 and 15, those aged above 15 in the law effective year were not exposed, those between 15 and 7 were partially exposed, and those below 7 were fully exposed. The detailed measurement of the level of exposure is shown in Equation 2.1. i represents the most educated child in the household and p represents the birth province of this child[22]. We add the province subscript p because the law effective year varies across provinces, as shown in Table 2.1. The number of observations for each exposure level is shown in Appendix Table B3.

$$exposure_{ip} = \begin{cases} = 0 & if \;\; age\ in\ law\ effective\ year_{ip} > 15 \\ = \frac{16 - age\ in\ law\ effective\ year_{ip}}{10} & if \;\; 7 \leq age\ in\ law\ effective\ year_{ip} \leq 15 \\ = 1 & if \;\; age\ in\ law\ effective\ year_{ip} < 7 \end{cases} \tag{2.1}$$

Intuitively, the impact of the CSL should be highly correlated with both the pre-law education level and the rural-urban divide. In districts where education levels were low or in rural regions, the effect of the law is expected to be greater compared to districts with relatively high education rates or urban areas. Thus, we construct the second IV as the interaction between exposure to the CSL and the provincial average years of schooling before law enforcement. The detailed pre-law education level is shown in Table 2.1.

As shown in the descriptive statistics table (Table 2.2), the mean for the mental health score is significantly lower than the median, which is consistent with the right-tailed distribution illustrated in Figure 2.1. Parents on average have 6 years of schooling while their highest-educated children have twice that amount of educational attainment. The average age of parents is roughly 55 and 92% of them are currently married. The marriage rate for their children is lower at around 60% due to a younger average

[22]For children whose birth province is missing (roughly 40% of the sample), we use their parents' current province as a proxy due to the substantial number of missing values for the children's current province. In addition, because CFPS lacks information on children's residency type at birth, it is proxied by their parents' current residency type. According to China's 2020 population census, the percentage of individuals moving between provinces was relatively low, at 8.84% of the total population.

age of 30 years old. Fathers make up almost half of the parents, while sons account for 65% among the most educated children. This dominant share of sons also reflects within-household education discrimination against girls in China (Lei et al., 2017). Approximately 80% of parents live in rural areas, and they have on average 2 children. In addition, more than 30% of parents smoked in the last month, and 18% of parents drank more than 3 times in the past month. 64% of parents have at least one type of pension, and 71% of them are currently employed. The CFPS also gathers information on parents' self-reported health conditions, rated from 1 (very unhealthy) to 5 (very healthy), as well as self-reported income levels and social status in comparison to others in the local area, ranging from 1 (very low) to 5 (very high). We also collect past mental health scores from CFPS 2010 and 2014, which are measured by Kessler psychological distress scale (K6) (Kessler et al., 2002). K6 is a different mental health measure compared to the CES-D scale but is also widely used in the literature (Cui et al., 2019). In our sample, the past mental health score for 2012 corresponds to the K6 score in CFPS 2010, while for 2016, 2018, and 2020, it is the average of the K6 scores in CFPS 2010 and 2014.

2.4 Empirical model

2.4.1 Identification strategy

Identifying the causal effect of children's education on parents' mental health is challenging due to a series of confounding factors. Two main issues widely discussed in the economic literature are reverse causality and omitted variable bias. First, the concern about reverse causality arises from the fact that parents' mental health also affects children's education. Recent health literature suggests that poor parental mental health has a detrimental impact on children's school enrolment in developing countries, as these children may need to enter the workforce early to cover their parents' health expenditures (Bratti & Mendola, 2014). Second, unobserved factors such as genes may simultaneously affect parents' mental health and their children's education (Ma, 2019). One possible solution is to use the twin fixed effects as in R. Guo et al. (2020),

however, the CFPS data we use lacks information on twins.

To address the endogeneity problems stated above, this paper uses the instrumental variable approach. Specifically, the years of schooling for the most educated children are instrumented by their exposure to China's compulsory schooling law (CSL), and its interaction with provincial average years of schooling before law enforcement. Such natural experiments have been extensively employed in investigating the causal effect of education in China (Ma, 2019; Liang & Dong, 2019; Cui et al., 2019; Du et al., 2021; Liang & Yu, 2022). To serve as a valid instrument, exposure to the CSL needs to satisfy the relevance condition and the exclusion restriction. The underlying mechanism of IV relevance is as follows. Conditional on all control variables, the later-born cohort is more educated only because they were exposed to the CSL. In other words, if the earlier-born cohort had been born later, they would have been exposed to the CSL and would be as educated as the later-born cohort. In addition, we use the interaction between exposure to the CSL and provincial average years of schooling before the law to capture the law intensity. This is because the impact of the CSL is expected to be stronger in less-educated provinces or rural areas (Ma, 2019). We use three methods to illustrate the relevance of exposure to the CSL and education.

Firstly, we regress children's education on the year dummies relative to CSL implementation. As shown in Figure 2.2, children with no exposure to the CSL are, on average, 0.5 to 1.5 years less educated than the first eligible cohort. In contrast, those who are partially exposed show a similar or slightly higher level of education, while those with full exposure are 1 to 2 years more educated. This gradual upward trend indicates a solid positive relationship between the CSL and education. Secondly, because the CSL requires a minimum of junior high school completion or a minimum of 9 years of schooling, it is supposed to have the largest impact on achieving at least 9 years of education. As shown in Figure 2.3, being fully exposed to CSL increases the chance of having at least 9 years of education by more than 20%, making it the most significant increase compared to other levels of education. The CSL also has a spillover effect on senior high school education, leading to a notable 10% increase in completion

rates. Thirdly, Appendix Table B3 shows that the average years of schooling increase with each level of exposure. Compared to the non-exposed cohort, the fully exposed group has an average of 3.37 more years of schooling, and the partially exposed cohort has nearly 1.2 years more. From a data-driven perspective, the relationship between exposure to the law and children's education is both positive and strong.

One crucial assumption underlying the exclusion restriction is that the instrumental variable should affect the dependent variable indirectly and solely through the endogenous variable, rather than affecting the dependent variable directly. Exposure to the education law clearly satisfies this condition as it has no direct impact on parents' mental health, however, concerns arise regarding its measurement. As mentioned earlier, the IV is defined by comparing children's birth year with the effective year of the education law. This implies that children exposed to the CSL were also potentially exposed to other social programs implemented during the same period that directly affect parents' mental health, such as the One Child Policy rolled out in 1978[23] (Ma, 2019). As a result, our IV measures exposure to a mix of valid policies, such as the CSL, which indirectly and solely affects mental health through education, and confounding policies, such as the One Child Policy, which directly impacts mental health. In this case, the exclusion restriction no longer holds.

To address this issue, we control for parents' mental health in the past, measured by parents' K6 score in 2010 or 2014 . This method validates our identification by eliminating the direct impact of confounding policies on the dependent variable. As confounding policies such as the One Child Policy were implemented considerably before 2010, after controlling for parents' mental health in 2010, they no longer have a direct impact on parents' mental health in 2012 or later (the dependent variable). Their impact is entirely captured by the coefficient of the parents' mental health in 2010. Thus, our instrumental variable becomes valid, as its confounding part has been

[23]From the perspective of children, having no siblings enables them to enjoy a greater share of educational resources and parental attention, which potentially leads to higher educational attainment. From the perspective of parents, raising only one child means fewer life burdens and thus greater mental health, or, worse mental health if there is an extreme preference for a son when their sole child is a girl.

excluded.

Another advantage of this approach is its ability to meet the exclusion restriction without adding children's birth cohort fixed effects, as used in Ma (2019) and Liang & Yu (2022). Due to the high correlation between our instrumental variable and children's birth year, the inclusion of birth cohort fixed effects not only results in perfect collinearity but also reduces the strength of the IV. One concern raised by not including children's birth cohort fixed effects is that the higher educational attainment is potentially caused by the time trend instead of the exposure to the CSL. Our paper addresses this issue by controlling for children's birth year trend, which effectively eliminates the impact of the time trend on education.

2.4.2 Estimation equation

To estimate the relationship between the educational attainment of the most educated children and the mental well-being of their parents, this study employs three approaches: the pooled two-stage least squares (2SLS), the correlated random effect two-stage least squares (CRE-2SLS), and the generalized quantile regression by Powell (2020).

In the context of pooled 2SLS, Equation 2.2 stands as the main equation, while Equation 2.3 corresponds to the first stage:

$$MH_{it} = C_2 + \beta\, ChildEdu_i + \eta\, X_{it} + \alpha_{2t} + u_{it} \tag{2.2}$$

$$ChildEdu_i = C_1 + \gamma_1\, expsoure_{ip} + \gamma_2\, expsoure_{ip} \times prelaw_{iph} + \delta\, X_{it} + \alpha_{1t} + e_{it} \tag{2.3}$$

where MH_{it} is the CES-D 8 score of the parent i in year t, $t = 2012, 2016, 2018, 2020$. $ChildEdu_i$ represents the years of schooling of the most educated child of parents i. As this study focuses on the education of adult children, $ChildEdu_i$ remains constant over the years[24]. Consequently, instead of using the fixed effects model, which

[24] As mentioned earlier, we exclude adult children whose educational attainment changes across waves, constituting approximately 9% of the sample size, to eliminate potential confounding effects of adult education or continuing education on parents' mental health. Besides, we only keep parents who have the same most educated child across waves.

would average out this time-invariant variable, we employ pooled 2SLS as our main specification despite having a panel setting[25].

X_{it} is a vector of control variables including parents' socioeconomic characteristics, such as own education, age and age squared, number of children, whether living in rural areas, smoking and drinking habits, past mental health scores (K6), whether having a pension, whether currently working, self-reported income level, social status, and health conditions. X_{it} also includes children's characteristics, such as their marital status, gender and birth year. We also control for the region dummies of parents' current location. To control for unobserved factors affecting parents and children across provinces and over time, province-specific birth cohort trends are included in X_{it}, which is in line with Stephens Jr & Yang (2014) and Ma (2019). α_t is the time fixed effects. By controlling for parents' age, time fixed effects and children's birth year, we are effectively accounting for the age of parents at the time of having children, as these variables are perfectly collinear. Not accounting for parents' age at the time of having children in this way could potentially invalidate our identification, as it is correlated with both the parents' mental health and their children's exposure to the CSL.

$expsoure_{ip}$ in Equation 2.3 is the first IV for children's education, which is measured by the level of exposure to the CSL for the most educated child of parents i born in province p. The detailed definition is provided in Equation 2.1. The second IV for children's education is $expsoure_{ip} \times prelaw_{iph}$, the interaction between children's exposure to the CSL and the average years of schooling before law enforcement in i's birth province p and residency type h ($h = urban, rural$). For example, if parent i lives in urban areas of Hebei Province, $prelaw_{iph}$ equals 8.252. If parent i lives in rural areas of Hebei Province, $prelaw_{iph}$ equals 9.921. The average years of schooling before the CSL for each province is presented in Table 2.1.

[25]In this case, we treat the same parent-children pair in 2016 as a different observation from the one in 2018. Because children's education is time-invariant, the variation of β comes entirely from parents' mental health.

One possible way to account for individual fixed effects while preserving the estimation of time-invariant variables is through the correlated random effects (CRE) estimator proposed by Wooldridge (2019). This estimator produces results for time-varying variables identical to those of the fixed effects estimator, while also enabling the estimation of coefficients for time-invariant variables, such as the key independent variable *ChildEdu* in our case. Extended from the Chamberlain–Mundlak approach introduced in Mundlak (1978) and Chamberlain (1982), the CRE approach also accommodates a two-stage least squares method (Joshi & Wooldridge, 2019; Auci & Pronti, 2023). The IV used in the CRE model is the same as the one used in the pooled 2SLS model. The estimated equation for the CRE approach is presented in Equation 2.4, where the unobserved individual heterogeneity is divided into two parts. The part that is correlated with X_{it} could be modelled as $\bar{X}_i$, while the part that is uncorrelated with X_{it} is denoted by r_i. In this way, the new error term $(r_i + v_{it})$ would be uncorrelated with X_{it} and thus can be estimated by the random effects estimator.

$$MH_{it} = C + \beta\, ChildEdu_i + \eta\, X_{it} + \gamma\, \bar{X}_i + \alpha_t + r_i + v_{it} \tag{2.4}$$

The two models above estimate the average effect of children's education on parents' mental health. However, they tend to overlook the potential variation of the impact across different parts of the parental mental health distribution. To uncover the potential heterogeneity in the impact of children's education, we apply the generalized quantile regression (GQR) approach proposed by Powell (2020)[26]. Following Baryshnikova et al. (2023); Jayawardana et al. (2023); Baryshnikova & Pham (2019), which use Powell's quantile methodology to assess the heterogeneous impact on mental health, the generalized quantile approach of Equation 2.2 is as follows, for all quantiles $\tau \in (0, 1)$:

$$Q_{MH_{it}}(\tau | ChildEdu_i, X_{it}) = C(\tau) + \beta(\tau)\, ChildEdu_i + \eta(\tau)\, X_{it} + \alpha_t(\tau) + \alpha_i(\tau) \tag{2.5}$$

where $\beta(\tau)$ represents the heterogeneous effect of children's education on parents' mental health. $\eta(\tau)$ captures the impact of control variables on parents' CES-D 8 score

[26]The related STATA command we use is *qregpd*, written by Matthew Baker, Travis A. Smith and David Powell.

as a function of quantile τ. $\alpha_t(\tau)$ and $\alpha_i(\tau)$ are the time fixed effects and unobserved individual fixed effects, respectively. One big advantage of the GQR approach is its ability to produce unconditional quantile treatment effects when the model includes additional covariates. When using the GQR estimator, adding covariates no longer changes the quantile function, and thus the interpretation of the estimated coefficient on the treatment variable remains unchanged. In addition, the GQR method extends the unconditional quantile regression models proposed by Firpo et al. (2009) and Frölich & Melly (2013) to allow for multiple endogenous treatment variables—whether binary or continuous—and employs a corresponding IV approach with either binary or continuous instrumental variables. The IV used in the quantile model is the same as the one used in previous linear models. Another important advantage of this quantile approach is its ability to incorporate individual fixed effects even when the main explanatory variable is time-invariant.

2.5 Empirical results

We will begin with empirical results for our main specification, which focuses on the effects of the most educated children.

2.5.1 First stage results

The first stage results of pooled 2SLS listed in Table 2.3 indicate a positive and significant impact of exposure to the education law on children's education, aligning with the upward trend observed in Figure 2.2. On average, the education law increases children's years of schooling by 0.43 to 1.41 years depending on residency type. For children born in rural regions, the effect ranges from 0.82 to 1.41, while for those born in urban areas, it drops to 0.43 to 0.82 years[27]. Though our results are not directly comparable to that of Ma (2019) due to differences in the share of children born in each residency type, they are still smaller than the 0.88 to 2.68 years reported

[27] These numbers are calculated by $2.061 - 0.144 \times prelaw_h, h = (rural, urban)$. Values for $prelaw_h$ are provided in Table 2.1.

in her study. Additionally, our results suggest that exposure to the CSL and its interaction with pre-law education levels are not a weak instrumental variable for the most educated children, with the Olea-Pflueger F-statistics exceeding the 5% critical value. We employ the Olea-Pflueger statistic here as it allows for heteroscedasticity errors (I. Andrews et al., 2018, 2019; Olea & Pflueger, 2013). As the specific diagnostic IV test for the quantile regression is unavailable, we only present the weak IV test for the linear model here (Jayawardana et al., 2023).

Next, we use the OLS regression to evaluate the reduced-form relationship between children's exposure to the CSL and parents' mental health. If the increased education of children, induced by exposure to the education law, decreases parental CES-D 8 scores (thereby alleviating mental health issues), then the reduced-form relationship is expected to be negative. Our results in Table 2.3 indeed suggest a negative reduced-form relationship for children born in rural areas, with parents' CES-D 8 score decreasing on average by -0.37 to -0.03 due to exposure to the CSL.

Finally, we add the endogenous variable to the reduced-form regression to examine the exclusion restriction. If our instruments affect parental mental health only through children's education, their coefficients should become insignificant after controlling for children's education. As shown in column (3) of Table 2.3, only the endogenous variable is significant while both IVs become insignificant. In addition, as mentioned in Section 2.4.1, we add parents' mental health in the past to eliminate the direct impact of confounding policies. If this strategy is effective, excluding past parental mental health from the regression should lead to significant IV coefficients, as our IVs now also capture other confounding factors that directly affect parents' mental health. Compared with column (3), IVs become significant in column (4), which indicates that adding past parental mental health effectively eliminates the direct impact of confounding policies.

2.5.2 The heterogeneous effect of children's education on parental mental health

The second stage results presented in Table 2.4 reveal statistically insignificant results for both the pooled 2SLS and the CRE models. This suggests that, on average, children's education has no effect on parents' mental health, which is consistent with findings of Ma (2019) in China and Cornelissen & Dang (2022) in Vietnam. This insignificant average effect, in turn, motivates an exploration of the distributional effect by applying a quantile approach (Ampofo, 2021). Although the CRE model suggests a slightly smaller magnitude of coefficient estimates compared to pooled 2SLS, the results are generally similar.

Contrasting with the linear model, our quantile regression model yields more informative results, revealing heterogeneity in the impact of children's education across the distribution of parents' mental health. As shown in Figure 2.4, children's education has little effect on parents who fall below the first 0.6 quantiles of the entire mental health distribution, that is, those experiencing mild to moderate mental health issues. Instead, the impact of children's education becomes significant and positive when parents are on the right or the higher end of the mental health distribution. The largest effect is observed at the 0.9 quantile, where a one-year increase in children's years of schooling reduces parents' CES-D 8 score by 0.118 points, translating to roughly a 2.1% decrease relative to the parents' average CES-D 8 score of 5.63. If we consider the upper bound impact of the education law, where children's education increases by nine years, parents' CES-D 8 score is expected to decrease by approximately 19%. This heterogeneous result also indicates that the insignificant average effect arises from the insignificance in the left side of the distribution.

The economic interpretation of this heterogeneous effect is that parents with poorer mental health appear to benefit more from improvements in their children's education. One plausible explanation is that poor mental health is often linked to lower socioeconomic status (Das et al., 2007). Parents with low socioeconomic status may rely more heavily on their children for both emotional and financial support. As discussed in

Section 2.7, better-educated children tend to provide more old-age support to parents, and their parents also perceive themselves to hold a higher social status relative to their neighbours.

It is difficult to compare our results with the literature directly, either because the literature reports an insignificant average effect of children's education on parental mental health, or because there are limited studies employing the quantile approach on this topic. However, we could compare our results with the average effect on other aspects of parents' health found in the literature. Ma (2019) has found the same positive impact of children's education on parents' life satisfaction in China, though the magnitude is nearly double ours. Liu (2021) suggests a one-year increase in children's education improves parents' good health status by 7.7 percentage points. Cornelissen & Dang (2022) find the same positive impact in Vietnam, where exposure of one child to the education reform enhances parents' self-reported health by 10 percentage points. By comparing our results with the literature, it appears that children's education has a larger positive impact on parents' physical health than on their mental health. In addition, our pooled 2SLS result[28] suggests that the age associated with the highest levels of depression in terms of the CES-D 8 scale is 57.6, which is slightly greater than the midlife nadir estimation of 53 years old for Germans in Cheng et al. (2017).

2.5.3 Heterogeneity by child gender, parents gender and residency type

The effect of children's education on parental mental health may vary across different subsamples. To investigate this heterogeneity, we split our sample into sons and daughters, fathers and mothers, and rural and urban. Analyzing these subsamples could also shed light on the potential driving forces behind the significant results found in the full sample.

In light of the substantial differences in cultural norms and parental expectations for boys and girls in China, it is interesting to further explore the differential impact of

[28] See Appendix Table B4.

sons' and daughters' education on parents' mental health. As illustrated in Figure 2.5, daughters' education has little effect on parents' mental health across almost all quantiles, except at the 0.1 and middle to upper-middle quantiles. In contrast, sons' education has a positive and significant impact across all quantiles, except at lower quantiles. In particular, for parents experiencing severe mental health issues (above the 0.75 quantile), only sons' education has a significant impact. Considering that a significant impact of children's education is observed only in quantiles above 0.7 in the full sample analysis, it could be inferred that the observed significance is driven by the education of sons, rather than daughters. This result is similar to the finding in Everding (2019), which suggests only sons' education has a significant impact on reducing parental depression in European countries.

One potential explanation for our results is the enduring and still-prevailing cultural norm of son preference in East and South Asian countries, particularly in China (Das Gupta et al., 2003). Coupled with the One Child Policy, the son preference of Chinese parents has even led to a high incidence of the 'missing girls' phenomenon through sex-selective abortion (Ebenstein, 2010; Chen et al., 2013). Because of this cultural norm, parents in China tend to place a greater emphasis on sons, resulting in a greater impact on parental mental health from son's education than from daughter's, as found in this paper. This conclusion is further supported by the data employed in this paper, which contains typical questions related to the level of son preference, such as *"do you agree that you need to have at least 1 son to continue family lineage"*, rated on a scale from 1 (strongly disagree) to 5 (strongly agree). As anticipated, nearly half of the respondents in our sample answered 5 to this question, and around 60% answered 4 or 5 (see Appendix Table B5).

Next, we evaluate whether fathers respond differently from mothers to children's education. As shown in Figure 2.6, the impact of children's education is similar for fathers and mothers when they have mild to moderate mental health problems (below the 0.7 quantile). However, the impact on mothers gradually surpasses the impact on fathers as they experience more severe depression. This result also indicates that the signifi-

cance observed in the right tail of the distribution in the full sample analysis is driven by the impact of children's education on mothers, rather than on fathers. One potential reason for this could be the difference in parenting behaviour between fathers and mothers. For example, mothers typically engage more in caregiving behaviours than fathers and tend to discuss emotions more during parent-child interactions (Möller et al., 2013; Schoppe-Sullivan et al., 2013). As a result, the higher educational attainment of children may reduce parental worry and is likely to provide greater relief from depression for mothers than for fathers, especially when parents suffer from severe mental health issues.

Finally, we explore the heterogeneity analysis of splitting the sample based on parents' current residency type (rural or urban), which is also called Hukou type in Chinese. As depicted in Figure 2.7, the impacts of children's education are almost insignificant at all quantiles for parents living in urban regions, in contrast to the significant and positive impacts observed at all quantiles (except at 0.1) for those living in rural areas. Given the considerable educational gap between rural and urban areas in China[29], our result exhibits a pattern of diminishing parental mental health returns to children's education. Previous studies have found evidence of the diminishing mental health returns of education within one's own generation (Cutler & Lleras-Muney, 2006; Bracke et al., 2013). Specifically, they argue that an additional improvement in education is less beneficial for own mental health if one is already highly educated. Our results extend this finding to the intergenerational context: improvements in children's education are less beneficial for parents' mental health if the parents live in urban regions (their children are more educated) rather than rural (their children are less educated) regions.

[29]In our sample, for parents who live in rural areas, the average years of schooling for their children is 10.66 years, whereas for parents who live in urban areas, this number is 13.58 years.

2.6 Comparative analysis

In this section, we conduct two comparative analyses (the most educated versus the least educated children and the first-born versus the second-born children) to gain a more complete understanding of the impact of children's education on parental mental health.

2.6.1 The most educated children versus the least educated children

There is established literature arguing that the most educated children have the highest impact on parents' physical health, such as hazard ratios of dying and survival probabilities, due to their greater capability to provide old-age support to elder parents compared to other less educated siblings (Zimmer et al., 2007; Ma, 2019). However, which child has a larger impact on parents' mental health remains unexplored. It is possible that the least educated children, who have lower social status and poorer living conditions than their more educated siblings, may have a greater influence on their parents' mental well-being due to increased parental worry, which potentially contrasts with the dynamics observed in physical health outcomes. The impact may also vary depending on the level of parental mental health.

To perform this comparative analysis, we regress parents' mental health on both years of schooling of the most and least educated children together. Because exposure to the education law and its interaction with the pre-law education level serve as weak instruments for the least educated children (with the Kleibergen-Paap F statistic equal to 1.629), we chose to run only the OLS regression and panel quantile regression without IV. Thus, our findings in this section are correlational rather than causal. As shown in Table 2.5, on average, the least educated children have a greater impact on parents' mental well-being. The results of the OLS regression are similar to those of the CRE model. In contrast to the linear model, the quantile regression yields more informative and interesting results. The non-depressive parents tend to place a higher value on the least educated children, while parents experiencing moderate to severe

mental health issues give more credit to their most educated children.

One rationale behind this result could be the different ways in which parents respond to mental health challenges of varying severity. On the one hand, depressive parents may tend to have a higher demand for healthcare and social support from their children. This potentially leads them to place greater emphasis on their most educated children, who are more capable of meeting such demands. On the other hand, non-depressive parents are typically more capable of self-care and tend to have lower healthcare and social support needs than depressive ones. They may even provide financial and emotional support to their children to some extent. In this case, parents are inclined to place greater value on their least educated children, who need more support and assistance from parents than their more educated siblings.

2.6.2 The first-born children versus the second-born children

Beyond educational attainments, birth order may also influence the effect of children's education on parents' mental health. To investigate this, we regress parents' mental health on both years of schooling of the first-born and the second-born children together. One key requirement for this analysis is the availability of information for all children. Two-child families, which account for 66% of all multi-child families, provide us with 4,078 observations, whereas this number reduces to only 1,040 if we use three-child families. As a result, we only use two-child families in this subsection. Since our IVs continue to be weak instruments when considering birth orders (with the Kleibergen-Paap F statistic below 1), we only run OLS regression and panel quantile regression without IV.

As depicted in Figure 2.8, the magnitudes of the impact for first-born children are greater than or equal to those of second-born children across all quantiles, except for the 0.9 quantile, where the latter's magnitude is only 0.01 year higher. The OLS regression suggests a similar result, with a one-year increase in the schooling of first-born children, on average, reducing parents' CES-D 8 score by 0.072 points, while

the effect for second-born children drops to 0.041 points and is weakly significant[30]. This generally greater educational impact of first-born children on parental mental health, compared to that of second-born children, implicitly reveals the prevalence of the first-born advantage among Chinese parents.

Among the established literature studying the birth order effect in China, papers employing data from younger generations suggest that there is a negative birth order effect, indicating that the earlier-born kids outperform their later-born siblings. This effect is primarily observed in academic achievements (J. H. Kim & Wang, 2023) as well as in cognitive ability and educational attainment of adult children (S. Zhang et al., 2023). Our findings implicitly demonstrate consistency with the observed negative birth order effect, or the first-born advantage, by showing that parental mental well-being is more strongly affected by the education of first-born children than that of second-born ones.

2.7 Mechanisms

Our results suggest that children's education significantly improves the mental well-being of parents suffering from severe mental health issues. In this section, we explore the potential mechanisms through which children's education improves parental mental well-being. The estimated equation is as follows:

$$Y_{it} = C_2 + \beta\, ChildEdu_i + \eta\, X_{it} + \alpha_{2t} + u_{it} \tag{2.6}$$

where Y_{it} is the three potential mechanisms being examined, including old-age support children provided to parents, parents' expenditures, and parents' subjective social status. $ChildEdu_i$ represents the years of schooling of the most educated child of parents i. Except for past mental health scores, X_{it} includes the same control variables as in Equation 2.2. The estimation strategy is the pooled 2SLS, as discussed in Section 2.4.2.

[30]The OLS results are available upon request.

2.7.1 Old-age support

Despite dramatic improvements in living standards and social benefits since the economic reform in the late 1970s, the social security system in China remains inadequate for the elderly population. According to the 2000 Chinese Census, only 8.2% of rural non-working individuals aged above 60 received pension benefits, while 82.6% of them relied on family support for livelihood (Wang, 2006). Although the pension insurance system expanded rapidly since the late 2000s and covered roughly 57% of Chinese residents in 2012, the benefits for rural individuals remained low, accounting for less than 11% of their average per capita income (Oliveira, 2016). In this circumstance, elderly parents tend to rely heavily on their children's support, and the inability of children to provide such support may negatively affect the parents' mental well-being. In addition, cultural norms like filial piety typically place an expectation on children to provide both monetary and in-kind support to their elderly parents (Wu & Li, 2014; R. Guo & Zhang, 2020). Failure to meet these expectations may also adversely affect parents' mental health.

As shown in Table 2.6, children's education has a positive impact on both monetary and in-kind support to parents. Magnitudes of the positive impact are considerable and could be translated to 54%, 11%, and 38% relative to the averages of financial support, visit frequency and housework frequency, respectively. Because a larger number of children transfer 0 to their parents, we also employ the Tobit model in the financial support regression to account for these corner solutions. The difference in magnitude may result from the Tobit model not using IV. Although our IVs appear to be weak instruments in this subsection with a reduced sample of fewer than 2,000 observations[31], they are not considered weak instruments for the most educated children in the full sample as shown previously in Table 2.3.

[31]In the CFPS, questions related to old-age support are available only after 2016 and focus only on parents aged 60 and above.

2.7.2 Expenditures

In middle- and low-income countries, where access to economic and environmental resources like private healthcare services and adequate sanitation is limited, increased consumption of these resources has been shown to significantly improve both physical and mental health (Ma, 2019). More educated children potentially enable their parents to spend more on these resources. In this study, we use the annual per capita household expenditure of parents to test this channel. As shown in Table 2.6, children's education greatly boosts parental expenditure. Specifically, each additional year of education leads to an 11.9% increase in parental expenditure.

2.7.3 Subjective social status

In recent decades, extensive studies reveal that subjective social status has a greater impact on one's physical and psychological health than objective social status (Hoebel & Lampert, 2020). One key channel through which subjective social status affects mental well-being is shame and the sense of inferiority created during social comparison processes (Wilkinson, 2002). By providing greater financial transfers and more access to economic and social resources, more educated children could potentially elevate their parents' social status relative to other parents with less educated children. The CFPS includes a question asking respondents to compare their social status with others in the local area, allowing us to examine this mechanism. As shown in the last column in Table 2.6, children's education significantly improves parents' subjective social status, with each additional year of children's education leading to about a 4% increase relative to the sample mean of 3.01.

2.8 Conclusion

Over the past few decades, rising awareness and concern have emerged regarding mental health. Worldwide depressive disorders prevalence has experienced a constant increase over the past 30 years, primarily driven by middle-aged and elderly indi-

viduals. This increasing trend is particularly significant in China, where its elderly population has encountered the second-highest rise in the prevalence of depressive disorders among 204 countries. Among all factors affecting the mental health of the elderly, concerns about their children's well-being emerge as a prominent one. Children's educational attainment as a key predictor of their well-being, therefore, should have a strong linkage with old parents' mental health.

This paper uses 2SLS and the CRE model to examine the average effect of children's education on parental mental well-being in China. Furthermore, we employ a generalized quantile regression (GQR) approach in an instrumental variable framework to explore the distributional effect, which enables us to study whether the effect of children's education varies along the distribution of parental mental health. Our results indicate an insignificant average effect but a significant distributional effect. Specifically, children's education does not affect parents experiencing mild to moderate mental health issues (below the 0.7 quantile) but significantly improves the mental well-being of those with severe issues (at the 0.7 quantile and above).

Our split-sample analysis suggests that a son's education has a greater positive impact on parental mental well-being than a daughter's. We also find the positive impact is greater for parents living in rural areas compared to those in urban areas. Our results from the comparative analysis indicate that when parents are not depressed, the schooling of the least educated children improves parents' mental well-being more. In contrast, when parents have moderate to severe mental health issues, the schooling of the most educated children has a greater impact. Additionally, the educational impact of first-born children is more substantial compared to that of second-born children.

Our results provide valuable insights into how intergenerational dynamics affect the mental well-being of middle-aged and elderly people in the developing world. Our results also offer practical implications for individuals and policymakers. From the perspective of parents, it is necessary and beneficial to increase education investments in children. Having a more educated child, in turn, could significantly mitigate future

mental health challenges. From the government's standpoint, there is a compelling need to augment investments in the education sector, with a particular emphasis on boosting funding for less-educated rural areas. In recent years, with a constant rise in the number of participants in both the college entrance examination and postgraduate entrance examination, prioritizing education investment remains crucial in policy formulation.

Figures for Chapter 2

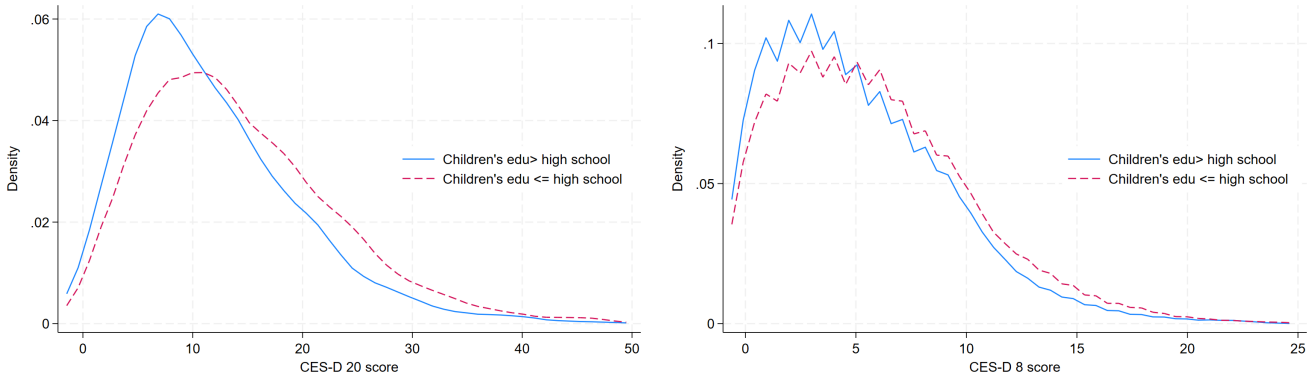


Figure 2.1: Distribution of parents' mental health scores by child's education level. *Notes:* A higher CES-D score represents more pronounced symptoms of depression. The solid blue line represents parents with children who have completed 2/3-year college or university. The red dashed line represents parents with children who have completed high school, middle school, primary school or illiterate. The children here are the most educated children.

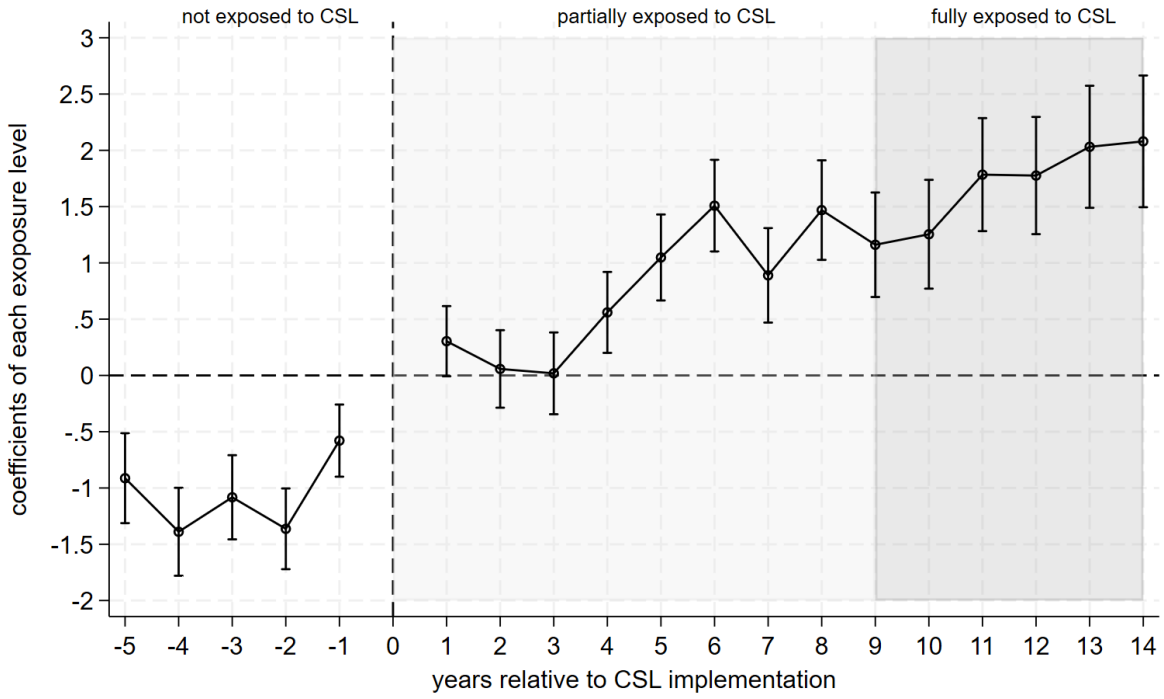


Figure 2.2: Impacts of each exposure year to CSL on children's education. *Notes:* The dependent variable is the years of schooling of the most-educated children. The reference group is the first eligible cohort for the CSL, for which $exposure_{ij}$ equals 0.1. Control variables include children's characteristics, birth year dummies, province-specific birth cohort trends, and province dummies. Presented above are point estimates with 95% confidence intervals.

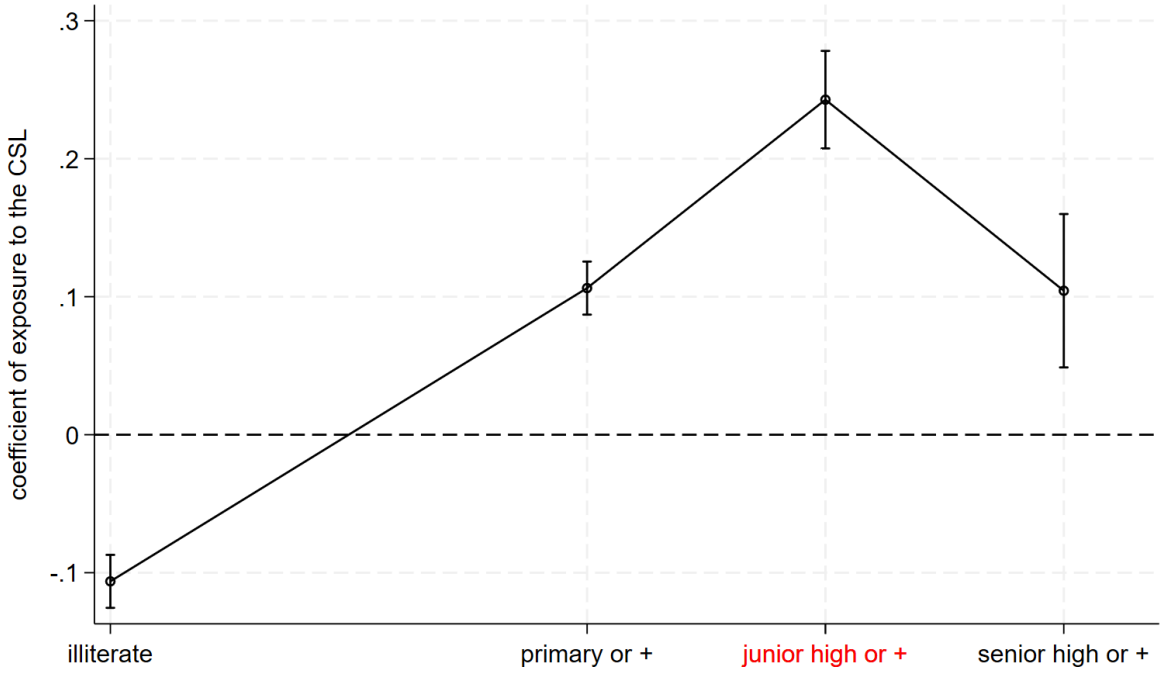


Figure 2.3: The impact of exposure to the CSL on children's educational attainment. *Notes:* The dependent variable is a dummy of children's educational attainment. Control variables include children's characteristics, birth year dummies, province-specific birth cohort trends, and province dummies. Presented above are point estimates with 95% confidence intervals.

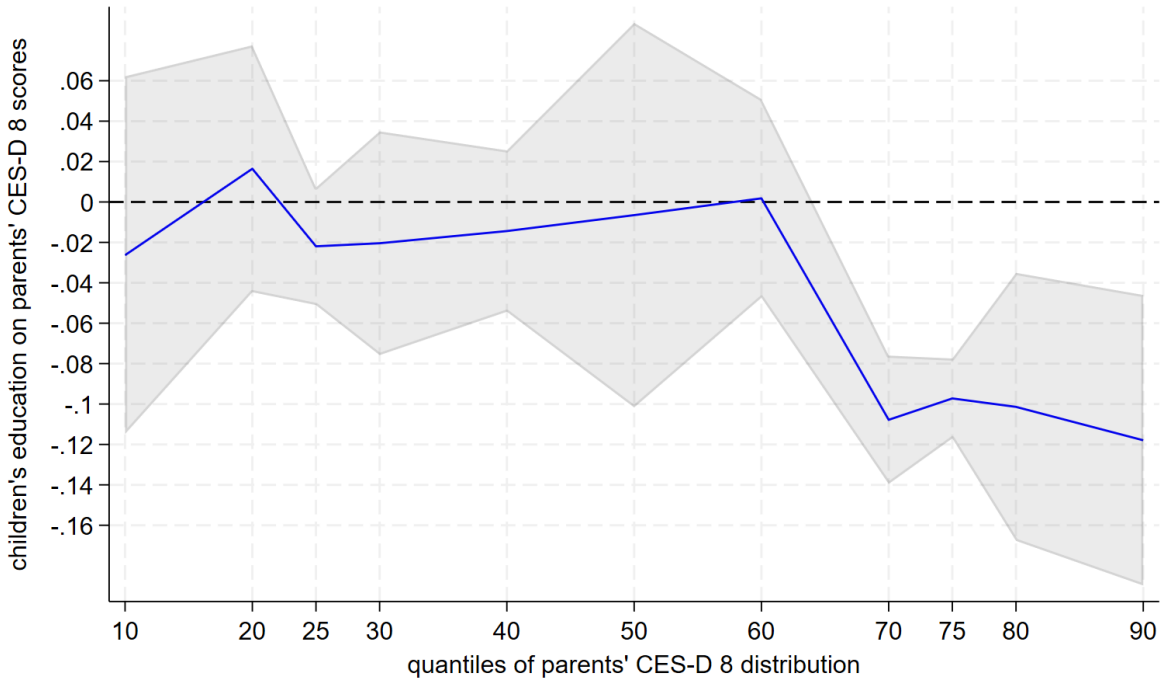


Figure 2.4: The heterogeneous impact of children's education on parents' mental health. *Notes:* The dependent variable is the parents' CES-D 8 score. A higher score indicates a more pronounced mental health issue. The results are based on panel quantile regression that uses exposure to the CSL and its interaction with the pre-law education level as IVs. All control variables inside X_{it} are added, including parents & children's characteristics and region dummies. Time fixed effects are also included. The confidence band displayed corresponds to the 95% confidence level.

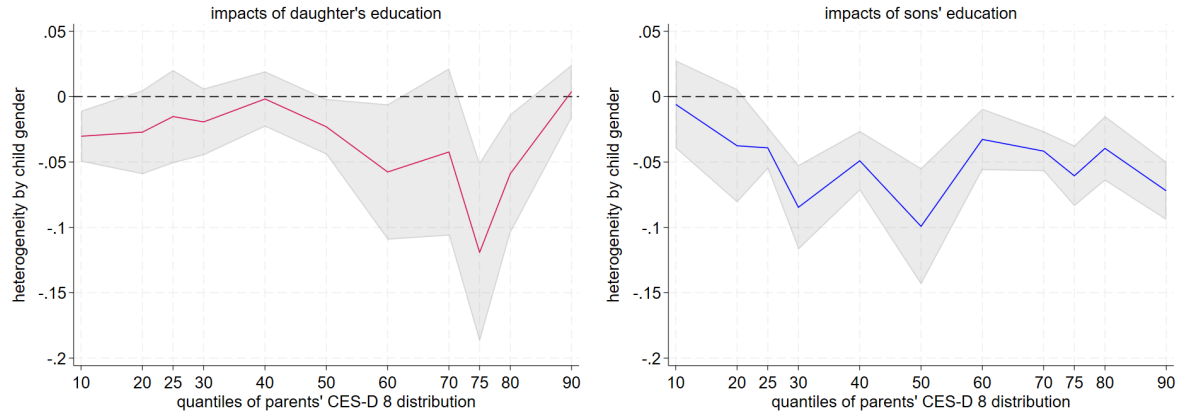


Figure 2.5: Children's education and parents' mental health (heterogeneity by child gender). *Notes:* The dependent variable is the parents' CES-D 8 score. A higher score indicates a more pronounced mental health issue. The results are based on panel quantile regression that uses exposure to the CSL and its interaction with the pre-law education level as IVs. All control variables inside X_{it} are added, including parents & children's characteristics and region dummies. Time fixed effects are also included. The confidence band displayed corresponds to the 95% confidence level.

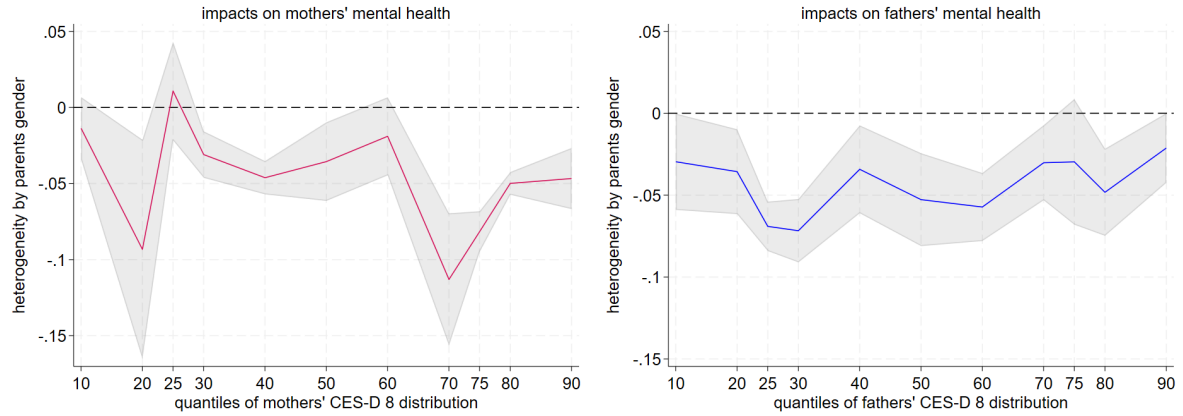


Figure 2.6: Children's education and parents' mental health (heterogeneity by parent gender). *Notes:* The dependent variable for the red line is the mothers' CES-D 8 score, while it is the fathers' CES-D 8 score for the blue line. A higher score indicates a more pronounced mental health issue. The results are based on panel quantile regression that uses exposure to the CSL and its interaction with the pre-law education level as IVs. All control variables inside X_{it} are added, including parents & children's characteristics and region dummies. Time fixed effects are also included. The confidence band displayed corresponds to the 95% confidence level.

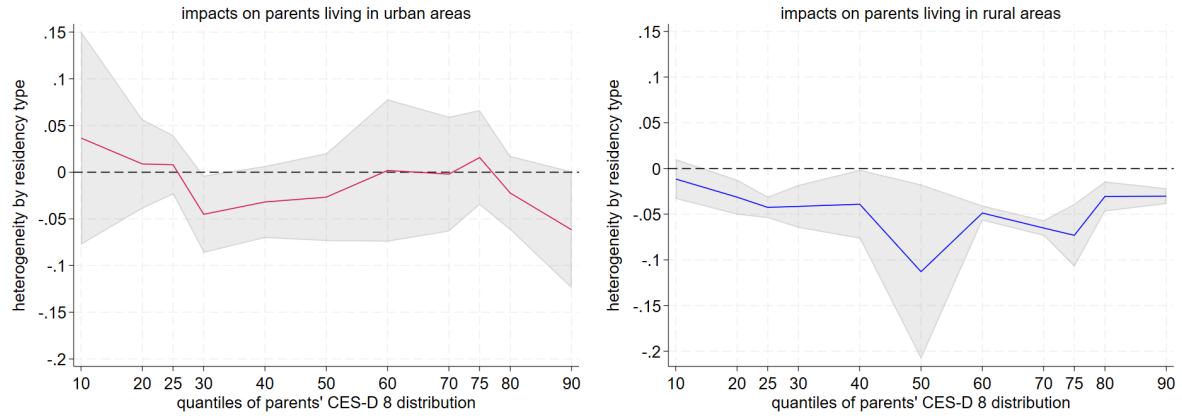


Figure 2.7: Children's education and parents' mental health (heterogeneity by residency type). *Notes:* The dependent variable is the parents' CES-D 8 score. A higher score indicates a more pronounced mental health issue. The results are based on panel quantile regression that uses exposure to the CSL and its interaction with the pre-law education level as IVs. All control variables inside X_{it} are added, including parents & children's characteristics and region dummies. Time fixed effects are also included. The confidence band displayed corresponds to the 95% confidence level.

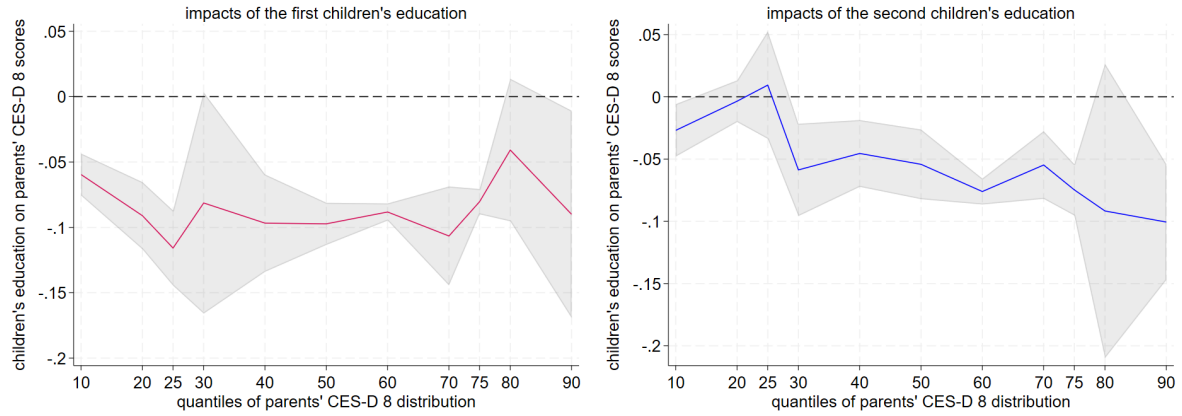


Figure 2.8: Comparative analysis: the first-born vs. the second-born children. *Notes:* The dependent variable is the parents' CES-D 8 score. A higher score indicates a more pronounced mental health issue. The results are based on panel quantile regression without IV. All control variables inside X_{it} are added, including parents & children's characteristics and region dummies. Time fixed effects are also included. The confidence band displayed corresponds to the 95% confidence level.

Tables for Chapter 2

Table 2.1: Law effective years and pre-law educational attainments for each province.

province	law effective year	first eligible cohort	pre-law average years of schooling	
			rural	urban
Beijing	1986	1971	8.625	11.31
Tianjin	1987	1972	7.921	10.304
Hebei	1986	1971	8.252	9.921
Shanxi	1986	1971	8.078	10.141
Inner Mongolia	1988	1973	7.305	9.906
Liaoning	1986	1971	7.883	10.203
Jilin	1987	1972	7.382	10.209
Heilongjiang	1986	1971	7.579	9.938
Shanghai	1987	1972	8.032	9.935
Jiangsu	1987	1972	8.209	9.752
Zhejiang	1986	1971	7.247	8.628
Anhui	1987	1972	7.202	9.143
Fujian	1989	1974	7.009	8.676
Jiangxi	1986	1971	6.859	9.602
Shandong	1987	1972	7.844	9.798
Henan	1987	1972	8.219	9.976
Hubei	1987	1972	7.411	9.844
Hunan	1991	1976	8.026	10.329
Guangdong	1987	1972	7.767	9.41
Guangxi	1991	1976	7.716	9.899
Chongqing	1986	1971	6.872	9.352
Sichuan	1986	1971	7.009	9.378
Guizhou	1988	1973	4.858	8.58
Yunnan	1987	1972	5.526	8.835
Shaanxi	1988	1973	7.44	10.134
Gansu	1991	1976	6.112	10.193
Qinghai	1988	1973	4.518	9.676
Xinjiang	1988	1973	7.233	10.031

Source: Table 1 in Huang (2025). *Notes:* Pre-law average years of schooling are calculated by Ma (2019) based on individuals who born between one to five years before the first cohorts eligible for the compulsory schooling law, using data from China's 2000 population census.

Table 2.2: Descriptive statistics

	mean	min	max	sd	N
parents					
CES-D 8 score	5.63	0	24	4.20	18244
years of schooling	6.03	0	19	4.48	18244
age	54.72	41	92	7.13	18244
married	0.92	0	1	0.27	18244
male(father)	0.49	0	1	0.50	18244
number of children	1.99	1	8	0.95	18244
agricultural (rural) hukou	0.79	0	1	0.41	18244
smoke	0.31	0	1	0.46	18244
drink alcohol	0.18	0	1	0.38	18244
pension	0.64	0	1	0.48	18244
work	0.71	0	1	0.46	18244
health level (=5 very healthy)	2.70	1	5	1.24	18244
relative income level (=5 very high)	2.60	1	5	1.11	18244
relative social status (=5 very high)	3.01	1	5	1.11	18244
previous mental health score (K6)	3.20	0	24	3.67	18244
log annual expenditure per capita (Chinese yuan)	9.20	4	14	0.92	17319
the most-educated child					
years of schooling	11.97	0	16	3.48	18244
exposure to the compulsory schooling law	0.91	0	1	0.25	18244
birth year	1986.55	1941	1997	6.67	18244
married	0.59	0	1	0.49	18244
male(son)	0.65	0	1	0.48	18244
age	29.67	23	71	5.79	18244
monthly financial support (Chinese yuan)	256.31	0	10000	684.69	1809
visit frequency (=6 every day)	4.61	0	6	1.61	1840
housework frequency (=6 every day)	1.48	0	6	2.30	1841
the least-educated child					
years of schooling	10.13	0	16	3.97	17262
exposure to the compulsory schooling law	0.85	0	1	0.31	17262
birth year	1984.94	1941	1997	7.42	17262
married	0.63	0	1	0.48	17262
male(son)	0.70	0	1	0.46	17262
age	30.76	23	71	6.45	17262
first-born child					
years of schooling	10.85	0	16	3.49	4078
exposure to the compulsory schooling law	0.93	0	1	0.20	4078
birth year	1985.86	1957	1996	4.85	4078
married	0.74	0	1	0.44	4078
male(son)	0.59	0	1	0.49	4078
age	31.53	23	60	4.62	4078
second-born child					
years of schooling	10.96	0	16	3.47	4078
exposure to the compulsory schooling law	0.97	0	1	0.13	4078

	mean	min	max	sd	N
birth year	1989.09	1960	1997	4.69	4078
married	0.55	0	1	0.50	4078
male(son)	0.62	0	1	0.49	4078
age	28.29	23	54	4.40	4078
regional fixed effects					
economically developed areas	0.40	0	1	0.49	18244
economically moderately developed areas	0.30	0	1	0.46	18244
economically underdeveloped areas	0.30	0	1	0.46	18244

Table 2.3: First-stage & reduced form results for the impact of CSL on education

	first stage	reduced form	IV+endogeous var	
	(1)	(2)	(3)	(4)
exposure to the CSL	2.061***	-0.751*	-0.659	-0.983**
	(0.482)	(0.453)	(0.453)	(0.471)
exposure × prelaw	-0.144***	0.084*	0.077	0.121**
	(0.051)	(0.050)	(0.050)	(0.053)
past parental mental health	-0.040***	0.295***	0.294***	
	(0.009)	(0.010)	(0.010)	
children's years of schooling			-0.045***	-0.061***
			(0.011)	(0.011)
Observations	18,244	18,244	18,244	18,244
parents & children controls	✓	✓	✓	✓
region dummies	✓	✓	✓	✓
time fixed effects	✓	✓	✓	✓
Individual fixed effects	No	No	No	No
Olea-Pflueger F^{eff}	12.17			
5% of Worst Case Bias	8.583			
p value of Hansen J statistic	0.281			

Notes: The dependent variable for column (1) is the years of schooling for the most educated child. The dependent variable for columns (2)-(4) is the parents' CES-D 8 score. A higher score indicates a more pronounced mental health issue. Prelaw refers to the provincial average years of schooling before law enforcement, with detailed information provided in Table 2.1. All control variables inside $X_{it}^{'}$ are included. Standard errors are clustered at the individual level as the same individual may be observed more than once in our sample. *** means p < 0.01, ** means p < 0.05, * means p < 0.1.

Table 2.4: Children's education on parents' mental health

	Pooled 2SLS	CRE-2SLS
children's years of schooling	-0.270	-0.251
	(0.206)	(0.160)
observations	18,244	18,244
parents & children controls	✓	✓
region dummies	✓	✓
time fixed-effects	✓	✓
individual fixed effects		✓

Notes: The dependent variable is the parents' CES-D 8 score. A higher score indicates a more pronounced mental health issue. All columns use exposure to the CSL and its interaction with the pre-law education level as the instrumental variable for years of schooling of the highest educated child. The Olea-Pflueger F^{eff} is 12.17. CRE-2SLS denotes the two-stage correlated random effects least squares model. Full results are presented in Appendix Table B4. Detailed quantile estimates shown in the graph are provided in Figure 2.4. Standard errors are clustered at the individual level as the same individual may be observed more than once in our sample. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Table 2.5: Comparative analysis: the most educated vs. the least educated children

	OLS	CRE	panel quantiles without IV			
			0.25	0.50	0.75	0.90
the most educated child	-0.008	-0.010	0.115*	-0.116**	-0.111***	-0.109*
	(0.021)	(0.021)	(0.064)	(0.055)	(0.031)	(0.058)
the least educated child	-0.057***	-0.059***	-0.138**	0.046	0.091**	-0.021
	(0.021)	(0.021)	(0.058)	(0.074)	(0.042)	(0.058)
observations	13,059	13,059	13,059	13,059	13,059	13,059
parents & children controls	✓	✓	✓	✓	✓	✓
region dummies	✓	✓	✓	✓	✓	✓
time fixed-effects	✓	✓	✓	✓	✓	✓
individual fixed effects		✓	✓	✓	✓	✓

Notes: The dependent variable is the parents' CES-D 8 score. A higher score indicates a more pronounced mental health issue. CRE refers to correlated random effects model. Standard errors are clustered at the individual level as the same individual may be observed more than once in our sample. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Table 2.6: Mechanisms: impacts of the most educated children's schooling on old-age support, subjective social status and expenditures of old parents

	old-age support				expenditure	subj. social status
	fin. support	fin. support (Tobit)	visit freq.	housework freq.		
children's years of schooling	139.336*	6.436*	0.492*	0.567*	0.119***	0.120**
	(74.340)	(3.641)	(0.262)	(0.335)	(0.045)	(0.055)
observations	1,809	1,809	1,840	1,841	17,319	18,244
Olea-Pflueger F^{eff}	4.023		3.468	3.409	13.338	12.590
5% of Worst Case Bias	8.586		8.783	8.786	8.804	8.583
p value of Hansen J	0.326		0.938	0.089	0.00	0.860

Notes: Fin. support refers to the monthly financial assistance that children provide to their parents, measured in Chinese Yuan. Visit frequency and housework frequency are measured as integers ranging from 0 (never) to 6 (every day). Expenditure is the logarithm of total annual household expenditures per capita, measured in 2010 Chinese Yuan. Subjective social status is an integer that ranges from 1 (very low) to 5 (very high). All columns, except for the Tobit model, present results from the 2SLS model, using exposure to the CSL and its interaction with pre-law education level as the instrumental variable for the years of schooling of the highest-educated child. Except for past mental health scores, all control variables inside X_{it} are added, including parents & children's characteristics and region dummies. Regressions for old-age support use age dummies in 5-year intervals as all parents are aged above 60. Standard errors are clustered at the individual level as the same individual may be observed more than once in our sample. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Appendix B

Table B1: Questionnaire of the original 20-item Center for Epidemiologic Studies Depression (CES-D) scale in Radloff (1977)

How often you experienced them in the past week:
1. I am annoyed by some trifles.
2. I don't want to eat and have a poor appetite.
3. I feel depressed even though I receive help from my family and friends.
4. I feel that I'm better than someone else.
5. I find it hard to focus on what I am doing.
6. I feel depressed.
7. I find it difficult to do anything.
8. I am hopeful about the future.
9. I feel that I have been a loser all the time.
10. I feel scared.
11. I have a poor sleep.
12. I am happy.
13. I talk less than usual.
14. I feel lonely.
15. I find people are unfriendly to me.
16. I have a happy life.
17. I have cried or want to cry.
18. I feel sad.
19. I find others dislike me.
20. I feel that I am unable to keep on with my life

Source: China Family Panel Studies 2012 Questionnaire. *Notes:* There are 4 possible answers for each question: 1. Almost never (less than one day) 2. Sometimes (1-2 days) 3. Often (3-4 days) 4. Most of the time (5-7days). The corresponding CES-D 20 score for each answer is 0, 1, 2, and 3 respectively. The scores for questions 4, 8, 12 and 16 are reversed (i.e. 3, 2, 1 and 0 respectively).

Table B2: Questionnaire of the 8-item Center for Epidemiologic Studies Depression (CES-D) scale in CFPS

How often you experienced them in the past week:
1. I feel depressed.
2. I find it difficult to do anything.
3. I have a poor sleep.
4. I am happy.
5. I feel lonely.
6. I have a happy life.
7. I feel sad.
8. I feel that I am unable to keep on with my life.

Source: China Family Panel Studies 2018 Full Questionnaires. *Notes:* The CES-D 8 questions are part of the CES-D 20 question. Questions 1 to 8 are corresponding to questions 6 7 11 12 14 16 18 20 in the CED-S 20 questionnaire. There are 4 possible answers for each question: 1. Almost never (less than one day) 2. Sometimes (1-2 days) 3. Often (3-4 days) 4. Most of the time (5-7days). The corresponding CES-D 8 score for each answer is 0, 1, 2, and 3 respectively. The scores for questions 4 and 6 are reversed (i.e. 3, 2, 1 and 0 respectively).

Table B3: Level of children's exposure to the compulsory schooling law

age in law effective year	exposure	mean schooling	frequency	%
>15	0	9.03	788	4.23
15	0.1	9.12	148	0.81
14	0.2	9.31	147	0.81
13	0.3	9.09	158	0.87
12	0.4	9.65	173	0.95
11	0.5	9.65	260	1.43
10	0.6	10.07	232	1.27
9	0.7	11.34	341	1.87
8	0.8	10.15	436	2.39
7	0.9	10.93	438	2.40
<7	1	12.40	15,123	82.89
Total		11.97	18,244	100.00

Notes: The exposure column corresponds to the variable $exposure_{ij}$ in Equation 2.1. Mean schooling is calculated as the average years of education for all children within each exposure level. Provinces began enforcing the law in 1986: Beijing, Chongqing, Hebei, Heilongjiang, Jiangxi, Liaoning, Shanxi, Sichuan, Zhejiang. Provinces began enforcing the law in 1987: Anhui, Guangdong, Henan, Hubei, Jiangsu, Jilin, Shandong, Shanghai, Tianjin, Yunnan. Provinces began enforcing the law in 1988: Guizhou, Inner Mongolia, Qinghai, Shaanxi, Xinjiang. Province began enforcing the law in 1989: Fujian. Provinces began enforcing the law in 1991: Gansu, Guangxi, Hunan.

Table B4: Children's education on parents' mental health

	Pooled 2SLS	CRE-2SLS
children's years of schooling	-0.270	-0.251
	(0.206)	(0.160)
parent's years of schooling	-0.006	0.014
	(0.037)	(0.043)
parent's age	0.15903***	0.02597
	(0.060)	(0.180)
parent's age squared	-0.00138 ***	-0.00146
	(0.000)	(0.001)
parent married	-1.058***	-2.200***
	(0.191)	(0.386)
children married	-0.517***	-0.230**
	(0.186)	(0.113)
parent male	-0.612***	-0.541***
	(0.104)	(0.099)
children male	-0.150	-0.022
	(0.202)	(0.844)
parent's number of children	0.193***	-0.042
	(0.040)	(0.204)
parent's self-reported health status	0.818***	0.470***
	(0.026)	(0.043)
parent rural hukou	0.216	0.500**
	(0.334)	(0.253)
parent smoke	-0.090	0.006
	(0.104)	(0.188)
parent drink	-0.234***	0.112
	(0.083)	(0.135)
province-specific birth cohort trend	-0.000	-0.000
	(0.000)	(0.000)
parent's past K6 score	0.285***	0.058
	(0.013)	(0.047)
parent's pension	0.060	-0.109
	(0.111)	(0.099)
parent currently work	-0.220***	-0.296***
	(0.074)	(0.108)
parent's self-reported income levels	-0.232***	-0.103**
	(0.032)	(0.042)
parent's self-reported social status	-0.189***	-0.100**
	(0.034)	(0.045)
children's birth year	0.022	0.011
	(0.021)	(0.015)
parent lives in economically developed areas	-0.709***	-0.152
	(0.224)	(1.203)

	Pooled 2SLS	CRE-2SLS
parent lives in economically moderately developed areas	-0.520***	-0.560
	(0.188)	(1.498)
year 2016	0.131	0.502
	(0.084)	(0.442)
year 2018	0.932***	1.336**
	(0.112)	(0.652)
year 2020	1.098***	1.651*
	(0.162)	(0.868)
observations	18,244	18,244
parents & children controls	✓	✓
region dummies	✓	✓
time fixed-effects	✓	✓
individual fixed effects		✓

Notes: This is the full results of Table 2.4. The dependent variable is the parents' CES-D 8 score. A higher score indicates a more pronounced mental health issue. All columns use exposure to the CSL and its interaction with the pre-law education level as the instrumental variable for years of schooling of the highest educated child. The Olea-Pflueger F^{eff} is 12.17. CRE-2SLS denotes the two-stage correlated random effects least squares model. To acquire a more precise turning point, the parent's age and its squared term are rounded to five decimal places, whereas the remaining variables are rounded to three. Standard errors are clustered at the individual level. *** means p $<$ 0.01, ** means p $<$ 0.05, * means p $<$ 0.1.

Table B5: Descriptive statistics on son preference

	at least 1 son to continue family lineage	men focus on career women focus on family	women marrying well is more important than doing well
strongly disagree	12.6%	3%	7.5%
disagree	9.2%	3.5%	8.4%
netrual	18.1%	15.7%	23.6%
agree	15.4%	22.5%	20.5%
strongly agree	44.7%	55.4%	40%
observations	16,277	16,284	16,247

Chapter 3

Birth order effects, gender composition and old-age support: new evidence from China

RUYI SHI

School of Economics and Public Policy

The University of Adelaide

Statement of Authorship

Title of Paper	Birth order effects, gender composition and old-age support: new evidence from China
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☒ Unpublished and Unsubmitted work written in manuscript style
Publication Details	

Principal Author

Name of Principal Author (Candidate)	Ruyi Shi
Contribution to the Paper	This is a sole-authored paper, I contribute to 100%.
Overall percentage (%)	100%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third part[illegible] clusion in this thesis. I am the primary author of this paper.
Signature	Date 12/03/2025

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author			
Contribution to the Paper			
Signature		Date	

Name of Co-Author			
Contribution to the Paper			
Signature		Date	

Please cut and paste additional co-author panels here as required.

Abstract

This paper provides rich new evidence on China's evolving birth order effects on education. We use family-by-wave fixed effects to explore within-family variations in birth order across siblings. A positive birth order effect on education is observed among older generations (born before 1967), while the birth order effect reverses to negative among younger generations (born after 1978). For the positive effect, school dropout among older siblings is the primary driver, whereas child labour laws and resource dilution are the key factors for the negative effect. Brothers have consistently attained higher levels of education than sisters, but this gap has narrowed over time. In addition, differences in old-age support levels are negligible both among siblings and between brothers and sisters. This implies that children who are expected to provide more old-age support to their parents in return for the favours they received during childhood do not, in fact, offer more support. The altruistic motive emerges as a potential reason why siblings, as well as brothers and sisters, provide equal levels of old-age support to their parents. Furthermore, we find no evidence that in-kind support and monetary support function as substitutes or complements. Our results offer a comprehensive understanding of the relationship between birth order, human capital investment, and intergenerational support in the face of China's rapid economic transition.

keywords: Birth order, Education, Old-age support, China

JEL Classification: D13, D64, I20, J13

3.1 Introduction

There is a large body of literature exploring the reasons behind differences in children's human capital between families, such as the famous quantity-quality model proposed by Becker (1960). While differences between families can be explained by variations in family environments, differences within the family are particularly noteworthy as all siblings grow up in the same family. The magnitude of within-family variations is also quite substantial. For example, they explain 51% of the total variation in completed education in the Philippines (Ejrnæs & Pörtner, 2004), 39% in birth weights in the US (Restrepo, 2016), and approximately 60% in children's cognition in China (S. Zhang et al., 2023). In this paper, within-family variations account for 41% of the overall variation in years of education[32]. Birth order, as an inherent and natural factor, is potentially a key contributor to differences within families. Investigating the relationship between birth order and human capital within the same family helps identify which siblings are more or less advantaged. This information can inform policy decisions and guide parents to reduce inequalities among siblings.

As a key component of human capital, education has been widely discussed in the birth order literature. In developed countries, earlier-born children generally outperform their later-born siblings in terms of years of education (Behrman & Taubman, 1986; Black et al., 2005; Kantarevic & Mechoulan, 2006; Booth & Kee, 2009; De Haan, 2010; Mechoulan & Wolff, 2015; Lehmann et al., 2018; Y.-J. Kim, 2020; Bagger et al., 2021; Hiriscau & Pintea, 2024). Since years of education decrease as birth order increases, this phenomenon is referred to as the negative birth order effect on education. On the contrary, earlier-born children tend to underperform their later-born siblings in years of education in the developing world, which indicates a positive birth order effect (Ejrnæs & Pörtner, 2004; Emerson & Souza, 2008; Tenikue & Verheyden, 2010; De Haan et al., 2014; Kanayama & Yamada, 2024). However, scholars find a mix of

[32]This analysis is based on an R-squared decomposition using a simplified model. To acquire the proportion of the total variation in education explained by within-family variations, we regress years of education on the family fixed effects only, which follows the approach in Restrepo (2016).

both phenomena in China. Weng et al. (2019) using a sample of older generations of Chinese individuals find a positive birth order effect on schooling, which aligns with the pattern seen in developing countries. In contrast, S. Zhang et al. (2023) and J. H. Kim & Wang (2023) observe a negative birth order effect on cognition, schooling years, and schooling performance in younger generations of Chinese. The negative pattern is consistent with developed countries.

This paper aims to explore the impact of birth order on years of education, commonly referred to as birth order effects on education, in the context of China. We seek to reconcile the divergent findings in the literature by examining whether birth order effects change over time during China's rapid economic transition in recent decades. We use family fixed effects to explore within-family variations in birth order across siblings. As birth order effects may reflect past within-family parental investment preferences[33], it is also interesting to probe whether siblings who were more favoured by their parents during childhood tend to provide more filial support to parents in adulthood than their less-favoured siblings. Therefore, the second research question of this paper is how birth order influences the old-age support siblings provide to their parents in China. In addition, we study the impact of children's gender composition on education and old-age support to examine how potential son preference plays a role in the intergenerational context.

To explore within-family differences in children's education, each family in our sample must have more than one child. One concern is that the One Child Policy (OCP), implemented between 1978 and 2013, may restrict couples from having a second child. However, while the OCP effectively curbed population growth in China, it did not strictly limit all households to having only one child. According to Gu et al. (2007), there were 22 exceptions under which couples could have a second or third child, including the notable 1.5 child policy, which allowed rural couples to have a second child if their first-born was a girl. Based on the 2000 population census, the authors report that there were 225.28 million Chinese with a fertility rate between 1.0 and

[33]See detailed explanation in Section 3.5.

1.3, 607.07 million between 1.3 and 1.5, 329.47 million between 1.5 and 2.0, and 77.81 million above 2.0. Additionally, in Appendix Figure C1, we use the China Family Panel Studies (CFPS) to show that a substantial number of Chinese people had more than one child during the OCP period. Thus, we could still obtain a sample with enough observations of families with more than one child[34].

China provides the best context for studying birth order effects and old-age support for two reasons. Firstly, China has experienced a rapid transformation from an underdeveloped nation to one approaching a developed status in recent decades. The rapid transformation of China provides a quasi-experimental setting to study the dynamics of birth order effects in a way that is harder to observe in slower-changing economies. Due to this swift transition, birth order effects in China could potentially evolve over time, as reflected by the contradictory results in the literature discussed above. Understanding how birth order effects evolve with economic transformation could inform broader theories on family dynamics, education, and intergenerational mobility in both developing and developed countries. This could also offer insights applicable to other developing nations that go through a similar transition.

Secondly, the cultural norm of filial piety combined with the underdeveloped pension and social security system makes Chinese elders rely heavily on children's support compared to other countries. The 2010 Chinese census indicates nearly half of Chinese elders aged above 65 rely on family support as their main income source, whereas this proportion is less than 5.5% in Japan (ILC-Japan, 2010) and less than 3% in the US (West et al., 2014). Given this heavy reliance, investigating children's old-age support for parents in China is more meaningful compared to other countries. Additionally, many developing countries, such as Japan, are facing growing pension burdens due to an ageing population (Coulmas, 2007). China's family-based elder care framework may offer these countries insights into shifting some of the heavy burden from public

[34]According to S. Zhang et al. (2023), selection bias may arise for families with a second child during the OCP period, particularly among urban families. This could be addressed to some extent by using the family fixed effects, which effectively eliminate such family-specific heterogeneity (see detailed explanations in Section 3.4).

to private care.

This paper makes three main contributions to the existing literature. Firstly, this study adds new evidence to the literature on birth order effects in China. Unlike previous studies that suggest a static birth order effect (Weng et al., 2019; J. H. Kim & Wang, 2023; S. Zhang et al., 2023), this paper uses a rich dataset spanning multiple generations (from those born in the 1960s to the present) to provide a more nuanced examination of how birth order effects evolve over time, particularly during a period of rapid socio-economic transformation. This methodological advantage allows us to reconcile the divergent findings and provide a clearer picture of the changing dynamics of birth order effects. We also suggest two important cutoff years in which the sign of birth order effects changes. The new evidence of China's evolving birth order effects not only facilitates the understanding of how within-family inequality changes during the rapid economic transition, but also provides valuable insights for other developing countries in their ongoing economic transitions.

Secondly, this paper fills the gap in the literature regarding how children's birth order affects the old-age support they provide to their parents. While existing studies mainly focus on the impact of the number of children (Oliveira, 2016; Rao & Zhang, 2024) or children's gender composition on old-age support (Ho, 2019; R. Guo & Zhang, 2020; Ho & McGarry, 2025), the impact of birth order has received less attention. Birth order is of particular interest for three reasons: (1) similar to gender, a person's birth order also significantly determines their responsibility for old-age support among siblings; and (2) combined with the results on birth order effects on education, investigating birth order differences in old-age support among siblings allows me to observe whether children who received more parental investment in the past tend to offer more old-age support now; and (3) based on differences in old-age support, parents may also adjust their preferences among their children. Understanding how parents invest in their children's human capital, and whether this investment aligns with the children who ultimately provide the most support, has implications for social policy and economic sustainability worldwide. It can inform global discussions on intergenerational equity,

pension reform, and the role of the family in aging societies—issues relevant to both developed and developing nations grappling with similar demographic transitions.

The third contribution of this paper is that it sheds light on whether monetary and in-kind support that children provide to their parents serve as substitutes or complements among siblings, a topic that remains underexplored (R. Guo & Zhang, 2020).

Using a 10-year longitudinal household survey from the China Family Panel Studies (CFPS), this paper finds that the birth order effects in China evolve over time. Specifically, a positive birth order effect is observed among individuals born before 1967, while a negative birth order effect is found among those born after 1978. Within the same family, educational attainment for sons is consistently higher than that of daughters, but the gap has narrowed over time. In addition, there is no significant difference in the levels of old-age support provided among siblings, nor between brothers and sisters. Furthermore, we find no evidence that in-kind support and monetary support serve as substitutes or complements.

The remainder of this paper is organised as follows. The next section reviews the literature related to birth order effects and old-age support. Section 3.3 describes the data source and presents the summary statistics, followed by the introduction of the empirical model in Section 3.4. Results on the birth order effects are presented in Section 3.5, followed by the corresponding mechanisms in Section 3.6. Results on old-age support and gender composition are discussed in Sections 3.7 and 3.8, respectively. We conclude this paper in section 3.9.

3.2 Literature review

3.2.1 Birth order effects on education

The literature on birth order effects on education presents contrasting findings, with different patterns observed in developed and developing countries. A negative birth order effect on years of education is mostly observed in developed countries[35], such

[35]See detailed literature review in S. Zhang et al. (2023) Table D.1.

as the United States (Behrman & Taubman, 1986; Kantarevic & Mechoulan, 2006; De Haan, 2010; Lehmann et al., 2018; Y.-J. Kim, 2020; Hiriscau & Pintea, 2024), Norway (Black et al., 2005), the United Kingdom (Booth & Kee, 2009), West Germany (Härkönen, 2014) , France (Mechoulan & Wolff, 2015), Denmark (Bagger et al., 2021), and Netherlands (Nuevo-Chiquero et al., 2022). A negative birth order effect is also found in two developing countries, Mexico (Esposito et al., 2020) and India (Fors & Lindskog, 2023), where the authors use children who are still in school rather than adults, and they focus on school enrolment and human capital scores instead of years of education.

Resource dilution is considered the predominant mechanism behind the negative birth order effect. According to this mechanism, within the same family, earlier-born children are more educated because they benefit from a greater share of parental time and financial resources during early childhood before the arrival of younger siblings. As family size increases, parental resources are distributed among more siblings, leading to a reduction in resources available for any one child (Downey, 2001). Thus, later-born children no longer benefit from undiluted parental resources during their childhood as they must share them with elder siblings (Lillehagen & Isungset, 2020).

In contrast, a positive birth order effect is found in developing countries, including Philhellenes (Ejrnæs & Pörtner, 2004), Brazil (Emerson & Souza, 2008), Sub-Saharan countries (Tenikue & Verheyden, 2010), Ecuador (De Haan et al., 2014), and Cambodia (Kanayama & Yamada, 2024). The literature attributes the positive birth order effect to credit constraints. Families suffering from credit constraints tend to pull the earlier-born children out of school and send them to the labour market to alleviate this hardship (Patrinos & Psacharopoulos, 1997; Edmonds, 2006; Seid & Gurmu, 2015). In this case, later-born children tend to benefit from this extra income and stay in school longer.

Birth order effects are ambiguous in the context of China. Weng et al. (2019) find a positive birth order effect in older generations of Chinese, whereas S. Zhang et al. (2023) and J. H. Kim & Wang (2023) observe a negative birth order effect in

younger generations. This conflicting result implies birth order effects are evolving during the course of rapid economic development in contemporary China. Oliveira (2019) provides additional evidence on the impact of gender composition on birth order effects in China. The author finds a positive birth order effect on females, whereas a negative one on males. Credit constraints and resource dilution are two major mechanisms behind birth order effects in China, consistent with those in other countries as mentioned earlier.

3.2.2 The old-age support Chinese children provide to parents

Chinese elders heavily rely on private care provided by their children due to the underdeveloped public care and social security system. According to 2010 Chinese census, the main income sources for Chinese elders aged above 65 were 48.8% from family support, 29.2% from pensions or the minimum living standard guarantee programme, and 19.7% from labour income. From the perspective of children, they are willing to offer old-age support to their parents, as filial support is considered a traditional virtue rooted in Confucian cultural norms (Rao & Zhang, 2024).

The existing literature offers limited insights into the differences in filial support provided by siblings in each birth order. Instead, most studies primarily explore the impacts of factors such as the number of children or children's gender composition on providing old-age support. Rao & Zhang (2024) find having more children significantly reduces parents' probability of working after retirement. They argue that the children's financial support effectively compensates for parents' reduced labour income in old age. Oliveira (2016) also suggests that an increase in the number of children raises total intergenerational transfers parents receive, despite the possibility of reduced transfers from each child. Furthermore, the author finds that an additional son transfers slightly more to parents (179 CNY per year) than an additional daughter.

Different from the above two studies, Ho (2019) examines how current old-age support is affected by past parental investment. She shows a negative correlation between past parental investment in the son's education and the son's current in-kind support. This

is because higher parental investment leads to higher income for the son, which, in turn, raises the opportunity cost of his time. In contrast, higher investment in the daughter's marriage is positively correlated with her in-kind support, as it serves as a return on the greater parental investment in her marriage. R. Guo & Zhang (2020) focus on the impact of children's gender composition on parental expected levels of old-age support. They reveal that if a daughter has a brother, her parents are less likely to expect financial transfers from her, to co-reside with her, or to instil filial piety during her childhood. This result sheds light on why parents tend to invest more in sons than in daughters. Ho & McGarry (2025) builds on R. Guo & Zhang (2020) by focusing on the actual old-age support children provide to their parents, rather than the expected support. Using the heteroscedastic-based IV approach, the authors find that patterns of old-age support differ depending on the child's gender and the type of support[36].

3.3 Data

We employ data from China Family Panel Studies (CFPS), a 10-year longitudinal household survey conducted every two years since 2010. This dataset was chosen for several reasons. Firstly, CFPS is countrywide representative, covering 25 out of the 33 administrative regions in China, with these surveyed provinces accounting for 95% of China's population. This data is also recognized as the first and most comprehensive large-scale panel household survey in China (Y. Xie et al., 2017). Secondly, CFPS includes basic information such as gender, education, and birth year for household members not living together, which is crucial for constructing an accurate birth order. Thirdly, CFPS covers information for children's support to their old parents, including in-kind support such as housework and visit frequency, as well as monetary support. This study uses all six waves of CFPS.

This paper uses two samples, each based on one of the two dependent variables. The

[36]In Ho & McGarry (2025), birth order is used as a control variable rather than the main dependent variable of interest.

first main dependent variable is children's education, measured by years of schooling. We use data from all six waves in the first sample. According to the CFPS user's manual, years of schooling are assigned based on educational degrees as follows: 0 for illiterate or semi-illiterate individuals, 6 for primary school, 9 for middle school (junior high school), 12 for high school (senior high school), 15 for 2 or 3-year college, and 16 for a bachelor's degree. As this study examines whether there are within-family differences in education, we restrict the age of the youngest child in the family to be above 22 for both samples to ensure that all siblings in the same household have completed basic education. Without applying this age restriction to all children, younger children below 22 who are still in school would be excluded, potentially leading to an underestimation of birth order effects. We set the threshold at above 22 because most Chinese students graduate from university at 22 or older. Additionally, we omit children who hold a master's or PhD degree, as their completion age is typically older than 22. This group makes up 1.1% of the first sample.

Followed by J. H. Kim & Wang (2023) and S. Zhang et al. (2023), we also remove families with twins from both samples, as they are likely to exhibit different parent-child dynamics compared to normal families. We also exclude single-child families, which constitute 33.62% of the sample, because each family must have at least two children to ensure there is within-family variation. We eventually obtain the first sample consisting of 15,853 observations from 6,598 families. As shown in Table 3.1, more than half of the sample is the two-child family.

The full sample in Figure 3.1 shows that later-born children are generally more educated than the earlier-borns, indicating a slightly positive birth order effect. However, since first children are typically born several years before their siblings, their lower level of education is likely due to the time trend. To filter out the effect of time trends, we also compute standardized years of schooling using a rolling window of three ages, with the corresponding descriptions shown in Appendix Figure C2. After removing the time trend, the full sample indicates a slightly negative birth order effect for two-child families.

To better capture the dynamics of the birth order effect over time, we split the sample by birth year, with cutoff years determined by the empirical results presented in section 3.5. Both education measurements yield the same result. In the sample of children born before 1967, two-child families exhibit a clear positive birth order effect, whereas the effect for three or more families is not clear. In contrast, the sample of children born after 1978 reveals a negative birth order effect for both family sizes. In general, both subsamples in Figure 3.1 provide descriptive evidence of the changing birth order effects in China.

The second main dependent variable is the old-age support children provide to their parents. This variable is divided into two categories: in-kind support and monetary support. In-kind support includes the frequency of the child's housework and visits to their parents, measured as integers ranging from 0 (never) to 6 (every day). Monetary support refers to whether the child provides financial assistance and, if so, the amount given per month. In addition, to assess the overall interaction between the child and parents, we include their relationship quality as another dependent variable, measured on a scale from 0 (distant) to 5 (close).

Because CFPS only surveyed parents aged above 60 for the old-age support questions, the second sample contains fewer observations, ranging between 11,400 and 11,702, from 4,532 to 4,638 families, depending on the type of support. The average ages of parents and children in this sample are approximately 64 and 37, respectively. In addition, we only use data from the CFPS 2016, 2018, and 2020 waves in this sample, as earlier waves lack information on the corresponding child ID for the support provided to parents, making it unclear which sibling is providing the support. As shown in Figure 3.2, the descriptive differences in old-age support across birth order are negligible, despite later-born children generally providing slightly more support to parents. Similarly, Figure 3.3 reveals comparable negligible differences in old-age support by children's gender, except for financial support and housework frequency.

As shown in the descriptive statistics table (Table 3.2), the average birth year of children in our sample is 1978, which is later than the average of 1966 in Weng et

al. (2019), who find a positive birth order effect, and earlier than that in S. Zhang et al. (2023), who find a negative birth order effect. Interestingly, the mean birth year for third-born or later-born children is earlier than that of second-born children. One reason is that the average birth year of children from families with more than two children is 1976, which is earlier than the average of 1980 for two-child families. The mean maternal age at the birth of children in our sample is 25.54. With an average age of nearly 39, 81% of children are married, and 69% of children are living in rural areas. Compared with the mean of 9.22 years of schooling of their children, parents in our sample have a lower mean of 5 years. The average number of children for parents is 2.75, while 31% of them live together with children. The mean annual income for parents is nearly 8,500 CNY, and 28% of them have low socioeconomic status.

3.4 Empirical model

To investigate the potential birth order effects on education, we estimate the following model with family fixed effects:

$$Edu_{ifwt} = C_1 + \beta_1 BirthOrder_{if} + \eta_1 X_{if} + \alpha_{fw} + \tau_t + u_{ifwt} \tag{3.1}$$

where Edu_{ifwt} represents years of schooling in wave w (w = 2010, 2012, 2014, 2016, 2018, 2020) for child i, born in year t, from family f. We use three measures for an individual's birth order, denoted as $BirthOrder_{if}$. The first one is the raw birth order, represented as an integer ranging from 1 to 7. This measurement aims to estimate the linear effect of the birth order. The second measurement is the relative birth order proposed by Ejrnæs & Pörtner (2004), which standardises an individual's raw birth order by their family size[37]. This method effectively reduces the correlation between the number of children and birth order from 0.54 (using raw birth order) to less than 0.01 (using relative birth order). Thirdly, we use a series of birth order dummies

[37] The relative birth order is given by $\frac{raw\ birth\ order-1}{number\ of\ children-1}$, which indicates an individual's birth position relative to their siblings. It equals 0 for the first-born children, while it equals 1 for the last-born children.

to capture the non-linear birth order effects on education, which is in line with the measurement in Weng et al. (2019) and J. H. Kim & Wang (2023). X_{if} is a vector of control variables including children's gender and maternal age dummies at birth. Because both $BirthOrder_{if}$ and X_{if} are fixed across waves, we omit wave w in the subscript. In this model, omitted children's specific characteristics (like their IQ) are less likely to bias the birth order effects, as they are unlikely to be correlated with birth order. τ_t represents the birth year fixed effects, capturing the cohort heterogeneity that affects education. To allow for serial correlation within the family, idiosyncratic errors u_{ifwt} are clustered at the family-by-wave level. This clustering strategy is in line with S. Zhang et al. (2023).

α_{fw} is family fixed effects. Unlike standard fixed effects, where each family appears multiple times across different waves, family fixed effects observe each family multiple times across different children within the same family. Consequently, Equation 3.1 investigates within-family variations in the birth order, examining how these variations affect educational outcomes. Using family fixed effects enables me to eliminate unobserved family-level heterogeneity that could potentially bias the estimation. This approach also differences out variables that are consistent across all children in the family, such as parents' education, age, income, religion, and so on. Given that each family may appear more than once in our sample, we treat the same family in different waves as a different family. Considering the same family across different waves as the same panel unit will fail to construct family fixed effects due to repeated time values within the panel. Therefore, α_{fw} is also the family-by-wave fixed effects. Our strategy of using family-by-wave fixed effects is consistent with S. Zhang et al. (2023). Equation 3.1 is estimated by the fixed-effects model.

To explore the relationship between birth order and old-age support, we estimate the following model:

$$Y_{ifw} = C_2 + \beta_2 BirthOrder_{if} + \eta_2 X_{ifw} + \alpha_{fw} + e_{ifw} \tag{3.2}$$

where Y_{ifw} is the old-age support child i provides to his or her parent f in wave w (w = 2016, 2018, 2020), which includes in-kind support (such as the frequency of the

child's housework and visits to their parents), monetary support (such as whether the child provides financial assistance and, if so, the amount given per month), and the quality of the relationship between parents and children. Same as Equation 3.1, the birth order denoted as $BirthOrder_{if}$ is measured in three ways. Different from Equation 3.1, X_{ifw} contains more control variables, including children's gender, age, marital status (yes or no), years of schooling, residency type (rural or urban), and whether the parent is co-resident with child i. α_{fw} is the family-by-wave fixed effects, and e_{ifw} is the idiosyncratic errors. Unlike Equation 3.1, we exclude the birth-year fixed effects, not only because they are expected to have much less impact on old-age support for parents than on education, but also to maintain parsimony. As Equation 3.2 is estimated by the fixed-effects model[38], its identification comes from within-family variations in the birth order across siblings.

3.5 Empirical results on birth order effects

Table 3.3 suggests a negative birth order effect on education in the full sample: as the raw birth order increases by one, years of schooling are expected to drop by 0.33 years on average. The relative birth order and birth order dummy specifications suggest a similar negative birth order effect. Within the same family, first-born and second-born children are, on average, 0.94 and 0.28 years more educated than third or later-born children, respectively. While the negative birth order effect observed in the full sample aligns with patterns seen in the developed world, it is interesting to explore the dynamics of birth order effects in China, given its rapid transition from extreme underdevelopment to a moderately prosperous society in recent decades. To achieve this, we split observations in the full sample into two groups: people born before a threshold year and people born in and after that year. We then investigate the birth order effect on both groups. To ensure sufficient observations in both groups, we use threshold years from 1964 to 1982. This approach allows me to observe not only the

[38]Because a larger number of children transfer 0 to their parents, we employ the Tobit model (without fixed effects) in the amount of monetary support regression to account for these corner solutions.

overall evolutionary trajectory of birth order effects but also the timing of when these effects have changed.

As shown in the left panel of Figure 3.4, a positive birth order effect is observed in older generations, however, it gradually weakens as younger generations are included and eventually reverses to a negative effect after including individuals born in 1979 or later. This result suggests that for older generations of Chinese, the pattern of birth order effects follows that seen in developing countries, while for younger generations of Chinese, the pattern aligns with developed countries. It's worth noting that individuals born before 1967 experience a positive birth order effect (see the left panel), while those born in or after that year face a negative birth order effect (see the right panel)[39]. This contrasting pattern indicates that it is primarily the younger generations who drive the negative birth order effect observed in the full sample.

In addition, our results suggest two important cutoff years in which the sign of birth order effects changes. As depicted in the left panel of Figure 3.4, 1967 is the last year a significant positive birth order effect is observed, while 1979 is the last year an insignificant negative birth order effect is observed. As a result, 1967 is the cutoff year when the positive birth order effect disappears, and 1979 is the cutoff year when the negative birth order effect begins to emerge. The two cutoff years play a crucial role in explaining the discrepancies in the birth order literature in China, as these discrepancies mainly stem from using samples taken from different generations. On the one hand, Weng et al. (2019) find a positive birth order effect in older generations with an average birth year of 1966, closely matching the 1967 cutoff indicated by our results. On the other hand, J. H. Kim & Wang (2023) and S. Zhang et al. (2023) report a negative birth order effect for younger generations born after 1995 and 1978, respectively, which are after or close to the 1979 cutoff year suggested by our findings.

Detailed results for subsamples based on the two cutoff years are presented in Table 3.3 columns (4) to (9). On average, for individuals born before 1967, each additional position in raw birth order is associated with an increase of nearly 1.5 years in educa-

[39]The same phenomenon is observed in all threshold years from 1964 to 1967.

tion, whereas for those born in or after 1979, it corresponds to a decrease of about 0.7 years in education. Results based on relative birth order and birth order dummy specifications are consistent with the raw birth order, except for column (5). Although it is statistically insignificant, the positive sign still suggests a positive birth order effect that aligns with the other two specifications.

While the results on birth order effects primarily reflect inequalities in years of schooling among siblings, they may also imply within-family disparities in parental investment. A more rigorous way is to regress parental investment on the birth order, however, CFPS lacks such information on childhood parental investment. Using children's education as a proxy is less rigorous but may indirectly reflect within-family inequalities in parental investment. This approach is in line with R. Guo & Zhang (2020), who use children's education and health as a proxy of child investment. In this context, our results also implicitly suggest that Chinese parents invest more in later-born children in older generations and more in earlier-born children in younger generations.

3.6 Mechanisms behind evolving birth order effects

This section explores the mechanism in an exploratory manner due to data limitations, such as the lack of precise information on school dropout timing and detailed data on parental investment during childhood.

3.6.1 School dropout among older siblings

A positive birth order effect is observed in many developing countries, where households often face liquidity constraints and rely on child labour. In this context, parents typically tend to encourage older siblings to leave school and work as child labourers, sending earnings back to support the family. Benefiting from this additional income, later-born children are more likely to stay in school longer (Emerson & Souza, 2008). Despite the scarcity of empirical evidence in China, scholars suggest that school dropout among older siblings is a main driving force behind the positive birth order

effect in other developing countries, such as Brazil (Emerson & Souza, 2008), Sub-Saharan countries (Tenikue & Verheyden, 2010), Ecuador (De Haan et al., 2014), and Cambodia (Kanayama & Yamada, 2024).

Evaluating whether school dropout leads to a positive birth order effect in China is challenging, as census data on child labour and school dropout rates from the 1940s to the 1990s are not available for public access. Given this limitation, we use high school completion rates from the CFPS and the Chinese General Social Survey (CGSS) as a proxy. The CGSS is the earliest national representative cross-sectional survey data, covering 11 waves from 2003 to 2018. The CGSS sample contains 94,622 individuals born between 1940 and 2000, while the CFPS sample includes 42,578 individuals. Figure 3.6 confirms that older generations in China have significantly lower high school completion rates than younger generations. The junior high school completion rate for people born in the 1960s is 25 percentage points lower than that for those born in the 1980s in the CFPS sample, which is similar to the difference in the senior high school completion rate. The CGSS sample suggests an even larger gap in high school education between the two groups. In addition, both samples indicate a general upward trend in the high school completion rate over time. This descriptive evidence suggests that school dropout was widespread among individuals born in the 1960s but became less common in younger generations. For a typical Chinese family with children born in the 1960s and 1970s, school dropout among older siblings could potentially lead to a positive birth order effect[40].

3.6.2 The Cultural Revolution and interrupted schooling

The second potential mechanism behind the positive birth order effect is the interrupted education caused by the Chinese Cultural Revolution. The Chinese Cultural Revolution, which started in 1966 and ended in 1976, led to more than six years of interrupted regular schooling (1966–1972) and eleven years of ceased university ed-

[40]This example demonstrates that the school dropout mechanism observed in other developing countries may also help explain the positive birth order effect in China. However, it may not explain the positive birth order effect results in section 3.5, as all observations in that sample are born before 1967.

ucation (1966–1977) nationwide[41]. The interrupted schooling is one potential factor that caused school dropout among older siblings born in the 1960s. Due to the lack of census data on education by each birth cohort, we use the Chinese General Social Survey (CGSS) to describe educational attainment for individuals affected by the Cultural Revolution. The first cohort affected by the Cultural Revolution consists of those born in 1960–1961, who turned six years old and began primary school in 1966, at the start of the Cultural Revolution.

Figure 3.5 shows that the Cultural Revolution indeed led to a substantial decrease in primary and junior high school completion rates. A decreasing trend in primary school completion rates is observed for people born since 1961, who belong to the first or second cohort affected by the Cultural Revolution. This drop in primary school completion rates also has a negative impact on junior high school completion. The lowest primary school completion rate is observed among people born in 1966 to 1967. This is because schools reopened in 1972, when the later-born cohorts reached the age of six (Meng & Gregory, 2002). This potentially explains why the positive birth order effect disappeared for people born in 1967 or later.

3.6.3 Child labour laws and resource dilution

As Emerson & Souza (2008) argue, the existence of a positive birth order effect crucially depends on the prevalence of child labour. While child labour was widespread in China during the 1970s and 1980s, as depicted in Figure 3.6, it has been significantly curbed since the 1990s due to the enactment of child labour laws. The Chinese government implemented the *Temporary provisions on State Businesses Employing Staff* in 1986 to restrict the minimum age of working in the public sector to 16. This was followed by the *Provisional Regulations of the People's Republic of China on Private Enterprises* in 1988, which extended this age restriction to private sectors. The Provisions on the *Prohibition of Using Child Labour* issued in 1991 further clarified penalties associated with the employment of child labour, including fines, business

[41] For a detailed description, see Deng & Treiman (1997) and Meng & Gregory (2002).

suspensions, and revocation of business licenses (Cui et al., 2019).

Since then, Chinese families have been less likely to pull their children out of school, despite strong incentives to do so, as the opportunity for children to work as labourers, which was abundant in the 1970s and early 1980s, has greatly diminished. Meanwhile, as most Chinese families still faced substantial credit constraints[42], parents tended to allocate the majority of resources to the earlier-born children to ensure they remained enrolled in school, while the resources available for later-born children became increasingly limited. The phenomenon described above is referred to as resource dilution, which is considered a major reason for the negative birth order effect in J. H. Kim & Wang (2023) and Fors & Lindskog (2023).

The ideal way to test the resource dilution mechanism is by examining whether a child's birth order affects the amount of parental investment they receive. However, CFPS lacks such information. In this case, we test this mechanism by splitting the sample based on the extent of resource dilution each family experiences. Intuitively, families tend to experience a high level of resource dilution if they have low socioeconomic status (SES), have more children, live in rural China, or lack a high income. In contrast, those with higher SES, fewer children, urban residency, or higher income often face a low level of resource dilution[43]. In this scenario, the magnitude of the birth order effect is expected to be greater in families experiencing high compared to those experiencing low resource dilution. In this paper, a family is defined as low SES if both parents live in rural areas and are illiterate. We define a family as high income if the annual income of the parents exceeds the 80th percentile in our sample, which is 18,000 CNY[44].

[42] According to the World Bank database, the average GDP per capita in the 1990s was 631 USD for China, while it was 29,766 USD for the United States.

[43] Our choice of whether a family experiences high or low resource dilution is in line with the literature. To test resource dilution, J. H. Kim & Wang (2023) split the sample by the number of children, while S. Zhang et al. (2023) interact birth order with household income. Oliveira (2019) uses SES as a proxy of resource constraints.

[44] Due to the larger number of missing values in the income variable since 2014 in the CFPS, we restrict the analysis to the years 2010 and 2012 for high-income and low-income subsamples.

As shown in Table 3.4, the magnitude of the birth order effect is greater for families that face a high level of resource dilution. An additional position in the raw birth order is expected to reduce children's years of education by 1.32 years in low SES families, which is approximately three times greater than the reduction of 0.42 years found in higher SES families. A comparable pattern is observed when comparing families with only two children to those with more than two children[45]. Furthermore, the negative birth order effect is observed only among families from rural areas and those with relatively low income. This implies that the result found in Table 3.3 column (7) is primarily driven by rural families and families with relatively low income. Our results in Table 3.4 confirm that resource dilution is one possible mechanism behind the negative birth order effect.

In addition, the resource dilution mechanism suggests that as the number of siblings increases, the resources allocated to each child decrease. First-born children, therefore, enjoy undiluted resources during their early childhood before the birth of other siblings. That is to say, the magnitude of the birth order effect is expected to be larger for first-born children. As shown in the right panel of Figure 3.7, within the same family, first-born children are more educated than second-born children, and this difference is statistically significant. This result reveals that parents indeed allocate more resources to their oldest children and leave fewer resources to later-born children, suggesting resource dilution is one potential mechanism underlying the negative birth order effects. Besides, it is worth noting that the difference in education between first and second-born children is negligible among older generations (see the left panel of Figure 3.7). This suggests resource dilution is unlikely to be the cause of the positive birth order effect, as parents tend to have a flat investment[46] among earlier-born children, who are likely to drop out of school and work as child labourers.

[45] We use relative birth order in this subsample to reduce the correlation between birth order and family size.

[46] Here, children's education is a proxy of parental investment. As mentioned earlier, this proxy is less rigorous but birth order effects on education may indirectly reflect the within-family inequality in parental investment.

3.7 Empirical results on old-age support

If the birth order effects implicitly reveal within-family parental investment preferences[47], it is interesting to explore whether siblings who were more privileged in childhood provide more old-age support to their parents in adulthood than their less-favoured siblings. As shown in Table 3.5 panel (a), in the full sample, differences in the old-age support provided by siblings are negligible, except for the monthly amount of monetary support, where later-born children tend to give more. This result is counter-intuitive. As the full sample results in Table 3.3 suggest a negative birth order effect (earlier-born advantage), earlier-born children—who are expected to provide more support to their parents in return for the favours they received during childhood—do not actually offer more support and, in fact, provide even less monetary support than their siblings.

A similar counter-intuitive phenomenon is found in subsamples. As shown in Table 3.5 panel (b), among older generations who experience a positive birth order effect (later-born advantage), later-born children—who are expected to offer more support to their parents in return for the benefits they received during childhood—actually provide similar levels of old-age support as their siblings and even have poorer relationship quality with their parents. Similar to the full sample result, Table 3.5 panel (c) suggests comparable old-age support levels among advantaged and disadvantaged siblings in younger generations. Therefore, from the perspective of efficient investment, parents do not achieve optimal results, as the returns to investing in the education of earlier-born children are no different from those of later-born children.

The altruistic motive emerges as a key reason for siblings providing equal levels of old-age support to their parents. In China, the altruistic motive is partly shaped by laws and cultural norms (Oliveira, 2016). On the one hand, Article 49 of the Chinese Constitution mandates that children are responsible for providing old-age support to elderly parents, with failure to do so leading to both moral condemnation and legal

[47]Specifically, a negative birth order effect implies parents invest more in earlier-born children. A positive birth order effect implies parents invest more in later-born children.

penalties, including imprisonment (Wu & Li, 2014). On the other hand, cultural norms such as Confucian filial piety (xiao) regard providing in-kind and monetary support to elderly parents as a traditional virtue that everyone is expected to uphold (Chen & Fang, 2021; Rao & Zhang, 2024). Constrained by both legal obligations and cultural expectations, even siblings who have received less favour from their parents during childhood tend to provide adequate old-age support to parents when they grow up. In addition, Table 3.5 panel (a) shows that more educated children tend to provide less in-kind support to parents. This is potentially attributed to the higher opportunity cost of time spent by the more educated children (Oliveira, 2019; Ho, 2019). This higher opportunity cost may also explain why privileged children offer the same level of in-kind support as their siblings, rather than a significantly higher one.

Our results also shed light on whether monetary and in-kind support function as substitutes or complements, a topic that remains underexplored (R. Guo & Zhang, 2020). If earlier-born children provide more monetary support to compensate for less in-kind support they give compared to later-born children, monetary and in-kind support should function as substitutes among earlier-born and later-born siblings. Alternatively, if earlier-born children provide both greater monetary and in-kind support than later-born children, monetary and in-kind support should function as complements among siblings. In other words, monetary and in-kind support act as substitutes if the coefficients of birth order for monetary and in-kind support have opposite signs (one positive and the other negative), or they act as complements if the coefficients have the same sign (both positive or both negative). However, the coefficients of birth order in Table 3.5 neither exhibit the same nor opposite signs, rather, they are not significant among siblings. This indicates monetary and in-kind support are neither substitutes nor complements among earlier-born and later-born siblings. Despite that, monetary and in-kind support serve as substitutes between more educated and less educated children, as the signs of coefficients of years of schooling are opposite for monetary and in-kind support (see Table 3.5 panel (a)).

3.8 Gender composition

Son preference has been a widespread phenomenon throughout Chinese history. During the period when China was an agricultural society, men's greater contribution to farming labour was a key reason parents preferred sons (Arnold & Zhaoxiang, 1986). This advantage of sons also led parents to perceive them as more capable than daughters in providing old-age support (Das Gupta et al., 2003). Such support was particularly crucial for the majority of Chinese elders, as it served as their primary source of living in the absence of a reliable public social security system and sufficient retirement income (Rao & Zhang, 2024). The traditional mindset of son preference has continued to influence parents in contemporary China, even though women's and men's contributions to labour are comparable in modern society. One typical example of son preference is the decision by parents to have a sex-selective abortion in order to have a son during the time of the One Child Policy, which led to the widespread phenomenon of 'missing girls' (Ebenstein, 2010; Chen et al., 2013).

Not surprisingly, our results in Table 3.3 suggest that parents have a strong preference for their sons' education over their daughters'. In the full sample, brothers are on average 0.65 years more educated than sisters (column (1)), with this difference increasing to over 1 year for older generations (column (4)) and decreasing to approximately 0.5 years for younger generations (column (7)). Although there is still a considerable educational gap between sons and daughters, it has been substantially reduced by half in the past few decades[48]. However, results in Table 3.5 show that brothers do not provide more old-age support than sisters, contrary to the expectation in the literature mentioned above. In the full sample, brothers are on average 9% less likely to give parents monetary support than sisters. They also tend to have a poorer relationship with their parents compared to sisters, with the magnitude of 0.11 corresponding to roughly a 2.6% worse relationship relative to the average relationship quality of 4.28 (see Table 3.2).

[48]This may be attributed to the expansion of education, which has led more people to work in non-agricultural sectors and adopt more progressive gender-role attitudes (J. Zhang & Meng, 2025).

This result provides new evidence from contemporary China: sisters provide no less old-age support to their parents than brothers. In addition to the previously mentioned laws and cultural norms, greater gender equality in the workplace is another reason for this phenomenon. Unlike the agricultural era, the labour force participation rate in China for individuals aged 15–64 in 2010 was 84% for males and 70% for females (International Labour Organization, 2021). This 14-percentage-point gap is narrower than the average gap of 20 percentage points across 104 countries in the dataset from International Labour Organization (2021), placing China 48th in the ranking. Such an increase in female empowerment provides daughters with more resources and greater bargaining power with their husbands, which ultimately enables them to offer filial support to parents.

In addition, full sample results in Table 3.5 also indicate monetary and in-kind support are neither substitutes nor complements between brothers and sisters. Our results focus on different perspectives compared to those of Ho (2019). She finds that the correlation between past parental investment and children's current in-kind support is negative for sons and positive for daughters. In contrast, our paper focuses on old-age support provided directly by children, rather than support driven by past parental investment.

3.9 Conclusion

There have been long-standing discussions about inequalities in children's human capital since the 1960s. In recent years, however, these discussions have gradually shifted their focus to the intra-household context. An individual's birth order and gender are innate factors that lead to inequalities within the family. In developing countries, earlier-born children often receive less parental investment, as parents tend to pull them out of school and encourage them to work as child labourers due to liquidity constraints and less stringent child labour laws. Whereas in the developed world, earlier-born children often receive more parental resources, as they enjoy undiluted attention and investment before the arrival of their younger siblings.

This paper investigates within-family differences in children's human capital in China, a country that has been experiencing dramatic economic development and is transitioning from a developing to an approaching developed status in recent decades. Specifically, we examine the impact of birth order on years of education in contemporary China. We find that within the same family, earlier-born children are less educated than later-born children among the older generations of Chinese born before 1967, while in younger generations born after 1978, earlier-born children are more educated. This result suggests that birth order effects in China evolve from positive to negative. High school dropout rates and the lack of child labour laws are two key reasons behind the positive birth order effect, while resource dilution serves as the mechanism behind the negative birth order effect. Results on birth order effects may also imply parental investment in siblings: a positive (negative) birth order effect implies younger (older) siblings receive more parental investment. Furthermore, brothers typically achieve more years of education than sisters, while the gap has narrowed over time.

This paper also reveals that differences in old-age support between older and younger siblings, as well as between brothers and sisters are negligible. This implies siblings who received more parental investment in the past do not offer more old-age support to their parents now. The altruistic motive driven by both law and cultural norms is likely the potential reason. In addition, there is no evidence that in-kind support and monetary support function as substitutes or complements between siblings, or between brothers and sisters.

By examining the timing of birth order effects and the contextual differences in family investment strategies, this paper offers a more comprehensive understanding of the relationship between birth order, human capital investment, and intergenerational support in the face of China's economic transition. For countries in Central and South Asia undergoing rapid economic transitions similar to China, policymakers should prioritise support for earlier-born siblings, who face a higher risk of dropping out of school during the initial phase of economic change. As the transition matures,

attention should shift to later-born children, who are likely to receive less parental investment due to resource dilution

Figures for Chapter 3

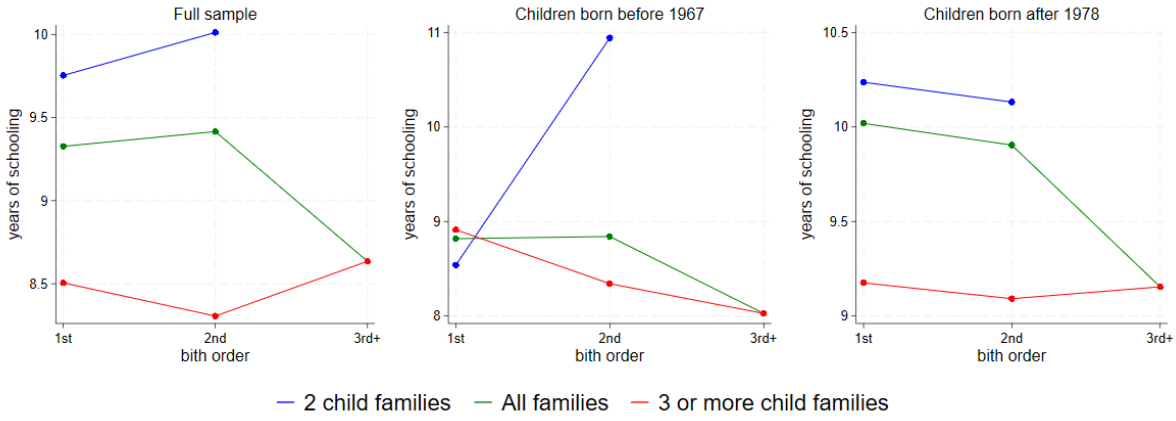


Figure 3.1: Descriptive birth order differences in children's education across family sizes. *Notes:* This figure plots the average years of schooling by birth order and family size. The sample size is 961 for children born before 1967 and 8,517 for those born after 1978.

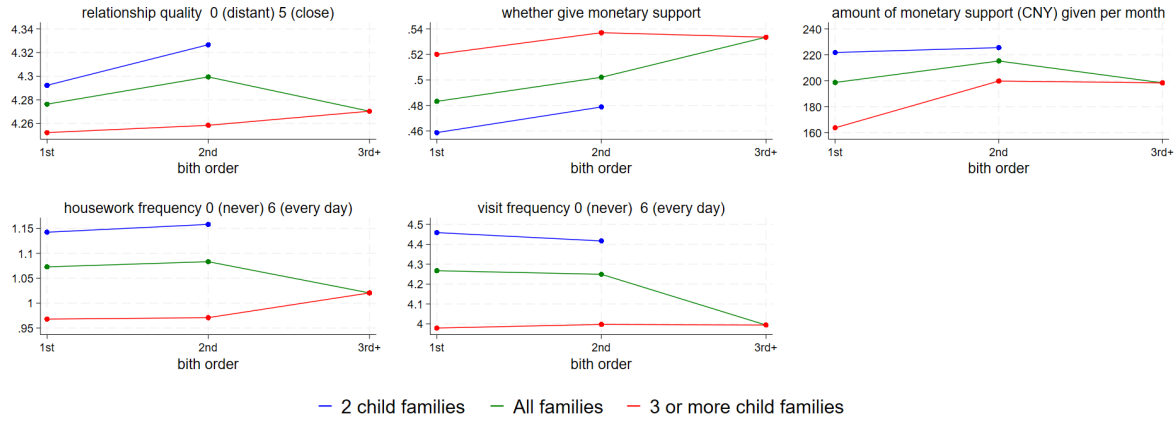


Figure 3.2: Descriptive birth order differences in old-age support across family sizes. *Notes:* This figure plots the average old-age support by birth order and family size. The sample sizes for each plot are 11,702, 11,689, 11,400, 11,673, and 11,668, respectively.

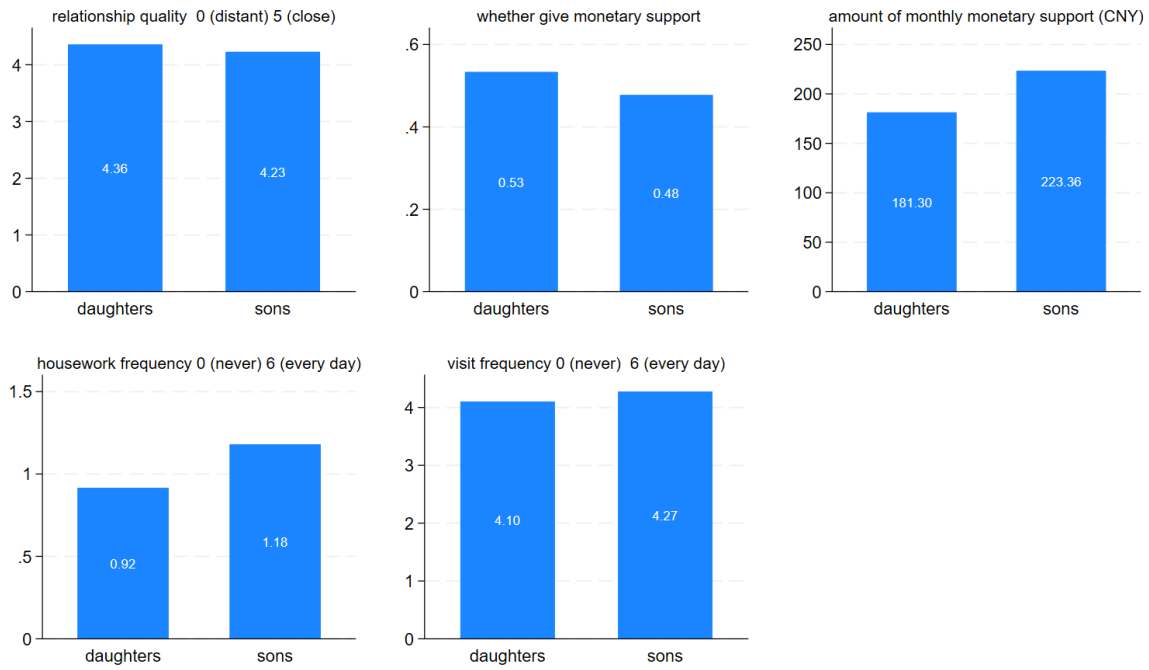


Figure 3.3: Descriptive differences in old-age support by children's gender. *Notes:* This figure plots the average old-age support by sons and daughters. The sample averages of relationship quality, probability of monetary support, amount of monetary transfer per month, housework frequency and visit frequency are 4.28, 0.5, 205.08, 1.07, and 4.2, respectively, as shown in Table 3.2.

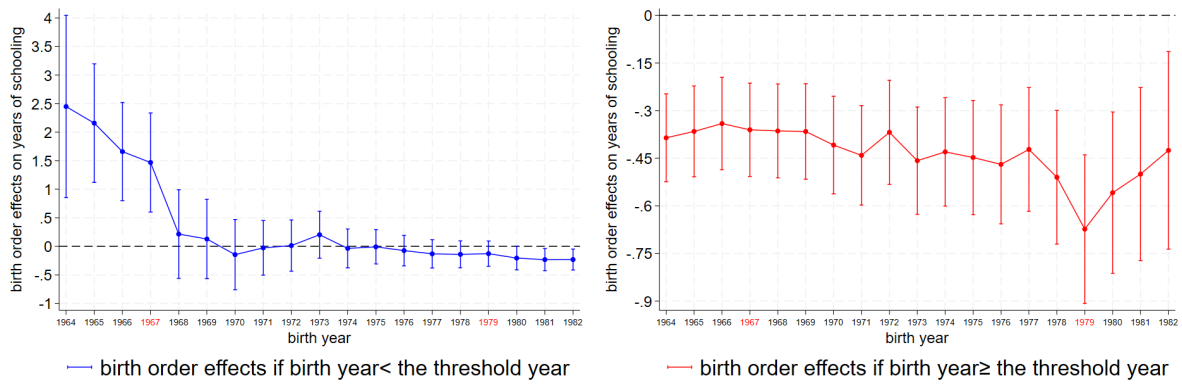


Figure 3.4: Birth order effects on years of schooling across birth years. *Notes:* Presented above are subsample results for Equation 3.1, which uses years of education as the dependent variable. The birth order variable is measured using the raw birth order. The blue subsample represents those born before the threshold year and the red subsample represents those born in or after that year. Results are point estimates of birth order effects with 95% confidence intervals. A positive (negative) coefficient indicates a positive (negative) birth order effect on education. To make the comparison easier, I combine the two subsamples shown above into a single chart, see Appendix Figure C3.

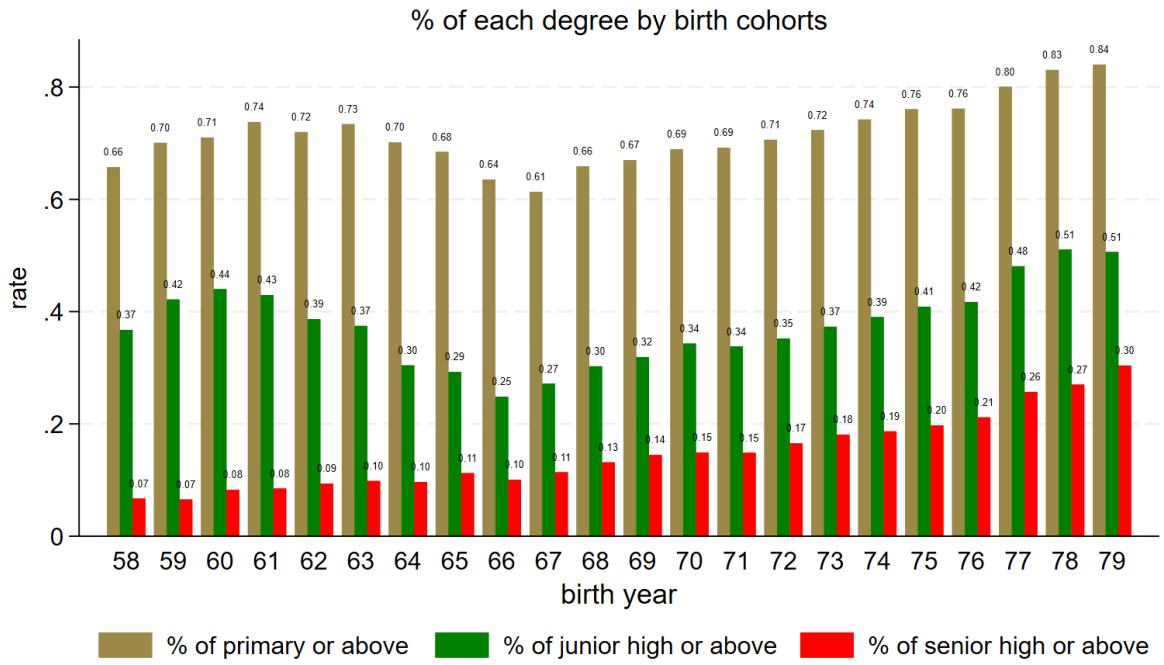


Figure 3.5: Percentage of each degree by birth cohorts. *Notes:* Data source: Chinese General Social Survey (CGSS). The sample size is 46,742. The bar chart illustrates the ratio of individuals in each birth cohort who have completed primary school, junior and senior high school or above to the total population of that cohort. The Chinese Cultural Revolution began in 1966.

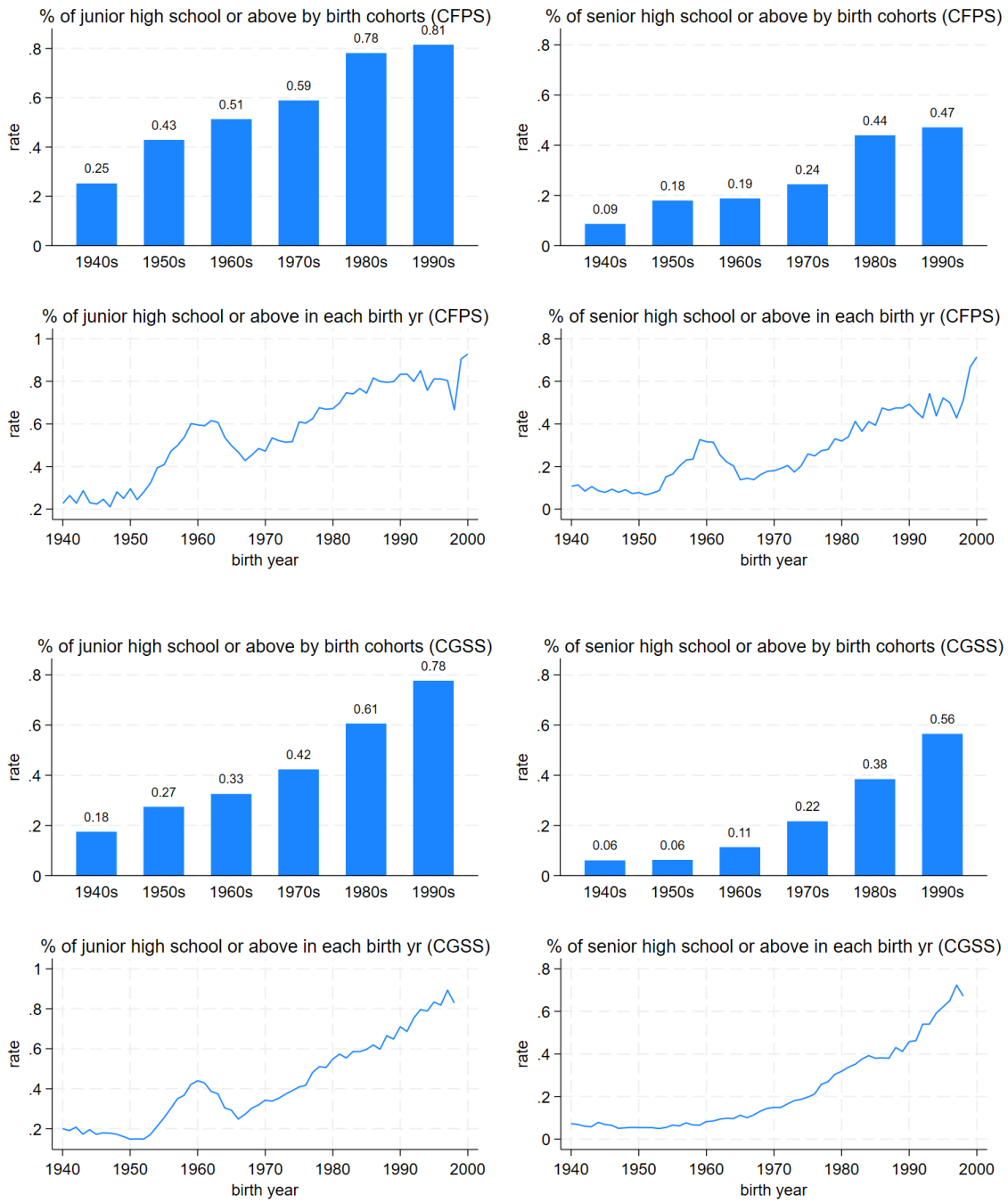


Figure 3.6: Junior and senior high school completion rates using CFPS and CGSS. *Notes:* The sample size is 42,578 for CFPS and 94,622 for CGSS. The bar chart illustrates the ratio of individuals in each birth cohort who have completed junior (senior) high school or above to the total population of that cohort. The line chart shows the ratio of individuals in each birth year who have completed junior (senior) high school or above to the total population born in that year.

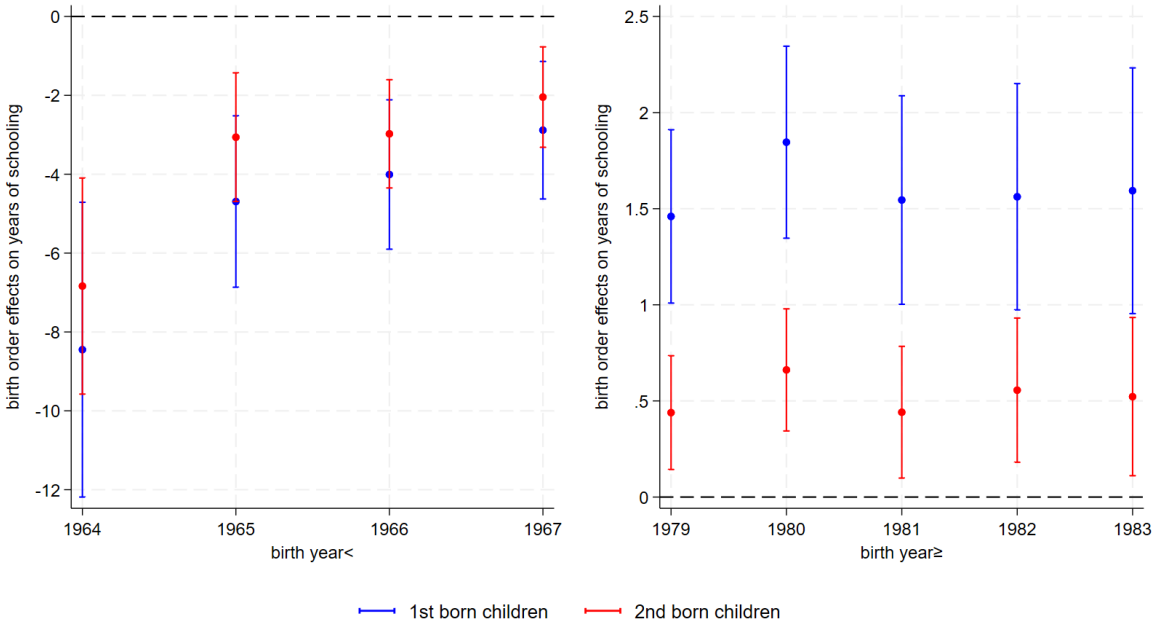


Figure 3.7: Mechanism behind the negative birth order effect: resource dilution. *Notes:* The dependent variable is the years of education. The birth order variable is measured using birth order dummies, with the reference group as third or later-born children. The left panel represents those born before a threshold year and the right panel represents those born in or after a threshold year. Results are point estimates of birth order effects with 95% confidence intervals.

Tables for Chapter 3

Table 3.1: Distribution of children's birth order and family size

Family size	birth order				observations
	1st child	2nd child	3rd child	4th+ child	
2 children	0.5	0.5			8,302
3 children	0.326	0.342	0.332		4,610
4 children +	0.223	0.222	0.224	0.331	2,941
observations	6,307	6,381	2,191	974	15,853

Notes: This is a cross-tabulation showing the percentage of each birth order for different family sizes. The maximum family size in the sample is 7.

Table 3.2: Descriptive statistics

	full sample		1st child		2nd child		3rd + child	
	mean	sd	mean	sd	mean	sd	mean	sd
children								
birth order	1.80	0.75	1.00	0.00	2.00	0.00	3.00	0.00
relative birth order	0.50	0.45	0.00	0.00	0.81	0.27	0.90	0.17
years of schooling	9.22	4.02	9.33	4.02	9.42	4.03	8.64	3.94
birth year	1978.27	6.8	1976.61	6.74	1979.75	6.64	1978.59	6.5
age in the survey year	38.77	7.92	40.38	7.85	37.17	7.78	38.76	7.7
maternal age at birth	25.54	4.29	23.08	3.2	26.23	3.84	29.07	4.1
married	0.81	0.4	0.84	0.36	0.77	0.42	0.79	0.41
agriculture (rural) Hukou	0.69	0.46	0.67	0.47	0.68	0.47	0.74	0.44
parents								
years of schooling	5.05	4.54	5.23	4.58	5.21	4.57	4.37	4.3
number of children	2.75	0.99	2.49	0.80	2.49	0.79	3.82	0.99
coresident	0.31	0.46	0.32	0.47	0.32	0.47	0.26	0.44
annual income (1000 CNY)	8.43	13.17	8.93	13.57	8.75	13.41	6.17	11.04
low SES	0.28	0.45	0.27	0.44	0.27	0.44	0.34	0.47
observations	15853		6307		6381		3165	
old-age support								
relationship	4.28	0.77	4.28	0.79	4.3	0.77	4.27	0.76
monetary support	0.5	0.5	0.48	0.5	0.5	0.5	0.53	0.5
monetary transfer per mth	205.08	597.54	198.68	556.66	215.25	665.72	198.4	536.72
housework frequency	1.07	1.94	1.07	1.94	1.08	1.96	1.02	1.92
visit frequency	4.2	1.61	4.27	1.62	4.25	1.57	3.99	1.64
observations	11702		4559		4562		2581	

Table 3.3: Birth order effects on years of schooling

	full sample			birth year < 1967			birth year ≥ 1979		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
raw birth order	-0.33***			1.47***			-0.67***		
	(0.07)			(0.44)			(0.12)		
relative birth order		-0.59***			0.65			-1.12***	
		(0.09)			(0.53)			(0.15)	
1st child			0.94***			-2.88***			1.85***
			(0.15)			(0.89)			(0.26)
2nd child			0.28***			-2.04***			0.66***
			(0.10)			(0.65)			(0.16)
male	0.65***	0.63***	0.63***	1.04***	1.09***	1.05***	0.48***	0.48***	0.47***
	(0.06)	(0.06)	(0.06)	(0.30)	(0.30)	(0.29)	(0.10)	(0.10)	(0.10)
family-by-wave FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
birth year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
observations		15,853			596			7,237	
number of families		6598			264			3317	
R-squared	0.728	0.728	0.729	0.787	0.782	0.789	0.720	0.722	0.723

Notes: Presented above are estimation results based on Equation 3.1, which uses the years of schooling as the dependent variable. All control variables inside $X_{if}^{'}$ are added, including children's gender and maternal age dummies at birth. Columns (1), (4), and (7) use raw birth order. Columns (2), (5), and (8) use the relative birth order. Columns (3), (6), and (9) use birth order dummies, with the reference group as third or later-born children. Standard errors shown in the parentheses are clustered at the family-by-wave level. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Table 3.4: Mechanism behind the negative birth order effect on education: resource dilution

	low SES		# of children		residency type		high income	
	yes	no	two	> two	rural	urban	yes	no
birth order	-1.32***	-0.42***	0.30**	0.66***	-0.76***	-0.52	-1.11	-1.33***
	(0.24)	(0.14)	(0.12)	(0.16)	(0.13)	(0.31)	(0.73)	(0.43)
male	1.11***	0.24**	-1.00***	-2.50***	0.58***	-0.08	-0.09	-0.31
	(0.21)	(0.11)	(0.18)	(0.41)	(0.11)	(0.32)	(0.65)	(0.42)
family-by-wave FE	✓	✓	✓	✓	✓	✓	✓	✓
birth year FE	✓	✓	✓	✓	✓	✓	✓	✓
observations	2,125	4,966	4,682	2,555	6,042	1,106	164	621
number of families	959	2291	2341	976	2752	525	78	295
R-squared	0.721	0.710	0.765	0.673	0.702	0.766	0.891	0.756

Notes: Results are based on the subsample of individuals born in or after 1979, which is the cutoff year when the negative birth order effect begins to emerge. The dependent variable is the years of schooling. All columns use raw birth order specification except for the number of children columns, which use the relative birth order. All control variables inside $X_{if}^{'}$ are added, including children's gender and maternal age dummies at birth. Standard errors shown in parentheses are clustered at the family-by-wave level. *** means p < 0.01, ** means p < 0.05, * means p < 0.1.

Table 3.5: Birth order and the old-age support children provide to their parents

	relationship	monetary support		in-kind support	
		yes or no	amount	visit freq.	housework freq.
Panel (a): full sample					
raw birth order	-0.00	0.00	9.27**	-0.03	0.02
	(0.01)	(0.01)	(4.15)	(0.03)	(0.04)
male	-0.11***	-0.09***	-10.24	0.05	-0.05
	(0.01)	(0.01)	(10.39)	(0.04)	(0.04)
years of schooling	0.01**	0.01***	10.17***	-0.02***	-0.02**
	(0.00)	(0.00)	(1.54)	(0.01)	(0.01)
observations	11,702	11,689	11,400	11,668	11,673
number of families	4,638	4,633	4,532	4,625	4,628
Panel (b): YOB $<$ 1967					
raw birth order	-0.21*	-0.01	-30.50	-0.08	0.14
	(0.11)	(0.08)	(35.35)	(0.26)	(0.42)
observations	415	414	404	415	415
number of families	185	185	181	185	185
Panel (c): YOB $\geq$ 1979					
raw birth order	-0.01	0.02	12.29	-0.05	0.01
	(0.03)	(0.02)	(10.86)	(0.07)	(0.08)
observations	3,756	3,754	3,656	3,748	3,746
number of families	1,690	1,689	1,645	1,686	1,686

Notes: Estimation results are based on Equation 3.2, which uses the old-age support children provide to their parents as the dependent variable. Children's relationship quality with their parents is measured on a scale from 0 (distant) to 5 (close). The monthly amount of monetary support is measured in CNY. The frequency of the child's visits and housework to their parents are measured as integers ranging from 0 (never) to 6 (every day). All control variables inside $X_{ifw}^{'}$ are added, including children's gender, age, marital status (yes or no), years of schooling, residency type (rural or urban), and whether the parent is co-resident with child i. All columns include family-by-wave fixed effects, except for the amount of monetary support, which uses a Tobit model. Parents' ages are all above 60. Standard errors shown in the parentheses are clustered at the family-by-wave level. *** means $p < 0.01$, ** means $p < 0.05$, * means $p < 0.1$.

Appendix C

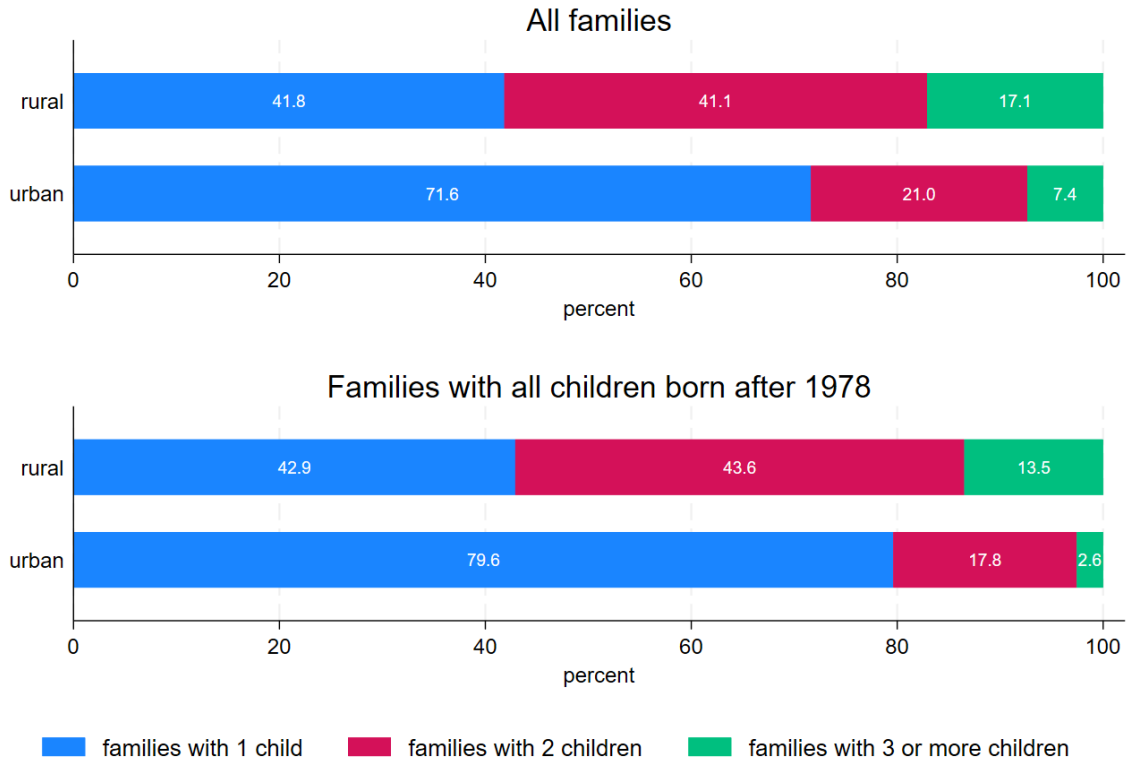


Figure C1: Proportion of families based on number of children. *Notes:* The data source is the China Family Panel Studies (CFPS). The One Child Policy period is from 1978 to 2013. All children were born before 2013. There are a total of 33,028 parent-child pairs, with approximately 71.6% living in rural areas. There are 27,309 parent-child pairs where all children were born after 1978, with 72.3% living in rural areas.

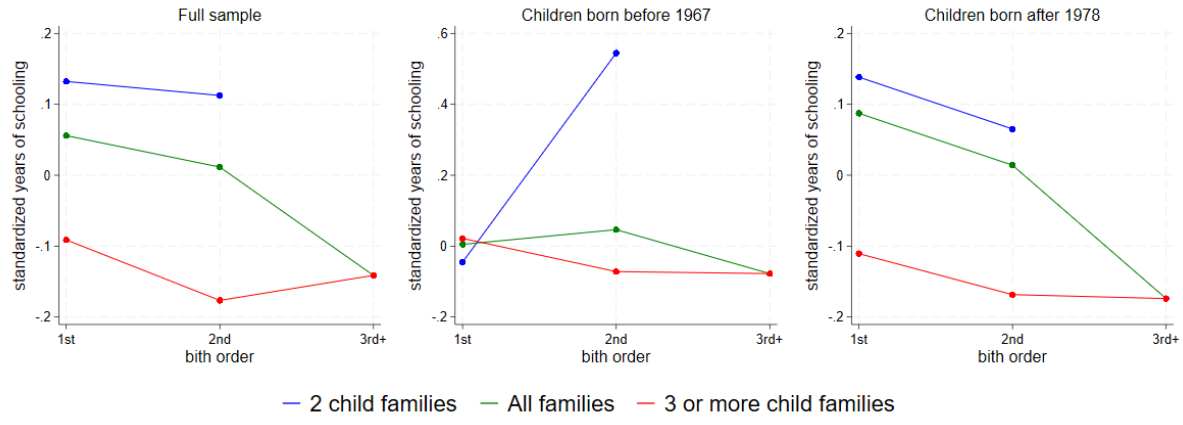


Figure C2: Descriptive birth order differences in children's standardized education across family sizes. *Notes:* This figure plots the average standardized years of schooling by birth order and family size. The sample size is 961 for children born before 1967 and 8,517 for those born after 1978.

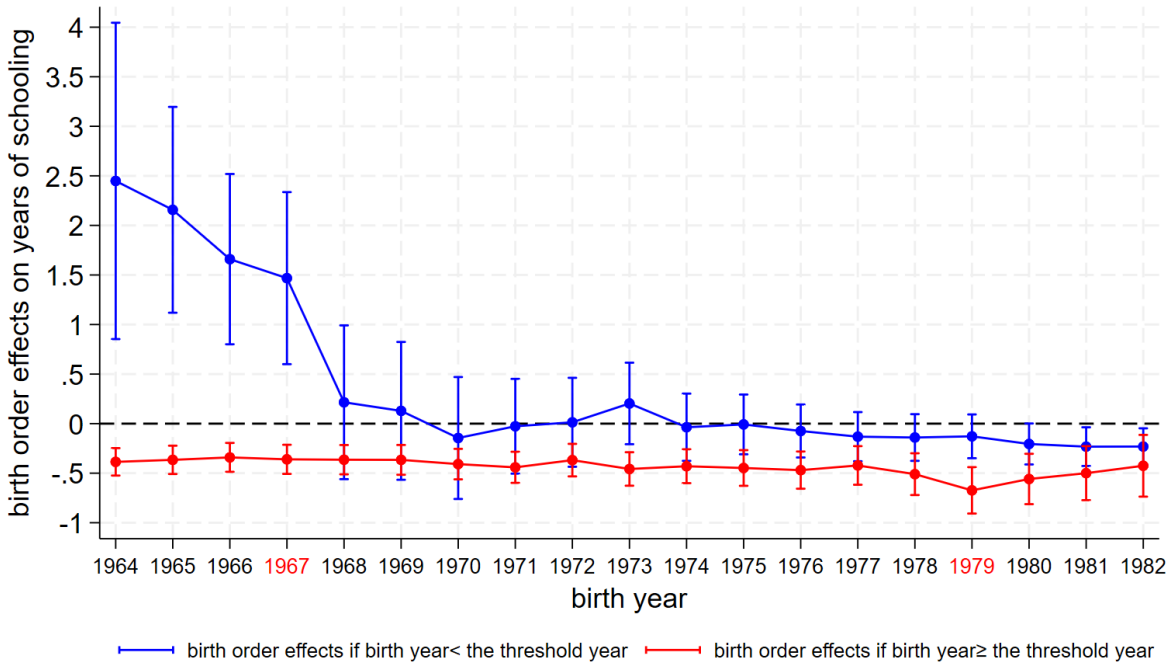


Figure C3: Birth order effects on years of schooling across birth years. *Notes:* Presented above are subsample results for Equation 3.1, which uses years of education as the dependent variable. The birth order variable is measured using the raw birth order. The blue subsample represents those born before the threshold year and the red subsample represents those born in or after that year. Results are point estimates of birth order effects with 95% confidence intervals. A positive (negative) coefficient indicates a positive (negative) birth order effect on education. This figure combines the two subsamples shown in Figure 3.4 into a single chart to facilitate comparison.

Conclusion

In recent decades, a growing concern has emerged regarding human capital development in developing countries. Helping less-advantaged groups create better pathways for their human capital development is a central objective for policymakers in these nations. One effective way to achieve this is by encouraging internal migration, which helps address skill mismatches by relocating the supply of labour to areas where the demand for specific skills is better matched. Intergenerational dynamics also play a critical role in the human capital development of both parents and children. For example, children with higher levels of education can positively influence their parents' mental health by providing more old-age support, increasing parental expenditures, and enhancing their parents' perceived social status. Conversely, disparities in parental investment among siblings during childhood can significantly influence educational opportunities, leading to long-term negative effects on the human capital development of less-favoured siblings.

In the first chapter, we investigate how effectively internal migration improves formal employment opportunities in Indonesia. Self-selection and reverse causality are two main issues in obtaining a causal relationship. Using 20 years of panel data from the Indonesia Family Life Survey (IFLS) and applying a panel fixed-effects estimator, we account for unobserved personal traits and effectively address the selection bias. We then use an instrumental variable approach to address the issue of reverse causality. Inspired by the significant role of agriculture in Indonesia, we exploit variations in precipitation to exogenously shift individual migration decisions, as fluctuations in rainfall have a substantial impact on household income and employment opportunities

in Indonesia. After controlling for a series of labour market indicators, such as nighttime lights and agricultural GDP, rainfall affects formal employment only through internal migration. Furthermore, to account for the binary nature of both formal employment (the dependent variable) and internal migration (the endogenous variable), we apply the random-effects probit regression model with endogenous covariates as an alternative econometric approach to the panel fixed-effects two-stage least squares model (FE-2SLS).

Our main result shows that internal migration increases the probability of formal employment by approximately 16 percentage points. This substantial increase in formal employment opportunities serves as strong evidence that internal migration effectively improves the human capital development of less-advantaged groups in Indonesia. In addition, we find the impact of migration rises to 25 percentage points for male migrants, while it drops to 7 percentage points for female migrants. Our estimation results are robust across all four different rainfall measurements. One limitation of this paper is the use of a random-effects probit regression model, which treats the unobserved individual heterogeneity as a random effect. While unobserved heterogeneity, such as risk aversion and internal locus of control, is likely to be correlated with migration decisions, the random-effects estimator could potentially be biased, although it provides results similar to those of the FE-2SLS.

We extend the investigation of human capital development to an intergenerational context in the second chapter, where we examine how children's education affects parents' mental health in China. To draw causality, it is important to account for omitted variable bias, such as genetic factors that affect both education and mental health, as well as reverse causality, where parental mental health may also influence children's education. To address the endogeneity issue, we use children's exposure to China's compulsory schooling law as an instrumental variable for their years of education. A common finding in the literature is that, on average, children's education has no impact on parental mental health (Ma, 2019; Cornelissen & Dang, 2022). This insignificant average effect is also observed in our paper. Under the circumstances

where the average effect is not informative, it is more meaningful to explore the impact across the distribution of parental mental health. To achieve this, we apply the Generalised Quantile Regression (GQR) approach proposed by Powell (2020). To the best of our knowledge, chapter two is the first paper to use a quantile approach to investigate the relationship between children's education and parental mental health.

We find a heterogeneous effect of children's education on elderly parents' mental health in China. Specifically, when parents have mild to moderate mental health issues (below the 0.7 quantile), children's education has no impact on their mental health. In contrast, when parents are depressed (at the 0.7 quantile and above), children's education has a positive and significant impact on their mental well-being. This heterogeneous result also indicates that the insignificant average effect arises from the insignificance in the left side of the distribution. Furthermore, our split sample analysis suggests that: (1) a son's education has a greater positive impact on parental mental health than a daughter's; and (2) the positive effect is greater for parents living in rural regions than those in urban regions. In addition, we find that when parents are not depressed, the schooling of the least educated children improves parents' mental well-being more. Whereas when parents have moderate to severe mental health issues, the schooling of the most educated children has a greater impact. Our results also show that the educational impact of first-born children is more substantial compared to that of second-born children.

In the third chapter, I continue exploring human capital development in an intergenerational context in China. I focus on the causes and consequences of within-family inequalities. Specifically, I study how birth order affects both one's years of education and the old-age support they provide to their parents. On the one hand, the impact of birth order on education implies parental investment disparities among siblings during children's childhood. Even among children from the same family, educational differences caused by this investment disparity can significantly affect their lifelong human capital development, particularly disadvantaging the less-favoured siblings. On the other hand, children's perceived differences in parental preferences among their sib-

lings during childhood may affect the old-age support they provide to their parents now. To the best of my knowledge, chapter three is the first paper to study the impact of birth order on old-age support.

This paper uses the family fixed effects to explore the within-family variations in birth order. I find that birth order effects on education in China evolve over time. A positive birth order effect is observed among older generations of Chinese, while a negative birth order effect is seen in younger generations. The prevalence of school dropout among earlier-born siblings is a key reason for the positive effect, while child labour laws and resource dilution are the main mechanisms behind the negative effects. Given China's tremendous economic growth in recent decades, the evidence of evolving birth order effects sheds light on how within-family inequality changes during the rapid economic transition. I also find that brothers have consistently outperformed sisters in years of education, although this gap has narrowed over time.

In addition, my results suggest differences in old-age support levels are negligible both among siblings and between brothers and sisters. This implies that children who are expected to provide more old-age support to their parents in return for the favours they received during childhood do not actually offer more support. The altruistic motive emerges as a key reason why siblings, as well as brothers and sisters, provide equal levels of old-age support to their parents. Furthermore, I find no evidence that in-kind support and monetary support function as substitutes or complements.

One limitation of chapter three is the absence of a theoretical model that explains birth order effects. In the literature, Wei et al. (2022) develop a model exploring how the endowments of first-born children influence parents' decisions to have a second child. Their model also explains that parents' unequal investments among siblings depend on the endowments of the first child and the predicted endowments of a potential second child. In the future, I could adapt their model to explain the changes in birth order effects in my paper. I could further extend their model to (1) explain how children's perceived differences in parental investments during childhood affect the old-age support they now provide to parents; and (2) determine the optimal parental

investment in each child, given that siblings will contribute equally to old-age support in the future.

In conclusion, this thesis investigates the impact of internal migration and intergenerational dynamics on human capital in the developing world. I particularly focus on using econometric tools such as the fixed-effects model and the instrumental variable approach to identify causal relationships. The results of this thesis highlight the following potential policy implications. From the perspective of policymakers, easing restrictions on regional mobility emerges as a powerful tool for improving employment opportunities and enhancing human capital. From the perspective of households, increasing educational investments and addressing within-family disparities benefit both parents' mental health and children's long-term human capital development. Society and families should work together to optimise resource allocation, focus on elders' mental well-being, and promote educational equality to build a better pathway for human capital accumulation.

References

Adams Jr, R. H., & Cuecuecha, A. (2010). The economic impact of international remittances on poverty and household consumption and investment in indonesia. *World Bank Policy Research Working Paper*(5433).

Alatas, S. (1993). Macro patterns of internal migration in indonesia, 1971-1990. *Majalah Demografi Indonesia*, *20*(40), 21–47.

Alatas, V., & Newhouse, D. (2010). Indonesia jobs report: Towards better jobs and security for all. *Washington, DC: Main Report, World Bank*.

Ampofo, A. (2021). Oil at work: natural resource effects on household well-being in ghana. *Empirical Economics*, *60*(2), 1013–1058.

Andrews, D., Caldera Sánchez, A., & Johansson, Å. (2011). Towards a better understanding of the informal economy. *OECD Economics Department Working Papers*. (No. 873, OECD Publishing, Paris, https://doi.org/10.1787/5kgb1mf88x28-en)

Andrews, I., Stock, J. H., & Sun, L. (2018). Weak instruments and what to do about them. *Lectures, NBER Summer Institute*.

Andrews, I., Stock, J. H., & Sun, L. (2019). Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, *11*(1).

Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, *106*(4), 979–1014.

Arnold, F., & Zhaoxiang, L. (1986). Sex preference, fertility, and family planning in china. *Population and Development Review*, 221–246.

Auci, S., & Pronti, A. (2023). Irrigation technology adaptation for a sustainable agriculture: A panel endogenous switching analysis on the italian farmland productivity. *Resource and Energy Economics*, *74*, 101391.

Bagger, J., Birchenall, J. A., Mansour, H., & Urzua, S. (2021). Education, birth order and family size. *The Economic Journal*, *131*(633), 33–69.

Baryshnikova, N. V., & Pham, N. T. (2019). Natural disasters and mental health: a quantile approach. *Economics letters*, *180*, 62–66.

Baryshnikova, N. V., Ploeckl, F., & Yunren, N. (2023). Does unsatisfactory subjective well-being of school children decrease their cognitive skill development? *Economic Record*, *99*, 50–66.

Bauernschuster, S., Falck, O., Heblich, S., Suedekum, J., & Lameli, A. (2014). Why are educated and risk-loving persons more mobile across regions? *Journal of Economic Behavior & Organization*, *98*, 56–69.

Bazzi, S., Gaduh, A., Rothenberg, A. D., & Wong, M. (2016). Skill transferability, migration, and development: Evidence from population resettlement in indonesia. *American Economic Review*, *106*(9), 2658–98.

Becker, G. (1960). An economic analysis of fertility. *Universities-National Bureau. Demographic and economic change in developed countries*.

Beegle, K., De Weerdt, J., & Dercon, S. (2011). Migration and economic mobility in tanzania: Evidence from a tracking survey. *Review of Economics and Statistics*, *93*(3), 1010–1033.

Behrman, J. R., & Taubman, P. (1986). Birth order, schooling, and earnings. *Journal of Labor Economics*, *4*(3, Part 2), S121–S145.

Bharati, T., Fakir, A. M., & Yoman, W. (2024). Internal migration and labor market outcomes in indonesia. *Economic Development and Cultural Change*, *72*(2), 997–1040.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). The more the merrier? the effect of family size and birth order on children's education. *The Quarterly Journal of Economics*, *120*(2), 669–700.

Booth, A. L., & Kee, H. J. (2009). Birth order matters: the effect of family size and birth order on educational attainment. *Journal of Population Economics*, *22*(2), 367–397.

Bracke, P., Pattyn, E., & Von dem Knesebeck, O. (2013). Overeducation and depressive symptoms: diminishing mental health returns to education. *Sociology of health & illness*, *35*(8), 1242–1259.

Brådvik, L. (2018). *Suicide risk and mental disorders* (Vol. 15) (No. 9). MDPI.

Bratti, M., & Mendola, M. (2014). Parental health and child schooling. *Journal of health economics*, *35*, 94–108.

Bryan, G., Chowdhury, S., & Mobarak, A. M. (2014). Underinvestment in a profitable technology: The case of seasonal migration in bangladesh. *Econometrica*, *82*(5), 1671–1748.

Caliendo, M., Cobb-Clark, D. A., Hennecke, J., & Uhlendorff, A. (2019). Locus of control and internal migration. *Regional Science and Urban Economics*, *79*, 103468.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, *69*(5), 1127–1160.

Castelló-Climent, A., Chaudhary, L., & Mukhopadhyay, A. (2018). Higher education and prosperity: From catholic missionaries to luminosity in india. *The Economic Journal*, *128*(616), 3039–3075.

Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of econometrics*, *18*(1), 5–46.

Chen, Y., & Fang, H. (2021). The long-term consequences of china's "later, longer, fewer" campaign in old age. *Journal of Development Economics*, *151*, 102664.

Chen, Y., Li, H., & Meng, L. (2013). Prenatal sex selection and missing girls in china: Evidence from the diffusion of diagnostic ultrasound. *Journal of Human Resources*, *48*(1), 36–70.

Cheng, T. C., Powdthavee, N., & Oswald, A. J. (2017). Longitudinal evidence for a midlife nadir in human well-being: Results from four data sets. *The Economic Journal*, *127*(599), 126–142.

Comola, M., & De Mello, L. (2011). How does decentralized minimum wage setting affect employment and informality? the case of indonesia. *Review of Income and Wealth*, *57*, S79–S99.

Conti, G., Heckman, J., & Urzua, S. (2010). The education-health gradient. *American economic review*, *100*(2), 234–238.

Cornelissen, T., & Dang, T. (2022). The multigenerational impacts of educational expansion: Evidence from vietnam. *Labour Economics*, *78*, 102243.

Coulmas, F. (2007). *Population decline and ageing in japan - the social consequences (1st ed.)*. Routledge. https://doi.org/10.4324/9780203962022.

Cui, Y., Liu, H., & Zhao, L. (2019). Mother's education and child development: Evidence from the compulsory school reform in china. *Journal of Comparative Economics*, *47*(3), 669–692.

Cutler, D. M., & Lleras-Muney, A. (2006). *Education and health: evaluating theories and evidence.* National bureau of economic research Cambridge, Mass., USA.

Das, J., Do, Q.-T., Friedman, J., McKenzie, D., & Scott, K. (2007). Mental health and poverty in developing countries: Revisiting the relationship. *Social science & medicine*, *65*(3), 467–480.

Das Gupta, M., Zhenghua, J., Bohua, L., Zhenming, X., Chung, W., & Hwa-Ok, B. (2003). Why is son preference so persistent in east and south asia? a cross-country study of china, india and the republic of korea. *The Journal of Development Studies*, *40*(2), 153–187.

De Haan, M. (2010). Birth order, family size and educational attainment. *Economics of Education Review*, *29*(4), 576–588.

De Haan, M., Plug, E., & Rosero, J. (2014). Birth order and human capital development: Evidence from ecuador. *Journal of Human Resources*, *49*(2), 359–392.

Deng, Z., & Treiman, D. J. (1997). The impact of the cultural revolution on trends in educational attainment in the people's republic of china. *American journal of sociology*, *103*(2), 391–428.

Downey, D. B. (2001). Number of siblings and intellectual development: The resource dilution explanation. *American psychologist*, *56*(6-7), 497.

Du, H., Xiao, Y., & Zhao, L. (2021). Education and gender role attitudes. *Journal of Population Economics*, *34*, 475–513.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American economic review*, *91*(4), 795–813.

Ebenstein, A. (2010). The "missing girls" of china and the unintended consequences of the one child policy. *Journal of Human resources*, *45*(1), 87–115.

Edmonds, E. V. (2006). Understanding sibling differences in child labor. *Journal of Population Economics*, *19*, 795–821.

Ejrnæs, M., & Pörtner, C. C. (2004). Birth order and the intrahousehold allocation of time and education. *Review of Economics and Statistics*, *86*(4), 1008–1019.

Elgin, C., Kose, M. A., Ohnsorge, F., & Yu, S. (2021). Understanding informality. (Available at SSRN: https://ssrn.com/abstract=3914265 or http://dx.doi.org/10.2139/ssrn.3914265)

Emerson, P. M., & Souza, A. P. (2008). Birth order, child labor, and school attendance in brazil. *World Development*, *36*(9), 1647–1664.

Esposito, L., Kumar, S. M., & Villaseñor, A. (2020). The importance of being earliest: birth order and educational outcomes along the socioeconomic ladder in mexico. *Journal of Population Economics*, *33*, 1069–1099.

Everding, J. (2019). *Heterogeneous spillover effects of children's education on parental mental health* (HCHE Research Paper No. 18). Hamburg. Retrieved from https://hdl.handle.net/10419/200978

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*(3), 953–973.

Fors, H. C., & Lindskog, A. (2023). Within-family inequalities in human capital accumulation in india. *Review of Development Economics*, *27*(1), 3–28.

Frase, R. T., Bauldry, S., Suitor, J. J., Gilligan, M., & Ogle, D. (2023). Adult children's education and mothers' psychological well-being: Do adult children's problems mediate this relationship? *The Journals of Gerontology: Series B*, *78*(3), 496–505.

Frölich, M., & Melly, B. (2013). Unconditional quantile treatment effects under endogeneity. *Journal of Business & Economic Statistics*, *31*(3), 346–357.

Funk, C. C., Peterson, P. J., Landsfeld, M. F., Pedreros, D. H., Verdin, J. P., Rowland, J. D., ... others (2014). A quasi-global precipitation time series for drought monitoring. *US Geological Survey data series*, *832*(4), 1–12.

Gu, B., Wang, F., Guo, Z., & Zhang, E. (2007). China's local and national fertility policies at the end of the twentieth century. *Population and development review*, 129–147.

Guo, M., Chi, I., & Silverstein, M. (2013). Sources of older parents' ambivalent feelings toward their adult children: The case of rural china. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *68*(3), 420–430.

Guo, R., Lin, L., Yi, J., & Zhang, J. (2020). The cross-spousal effect of education on health. *Journal of Development Economics*, *146*, 102493.

Guo, R., & Zhang, J. (2020). The effects of children's gender composition on filial piety and old-age support. *The Economic Journal*, *130*(632), 2497–2525.

Halimatussadiah, A., Cesarina, A., Siregar, A. A., Hanum, C., Wisana, D., Rahardi, F., ... others (2020). Thinking ahead: Indonesia's agenda on sustainable recovery from covid-19 pandemic. *Institute for Economic and Social Research Faculty of Economics and Business, Universitas Indonesia (LPEM FEB UI) and Ministry of National Development Planning/National Development Planning Agency (BAPPENAS)*.

Hamory, J., Kleemans, M., Li, N. Y., & Miguel, E. (2021). Reevaluating agricultural productivity gaps with longitudinal microdata. *Journal of the European Economic Association*, *19*(3), 1522–1555.

Han, S., & Lee, S. (2019). Estimation in a generalization of bivariate probit models with dummy endogenous regressors. *Journal of Applied Econometrics*, *34*(6), 994–1015.

Hardjono, J. (1988). The indonesian transmigration program in historical perspective. *International Migration (Geneva, Switzerland)*, *26*(4), 427–439.

Härkönen, J. (2014). Birth order effects on educational attainment and educational transitions in west germany. *European sociological review*, *30*(2), 166–179.

Harriss-White, B. (2010). Work and wellbeing in informal economies: The regulative roles of institutions of identity and the state. *World development*, *38*(2), 170–183.

Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American economic review*, *102*(2), 994–1028.

Hiriscau, A., & Pintea, M. (2024). Birth order, socioeconomic background and educational attainment. *Education Economics*, *32*(3), 395–412.

Ho, C. (2019). Child's gender, parental monetary investments and care of elderly parents in china. *Review of Economics of the Household*, *17*, 741–774.

Ho, C., & McGarry, K. (2025). Brothers, sisters, and support to older parents: separate spheres across and within support types? *Journal of Population Economics*, *38*(1), 25.

Hoebel, J., & Lampert, T. (2020). Subjective social status and health: Multidisciplinary explanations and methodological challenges. *Journal of Health Psychology*, *25*(2), 173–185.

Hsieh, C.-R., & Qin, X. (2018). Depression hurts, depression costs: The medical spending attributable to depression and depressive symptoms in china. *Health Economics*, *27*(3), 525-544. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.3604 doi: https://doi.org/10.1002/hec.3604

Huang, J. (2025). Does better parental education improve child health outcomes? evidence from the compulsory schooling law in china. *Economic Development and Cultural Change*, *73*(2), 835-873. Retrieved from https://doi.org/10.1086/729062 doi: 10.1086/729062

ILC-Japan. (2010). *A Profile of Older Japanese.* Retrieved from https://www.ilcjapan.org/agingE/POJ10.html

International Labour Organization. (2021). *Labour Force Statistics, LFS [database].* (Available from https://rshiny.ilo.org/dataexplorer59/?lang=en&id=EAP_DWAP_SEX_AGE_RT_A)

Jayawardana, D. (2022). Happily ever after? mental health effects of early marriage in indonesia. *Feminist Economics*, 1–25.

Jayawardana, D., Baryshnikova, N. V., & Cheng, T. C. (2023). The long shadow of child labour on adolescent mental health: a quantile approach. *Empirical Economics*, *64*(1), 77–97.

Jayawardana, D., Baryshnikova, N. V., & Pham, N. T. A. (2021). Can unconditional in-kind transfers keep children out of work and in school? evidence from indonesia. *The BE Journal of Economic Analysis & Policy*, *21*(3), 1035–1065.

Joshi, R., & Wooldridge, J. M. (2019). Correlated random effects models with endogenous explanatory variables and unbalanced panels. *Annals of Economics and Statistics*(134), 243–268.

Kanayama, Y., & Yamada, H. (2024). Changing effects of birth order on education over time: Evidence from cambodia. *The Journal of Development Studies*, 1–23.

Kantarevic, J., & Mechoulan, S. (2006). Birth order, educational attainment, and earnings: An investigation using the psid. *Journal of human resources*, *41*(4), 755–777.

Karim, J., Weisz, R., Bibi, Z., & ur Rehman, S. (2015). Validation of the eight-item center for epidemiologic studies depression scale (ces-d) among older adults. *Current Psychology*, *34*, 681–692.

Katungi, D., Neale, E., & Barbour, A. (2006). People in low-paid informal work:'need not greed'.

Kebschull, D. (1986). *Transmigration in indonesia: An empirical analysis of motivation, expectations and experiences*. Transaction Publishers.

Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S.-L., ... Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological medicine*, *32*(6), 959–976.

Kim, J. H., & Wang, S. (2023). Birth order effects and parenting behaviors. *China Economic Review*, *79*, 101950.

Kim, Y.-J. (2020). Born to be more educated? birth order and schooling. *Review of Economics of the Household*, *18*(1), 165–180.

Kleemans, M., & Magruder, J. (2018). Labour market responses to immigration: Evidence from internal migration driven by weather shocks. *The Economic Journal*, *128*(613), 2032–2065.

Kocornik-Mina, A., McDermott, T. K., Michaels, G., & Rauch, F. (2020). Flooded cities. *American Economic Journal: Applied Economics*, *12*(2), 35–66.

Lagakos, D., Marshall, S., Mobarak, A. M., Vernot, C., & Waugh, M. E. (2020). Migration costs and observational returns to migration in the developing world. *Journal of Monetary Economics*, *113*, 138–154.

Lee, C., Glei, D. A., Goldman, N., & Weinstein, M. (2017). Children's education and parents' trajectories of depressive symptoms. *Journal of Health and Social Behavior*, *58*(1), 86–101.

Lee, C. K. C., & Morrish, S. C. (2012). Cultural values and higher education choices: Chinese families. *Australasian Marketing Journal*, *20*(1), 59–64.

Lehmann, J.-Y. K., Nuevo-Chiquero, A., & Vidal-Fernandez, M. (2018). The early origins of birth order differences in children's outcomes and parental behavior. *Journal of Human Resources*, *53*(1), 123–156.

Lei, X., Shen, Y., Smith, J. P., & Zhou, G. (2017). Sibling gender composition's effect on education: evidence from china. *Journal of population economics*, *30*, 569–590.

Levy, S. (2010). *Good intentions, bad outcomes: Social policy, informality, and economic growth in mexico*. Brookings Institution Press.

Liang, Y., & Dong, Z. (2019). Has education led to secularization? based on the study of compulsory education law in china. *China Economic Review*, *54*, 324–336.

Liang, Y., & Yu, S. (2022). Does education help combat early marriage? the effect of compulsory schooling laws in china. *Applied economics*, *54*(55), 6361–6379.

Lillehagen, M., & Isungset, M. A. (2020). New partner, new order? multipartnered fertility and birth order effects on educational achievement. *Demography*, *57*(5), 1625–1646.

Liu, Z. (2021). Children's education and parental health: Evidence from china. *American Journal of Health Economics*, *7*(1), 95–130.

Loayza, N. V., & Rigolini, J. (2011). Informal employment: safety net or growth engine? *World Development*, *39*(9), 1503–1515.

Ma, M. (2019). Does children's education matter for parents' health and cognition? evidence from china. *Journal of Health Economics*, *66*, 222–240.

Maccini, S., & Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review*, *99*(3), 1006–26.

Mamo, N., Bhattacharyya, S., & Moradi, A. (2019). Intensive and extensive margins of mining and development: Evidence from sub-saharan africa. *Journal of Development Economics*, *139*, 28–49.

Mechoulan, S., & Wolff, F.-C. (2015). Intra-household allocation of family resources and birth order: evidence from france using siblings data. *Journal of Population Economics*, *28*, 937–964.

Meng, X., & Gregory, R. G. (2002). The impact of interrupted education on subsequent educational attainment: A cost of the chinese cultural revolution. *Economic Development and Cultural Change*, *50*(4), 935–959.

Mirowsky, J. (2017). *Education, social status, and health*. Routledge.

Missinne, S., Vandeviver, C., Van de Velde, S., & Bracke, P. (2014). Measurement equivalence of the ces-d 8 depression-scale among the ageing population in eleven european countries. *Social science research*, *46*, 38–47.

Möller, E. L., Majdandžić, M., de Vente, W., & Bögels, S. M. (2013). The evolutionary basis of sex differences in parenting and its relationship with child anxiety in western societies. *Journal of Experimental Psychopathology*, *4*(2), 88–117.

Mondragón-Vélez, C., Peña, X., Wills, D., & Kugler, A. (2010). Labor market rigidities and informality in colombia [with comment]. *Economía*, *11*(1), 65–101.

Mueller, V., Sheriff, G., Dou, X., & Gray, C. (2020). Temporary migration and climate variation in eastern africa. *World development*, *126*, 104704.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica: journal of the Econometric Society*, 69–85.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the us labor market. *The Quarterly Journal of Economics*, *118*(2), 549–599.

Mustafa, A., Jungari, S., & Shekhar, C. (2023). Relationship between children's educational attainment and parental depression among indian elderly parents. *Asian Population Studies*, 1–21.

Nguyen, L. D., Raabe, K., & Grote, U. (2015). Rural–urban migration, household vulnerability, and welfare in vietnam. *World Development*, *71*, 79–93.

Nuevo-Chiquero, A., Vidal-Fernandez, M., & Lehmann, J.-Y. (2022). The birth order effect: A modern phenomenon? *Life Course Centre Working Paper*(2022-18).

Olea, J. L. M., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, *31*(3), 358–369.

Oliveira, J. (2016). The value of children: Inter-generational support, fertility, and human capital. *Journal of Development Economics*, *120*, 1–16.

Oliveira, J. (2019). Birth order and the gender gap in educational attainment. *Review of Economics of the Household*, *17*, 775–803.

Patrinos, H. A., & Psacharopoulos, G. (1997). Family size, schooling and child labor in peru–an empirical analysis. *Journal of population economics*, *10*, 387–405.

Pei, Y., Cong, Z., & Wu, B. (2020). Education, adult children's education, and depressive symptoms among older adults in rural china. *Social science & medicine*, *253*, 112966.

Peng, S., Bauldry, S., Gilligan, M., & Suitor, J. J. (2019). Older mother's health and adult children's education: Conceptualization of adult children's education and mother-child relationships. *SSM-Population Health*, *7*, 100390.

Pérez, J. P. (2020). The minimum wage in formal and informal sectors: Evidence from an inflation shock. *World Development*, *133*, 104999.

Perry, G. (2007). *Informality: Exit and exclusion.* World Bank Publications.

Piracha, M., & Vadean, F. (2010). Return migration and occupational choice: Evidence from albania. *World Development*, *38*(8), 1141–1155.

Powell, D. (2020). Quantile treatment effects in the presence of covariates. *Review of Economics and Statistics*, *102*(5), 994–1005.

Radloff, L. S. (1977). The ces-d scale: A self-report depression scale for research in the general population. *Applied psychological measurement*, *1*(3), 385–401.

Rao, Z., & Zhang, Y. (2024). Rely on children or work longer? the impact of fertility and child gender on old-age labor supply. *World Development*, *181*, 106661.

Restrepo, B. J. (2016). Parental investment responses to a low birth weight outcome: who compensates and who reinforces? *Journal of Population Economics*, *29*(4), 969–989.

Rothenberg, A. D., Gaduh, A., Burger, N. E., Chazali, C., Tjandraningsih, I., Radikun, R., ... Weilant, S. (2016). Rethinking indonesia's informal sector. *World Development*, *80*, 96–113.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological monographs: General and applied*, *80*(1), 1.

Sabater, A., & Graham, E. (2016). The role of children's education for the mental health of aging migrants in europe. *GeroPsych*.

Schoppe-Sullivan, S. J., Kotila, L. E., Jia, R., Lang, S. N., & Bower, D. J. (2013). Comparisons of levels and predictors of mothers' and fathers' engagement with their preschool-aged children. *Early child development and care*, *183*(3-4), 498–514.

Seid, Y., & Gurmu, S. (2015). The role of birth order in child labour and schooling. *Applied Economics*, *47*(49), 5262–5281.

Stephens Jr, M., & Yang, D.-Y. (2014). Compulsory education and the benefits of schooling. *American Economic Review*, *104*(6), 1777–1792.

Strauss, J., Witoelar, F., & Sikoki, B. (2016). *The fifth wave of the indonesia family life survey: overview and field report* (Vol. 1). Rand Santa Monica, CA, USA.

Sukamdi, & Mujahid, G. (2015). *Internal migration in indonesia*. UNFPA Indonesia.

Tenikue, M., & Verheyden, B. (2010). Birth order and schooling: Theory and evidence from twelve sub-saharan countries. *Journal of African Economies*, *19*(4), 459–495.

The Lancet Global Health. (2020). Mental health matters. *The Lancet. Global Health*, *8*(11), e1352.

Tirtosudarmo, R. (1997). Economic development, migration, and ethnic conflict in indonesia: A preliminary observation. *Sojourn: Journal of Social Issues in Southeast Asia*, 293–328.

Torssander, J. (2013). From child to parent? the significance of children's education for their parents' longevity. *Demography*, *50*(2), 637–659.

Van de Velde, S., Levecque, K., & Bracke, P. (2009). Measurement equivalence of the ces-d 8 in the general population in belgium: a gender perspective. *Archives of Public Health*, *67*, 1–15.

Van Lottum, J., & Marks, D. (2012). The determinants of internal migration in a developing country: quantitative evidence for indonesia, 1930–2000. *Applied Economics*, *44*(34), 4485–4494.

Vos, T., Lim, S. S., Abbafati, C., Abbas, K. M., Abbasi, M., Abbasifard, M., ... others (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The Lancet*, *396*(10258), 1204–1222.

Wahhaj, Z. (2022). The economics of early marriage: causes, consequences, and policy solutions. In *Handbook of labor, human resources and population economics* (pp. 1–26). Springer.

Wang, D. (2006). China's urban and rural old age security system: challenges and options. *China & World Economy*, *14*(1), 102–116.

Wei, X., Zhou, Y., & Zhou, Y. (2022). Signaling of earlier-born children's endowments, intra-household allocation, and birth-order effects. *Economic Modelling*, *108*, 105754.

Weng, Q., Gao, X., He, H., & Li, S. (2019). Family size, birth order and educational attainment: Evidence from china. *China Economic Review*, *57*, 101346.

West, L. A., Cole, S., Goodkind, D., & He, W. (2014). *65+ in the united states: 2010.* US Census Bureau, US Government Printing Office Washington DC.

Wilkinson, R. G. (2002). *Unhealthy societies: the afflictions of inequality*. Routledge.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wooldridge, J. M. (2019). Correlated random effects models with unbalanced panels. *Journal of Econometrics*, *211*(1), 137–150.

Wu, X., & Li, L. (2014). The motives of intergenerational transfer to the elderly parents in china: consequences of high medical expenditure. *Health economics*, *23*(6), 631–652.

Xiao, Y., Li, L., & Zhao, L. (2017). Education on the cheap: The long-run effects of a free compulsory education reform in rural china. *Journal of Comparative Economics*, *45*(3), 544–562.

Xie, S., & Mo, T. (2014). The impact of education on health in china. *China Economic Review*, *29*, 1–18.

Xie, Y., Zhang, X., Tu, P., Ren, Q., Sun, Y., Lv, P., ... Wu, Q. (2017). China family panel studies user's manual. *Peking University: Beijing, China*.

Xu, A., & Xia, Y. (2014). The changes in mainland chinese families during the social transition: A critical analysis. *Journal of comparative family studies*, *45*(1), 31–53.

Yahirun, J. J., Sheehan, C. M., & Hayward, M. D. (2017). Adult children's education and changes to parents' physical health in mexico. *Social Science & Medicine*, *181*, 93–101.

Yahirun, J. J., Sheehan, C. M., & Mossakowski, K. N. (2020). Depression in later life: The role of adult children's college education for older parents' mental health in the united states. *The Journals of Gerontology: Series B*, *75*(2), 389–402.

Zhang, J., & Meng, C. (2025). Educational expansion and attitudes toward son preference in china. *Journal of Asian Economics*, 101960.

Zhang, S., Guo, N., & Zhang, J. (2023). Reexamining the effect of birth order on cognitive and non-cognitive abilities: New evidence from china. *Economics of Education Review*, *97*, 102476.

Zimmer, Z., Martin, L. G., Ofstedal, M. B., & Chuang, Y.-L. (2007). Education of adult children and mortality of their elderly parents in taiwan. *Demography*, *44*, 289–305.