ISSN: 2424-8150 (Print)
ISSN: 2424-8606 (Online)
Volume 3 · Issue 2
December 2017

INTERNATIONAL JOURNAL OF POPULATION STUDIES

WHIOCE PUBLISHING PTE. LTD.

International Journal of Population Studies

Print ISSN: 2424-8150 Online ISSN: 2424-8606

ABOUT THE JOURNAL

International Journal of Population Studies (IJPS) is an open access, multidisciplinary journal that publishes high quality original research and timely reviews of recent advances and emerging issues in population processes; dynamics of fertility, mortality, and migration; and linkages with socioeconomic and environmental change across times, spaces, and cultures. The journal aims to provide a platform for researchers worldwide to promote and share cutting-edge knowledge and advances in different areas of population research. Article formats include editorials, research articles, review articles, letters to editors, commentaries, perspectives, reports, and book reviews that address demography and population-related issues. The journal also offers special issues arising from conferences and other meetings.

ABOUT THE PUBLISHER

Whioce Publishing is a publishing company based in Singapore. We have in our portfolio a range of high-quality, open-access, peer-reviewed journals and books from a broad spectrum of science, technology, biomedicine, engineering, education, social sciences and many more.

Whioce Publishing
7030 Ang Mo Kio Avenue 5
#04-15 Northstar@AMK
Singapore 569880

Volume 3 • Issue 2 • December 2017
ISSN 2424-8150 (print) ISSN 2424-8606 (online)

INTERNATIONAL JOURNAL OF POPULATION STUDIES

Editor-in-Chief

Danan Gu

United Nations, New York, United States

WHIOCE PUBLISHING PTE. LTD.

International Journal of Population Studies

ISSN: 2424-8150 (print)

ISSN: 2424-8606 (online)

Editorial and Production Credits

Publisher: Whioce Publishing
Article Layout and Typeset: In-house Production

For all advertising queries, contact info@whioce.com.

Disclaimer

Whioce Publishing is not liable to the statements, perspectives, and opinions contained in the publications. The appearance of advertisements in the journal shall not be construed as a warranty, endorsement, or approval of the products or services advertised and/or the safety thereof. Whioce Publishing disclaims responsibility for any injury to persons or property resulting from any ideas or products referred to in the publications or advertisements. Whioce Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

International Journal of Population Studies

Editorial Board

Volume 3 | Issue 2 | December 2017 | Page 1-56 |

International Journal of Population Studies

CONTENTS

International Journal of Population Studies

RESEARCH ARTICLE

Household structure and child education in Cambodia

Patrick Heuveline* and Savet Hong

California Center for Population Research (CCPR), University of California, Los Angeles (UCLA)

Abstract: We analyze the effects of household structure on education in Cambodia. Consistent evidence documents that residence with both biological parents benefits children's education in Western countries. Elsewhere, the issue is gaining more attention with the growing number of "left-behind children" due to adult migration and, possibly, changes in family behavior. The extant record is both thinner and more contrasted, however. Controlling for the presence of grandparents and some household characteristics, we find children residing with both biological parents are more likely to be enrolled in school, in the appropriate grade for their age, and literate than those living with only one parent. The effect sizes appear comparable to those in most Western countries, but the effects shrink or even disappear when grandparents are present. The results for children not residing with either parent are mixed, possibly resulting from negative effects for some children and positive selection for some others.

Keywords: *education;family demography; global/international; household structure; single parents*

ARTICLE INFO
Received: March 9, 2017
Accepted: April 20, 2017
Published Online: April 26, 2017

***CORRESPONDING AUTHOR**
Patrick Heuveline
California Center for Population Research (CCPR)
4284 Public Affairs Bldg
University of California, Los Angeles (UCLA)
Los Angeles, LA 90095
heuveline@soc.ucla.edu

CITATION
Heuveline P and Hong S (2017). Household Structure and Child Education in Cambodia. *International Journal of Population Studies*, 3(2): 1-15. doi: 10.18063/ijps.v3i2.309.

1. Introduction

In this paper, we analyze the effects of household structure and children's living arrangements on their educational outcomes in Cambodia. Such effects have been extensively and fairly consistently documented in high-income Nations. With the notable exception of a body of research on the effects of orphanhood and child fosterage in sub-Saharan Africa (*e.g.*, Bledsoe, 1990; Isiugo-Abanihe, 1985; Madhavan, 2004; Nyamukapa and Gregson, 2005), there had been comparatively little research on the topic in medium- and low-income Nations, until recently at least. A growing interest in the effects of children's living arrangements worldwide follows what may appear to be the emergence of a "Second Demographic Transition" in non-Western Nations. Whether or not their recent demographic trends fit the pattern first identified in Europe (van de Kaa, 1997), it suffices to say that non-marital partnership formation and parental divorce have become more common in some South American and East Asian countries (Esteve, Lesthaeghe and López-Gay, 2012; Raymo, Hyunjoon, Yu *et al*., 2015). Substantial numbers of children residing with only one or none of their biological parents have thus emerged in these regions. Moreover, work opportunities in urban and peri-urban areas, or even abroad, attract an increasing number of short-term and "circular" migrants. Among them, some parents temporarily, though durably sometimes, must leave their children behind (Collinson, Tollman, Kahn *et al*., 2006; Parreñas, 2015). Parents may then leave their children with family members, grand-

parents in particular (Safman, 2003).

In the next section, we briefly review this growing literature, with an emphasis on Southeast Asia. We then describe recent demographic trends in Cambodia and the country's educational system. From this review and assessment of the current situation in Cambodia, we derive hypotheses about the direction and size of the effects of not residing with a biological parent, as well as the effects of living in a multi-generational household. The subsequent section describes the data and methods used to test these hypotheses. The penultimate section describes our results. As discussed in the final section, they appear fairly consistent overall with the body of research conducted in high-income Nations, which, as discussed below, has not always been the case in medium- and low-income Nations. However, as did a few previous studies in these settings, we note some "anomalies" with respect to children residing with neither of their biological parents.

1.1 Literature Review

A central concern of the expansive literature on the effects of growing up with a single biological parent is the endogeneity of parents' living-arrangement decisions. Parents who have a child outside of marriage and those who divorce can be expected to differ from those who remain married on many characteristics. Some of those may very well affect their child's wellbeing regardless of whether parents are married or not. Consequently, research on these effects in Western societies has become increasingly sophisticated in accounting for differences between households with children and both of their biological parents vs. other types of households in which children live. Nonetheless, the overall conclusion seems to stand that these differences, economic in particular, account for some—maybe half—but not all of the differences in various wellbeing indicators between children living with both of their biological parents and those who do not (among many reviews of this large literature, see, for instance, McLanahan, Tach and Schneider, 2013).

A strain of comparative work across Western societies has also documented institutional effects on how children fare across different living arrangements (Cooke and Baxter, 2010). Large international survey programs such as the Trends in International Mathematics and Science Study (TIMSS) or the Programme for International Student Assessment (PISA) have provided opportunities for cross-national research on these effects with respect to child educational performance. These studies have found that, in all countries, the average educational performance of children who live with both of their biological parents is higher than the average educational performance of those who do not—a typical pattern referred to as the "educational gradient" thereafter. Within countries, this educational gradient tends to increase slightly with age, but its magnitude varies markedly between countries. In multiple studies, gradients were found to be larger in the U.S.A. than in any other country included in the analyses, for instance, and to become almost negligible in countries with the most generous welfare provisions. The magnitude of the gradients, typically adjusted for parental characteristics, is difficult to compare across studies that may use different parental characteristics as control variables. Using data on TIMSS "population1" students (centered on 9 year-olds), the unadjusted U.S. gradients were estimated to amount to 20 to 29 points in Math and 17 to 33 points in Sciences (Pong, Dronkers and Hampden-Thompson, 2003). (In both TIMSS and PISA surveys, achievement scores are "curved" so that a standard deviation (SD) is close to 100 points). For "population-2 students" (centered on 13 year-olds), researchers reported U.S. unadjusted gradients equivalent to 35 points in Math and 36 points in Sciences again analyzing TIMSS data (Heuveline, Yang and Timberlake, 2010). Using PISA data on 15 year-old students, unadjusted gradients in the U.S.A. have been estimated to range from 36 to 50 points in reading, 42 to 53 points in Math, and 37 points in Science (Garib, Martin Garcia and Dronkers, 2007; Hampden-Thomson, 2013; Marks, 2006).

In a systematic analysis of TIMSS data in non-Western societies, Schiller, Khmelkov and Wang, (2004) find that educational gradients are related to a country's Gross Domestic Product, with the poorest countries exhibiting the smallest effects of living arrangements on children's education. Several country-specific analyses have replicated results from high-income countries regarding the advantage of children living with both parents (*e.g.*, most recently, Chae, 2016). Other studies have obtained mixed results or even produced findings to the contrary. Evidence of a reversed gradient indicates that the cultural context matters with respect to, for instance, whether and to which extent some living arrangements are stigmatized, and which children may be selected into particular living arrangements. Without delving into this country-specific literature (see DeRose, Corcuera Garcia, Salazar *et al*., 2014 for a comprehensive review), we should note two issues that have received more attention in non-Western than in Western countries. The first one concerns the different effects associated with the different pathways into residing with only one or no biological parent. In the West, researchers may have paid attention to the differences between non-marital childbearing and divorce, or more rarely, between divorce and widowhood in the West (*e.g.*, Biblarz and Gottainer, 2000). In many low- and medium-income countries, parental mortality remains high (Beegle, De Weerdt and Dercon, 2009; Birdthistle, Floyd, Nyagadza *et al*., 2009; Case and Ardington, 2006; Evans and Miguel, 2007; Gertler, Levine and Ames, 2004) and parental migration is becoming more

prevalent (Kandel and Kao, 2001; Kuhn, 2006; Yao and Treiman, 2011; Nobles, 2011; Townsend, Madhavan, Tollman *et al*., 2002). These different pathways into single-parent-headed households may produce different effects on the children in these households. A recent article argues for negative selection among children with a deceased parent, but positive selection for the children left behind by their migrant parents in India (Das, 2016). For Southeast Asia, Pong (2006) reports that in Malaysia's collectivist culture divorce and separation have a negative effect on children's education, but widowhood does not. In Indonesia and Thailand, Park (2007) finds students in single-parent families to outperform their peers in intact families. Strong norms against never-married or divorced parents imply that a higher proportion of single parents are widows who may receive more social and institutional support than single parents elsewhere. In Vietnam, however, Loenzien (2016) reports lower enrollment and attainment levels for children of lone mothers, regardless of whether they are never married, divorced, separated or widowed.

Prior research in the U.S.A. has shown that children living with a single mother may benefit from the additional presence of grand-parents (*e.g*., DeLeire and Kalil, 2002). Huisman and Smits (2009) obtain similar results in their comparative study of low- and middle-income countries. Reduced when one of the biological parents is absent, a child's enrollment chances appear to increase when living in an extended family, especially one with grandparents. For Thailand, Mahaarcha and Kittisuksathit, (2009) also showed the positive effect of having grandparents, in the household, on school enrollment for adolescents. These results are consistent with the additional resources grand-parents contribute, both economic and social (*i.e*., Coleman's (1988) inter-generational closure). In Japan, however, Shirahase and Raymo (2014) found children of single mothers to fare worse in multi-generational than in single-parent households. The authors tie these results to strong norms favoring nuclear households, since such norms imply highly-negative selection into multi-generational households.

1.2 The Cambodian Context

1.2.1 Household Structure and Children's Living Arrangements

There is a strong norm against pre-marital childbearing in Cambodia. Anecdotally, data collectors on a fertility study in the late 1990s were found to skip fertility histories for never-married women for fear of offending study participants. In Phnom Penh at least—as in Bangkok (Esara, 2012), Manila (Xenos and Kabamalan, 2007) and probably other capital cities in Southeast Asia—pre-marital cohabitation is beginning to appear and may result in pre-marital conceptions and even births. Lacking national data on mothers' marital status at birth, however, it is impossible to know whether the phenomenon is limited to the capital city or even to its few highly "globalized" neighborhoods, where Westerners and Cambodian youths interact (Hoefinger, 2013).

Trends in divorce are easier to track in survey data. Marriage stability among recent cohorts is markedly lower than among earlier ones, divorce remains rare. Among the late 1990s cohorts, only 6% had ended in divorce within 5 years of marriage (Heuveline and Poch, 2006). Even though adult mortality has declined from the dramatically high levels of the late 1970s, parental death likely remains the most common reason for a child not to live with both biological parents. Heuveline and Hong (2016) report that parental mortality actually accounts for nearly half of children (46.2%) residing with only one of their biological parents.

Another increasingly common reason for parental absence appears to be work-related parental mobility. The majority of Cambodian households remains engaged in farming, and rice farming in particular. Towards the end of the dry season when the demands of agricultural work slow down considerably, farmers have commonly sought temporary work in the cities—men in construction, for instance, and women in street-food vending. With the relative decline of farming revenues relative to other sectors, such temporary migration has only become more frequent or more permanent (National Committee for Population and Development, 2009). In particular, the rapid development of a garment industry (Chea and Sok, 2001; Ear, 2011) has fueled the migration of young female workers to the outskirts of the capital city (Derks, 2008). At the outset, garment factories were recruiting almost exclusively never-married women intending to save up some money before marriage. Over time, however, with wages substantially higher than the income that farming may generate, it has become more common for women to return to the factories after marriage.

A widely recognized normative sequence of living arrangements begins with newlyweds residing with the brides' parents, but only temporarily until they build up the desired resources to eventually settle their own independent household nearby (Ebihara, 1968; Heuveline and Hong, 2016). Correspondingly, the dominant living arrangement is nuclear, with a preference for uxorilocality. However, Cambodian households can be quite pragmatic in their living arrangements and young couples routinely depart from the uxorilocal norm, for instance, if economic opportunities are available near the groom's parents (Demont and Heuveline, 2008). The prevalence of multigenerational households (in which, as will be shown below, nearly a quarter of all rural children under 18 years of age live), however, clearly exceeds what it would be were married couples to live in this arrangement only for a few years after marriage. Another indication

of this pragmatism is that nuclear households are less common in urban areas than in rural areas (Heuveline and Hong, 2016), as urban households are more frequently solicited to take in rural relatives who want to pursue work or education opportunities in the city.

1.2.2 Education and Social Welfare

The Royal Government of Cambodia's (RGC) official target is for all children to receive nine years of basic education from the first-grade enrollment age of 6 years to age 15 years (Ayres, 2000). School enrollment among 6-to-14 year-olds has been gradually increasing, reaching 88.5% for girls and 86.9% for boys in 2014, up from 84.5% and 83.9% respectively in 2009 (National Institute of Statistics, 2015). As public education in Cambodia follows a 6 + 3 + 3 model (6 years of primary education, 3 years of lower- and 3 years of upper-secondary education), the 9-year target corresponds in theory to lower-secondary school completion. In reality though, late enrollment in first grade and grade repetition are both common. National statistics suggest that less than half of the 7th graders are of the expected age of 12 years or younger (Ministry of Education, Youth, and Sports, 1997). The proportion of 18-to-24 year-olds who have actually completed lower-secondary school has increased from 27.3% for females and 37.3% for males in 2009 to 41.1% for females and 43.0% for males in 2014 (National Institute of Statistics, 2015). Nevertheless, more than half of the youths in these cohorts still fail to meet the official basic-education target.

The RGC has been aggressively attacking one of the barriers to universal school attendance—distance from the closest school in rural areas. In the past, parents might have been particularly reluctant to send young girls far away from their home village. Gender differences in attendance have declined as the RGC is moving closer to its stated goal of one primary school per village and one lower-secondary school per commune—an administrative unit typically consisting of 5 to 15 villages, depending on their size (Heuveline and Hong, 2016).

Unfortunately, the RGC has not made a similar commitment to supporting these schools' operating costs and their teachers' salaries (Ayres, 2000; Brehm, 2016). Public schools' tuition and fees are low, but their students are expected to buy their own textbooks and supplies, as well as to contribute to some of the school's running expenses, such as building maintenance funds, and to "give" to their teachers a small amount of money daily—in effect, a salary supplement. Another common strategy for public-school teachers to augment their salary is to provide additional lessons for a fee, often in public-school classrooms (Brehm, 2016; Nguon, 2012). While less than 2% of primary school and lower-secondary school students were attending a private school, 13.6% of primary school students and 46.9% of lower-secondary school students were taking private lessons in 2014 (National Institute of Statistics, 2015). Overall, in 2014, the estimated annual educational costs were $78.5 per primary school student, $152 per lower-secondary school student, and $303.5 per upper-secondary school student. This compares to an average income per capita just under $1,000 per year, of which, on average again, $700 is spent on food and housing plus utilities alone. In terms of national averages, educational expenses may appear modest, but for many households with school-age children, they represent more than a trivial share of what is left of their income after paying for basic necessities. In 2014, less than 5% of parents cited distance from school as a reason for a 6-to-17 year-old not to attend school, but 11.0% cited being too poor, another 29.1% the need for the child to contribute to the household income, and yet another 6.3% the need to contribute to the household chores (National Institute of Statistics, 2015).

Until recently, the RGC only provided welfare payments to specific groups, most notably its retired civil servants and veterans of the armed forces and the national police. A large number of un-coordinated, donor-funded initiatives, projects and activities have sprung up to provide additional support, but these are often limited in geographical scope and duration (International Labor Office, 2012). Two of the best coordinated social protection programs concern health and education. Introduced in 2000, Health Equity Funds (HEF) have been funded through the RGC Ministry of Health by international donors and NGO, which select HEF beneficiaries and compensate their healthcare providers for lost fee revenues. They have spread widely since, though still substantially shy of the RGC's target of nationwide coverage by 2013 (Flores, Por, Chean *et al*., 2013). Following a similar model, international donors and NGO began financing scholarships for poor children in 2003. Working with the RGC Ministry of Education, the scholarship program came to identify the transition from primary to lower-secondary school as a critical period, thus targeting children in the 6th grade. All 6th graders in any of the primary schools that feed lower-secondary schools selected by the Ministry of Education automatically apply to these scholarships. The program was found to have a significant impact on enrollment, but there was no evidence that it also affected academic achievement (Filmer and Schady, 2009). Moreover, the program remains relatively small, with less than 15% of students receiving a scholarship to one of the selected schools, which themselves only represent about 1/8th of the lower-secondary schools in the country.

The RGC has recently taken steps towards an integrated social protection system. In 2007, the RGC gradually implemented an official poverty targeting system known as the IDPoor, with households identified as poor (IDPoor 1) or extremely poor (IDPoor 2) receiving an IDPoor Card. By 2012, almost all rural areas were covered. In 2011, the RGC also created a National Social Protection Strategy (NSPS) for 2011–15, and in 2013, pilot programs started to be implemented in order to experiment with the design and delivery mechanisms for safety nets. Despite the availability of this targeting mechanism, limited fiscal revenues seriously constrain the extent of social protection that national institutions and government agencies can provide. In 2013, the World Bank estimated that the coverage of safety nets remained at only 2% of the poorest quintile of the population (World Bank, 2014). A high share of the population thus continues to face serious vulnerabilities that may induce a temporary inability to face education and health expenditures. To cope, most households continue to rely primarily on their extended kin network (Kim, 2011; Heuveline and Hong, 2016).

1.3 Conceptual Framework and Hypotheses

Based on the framework developed in Heuveline, Yang and Timberlake, (2010) to study international differences in the effects of children's living arrangements on their educational outcomes, we conceptualize living arrangements as operating through the "quality" and "quantity" of parenting available to children. Even though the involvement of the non-residential parent(s) may vary, living with only one or no biological parent is expected to reduce the quantity of parenting. Single parents' own parenting might be affected, as research documents the "time deficit" they experience when parenting alone (Hill, Yeung and Duncan, 2001; Bianchi, Robinson and Milkie, 2006). The additional stress associated with single parenting and poor communication between parents living apart might also affect the quality of parenting. Accordingly, and as repeatedly reported across diverse contexts in the literature reviewed above, we expect children residing with both biological parents to have better educational outcomes than children who do not, even after controlling for observed differences in household resources (Hypothesis 1).

However, this conceptual framework also emphasizes that parents draw the resources their children need from a larger environment that includes not just the labor market but also governmental programs that may be in place to support families with children. Following Pong, Dronkers and Hampden-Thompson (2003), we expect that family policies that equalize resources between different types of families reduce the educational gradient. Moreover, public funding for education may reduce the impact that differences in resources have on educational outcomes. As discussed in the previous section, at the time data used in this study were collected, Cambodia still very much lacked the type of welfare support that might be available in Western or East Asian societies. This would lead us to hypothesize that the educational gradient in Cambodia should be larger than average in Western societies (Hypothesis 2a).

Another important aspect of the larger context in which households operate, albeit difficult to measure precisely, is the cultural environment which affects, in particular, the extent to which less traditional households are stigmatized or supported by their communities. Results reviewed above from Malaysia, Indonesia, or Thailand showed smaller than expected differences in outcomes that might be linked to the relatively large share of non-stigmatized widows among the single parents with children. Similarly for Cambodia, one may alternatively hypothesize that the educational gradient should be smaller than average in Western societies (Hypothesis 2b).

The most direct form of communal support is arguably multi-family household formation, allowing for different families to share a substantial amount of resources and time, and for parenting to be provided by adult co-residents other than the biological parents. As documented in Safman (2003) for Thailand, and equally valid for Cambodia, grand-parents are parents' preferred "social" parents in the event they cannot take care of their children themselves. As reviewed above, studies have generally found that children not living with one or both of their biological parents fare better in multi-generational households (with at least one grand-parent present). Results to the contrary have been attributed to a strong negative selection into such households when these represent a strong deviation from the norm. As a great deal of tolerance and pragmatism with respect to living arrangements has been reported in Cambodia, we expect that more favorable educational outcomes for children living in multi-generational households than for those of children living in other household structures (Hypothesis 3).

2 Data and Methods

This paper utilizes survey data from the Mekong Integrated Population-Registration Areas of Cambodia (MIPRAoC) project. The MIPRAoC project grew out of The Mekong Island Population Laboratory (MIPopLab, 2000–2006; ICPSR36601-v1). Both projects include occasional, topical, "rider" social-science surveys built on a longitudinal health and demographic surveillance system (HDSS). The analyses presented in this paper are based on the baseline survey (2008) of the MIPRAoC HDSS. We first describe this survey, then describe the variablesconstructed from these data and our

analytical strategy.

2.1 Study Population

The target population of the baseline survey consisted of residents from six new Population-Registration Areas (PRA), in addition to all residents from the original MIPopLab site.

All seven PRA are located along the Mekong River, which flows through Eastern Cambodia and the capital city, Phnom Penh. To be enumerated as a resident, a seasonal migrant needs to have spent less than three months away from the household in the preceding six months and a circular migrant, such as a child attending school, needs to have spent less than 4 nights a week away. The project was only designed to provide representative data for the population of the contiguous districts along the Mekong River, where 20% of the rural households in the country resided at the time of the 1998 General Population Census (National Institute of Statistics, 1999). However, comparisons between nationally-representative and MIPopLab or MIPRAoC data have repeatedly shown little differences in marital, reproductive, or household composition levels and trends (Heuveline and Hong, 2016).

The combined population—at initial registration in the six new PRA and after demographic update in the original MIPopLab site—was close to 60,000. Due to the recent establishment of several garment factories, the population of one of the PRA, located at the outskirts of the Phnom Penh, was much larger at the time of the baseline survey (Round 1, 2008) than in the sampling frame from the 1998 General Population Census. The response rate was 96.3% in this PRA and 99.1% elsewhere (averaging 98.4%). The seemingly high response rate is in fact typical of social surveys conducted in the country (e.g., National Institute of Statistics, 2015).

Overall, the MIPRAoC is representative of the Mekong River Valley households in Cambodia and follows the national trends. Unlike existing nationally representative surveys, the MIPRAoC is designed to be a longitudinal data collection project, where the ongoing data collection will enable researchers to examine the long-term impact of household compositions. This paper examines the initial census of the seven PRA, which contained 14,989 children under the age of 18 across 7,205 households.

2.2 Analytical Strategy

As we are concerned with the academic achievements of children, the analytic sample has been restricted to school age children (ages 6 to 17 years). Because the sample might thus include several siblings from the same household, we use Hierarchical Linear Models (HLM) with each child's characteristics as level-1 independent variables and (possibly shared) household characteristics as level-2 independent variables. Among these variables are the parental characteristics described below. Because parental characteristics are only available when the parent co-resides with her child, models with parental characteristics are estimated on two sub-samples: a sub-sample of children co-residing with their mother (with or without their father) and a sub-sample of children co-residing with their father (with or without their mother).

2.3 Measurements

At the time of initial registration into the MIPopLab/MIPRAoC HDSS (benchmark census or subsequent in-migration), each household head provides, for each resident household member, their name (later replaced by a unique identifier), gender, birth date, relationship to the head, parental information (on survival, current residence, or timing of death), marital status, literacy, education, and occupation.

2.3.1 Outcomes

For dependent variables, we consider three measures of academic achievements reported by the head of the household or another adult household resident: 1) literacy (self-reported ability to read and write in any language), 2) school attendance (attending any formal educational institution), and 3) child's grade level relative to the expected grade for age at the time of the baseline survey.

2.3.2 Individual-Level Covariates

Level-1 independent variables include age groups, gender, and parental survival. We use the age groups 6 to 8, 9 to 11, 12 to 14, and 15 to 17 years, corresponding to the first and second half of primary school, lower secondary, and upper secondary school respectively. The proportion of MIPRAoC children (17 years and under) who are of school age (6 years and over) is 68%, and among those a majority (87%) indeed attended school, but only 73% were literate (see **Table 1**).

2.3.3 Household-Level Covariates

Table 1. Child, parents and household characteristics, descriptive statistics

	All children		Ages 6 to 17		Ages 6 to 14[a]		
Variables	*M*	*SD*	*M*	*SD*	*M*	*SD*	Range
Female child	.49	.50	.49	.50	.50	.50	0 - 1
Under age 6	.32	.47	--	--	--	--	0 - 1
Ages 6 to 8	.16	.37	.24	.43	.26	.44	0 - 1
Ages 9 to 11	.16	.37	.24	.42	.34	.47	0 - 1
Ages 12 to 14	.18	.38	.26	.44	.40	.49	0 - 1
Ages 15 to 17	.18	.38	.26	.44	--	--	0 - 1
Child attending school	--	--	.87	.33	1.00	.00	0 - 1
Actual minus expected grade	--	--	--	--	-0.98	1.75	-8 - 12
Child literacy	--	--	.73	.44	--	--	0 - 1
Nuclear household	.61	.49	.63	.48	.63	.48	0 - 1
Multigenerational household	.23	.42	.22	.41	.22	.41	0 - 1
Two co-resident parents	.84	.37	.81	.39	.82	.38	0 - 1
Only one co-resident parent	.11	.31	.13	.33	.12	.32	0 - 1
Co-resident mother: Literate	.64	.48	.63	.48	.65	.48	0 - 1
Employed in farming, hunting or fishing	.38	.49	.40	.49	.40	.49	0 - 1
Employed in crafts	.11	.32	.12	.32	.13	.33	0 - 1
Employed in industry	.04	.19	.03	.17	.03	.17	0 - 1
Employed in service sector	.18	.38	.18	.38	.18	.39	0 - 1
Employed in civil service	.02	.13	.02	.13	.02	.14	0 - 1
Co-resident mother employed in farming, hunting or fishing, or in crafts and who is: Owner	.36	.48	.38	.49	.39	.49	0 - 1
User for free	.02	.13	.02	.12	.02	.12	0 - 1
User for fee/rent	.06	.23	.06	.23	.06	.24	0 - 1
Paid laborer	.06	.24	.06	.24	.06	.24	0 - 1
Co-resident father: Literate	.69	.46	.66	.47	.69	.46	0 - 1
Employed in farming, hunting or fishing	.45	.50	.45	.50	.45	.50	0 - 1
Employed in crafts	.03	.18	.03	.16	.03	.18	0 - 1
Employed in industry	.05	.21	.04	.19	.04	.19	0 - 1
Employed in services	.21	.10	.18	.39	.19	.39	0 - 1
Employed in civil service	.09	.29	.09	.29	.10	.30	0 - 1
Co-resident father employed in farming, hunting, fishing, or in crafts and who is: Owner	.38	.48	.38	.49	.39	.49	0 - 1
User for free	.02	.13	.01	.11	.01	.11	0 - 1
User for fee/rent	.07	.25	.07	.25	.07	.25	0 - 1
Paid laborer	.03	.16	.02	.14	.02	.14	0 - 1
N	22,032		15,010		9,491		

Source: Authors' calculations

Note: [a] School-Attending children only

With regard to household composition, MIPRAoC captures the relationship of individuals to household heads as well as some information on their biological parents. This enables us to construct two sets of household variables. First, an indicator of the number of co-resident biological parents is created. Over 80% of children in the study lived with both biological parents. As children age, the cumulative risk of parental mortality or separation increases, and so does the prevalence of children residing with only one biological parent. However, the prevalence of co-residence with only one parent remains relatively low at 13% for children aged 6 to 17 years (**Table 1**).

The second level-2 variable is categorical and describes household structure. The first category is *nuclear*: households consisting only of (single or two) parents and (biological, step, or adopted/foster) children. The second category is multigenerational and includes households with at least one grandparent in addition to parents and children. The residual category includes all households with other members, such as extended family members or non-relatives, and four-generational households. This residual category only accounts for 15% of all the households in the sample, whereas the nuclear households are the most common, as they are at the national level, followed by three-generational households (Heuveline and Hong, 2016).

Level-2 covariates also included parental characteristics. With respect to the determinants of child literacy, we first use a dichotomous variable for each parent's own literacy. In the MIPRAoC data, 65% of co-residing fathers and 63% of co-residing mothers of school aged children are literate (see **Table 1**). Several additional socio-economic variables were considered to capture some of the differences among the households children live in. The first one is the type of

employment of all previously or currently employed parents. We created five employment sectors, namely farming (including fishing, hunting, forestry, and plantation), crafts, industry, civil service, and white-collar or service jobs. All individuals employed in farming or crafts were also asked whether they owned the land or resources needed for their activity (*e.g.*, boat for fishing, loom for weaving, etc.). We created four categories for property: ownership, renting in exchange of payment, free usage (*e.g.*, lending by kin), or being a laborer.

3. Results

3.1 School Enrollment

We first consider school enrollment among 6- to 17-year olds. As shown in **Table 2**, we find the odds of being enrolled in school to be 34% lower for children residing with only one biological parent compared to children residing with both

Table 2. Summary of hierarchical regression analysis for variables predicting school enrollment for all children ages 6 to 17 (n = 14,016), those residing with their mother (n = 13,776) and those residing with their father (n = 12,315)

	All children (Ages 6 to 17)			**Children residing with their mother**			**Children residing with their father**		
Predictor	*β*	*SE*	*OR*	*β*	*SE*	*OR*	*β*	*SE*	*OR*
Female child	0.06	0.08	1.07	0.07	0.08	1.07	0.14	0.09	1.16
Ages 9 to 11	3.02**	0.13	20.49**	3.01**	0.14	20.21**	3.16**	0.15	23.53**
Ages 12 to 14	3.65**	0.15	38.46**	3.63**	0.15	37.86**	3.82**	0.17	45.49**
Ages 15 to 17	3.16**	0.14	23.51**	3.11**	0.14	22.53**	3.34**	0.16	28.21**
Deceased mother	0.49	0.32	1.64	-	-	-	-0.30	0.81	0.74
Deceased father	-0.08	0.23	0.93	-0.37	0.29	0.69	-	-	-
Multi-generational household	0.19	0.14	1.21	0.04	0.15	1.04	0.12	0.15	1.13
Other households	0.53**	0.15	1.69**	0.47*	0.16	1.59*	0.54*	0.19	1.71*
No co-resident parent	-0.24	0.22	0.78	-	-	-	-	-	-
Only 1 co-resident parent	-0.41†	0.20	0.66†	-0.52†	0.24	0.60†	-0.58	0.71	0.56
Parent is: Literate				1.47**	0.11	4.37**	1.55**	0.14	4.73**
Employed in crafts				1.31**	0.20	3.71**	1.25**	0.37	3.49**
Employed in industry				0.97	0.57	2.63	1.00	0.51	2.73
Employed in services				1.03†	0.51	2.80†	0.75	0.46	2.11
Employed in civil service				2.67**	0.76	14.45**	1.46*	0.49	4.29*
Sector unknown				0.87	0.52	2.38	-0.06	0.53	0.94
Parent is: User for free				-0.01	0.40	0.99	0.01	0.45	1.01
User for fee/rent				0.17	0.23	1.19	0.19	0.22	1.20
Paid laborer				-0.79**	0.20	0.46**	-0.67†	0.33	0.51†
Ownership unknown				-0.76	0.51	0.47	-0.36	0.46	0.70
Hh. level variance (ψ)		6.16			5.44			5.94	
Intra-class correlation (ρ)		0.65			0.62			0.64	
Model fit (*-2LL*)		4,515.62			3,987.02			3,557.42	

Source: Authors' calculations.

Note: The parental characteristics are those of the child's mother in Column 2 and of the child's father in Column 3. All models also include a non-significant interaction term between household types and parental survival. The reference categories are male child, ages 6 to 8, living in a nuclear household with both biological parents, parent is not literate, employed in farming, hunting, fishing or in crafts and is the owner of the land/crafts equipment. Ownership is only assessed for parents employed in farming, hunting, fishing or in crafts.

$^{\dagger}p<0.1$. $^{*}p<0.05$. $^{**}p<0.01$.

(difference significant at the 10% level only). A more surprising result is the odds ratio for children residing with neither of their biological parents (odds ratios of 0.78 relative to children co-residing with both biological parents, and not significantly different from one). This might be due to the fact that some children reside away from the parental home precisely to attend a somewhat distant school. Consistent with this scenario, we find that children who live in a household structure other than nuclear or multi-generational have much higher odds to be in school (+69%, **Table 2**). Living in a multi-generational household—a more common arrangement for children who live with only one parent rather than with both—also increases the odds of school enrollment compared to nuclear households, but only by 21% (non-significant difference).

To further control for parental characteristics (literacy, occupation, land/craft-tool ownership), we limited the next analyses to children residing with at least one parent. **Table 2** also presents estimates of the odds of school enrollment controlling for maternal characteristics for children residing with their biological mother, with or without a co-resident biological father; and similar estimates controlling for paternal characteristics for children residing with their biological father, with or without a co-resident biological mother. Controlling for maternal characteristics, the odds ratio of school enrollment for children residing without their biological father is slightly reduced compared to children residing with their biological father—from 0.66 to 0.60, difference still significant at the 10% level only. For children residing with their biological father, the odds ratio of school enrollment for children residing without relative to children residing with their biological mother is even lower (0.56, but the difference is not significant due to the small number of children not living with their mother). The estimated odds ratios corresponding to household structures other than nuclear or multi-generational remain largely unchanged by the addition of parental characteristics in the model (from 1.69 for all children to 1.59 for children living with their mother and 1.71 for children living with their father). Among the parental characteristics accounted for in these models, literacy and non-farming occupations have strong positive association with children's school attendance. Children whose parents are employed in farming are less likely to be in school, especially if their parents are paid laborers, rather than land owners or tenants. Also notable across models is the absence of significant gender differences in school attendance.

3.2 Grade for Age

For children who were attending school, we next examined whether they were at the expected grade for their age. With our dependent variable—the difference between the actual and the expected highest grade completed by the child's age—being now continuous rather than dichotomous, we followed the same strategy as for school attendance. Namely, we first compared children co-residing with both biological parents to those co-residing with only one biological parent and to those not residing with their parents; then we compared children residing with both parents to children residing with their biological mother and children residing with their biological father. In **Table 3**, we observe similar differences by parental co-residency as we do for school enrollment. Children residing with only one parent are on average 0.23 grade below those residing with both parents at the same age, and those not residing with any parent a little lower still (0.28 grade

Table 3. Summary of hierarchical regression analysis for variables predicting actual minus expected grade for age for all children aged 6 to 14 (n = 9,399), those residing with their mother (n = 8,705) and those residing with their father (n = 7,863)

	All Children (Ages 6 to 17)		**Children Residing With Their Mother**		**Children Residing With Their Father**	
Predictor	B	SE(B)	B	SE(B)	B	SE(B)
Female child	0.23**	0.03	0.23**	0.03	0.25**	0.03
Ages 9 to 11	-1.01**	0.03	-1.00**	0.03	-0.98**	0.03
Ages 12 to 14	-1.92**	0.03	-1.88**	0.03	-1.86**	0.03
Deceased mother	0.27†	0.12	--	--	0.28	0.29
Deceased father	0.02	0.08	-0.07	0.10	--	--
Multi-generational household	0.16**	0.05	0.13*	0.05	0.12†	0.05
Other households	0.25**	0.05	0.13*	0.05	0.12†	0.05
No co-resident parent	-0.28**	0.08	--	--	--	--
Only one co-resident parent	-0.23**	0.07	-0.24*	0.08	-0.53†	0.26
Parent is: Literate			0.73**	0.04	0.70**	0.05
Employed in crafts			0.72**	0.05	0.44**	0.09
Employed in industry			-0.10	0.19	0.00	0.17
Employed in services			0.11	0.17	0.09	0.15
Employed in civil service			0.83**	0.21	0.57**	0.16
Sector unknown			-0.05	0.18	-0.18	0.18
Parent is: User for free			-0.00	0.14	0.03	0.15
User for fee/rent			-0.23*	0.07	-0.15†	0.07
Paid laborer			-0.02	0.07	-0.39**	0.12
Ownership unknown			0.38†	0.17	0.33†	0.15
Hh level variance (ψ)	0.94		0.73		0.74	
Intra-class correlation (ρ)	0.45		0.40		0.40	
Model fit (-2LL)	16,289.57		14,678.69		13,238.50	

Source: Authors' calculations.

Note: See footnote to Table 2.

below on average). However, we again find that these effects might be counter-balanced by residing in a non-nuclear household structure. Living in a multi-generational household raises the expected grade by 0.16 and the expected increase is even higher—0.25 of a grade—when living in another type of non-nuclear household structure (*i.e.*, Other households type). Combining the estimated effects of parental absence and of living in a non-nuclear household, children residing with only one parent in these household structures are expected to be very nearly in the same grade on average as children residing with both parents in a nuclear household.

Once we control for the characteristics of the resident parent for children living with at least one of their parents, we see the same pattern as with school attendance. For children residing with their biological mother, the estimates are very similar to those of the previous model (0.24-grade difference between those who also reside with their father and those who do not). For children residing with a biological father, the difference is much more substantial, more than half a grade (0.53-grade difference between those who also reside with their mother and those who do not), albeit significant only at the 10% level due to the small number of children not living with their mother. As in the school-attendance models, the effects of household structures are attenuated by the inclusion of parental characteristics. The coefficients for multi-generational household structures hardly change though, while the coefficients for other household structures are reduced in half. (Again, coefficients in the model with co-resident fathers are only significant at the 10% level). Adding the characteristics of the resident parent, we now find that children residing with only one of their parent in a multi-generational household are in lower grade on average that those residing with both parents in a nuclear-family household, especially if the absent parent is the mother.

The literacy of the co-resident parent has equally strong positive effects on the child's grade-for-age in the models with mother present and in those with father present. We also observe strong positive effects of the resident parent's occupation. As with school-attendance models, children whose parents are engaged in crafts or civil service fare better that those whose parents are employed in the agricultural sector. As in those models, children whose fathers are paid laborers also fare worse, but we no longer observe the same disadvantage for children whose mothers are paid laborers.

Another notable difference with the school-attendance models is that we now find a gender difference in grade for age, with girls being about a quarter of a grade above same-age boys across the three models. Overall, we also confirm how common grade repetition is in Cambodia, with 9 to 11 year-olds being on average one year below their expected grade compared to 6 to 8 year-olds, and 12 to 14 year-olds being an additional year below their expected grade. In other words, the average child only moves up two grades every three years.

Table 4. Summary of hierarchical regression analysis for variables predicting literacy for all children ages 6 to 17 (n = 14,008), those residing with their mother (n = 13,778) and those residing with their father (n = 12,308)

	All children(Ages 6 to 17)			**Children residing with their mother**			**Children residing with their father**		
Predictor	*β*	*SE*	*OR*	*β*	*SE*	*OR*	*β*	*SE*	*OR*
Female child	0.18*	0.07	1.20*	0.21*	0.07	1.24*	0.24*	0.08	1.27*
Ages 9 to 11	3.41**	0.11	30.32**	3.46**	0.12	31.72**	3.60**	0.13	36.44**
Ages 12 to 14	5.19**	0.14	179.37**	5.29**	0.16	198.34**	5.49**	0.17	242.29**
Ages 15 to 17	5.17**	0.15	176.34**	5.27**	0.16	194.29**	5.55**	0.18	256.82**
Deceased mother	0.42	0.27	1.52	-	-	-	-0.36	0.72	0.70
Deceased father	0.14	0.19	1.14	-0.19	0.24	0.82	-	-	-
Multi-generational household	0.15	0.12	1.16	0.09	0.12	1.09	0.12	0.13	1.13
Other households	0.15	0.12	1.17	-0.03	0.13	0.98	0.02	0.14	1.02
No co-resident parent	0.07	0.18	1.07	-	-	-	-	-	-
Only one co-resident parent	-0.23	0.17	0.80	-0.27	0.20	0.76	-0.61	0.64	0.54
Parent is: Literate				1.83**	0.10	6.26**	1.85**	0.13	6.39**
Employed in crafts				1.13**	0.15	3.11**	1.04**	0.27	2.84**
Employed in industry				0.07	0.47	1.07	0.60	0.44	1.82
Employed in services				0.34	0.42	1.41	0.14	0.40	1.16
Employed in civil service				1.61*	0.55	5.02*	0.83†	0.42	2.30†
Sector unknown				0.23	0.43	1.26	0.03	0.47	1.03
Parent is: User for free				0.43	0.34	1.54	0.08	0.39	1.08
User for fee/rent				0.10	0.18	1.10	0.28	0.18	1.32
Paid laborer				0.14	0.18	1.15	-0.48	0.29	0.62
Ownership unknown				-0.05	0.42	0.95	0.24	0.40	1.27
Hh level variance (ψ)		5.12			4.44			5.00	
Intra-class correlation (ρ)		0.61			0.57			0.60	
Model fit (-2LL)		6,103.84			5,330.64			4,774.47	

Source: Authors' calculations.

Note: See footnote to Table 2.

3.3 Literacy

Last, we consider the educational proficiency of all children between the ages of 6 and 17 years as assessed by their literacy. Unfortunately, literacy is only reported as a dummy variable and our models are thus similar to those for school enrollment. In these models, however, the effects of living arrangements and household structure are generally smaller than in the previous two sets of models and none of them are significant. For the model including all children, for instance, the odds of literacy for children co-residing with only one of their biological parents is 20% lower than for those co-residing with both (**Table 4**). More surprisingly in this model, children who reside with neither of their parents, have the better odds of being literate than those residing with both, even though, again, the odds ratio is not significantly different from one. This may either indicate no effect or a combination of negative effects for some children and positive effects for some other children. While we typically expect negative selection for not residing with any parent, some positive selection is possible as well. For instance, parents may send those children that are doing better in school, to stay with relatives in order to facilitate their further studies.

Controlling for parental characteristics, we find again that parental literacy and employment in either crafts or civil service are the parental characteristics most associated with higher odds of literacy. We again find gender differences in favor of girls, whose odds of literacy are 20% to 27% higher than those of boys across models. Also of note, the age pattern only flattens after age 11 (between the age groups 12 to 14 & 15 to 17), indicating a protracted process of becoming literate. The odds of being literate among 9 to 11 year-olds are only one sixth of the odds among 12 to 17 year-olds.

4. Discussion

Our analyses of the effects of residing with only one or no biological parent on children's school enrollment, grade-for-age, and literacy reveal some similarities with the effects that have been well documented in high-income Nations. Consistent with our Hypothesis 1, we find the odds of being in school for children residing with their biological mother but not with their biological father to be 40% lower than for those residing with both biological parents; and when they are in school, the former are enrolled in a lower grade on average. Their odds of being literate are also 24% lower than their peers residing with both parents, but this difference is not significant. Compared with children residing with both parents, the observed disadvantage of residing with only one parent appears to be up to twice larger across these indicators when the absent parent is the child's mother rather than the child's father—a finding consistent with earlier research (*e.g.*, Llyod, and Blanc, 1996). Relatedly, we find school enrollment increases when mothers are literate and employed in a sector other than farming, which might correlate with a greater influence on familial decisions.

With respect to Hypotheses 2a & 2b, we estimated differences in grade-for-age between Cambodian children residing with only one versus both of their parents amounting to 0.13 of a SD (0.23 over 1.75, **Tables 1** and **3**) in a model without parental characteristics, and barely changed (0.14 of a SD, 0.24 over 1.75) when some maternal characteristics are added. The literacy gradient is comparatively larger, amounting to roughly half of a SD (0.23 over .44, **Tables 1** and **4**) before and nearly two thirds of a SD (0.27 over .44) after maternal characteristics are introduced. By this metric, differences in Cambodia appear to be within the range of those found in a number of high-income Nations as reviewed in our Background section. Contrary to what has been suggested elsewhere in Southeast Asia (Malaysia, Indonesia and Thailand), we do not find that living with only one parent due to the death of the other parent reduces the educational gradient. In fact, none of the coefficients for parental survival are significant and if anything the odds of being in school and literate are lower for children not living with their father when the father is deceased. While educational gradients have been found to be smaller in poor Nations than in rich ones, our results are more consistent with the hypothesis that the relative lack of welfare support at the time of the survey contributed to produce relatively large educational gradients in this low-income Nation. In this respect, the future impact of the RGC's steps towards a more comprehensive and integrated social protection system will deserve further analysis.

As indicated in the background section, educational expenses may account in some households for a sizable portion of the household's disposable income left after purchasing basic necessities. Although we are limited in our ability to assess household's financial situation, we find that socio-economic indicators are positively associated with children's school attendance, grade-for-age, and literacy (parents employed in non-agricultural sectors and when farming, not as paid laborers). This likely explains in part the evidence we find in support of Hypothesis 3 that, consistent with results in the U.S.A., living in a multi-generational households substantially improves children's educational outcomes. Anecdotal evidence suggests that parents may move into a multi-generational living arrangement to increase their own mobility and ability to respond to more remote work opportunities. On some measures, we even found that children not residing with either parent fared better than those residing with both parents, which may also be attributed to the opportunistic

placement of children away from home to facilitate their school attendance.

A better understanding of the selection mechanisms into the various types of household structure would clearly be needed to make causal claims about the effect of parental absence in Cambodia. The cross-sectional distribution of household structures in Cambodia may seem more similar to the same distribution in the U.S.A. than expected. The nuclear household structure dominates in Cambodia too. The main difference is a high level of flexibility and fluidity, for children as well as for adults. Whereas in the West, except in rare circumstances, parents' identical treatment of siblings seems to be the norm, many Cambodian parents readily admit to providing differently to children that they perceive as having different abilities and personalities from birth (Smith-Heffner, 1999). In Cambodia, accounting for the endogeneity of children's living arrangements would thus require controlling not just for parents' characteristics, as is common in studies in the West, but for children's characteristics as well—a point that had also emerged from the earlier literature of child fosterage. Not yet available at the time of this writing, additional data from subsequent MIPRAoC rounds will allow longitudinal analyses to account for some of the effects of child characteristics and, in particular, for the strategic placement of certain children away from the parental household.

Acknowledgments

The authors benefited from facilities and resources provided by the California Center for Population Research at UCLA (CCPR), which receives core support from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). The authors also acknowledge direct support from NICHD through a 2007–2014 research grant to the first author. This research grant has provided major support to the Mekong Integrated Population-Registration Areas of Cambodia (MIPRAoC) project, which provided the data analyzed in this paper. MIPRAoC integrates the activities and extant data of The Mekong Island Population Laboratory (MIPopLab), which received major support from NICHD through a 2002–2005 research grant to the first author. The transition from MIPopLab to MIPRAoC was supported by a 2005 seed grant to Alan Kolata and the first author from the Population Research Center, NORC and the University of Chicago, through its NICHD infrastructure grant. Additional support has been provided by a 2014–2015 seed grant to Adriana Lleras-Muney, Manisha Shah and the first author from the CCPR, a 2015–2016 pilot project award to the first author from the University of Southern California/University of California-Los Angeles Biodemography Center, which receives core support from the National Institute on Aging (NIA), and a 2017 fellowship to the first author from the Center for Khmer Studies.

In addition to the above institutions and individuals, the first author gratefully acknowledges advising from Colm O'Muircheartaigh for the design of MIPRAoC. From 2005 to 2014, Poch Bunnak has served as the MIPRAoC project manager. Since 2007, Samuel J. Clark oversees the database management of the project, with the assistance of Jeffrey Eaton and Benjamin Clark. Dwight Davis also contributed to the coding and cleaning of the data files.

While they remain solely responsible for the content of this paper, the authors also wish to thank Wei-Jun Jean Yeung, Hyunjoon Park, Shannon Cavanagh, Rachel Goldberg, and the participants of the UCLA Sociology Family Working Group for suggestions made on an earlier version of this paper.

Authors' Contribution

Patrick Heuveline designed the study, supervised the analyses and drafted the first draft of the manuscript. Savet Hong prepared the data and conducted the analyses. The authors jointly revised the manuscript.

Conflict of Interest

The authors report no conflict of interest.

Ethics Approval

Ethics approval for the MIPRAoC Project has been obtained by the Internal Review Board of the University of California, Los Angeles (IRB#11-002684) and from the National Ethics Committee for Health Research of the Ministry of Health, Royal Government of Cambodia (#83NECHR). Consent was obtained from all human subjects in the study.

Grant Numbers and Funding Information

Research grant R01HD054618 (Heuveline, P.I.) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD); California Center for Population Research, UCLA infrastructure grant R24HD041022 (Pebley, P.I.) from NICHD; University of Southern California/University of California-Los Angeles Biodemography Center infrastructure grant P30AG017265 (Seeman and Crimmins, P.I.) from the National Institute on Aging (NIA);

Senior Fellowship from the Center for Khmer Studies.

References

Ayres DM (2000). Tradition, modernity, and the development of education in Cambodia. *Comparative Education Review*, 44(4): 440–462. http://doi.org/10.1086/447629.

Beegle K, De Weerdt J and Dercon S (2009). Orphanhood and human capital destruction: Is there persistence into adulthood? *Demography*, 47(1): 163–180. http://doi.org/10.1353/dem.0.0094

Bianchi SM, Robinson JP and Milkie MA (2006). Changing Rhythms of American Family Life. New York: Russell Sage.

Biblarz TJ and Gottainer G (2000). Family structure and children's success: A comparison of widowed and divorced single-mother families. *Journal of Marriage and the Family*, 62(2): 533–548. http://doi.org/10.1111/j.1741-3737.2000.00533.x.

Birdthistle I, Floyd S, Nyagadza A, *et al*. (2009). Is education the link between orphanhood and HIV/HSV-2 risk among female adolescents in urban Zimbabwe? *Social Science and Medicine*, 68(10): 1810–1818. http://doi.org/10.1016/j.socscimed.2009.02.035

Bledsoe C (1990). The social management of fertility: Child fosterage among the Mende of Sierra Leone. In Handwerker WP (editors). *Births and power: Social change and the politics of reproduction* (pp. 81–101). Boulder, CO: Westview Press.

Brehm WC (2016). The contemporary landscape of education in Cambodia. In Brickell K and Springer S (editors). *The handbook of contemporary Cambodia* (pp.271–282) Abingdon, United Kingdom, and New York: Routledge.

Case A and Ardington C (2006). The impact of parental death on school outcomes: Longitudinal evidence from South Africa. *Demography*, 43(3): 401–420. http://doi.org/10.1353/dem.2006.0022.

Chae S (2016). Parental divorce and children's schooling in rural Malawi. *Demography*, 53(6): 1743–1770. http://doi.org/10.1007/s13254-016-0521-7

Chea H and Sok H (2001). The Cambodian garment industry. *Cambodia Development Review*, 5(3): 1–8.

Coleman J (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 95: S95–S120. http://doi.org/10.1086/228943.

Collinson MA, Tollman SM, Kahn K, *et al*. (2006). Highly prevalent circular migration: Households, mobility and economic status in rural South Africa. In Tienda M, Findley SE, Tollman SM, *et al*. (editors.). *Africa on the move: African migration and urbanisation in comparative perspective* (pp. 194–216). Johannesburg, South Africa: University of the Witwatersrand Press.

Cooke LP and Baxter J (2010). "Families" in international context: Comparing institutional effects across Western societies. *Journal of Marriage and Family*, 72(3): 516–536. http://doi.org/10.1111/j.1741-3737.2010.00716.x.

Das D (2016). "Enrollment, educational expenditures and work among one-parent children in India." *Marriage and Family Review*, 52(1–2): 196–215. http://doi.org/10.1080/01494929.2015.1073652.

DeLeire T and Kalil A (2002). Good things come in 3's: Multigenerational coresidence and adolescent adjustment. *Demography*, 39(2): 393–413. http://doi.org/10.1353/dem.2002.0016.

Demont F and Heuveline P (2008). Diversity and change in Cambodian households, 1998-2006. *Journal of Population Research*, 25(3): 287–313. http://doi.org/10.1007/BF03033892.

Derks A (2008). Khmer women on the move: Exploring work and life in urban Cambodia. Honolulu: University of Hawaii Press.

DeRose L, Corcuera Garcia P, Salazar A, *et al*. (2014). Household structure and school attendance in 67 Countries: Why children with absent fathers do better in some places. Paper presented at the 2014 meetings of the Population Association of America, May 1–3, Boston, MA.

Ear S (2011). Growth in the Rice and Garment Sectors. In Hughes C and Un K (editors) *Cambodia's economic transformation* (pp.70–93). Copenhagen, Denmark: Nordic Institute of Asian Studies.

Ebihara M (1968). Svay : A Khmer village in Cambodia, doctoral thesis. New York: Columbia University Press.

Esara P (2012). Moral scrutiny, marriage inequality: Cohabitation in Bangkok, Thailand. *The Asia Pacific Journal of Anthropology*, 13(3): 211–227. http://doi.org/10.1080/14442213.2012.680486.

Esteve A, Lesthaeghe R and López-Gay A (2012). The Latin American cohabitation boom, 1970–2007. *Population and Development Review*, 38(1): 55–81. http://doi.org/10.1111/j.1728-4457.2012.00472.x.

Evans D K and Miguel E (2007). Orphans and schooling in Africa: A longitudinal analysis. *Demography*, 44(1): 35–57. http://doi.org/10.1353/dem.2007.0002.

Filmer D and Schady N (2009). School enrollment, selection and test scores. Washington, D.C.: World Bank.

Flores G, Por I, Chean RM, *et al*. (2013). Financial protection of patients through compensation of providers: The impact of Health Equity Funds in Cambodia. *Journal of Health Economics*, 32(6): 1180–1193. http://doi.org/10.1016/j.jhealeco.2013.09.012.

Garib G, Martin Garcia T and Dronkers J (2007). Are the effects of different family forms on children's educational performance related to the demographic characteristics and family policies of modern societies? In Moerbeek H, Niehof A and van Ophem J (editors). *Changing families and their lifestyles* (pp.27–50). Wageningen, The Netherlands: Wageningen Academic.

Gertler P, Levine DI and Ames M (2004). Schooling and parental death. *The Review of Economics and Statistics*, 86(1): 211–225. http://doi.org/10.1162/003465304323023769.

Hampden-Thompson G (2013). Family policy, family structure, and children's educational achievement. *Social Science Research*, 42(3): 804–817. http://doi.org/10.1016/j.ssresearch.2013.01.005.

Heuveline P and Hong S (2016). One-parent families in contemporary Cambodia. *Marriage & Family Review*, 52(1–2): 216–242. http://doi:10.1080/01494929.2015.1076553.

Heuveline P and Poch B (2006). Do marriages forget their past? Marital stability in post-Khmer-Rouge Cambodia. *Demography*, 43(1): 99–125. http://doi.org/10.1353/dem.2006.0005.

Heuveline P, Yang H and Timberlake JM (2010). It takes a village (or perhaps a Nation): Family structures, State policies and children's educational achievement. *Journal of Marriage and Family*, 72(5): 1362–1376. http://doi.org/10.1111/j.0022-2445.2004.00088.x.

Hill M, Yeung W and Duncan G (2001). Childhood family structure and young adult behaviors. *Journal of Population Economics*, 14: 271–279. http://doi.org/ 10.1007/s001480000.

Hoefinger H (2013). Sex, love and money in Cambodia: Professional girlfriends and transactional relationships. Abington, United Kingdom and New York: Routledge.

Huisman J and Smits J (2009). Effects of household-and district-level factors on primary school enrollment in 30 developing countries. *World Development*, 37(1): 179–193. http://doi.org/10.1177/2158244015609666.

International Labour Office (2012). Cambodia: Social protection expenditure and performance review. Geneva, Switzerland: International Labour Office.

Inter-university Consortium for Political and Social Research [distributor]. 2017. The Mekong Island Population Laboratory (MIPopLab), A Demographic Surveillance System in Rural Cambodia (2000-06). ICPSR36601-v1. Ann Arbor, MI: ICPSR, 2017-04-13. http://doi.org/10.3886/ICPSR36601.v1

Isiugo-Abanihe UC (1985). Child fosterage in West Africa. *Population and Development Review*, 11(1): 53–74.

Kandel W and Kao G (2001). The impact of temporary labor migration on Mexican students' educational aspirations and performance. *International Migration Review*, 35(4): 1205–1231. http://doi.org/10.1111/j.1747-7379.2001.tb00058.x.

Kim S (2011). Reciprocity: Informal patterns of social interaction in a Cambodian Village. In J. Marston (Ed.) *Anthropology and community in Cambodia: Reflections on the work of May Ebihara* (pp.153–69). Victoria, Australia: Monash Asia Institute.

Kuhn R (2006). The effects of fathers' and siblings' migration on children's pace of schooling in rural Bangladesh. *Asian Population Studies*, 2(1): 69–92. http://doi.org/10.1080/17441730600700572.

Lloyd CB and Blanc AK (1996). Children's schooling in sub-Saharan Africa: The role of fathers, mothers, and others. *Population and Development Review*, 22(2): 265–298.

Loenzien M de (2016). Lone motherhood and its educational outcomes for children in Vietnam. *Marriage and Family Review*, 52(1–2): 162–195. http://doi.org/10.1080/01494929.2015.1136859.

Madhavan S (2004). Fosterage patterns in the age of AIDS: Continuity and change. *Social Science & Medicine*, 58(7): 1443–1454. http://doi.org/10.1016/S0277-9536(03)00341-1.

Mahaarcha W and Kittisuksathit S (2009). Impact of family structure, parental migration, and parental divorce on an adolescent's educational enrollment: Evidence from a longitudinal study in Kanchanaburi province, Thailand. *Journal of Population and Social Studies*, 18(1): 1–22.

Marks GN (2006). Family size, family type and student achievement: Cross-national differences and the role of socioeconomic background and schools. *Journal of Comparative Family Studies*, 31(1): 1–24.

McLanahan S, Tach L and Schneider D (2013). The causal effects of father absence. *Annual Review of Sociology*, 39: 399–427.

Ministry of Education, Youth and Sports. 1997. Education statistics, 1996–1997. Phnom Penh, Cambodia: RGOC, UNESCO, and

UNDP.

National Committee for Population and Development (2009). *Migration and development in Cambodia*. Phnom Penh, Cambodia: The Office of the Council of Ministers and United Nations Population Fund.

National Institute of Statistics (1999). *General population census of Cambodia 1998: Final census results*. Phnom Penh, Cambodia: Ministry of Planning.

Nguon S. 2012. Parental involvement and students' achievement in Cambodia: Focusing on parental resourcing of public schooling. *International Journal of Educational Research*, 53: 213–24. http://doi.org/10.1016/j.ijer.2012.03.011.

Nobles J (2011). Parenting from abroad: Migration, nonresident father involvement, and children's education in Mexico. *Journal of Marriage and Family*, 73(4), 729-746. http://doi.org/10.1111/j.1741-3737.2011.00842.x.

Nyamukapa C and Gregson S (2005). Extended family's and women's roles in safeguarding orphans' education in AIDS-afflicted rural Zimbabwe. *Social Science and Medicine*, 60(10): 2155–2167. http://doi.org/10.1016/j.socscimed.2004.10.005

Park H (2007). Single parenthood and children's reading performance in Asia. *Journal of Marriage and Family*, 69(3): 863–877. http://doi.org/10.1111/j.1741-3737.2007.00410.x.

Parreñas RS (2005). Children of global migration: Transnational families and gendered woes. Stanford: Stanford University Press.

Pong SL (1996). School participation of children from single-mother families in Malaysia. *Comparative Education Review*, 40(3): 231–249. http://doi.org/10.1086/447383.

Pong SL, Dronkers J and Hampden-Thompson G (2003). Family policies and children's school achievement in single-versus two-parent families. *Journal of Marriage and Family*, 65(4): 681–699. http://doi.org/10.1111/j.1741-3737.2003.00681.x.

Planning MO (2015). Cambodia socio-economic survey 2014. Phnom Penh, Cambodia: Ministry of Planning.

Raymo JM, Hyunjoon P, Yu X, *et al*. (2015). Marriage and family in East Asia: Continuity and change. *Annual Review of Sociology*, 41: 471–492. http://doi.org/10.1146/annurev-soc-073014-112428.

Safman RM (2003). Assessing the impact of orphanhood on Thai children affected by AIDS and their caregivers. *AIDS Care*, 16(1): 11–19. http://doi.org/10.1080/0954012031000163393O.

Schiller KS, Khmelkov VT and Wang XQ (2004). Economic development and the effects of family characteristics on mathematics achievement. *Journal of Marriage and Family*, 64(3): 730–742. http://doi.org/10.1111/j.1741-3737.2002.00730.x.

Shirahase S and Raymo JM (2014). Living arrangements and poverty among single mothers in Japan. *Social Forces*, 93(2): 545–569. http://doi.org/10.1093/sf/sou077.

Smith-Hefner NJ (1999). Khmer Americans: Identity and moral education in a diasporic community. Berkeley: University of California Press.

Townsend N, Madhavan S, Tollman S, *et al*. (2002). Children's residence patterns and educational attainment in rural South Africa, 1997. *Population Studies*, 56(2) : 215–225. http://doi.org/10.1080/00324720215925.

van de Kaa DJ (1987). Europe's second demographic transition. *Population Bulletin* 42(1): 1–59. Washington, D.C.: Population Reference Bureau.

World Bank (2014). *Where have all the poor gone? Cambodia poverty assessment 2013*. Washington, D.C.: World Bank.

Xenos P and Kabamalan MM (2007). Emerging forms of union formation in the Philippines. *Asian Population Studies*, 3(3): 263–286. http://doi.org/10.1080/17441730701746417.

Yao L and Treiman DJ (2011). Migration, remittances and educational stratification among blacks in apartheid and post-apartheid South Africa. *Social Forces*, 89(4), 1119–1143. http://doi.org/10.1093/sf/89.4.1119.

International Journal of Population Studies

RESEARCH ARTICLE

Consequences of forced migration during early childhood on cognitive well-being in later childhood in Andhra Pradesh, India

Ashish Kumar Upadhyay*, Swati Srivastava and Chhavi Paul

International Institute for Population Sciences, Mumbai, India

Abstract: Unlike its short-term impact on consumption and income, forced migration is expected to deliver a permanent shock to the overall well-being of households, specifically children in the stage of infancy. Studies on the effect of forced migration on child cognitive well-being are few in number. Therefore, the present study is intended to examine the consequences of forced migration during infancy on child cognition at later age. We hypothesized that the effect of forced migration on child cognitive well-being can be mitigated by social support. The study used longitudinal data from three waves of the Young Lives Study (YLS) conducted in 2002, 2006–2007, and 2009 in the state of Andhra Pradesh, India. We used bivariate and multivariate regression models to analyze the consequences of forced migration in early childhood on the cognitive well-being in later childhood. The information on forced migration was collected in Wave 1 (at age 1), whereas the information on the cognitive well-being of the children was collected in Wave 3 (at age 8). Child cognitive well-being was measured using scores obtained by the children on the Peabody Picture Vocabulary Test (PPVT), math, Early Grade Reading Assessment (EGRA), and memory tests. The results of the bivariate analysis show that the mean PPVT, math, EGRA, and memory scores obtained by children from the migrated households were lower than those obtained by children from the non-migrated households. Results of the multivariate linear regression models also show that children from the migrated households were statistically less likely to achieve higher scores on math (coefficient: -2.008, 95% C.I.-3.108, -0.908), EGRA (coefficient: -0.746, 95% C.I.-1.366, -0.126), and memory (coefficient: -0.503, 95% C.I. -0.834, -0.173) as compared to children from the non-migrated households. Our findings also indicate that the effect of forced migration on child cognitive well-being was not mitigated by social support. Findings of this study conclude that forced migration during infancy has a significant effect on child cognitive well-being at later age. Therefore, interventions should be made, paying attention to the most vulnerable children who were displaced during critical development ages.

Keywords: forced migration; cognitive well-being; Peabody Picture Vocabulary Test (PPVT); Early Grade Reading Assessment (EGRA); memory scores; social support; India

ARTICLE INFO
Received: April 27, 2017
Accepted: June 18, 2017
Published Online: June 25, 2017

***CORRESPONDING AUTHOR**
Ashish Kumar Upadhyay, International Institute for Population Sciences, Govandi Station Road, Deonar, Mumbai– 400088, India; ashu100789@gmail.com

CITATION
Upadhyay AK, Srivastava S and Paul C (2017). Consequences of forced migration during early childhood on cognitive well-being in later childhood in Andhra Pradesh, India. *International Journal of Population Studies*, 3(2): 16-28. doi: 10.18063/ijps.v3.i2.348

1. Introduction

Forced migration poses challenges that modulate its influence based on the intensity. The last few decades have witnessed an increase in the inclination to study the impact of forced migration on population development. An estimated number of 230 million people are currently living as international migrants, and the number is projected to surpass 400 million by 2050 (Martin，2013). In addition to the people who cross international borders, more than two to three times as many probably are internal migrants, people who have moved within their own countries (Esipova *et al*., 2013). The experts in the study of forced displacement have projected that between 25 million to one billion people are expected to displace from their current environment to a new point of destination over the next 40 years (OCHA/IDMC, 2009). The last two decades have seen an upsurge in the occurrence of the sudden-onset climate-related natural disasters, resulting in the displacement of more than 20 million people in 2008 alone (OCHA/IDMC, 2009). The less developed regions are more likely to face forced migration-related challenges due to their higher dependency on climate-sensitive factors and lower adaptive capacity in terms of human, financial and natural resources as well as restricted institutional and technological capabilities (Kniveton *et al*.， 2008).

Sudden forced displacement can have substantial repercussions on a child's overall development. Unlike its short-term impact on consumption and income, forced migration can deliver a more enduring shock that affects the overall well-being of households, specifically children in the infancy stage. Forced migration during infancy may affect a child's cognitive well-being in several ways. First, it may result in poor mental health of the mother. A number of previous studies have reported that poor mental health of the mother was negatively associated with the child's growth and cognitive development (Bennett *et al*., 2015). Second, it may be associated with a lack of nutrient intake, low immunization, childhood diseases and lower access to health care, which may lead to poor development of children at later ages.

Kondylis (2005) reported that conflict-induced migration has a negative impact on returnees' stock of human capital (Kondylis，2005). Rousseau *et al*., (1999) also examined the intergenerational effect of displacement on mental health. They reported that trauma experienced by parents before displacement was positively associated with risk behavior and school failure in boys. However, in the case of girls, trauma was associated with positive social adjustment. In a study in Colombia, the likelihood of chronic malnutrition due to forced displacement was found to range between 12.6% and 18.1% (Ortiz Becerra, 2014). Various studies have examined child immunization among refugee and displaced populations. Children born during increased hostilities in Sierra Leone were found lacking in age-appropriate immunization (Senessie *et al*., 2007). Hildebrandt *et al*., (2005) showed that migration may reduce the likelihood of breastfeeding and vaccinations (Hildebrandt *et al*., 2005). Some studies have shown that as a consequence of forced and environmental displacement, dropout rates were observed among children aged 12 years and above increased due to the emergence of alternative sources of household income (Ibáñez *et al*., 2010; World Bank, 2013). A study using the Young Lives data from Peru found maternal migration to be positively associated with child nutritional status, but found no significant effect of migration on cognitive development at the age of 5 years (Flores *et al*., 2009).

Pieces of evidence confirm the immediate and the long-term health implications of forced migration on the disadvantaged groups particularly in terms of child outcomes (Avogo *et al*., 2010; Ortiz Becerra, 2014; Rossi, 2008). Forced migration directly impacts childhood mortality through malnourishment and physical injuries. The less prominent impact includes psychological distress affecting a child's cognitive ability due to forced exit from a familiar environment to an unusual destination. Forced migrants are, therefore, more exposed to child developmental vulnerabilities as against non-forced migrants, who are better equipped to deal with the immediate as well as long-term disturbances arising out of migration (Agadjanian *et al*., 2003; Avogo *et al*., 2010; Doocy *et al*.， 2007; Guha Sapir *et al*.， 2004; O'Hare *et al*., 2007).

The availability of a social support system during forced migration helps to build a social environment that enables the displaced people to organize their life at the new destination. Displaced pregnant women seek prenatal care in order to ensure safe and healthy delivery. Support and referral services for such women are both directly and indirectly linked to child survival as well as development. In the absence of support, disturbances at the destination can result in abnormalities among the displaced children due to variations in the genetic, cognitive, physical, family, cultural, nutritional, educational, and environmental factors.

There is ample international literature on the impact of forced migration on early childhood nutritional development, schooling, household income, food consumption and adult human capital. However, there is no empirical study related to the effect of early childhood internal forced migration on children's cognitive well-being during later childhood in India. This may be on account of lack of data about migrant people. The present study aims to fills this gap by examining

the effect of forced migration during infancy on the cognitive well-being of the children at the age of 8 years, using longitudinal data from the first, second and third waves of the Young Lives Study from Andhra Pradesh, India.

2. Data and Methods

2.1 Data

We used data from the first, second, and third waves of the Young Lives Study (YLS), conducted in the state of Andhra Pradesh in India in 2002, 2006–2007 and 2009, respectively. The Young Lives Study is an international longitudinal study investigating the changing nature of childhood poverty. About 12,000 children are being followed in four countries: Ethiopia, Peru, Vietnam and India (Andhra Pradesh). Each country has two cohorts: younger cohort and older cohort. The younger cohort consists of about 2,000 children born in 2001–2002, and the older cohort consists of about 1,000 children born in 1994–1995 to be followed over a period of 15 years. The YLS is conducted every three/four years to collect data on a range of indicators related to the growth and development of children. The YLS also collects information on child welfare outcomes including nutritional status, growth, physical health, cognitive development, social and emotional well-being, and educational development.

A multistage sampling design was adopted by the YLS in India. In the first stage, two districts were selected from each of the three geographic regions (Coastal, Rayalaseema and Telangana) of the state of Andhra Pradesh. The selection of the districts was based on their relative rankings on the economic, human, and infrastructure development fronts. In the second stage, 19 (15 from rural areas and 4 from urban areas) sentinel sites (administrative blocks or 'mandals') were selected from the six sampled districts. In addition, one sentinel site was selected from the urban slum of the city of Hyderabad. In the third stage, villages were selected from the sampled rural sentinel sites and wards were selected from the sampled urban sites. Each sentinel site was divided into four contiguous geographical areas, and one village was randomly selected from each area. All the households with a one-year-old child (born in 2001–2002) or an eight-year-old child (born in 1994–1995) in the selected villages and wards were included in the YLS. Overall, 2011 households (with 2011 children) in the younger cohort and 1008 households (with 1008 children) in the older cohort were included in the first wave of the YLS, which was conducted in 2002 (for details of the YLS sampling design, please see Kumra, 2008).

The second wave took place between the late 2006 and the early 2007. Of the 2011 children (younger cohort) surveyed in the first wave, 32 children had died, 7 (households) refused to continue with the study and 22 children were untraceable. Hence, the second wave included only 1950 children. The third wave took place in 2009 and included 1930 children (of the 1950 children in Wave 2, 4 children had died, 5 (households) refused to continue with the study and 11 were untraceable) in the younger cohort. The attrition rate between Waves 1 and 2 was about 3% and between Waves 2 and 3 about 1% (Barnett *et al*., 2012).

We used data from each of the three waves of the YLS to examine the effect of forced migration on child cognitive well-being. Of the 2011 children who were surveyed in the first wave, information on the migration status was available for 1,913 children in Wave 3. However, the analytical sample size to investigate the effect of forced migration on child cognitive well-being in terms of Peabody Picture Vocabulary Test (PPVT), maths, Early Grade Reading Assessment (EGRA), and memory scores was 1,845, 1,861, 1,864 and 1,875, respectively. Of the 1,913 children (for whom the information on forced migration was available), the PPVT, maths, EGRA and memory tests could not be conducted for 29, 26 and 15 children respectively.

2.2 Outcome Variables

The outcome variables of interest were the scores attained by children on the Peabody Picture Vocabulary Test (PPVT), Maths Achievement Test, EGRA, and memory test. The information on each of the outcome variables–the PPVT, math, EGRA, and memory scores–were collected during the third wave of the YLS (when the children were at age 8 years).

The PPVT is a very common and widely accepted assessment test for identifying the verbal abilities, learning disabilities and scholastic aptitude among school-going children. The YLS uses version-III (204 items, Dunn *et al*., 1997) to assess PPVT scores in India. We used the PPVT score in natural logarithmic units to model a potential non-linear association between PPVT score and forced migration. Mathematics achievement test was used to identify children's sense of numbers. The maths test included 20 items on addition, subtraction, multiplication and division. The EGRA was used to assess the ability to recognize letters of the alphabet, read simple words and understand sentences and paragraphs and to assess listening comprehension. For details of the PPVT, maths test and the EGRA, see Cueto and Leon, 2012.

2.3 Key Exposure Variables

The key exposure variables of interest were household forced migration (yes versus no) and the social support (low versus medium/high) received by the households in Wave 1 (children aged 1 year).

The variable of forced migration was created using two questions asked in the first wave of YLS. The survey asked "Since pregnancy of the index child, have there been any big changes or events (natural disaster; moved/migrated/fled; decrease in food availability; deaths of livestock; crop failure; deaths of household members; severe illness/injury and victim of crime) that decreased the economic welfare of your household? If the answer was yes, then the survey further asked "What did the households do in response to the big changes/events?". Out of the 2,011 households/children recruited in Wave 1, the information on the migration status was available only for 1,913 households/children in wave 3. A total of 6.2% households (n = 119) reported that they had migrated since the pregnancy of the index child due to the aforementioned catastrophic events in Wave 1. It is worth noting that the present study used the term forced migration as defined by the International Organization for Migration (IOM) (Laczko and Aghazarm, 2009).

The information on social support (economic support and emotional support or assistance) was also collected in the survey. The survey asked the women in the previous 12 months they had received any kind of economic or emotional help or assistance from a work-related/trade union (yes, no), community association/co-op (yes, no), women's group (yes, no), political group(yes, no), religious group (yes, no), credit or funeral group (yes, no), sports group (yes, no), family (yes, no), neighbourhood (yes, no), friends (yes, no), community leaders (yes, no), religious leaders (yes, no), politicians (yes, no), government officials/civil service (yes, no), charitable organizations/NGOs (yes, no), and any other group (yes, no). If a woman had received any kind of help, her reply was coded as '1'; otherwise, it was coded as '0'. Help received from each group/person was added so that the amount of help received ranged from 0 to 16. If a woman reported no help or assistance, it was regarded as 'low' social support. Support ranging from 1 to 4 was considered as 'medium' social support and from 5 to 16 as 'high' social support. The details of the social support measurement are presented elsewhere (Galab *et al*., 2003).

2.4 Other Variables

Past studies on the subject of child cognitive well-being have included a number of other socio-economic, demographic and residence-related variables. Accordingly, we included birth size (below average, average and above), preterm birth (full term, preterm), sex of child (male, female), ever breastfed (no, yes), serious illness/injury in Wave 1 (no, yes), stunting at Wave 1(not stunted, stunted), child immunization (no/partial, full), pre-schooling(no, yes), type of school (private, government), mother's height (in cm), mother's education (below primary, primary and above), mother's age at birth of child (<18 years, 18–24 years, 25–29 years, ≥30 years), mother's working status (not working, agricultural work, other), household head education (below primary, primary and above), sex of household head (male, female), drinking water (improved, non-improved), toilet facility (improved, non-improved), wealth index in Waves 1 and 3 (poor, middle, rich), religion (Hindu, Muslim, other), caste (Scheduled Tribe (ST), Scheduled Caste (SC), other backward classes (OBC), other), and place of residence (rural, urban).

The survey collected data on the respondent's (mother/caregiver) perception of the size of the baby at birth. The YLS asked the respondent if the child was very small, small, average, large or very large at birth? Very small or small size at birth was coded as 'below average', and average, large or very large size at birth was coded as 'average and above'.

During the first wave of the survey, the mothers/caregivers were asked whether the child had suffered from any serious illnesses or injury since birth which led the mother/caregiver to think that the child might die (Yes/No/Don't know). If the mother replied with a 'yes', then serious illness was coded as 'yes'; otherwise it was coded as 'no'.

In the second wave of the YLS, the survey asked to the mothers whether the child had received vaccinations for tuberculosis (BCG); diphtheria, whooping cough (pertussis) and tetanus (DPT); poliomyelitis (Polio); and measles. Children who had received all the afore-mentioned vaccines were coded as 'fully immunized'. The remaining children were coded as 'no/partially immunized'.

The wealth index was calculated using the wealth scores, which are already computed and given in the YLS dataset. The wealth scores were generated through principal component analysis conducted on a set of variables including household assets (including radio, refrigerator, bicycle, television, motorbike/scooter, car, pump, sewing machine, mobile, phone, landline telephone, fan, almirah, clock, table, chair, sofa, bedsheet and animals), household quality (including

wall, roof and floor) and services (including electricity, drinking water, toilet facility). The lowest 33.3% households were coded as poor, the next 33.3% as middle, and the remaining 33.3% as rich.

The YLS has also collected information on the main source of drinking water. Children were classified into two categories according to whether the households they lived in used safe or unsafe water for drinking. Households having piped water in a dwelling/yard/plot or using a public tap/standpipe or a tube well/borehole or protected dug well were considered as using safe drinking water. Other households were categorized as using unsafe drinking water. Information on the type of toilet facility used by the households was also gathered in each of the three waves of the YLS. Improved toilet facilities included flush toilet/pit latrine connected to a septic tank. Non-improved toilet facilities included public/shared facility, simple latrine, toilet in a health post or defecating in a/an forest/field/open place (WHO and UNICEF, 2012 update).

2.5 Statistical Methods

Bivariate analysis was performed to compare the characteristics of migrant and non-migrant households using cross tabulation. Furthermore, a series of multivariate linear regressions were used to examine the effect of forced migration on child cognitive well-being. Each of the four outcome variables–PPVT score, maths score, EGRA score and memory score (collected in the third wave)–were regressed on household forced migration (collected in the first wave) and included all other variables listed in the 'other variables'. Adjusted coefficients and 95% confidence interval were reported. To assess whether social support mitigates the effects of forced migration on child cognitive well-being, we used the recommended procedure (Baron and Kenny, 1986). Variables were included into the multivariate model based on previous studies and based on their association with cognitive well-being in the bivariate analysis. All the variables were tested for multicollinearity using variance inflammation factor (VIF) before being included in the regression models. All the statistical computations were done in STATA 13.0.

3. Results

Figure 1 describes the mean PPVT-score, math-score, EGRA-score and memory-score obtained by children from migrant and non-migrant households. The children of non-migrant households outperformed than the children of migrant households on all tests. The mean PPVT, math, EGRA and memory scores of children from migrant households were 48.8, 9.1, 4.4 and 3.3 respectively, which were substantially lower than the scores of 59.3, 12.2, 5.4 and 3.8, respectively,

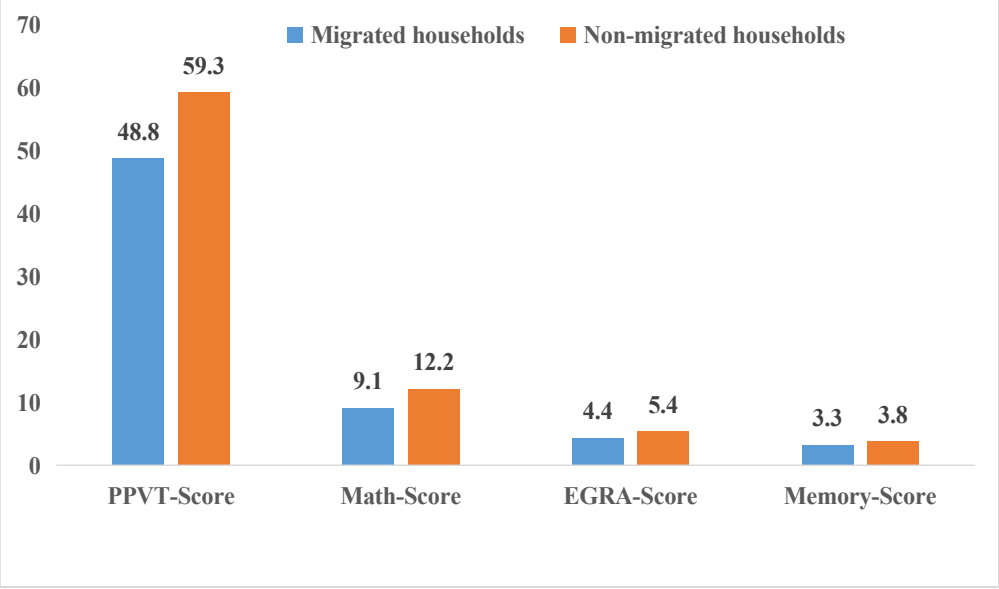


Figure 1. Mean PPVT-score, Math-score, EGRA-score and Memory-Score at the age of 8 years obtained by children from migrant and non-migrant households

obtained by children from non-migrant households.

Table 1 describes the bivariate analysis of child characteristics and maternal/household characteristics according to the status of forced migration in Andhra Pradesh (India). Results show that about 6.2% households had migrated since the pregnancy of the index child due to some catastrophic event following Wave 1. The information on the migration status of 1913 children interviewed in Wave 3 showed that around 29.4% of the children from the migrant households were below average size at birth, whereas only 25.5% of the children from the non-migrant households were below average size at birth. 87.4% of the migrant households and 79.7% of the non-migrant households had received medium/high social support. A higher percentage of children of non-migrant mothers had been breastfed compared to those of migrant mothers. Around 28.6% children from migrant households had suffered a serious illness/injury in wave 1, whereas only 21.7% children had done so in case of non-migrant households. About 33.9% migrant children were stunted, while only 29.7% non-migrant children experienced stunting. Migrant and non-migrant households enjoyed equal coverage of immunization and pre-schooling. The inter-class comparison of mother's education for both migrant and non-migrant households revealed that around 85.7% of the migrant mothers had below primary level education compared to only 58.7% of the non-migrant mothers. The migrant mothers were primarily engaged in agricultural work, while the majority of the non-migrant mothers were non-working. The majority of the migrant households lacked access to improved drinking water and improved toilet facility. About 94.1% of the migrant households belonged to a rural area compared to a lower figure of 72.4% in the case of non-migrant rural households.

Table 2 describes the results of multivariate linear regression analysis examining the effect of forced migration on child cognitive well-being in Andhra Pradesh, India. Model I presents the adjusted regression coefficients of all the scores for all the predictor variables except social support. Findings of Model I show that children born to migrant households were significantly less likely to achieve higher math (Coefficient: -2.008, 95% CI: -3.108, -0.908), EGRA (Coefficient: -0.746, 95% CI: -1.366, -0.126) and memory scores (Coefficient: -0.503, 95% CI: -0.834, -0.173) compared to the children of non-migrant households. Model II of **Table 2** presents the adjusted regression coefficients of all the scores for all the predictor variables including social support. The effect of social support was visible only in the case of PPVT scores. The children from households that had received medium/high social support were significantly more likely to achieve higher PPVT scores (Coefficient: 0.081, 95% CI: 0.030, 0.133) compared to those from households that had received low social support. Notably, in this model also, the children born to migrant households were significantly less likely to achieve higher math, EGRA and memory scores compared to the children of non-migrant households. By comparing Model I and Model II, we found that the magnitude of the effect of forced migration on child cognitive well-being remained unchanged even after controlling the social support.

A number of other variables were found to be statistically associated with the cognitive well-being of children. Preterm born children were significantly less likely to achieve higher PPVT scores (Coefficient: -0.101, 95% CI: -0.173, -0.030) compared to the children of full-term pregnancy. Age of children was also found to be significantly associated with the cognitive well-being, with the older children being more likely to achieve higher PPVT, math, EGRA and memory scores. Interestingly, female children were significantly less likely to achieve higher PPVT scores (Coefficient: -0.100, 95% CI: -0.140, -0.059), math scores (Coefficient: -0.539, 95% CI: -1.070, -0.008) and memory scores (Coefficient: -0.198, 95% CI: -0.358, -0.037) compared to the male children. Children who had suffered from a serious illness/injury during Wave 1 were significantly less likely to achieve higher PPVT scores (Coefficient: -0.062, 95% CI: -0.111, -0.014), math scores (Coefficient: -0.647, 95% CI: -1.283, -0.012) and EGRA scores (Coefficient: -0.388, 95% CI: -0.746, -0.029) than their counterparts. Stunted children during Wave 1 were significantly less likely to achieve higher PPVT, math, EGRA and memory scores compared to the non-stunted children during the same wave. It is worth noticing that the fully immunized children were significantly more likely to obtain higher math scores (Coefficient: 1.768, 95% CI: 0.757, 2.780) and EGRA scores (Coefficient: 0.577, 95% CI: 0.005, 1.149) compared to the not/partially immunized children. Children studying in government schools were significantly more likely to obtain higher EGRA and memory scores compared to the ones studying in private schools. Children whose mothers had primary and higher levels of education were significantly more likely to achieve higher PPVT, math and EGRA scores than those children whose mothers had below primary level education. The math, EGRA and memory scores were significantly higher among children of urban settings than those belonging to rural settings.

The estimates obtained from multivariate linear regression may have a potential selection bias due to the differences in

Table 1. Descriptive statistics of child characteristics and maternal/household characteristics according to status of forced migration in Andhra Pradesh (India)

Characteristics	Percentage	Sample Size (N = 1,913)[#]	Migrated (N = 119)	Not-Migrated (N = 1,794)	X^2(p-value)
Birth size					
Below average	25.7	492	29.4	25.5	0.906(0.341)
Average and above	74.3	1,421	70.6	74.5	
Social Support					
Low	19.7	373	12.6	20.3	4.108(0.043)
Medium/high	80.2	1,514	87.4	79.7	
Preterm birth					
Full term	91.5	1,750	92.4	91.4	0.149(0.699)
Preterm	8.5	163	7.6	8.6	
Sex of child					
Male	53.6	1,025	56.3	53.4	0.378(0.539)
Female	46.4	888	43.7	46.6	
Ever breastfed					
No	2.5	47	3.4	2.4	0.421(0.516)
Yes	97.5	1,842	96.6	97.6	
Serious illness/injury in Wave-1					
No	77.9	1,490	71.4	78.3	3.074(0.080)
Yes	22.1	423	28.6	21.7	
Stunting in Wave-1					
Not stunted	70.1	1,326	66.1	70.3	0.933(0.334)
Stunted	29.9	567	33.9	29.7	
Immunization					
No/partial immunization	7.6	145	7.6	7.6	0.001(0.994)
Full immunization	92.4	1,768	92.4	92.4	
Attended pre-schooling since age of 3-years					
No	12.9	246	13.5	12.8	0.0389(0.844)
Yes	87.1	1,667	86.5	87.2	
Type of school					
Private	44.2	846	32.8	45.0	6.745(0.009)
Government	55.8	1,067	67.2	55.0	
Mother's education					
Below primary	60.4	1,156	85.7	58.7	33.928(0.000)
Primary and above	39.6	757	14.3	41.3	
Mother's age at birth of child					
<18 year	8.3	159	16.0	7.8	9.810(0.007)
18–24 year	69.7	1,333	64.7	70.0	
>=25 year	22.0	421	19.3	22.2	
Mother's working status					
Not working	49.1	940	19.3	51.1	45.286(0.000)
Agricultural work	40.7	778	65.6	39.0	
Other work	10.2	195	15.1	9.9	
Household head education					
Below primary	59.1	1,130	77.3	57.9	17.464(0.000)
Primary and above	40.9	783	22.7	42.1	
Drinking water					
Improved	24.0	460	21.0	24.3	0.641(0.423)
Non-improved	76.0	1,453	79.0	75.6	

Table 1. Continued.

Characteristics	Percentage	Sample Size (N = 1,913)[#]	Migrated (N = 119)	Not-Migrated (N = 1,794)	X^2(p-value)
Toilet facility					
Improved	18.1	346	4.2	19.0	16.512(0.000)
Non-improved	81.9	1,567	95.8	81.0	
Wealth index in Wave-1					
Poor	33.4	639	51.3	32.2	39.237(0.000)
Middle	33.6	642	41.2	33.1	
Rich	33.0	632	7.5	34.7	
Wealth index in Wave-3					
Poor	32.9	629	45.4	32.0	13.498(0.001)
Middle	33.9	648	35.3	33.8	
Rich	33.2	636	19.3	34.2	
Place of residence					
Rural	73.7	1,410	94.1	72.4	27.279(0.000)
Urban	26.3	503	5.9	26.6	
Total	100	1,913	6.2	93.8	

Note: [#] number of cases may vary slightly based on missing observation of background characteristics

the socioeconomic development between migrant and non-migrant households. To partially overcome this bias, we used Propensity Score Matching (PSM) analysis in addition to multivariate regression models. PSM is a statistical technique that reduces the bias due to confounding variables, which in the case of the present study could be found in estimates for forced migration obtained from simply comparing outcomes for non-migrant and migrant households. Results obtained from PSM were similar to the results obtained by using multivariate linear regression analysis (see supplementary table S1).

4. Discussion

Using longitudinal data from three waves of the Young Lives Study conducted in 2002, 2006-2007 and 2009, we examined the consequences of forced migration during the early life of children on the cognitive well-being of the children at later age in Andhra Pradesh, India. Our study showed that migrant children were statistically less likely to have higher math, EGRA, and memory scores compared to the non-migrant children. These results hold even after adjusting for some of the well-known confounders of child cognitive well-being. Findings of this study are consistent with previous studies that have shown the adverse effect of forced migration on child outcomes in other countries (Avogo *et al*., 2010; Ortiz Becerra, 2014; Stevens *et al*., 2008). However, one particular study that used the YLS data from Peru, reported no significant effect of maternal migration on child cognitive well-being (Flores *et al*., 2009). Notably, the study from Peru had only taken the PPVT scores to measure child cognitive well-being. By contrast, the present study used some other test scores (math, EGRA and memory) to measure the child cognitive well-being. It is important to note that similar to the study from Peru, our study also did not find any significant effect of forced migration on child PPVT score.

Migrants in our study included those who had experienced economic loss due to drought, flood, earthquake, crime, crop failure, and so on. Therefore, it is possible that the observed effect of forced migration on cognitive well-being may be because of the economic loss at the household level. To ensure the robustness of our estimates, we ran another regression by taking interaction between economic shock and migration (results not shown) and comparing the cognitive well-being of the children for three groups: 1) household experienced economic shocks but did not migrate, 2) household experienced economic shock and migrated, and 3) non-migrated households. While comparing the cognitive well-being of children from these groups, we found that children belonging to households who had experienced an economic shock and had migrated were statistically less likely to get higher math, EGRA and memory scores than children from the households who had experienced an economic shock but had not migrated. No significant difference in child cognitive well-being was

Table 2. Results of multivariate linear regression analysis to examine the effect of forced migration on child cognitive well-being in Andhra Pradesh, India (2002-2009)

	PPVT-Score		Math-Score		EGRA-Score		Memory-Score	
	I	II	I	II	I	II	I	II
Migrant household (vs. non-migrant)	-0.074 (-0.157,0.009)	-0.078 (-0.162,0.005)	-2.008* (-3.108,-0.908)	-2.033* (-3.132,-0.934)	-0.746* (-1.366,-0.126)	-0.728* (-1.347,-0.109)	-0.503* (-0.834,-0.173)	-0.505* (-0.836,-0.175)
Social Support medium/high (vs. low)		0.081* (0.030,0.133)		0.419 (-0.257,1.095)		-0.029 (-0.410,0.352)		0.114 (-0.091,0.320)
Birth size average or above (vs. below average)	0.024 (-0.023,0.072)	0.020 (-0.028,0.067)	0.216 (-0.403,0.834)	0.194 (-0.427,0.816)	-0.142 (-0.492,0.207)	-0.168 (-0.519,0.183)	-0.001 (-0.188,0.187)	-0.017 (-0.206,0.171)
Preterm birth (vs. full term)	-0.101* (-0.173,-0.030)	-0.096* (-0.167,-0.024)	-0.356 (-1.291,0.579)	-0.332 (-1.272,0.609)	-0.292 (-0.820,0.235)	-0.344 (-0.874,0.186)	-0.175 (-0.458,0.107)	-0.172 (-0.457,0.113)
Age of child (in months)	0.010* (0.009, 0.019)	0.015* (0.010, 0.020	0.342* (0.275, 0.408)	0.344* (0.278, 0.411)	0.134* (0.097, 0.172)	0.134* (0.096, 0.171)	0.032* (0.012, 0.053)	0.033* (0.013, 0.054)
Female child (vs. male)	-0.100* (-0.140,-0.059)	-0.102* (-0.143,-0.061)	-0.539* (-1.070,-0.008)	-0.563* (-1.096,-0.030)	-0.067 (-0.367,0.232)	-0.038 (-0.339,0.262)	-0.198* (-0.358,-0.037)	-0.181* (-0.343,-0.020)
Serious illness/injury in Wave-1 (vs. no)	-0.062* (-0.111,-0.014)	-0.062* (-0.110,-0.013)	-0.647* (-1.283,-0.012)	-0.606 (-1.244,0.032)	-0.388* (-0.746,-0.029)	-0.365* (-0.725,-0.005)	-0.149 (-0.342,0.043)	-0.152 (-0.346,0.041)
Stunting at Wave-1 (vs. no)	-0.124* (-0.169,-0.079)	-0.126* (-0.172,-0.081)	-2.228* (-2.821,-1.635)	-2.213* (-2.808,-1.617)	-1.009* (-1.343,-0.674)	-0.982* (-1.318,-0.647)	-0.425* (-0.605,-0.245)	-0.430* (-0.610,-0.249)
Full immunization (vs. no/partial)	0.064 (-0.013,0.141)	0.064 (-0.013,0.142)	1.768* (0.757,2.780)	1.817* (0.802,2.832)	0.577* (0.005,1.149)	0.614* (0.041,1.188)	0.220 (-0.086,0.527)	0.231 (-0.077,0.539)
Enrolled a in public school (vs. private school)	0.012 (-0.038,0.062)	0.009 (-0.042,0.059)	0.620 (-0.036,1.276)	0.598 (-0.062,1.258)	0.704* (0.333,1.075)	0.715* (0.342,1.088)	0.270* (0.072,0.469)	0.272* (0.072,0.472)
Respondent's education is primary or above (vs. below primary)	0.090* (0.039,0.141)	0.086* (0.034,0.137)	2.611* (1.938,3.284)	2.554* (1.877,3.231)	1.119* (0.739,1.500)	1.093* (0.710,1.475)	0.202 (-0.002,0.406)	0.196 (-0.010,0.402)
Respondent's place of residence is urban (vs. rural)	0.009 (-0.059,0.078)	0.006 (-0.062,0.075)	1.732* (0.832,2.632)	1.706* (0.802,2.610)	1.130* (0.623,1.636)	1.095* (0.586,1.603)	0.369* (0.098,0.640)	0.339* (0.067,0.612)

observed between children from non-migrated households irrespective of whether had experienced an economic shock or not. Therefore, the significant effect in our study can be safely considered as an actual effect of forced migration on child cognition. Findings of the present study also indicate that the effects of forced migration on child cognitive well-being were not mitigated by social support. One of the possible reasons for the lack of association between social support and childhood cognitive well-being may be reverse causality, whereby households who were in a more adverse condition were more likely to receive support from other individuals within their community or from relatives and friends. Kawachi and Berkman (2001) reported that the protective effect of social support may not be uniform across society. To the best of our knowledge, this is the first study that examined the causal association between forced migration and child cognitive well-being in India.

When interpreting our findings, the following limitations should be taken into account. First, the effect of forced migration on child cognitive development may be influenced by the duration of migration (temporary, permanent, and returned migrants), the type of migration (rural to urban, urban to rural, and so on), and the urban-rural sampling structure. On our part, we were unable to assess this due to the unavailability of such information in the YLS. Also, comparisons of child cognition must be done at the sending places, particularly if the levels of socioeconomic development between the sending and receiving places are very different. Again, we were unable to split the sample according to their current migration status due to the unavailability of such information. Therefore, the negative effect of forced migration on child cognition in our study may be due to the socioeconomic differentials between the migrant and non-migrant households. However, the results obtained from PSM analysis also support the findings of the multivariate regression analysis. Consequently, estimates obtained from the multivariate regression analysis can be safely taken as the actual effect of forced migration on child cognition.

Second, we could not control the respondent/household characteristics prior to the forced migration due to the unavailability of data. We did, nevertheless, control the following variables: respondent's height, schooling of the respondent, respondent's age at birth of child and schooling of the head of household. We added respondent's height as a way of capturing genetic factors. We included the educational attainments of the respondent and the household head as a proxy for wealth (that is, household economic status prior to migration). Although, a number of previous studies have reported a high level of correlation between education and economic status, we must acknowledge the fact that all the variations in wealth may not be captured by education and, thus, some care is required while interpreting the findings of our study.

Third, the magnitude of the effect of forced migration on child cognitive well-being may be lower than expected due to mortality selection among children from migrant households. Some previous studies have reported that forced migration is significantly associated with higher infant and child mortality (Avogo *et al*., 2010). However, the effect of mortality selection on child cognitive development should be minimal due to the fact that only very few deaths occurred betwwen Wave 1 and Wave 3 in the YLS sample. Fourth, the observed differences in the cognitive well-being of migrant and non-migrant children may be due to a potential selection bias resulting from attrition between the first and third waves of the YLS. The attrition rate between waves 1 and 2 was about 3% and between waves 2 and 3 about 1% (Barnett *et al*., 2012). Dercon *et al*., (2008) found limited evidence of attrition bias in the YLS and argued that the attrition in the YLS samples were highly unlikely to cause a bias in research inferences.

Despite these limitations, our study has some strength. First, a large cohort of children was included in the analysis, representing children from a wide range of family backgrounds. Second, the YLS is the only large-scale available dataset that provides information on forced migration, social support, and child cognitive well-being in India. Third, the YLS uses a child-focused mixed sampling approach, allowing for an examination of the complex interrelationship between forced migration, social support, and child cognitive development in India. Fourth, the study included both rural and urban areas, representing a range of regions, policy context, and living conditions that reflect the ethnic, geographical and religious diversity of the country. Another key strength of this study is that the information on forced migration pertained to the period between pregnancy of the index child and attaining one year of age. Some studies have reported that the first 1,000 days of the child (including the duration of pregnancy) is a very crucial period for child cognitive well-being at later ages (Black *et al*., 2013). However, the majority of the previous studies have been unable to control for forced migration in this crucial period in their analyses (Flores *et al*., 2009; Rossi, 2008). Lastly, our study came out with some findings that may either lead to the formulation of new policies or may lead to the strengthening of the existing policies and programmes.

According to the United Nations High Commissioner for Refugees, the total number of displaced persons worldwide

was about 42 million in 2008, and about 44% of these migrants were children below age 18 (UNHCR, 2009). Findings of the present study call for immediate interventions from government and non-government organizations. Given the evidence about the effect of early childhood cognitive development on education, productivity, and job performance, children from migrant households need special interventions that attenuate the long-term effect of early childhood cognitive development on human capital. Interventions should pay attention to the most vulnerable children who were displaced during critical development ages. Social programmes executed in other developing countries have shown a positive effect of the intervention programmes on education and health of displaced children (Bernal *et al*., 2009). Further studies are needed to compare cognitive development among children born in a household during migration and those born after migration in the same household. Moreover, a follow-up of these children is necessary to assess the relationship between cognitive development and income during adulthood.

5. Conclusions

A number of studies have examined the adverse effect of forced migration on child nutritional status, childhood immunization and school dropout rates in developing countries. However, research on the effect of forced migration on child cognitive well-being is still lacking in developing countries including India. Therefore, using longitudinal data from three waves of the Young Lives Study conducted during 2002, 2006–2007 and 2009, the present study examined the effect of forced migration during early childhood on cognitive well-being at later age. We found that forced migration during early childhood was significantly associated with poor cognitive well-being at later childhood. The study also found that the adverse effect of forced migration on child cognition was not mitigated by maternal social support. The findings of this study have implications for intervention programmes that should pay attention to the most vulnerable children who were displaced during critical development ages.

Authors' Contributions

AKU conceived the idea, designed the experiment and analysed the data; AKU, SS and CP drafted the manuscript. All authors read and approved the final manuscript.

Ethical Approval

Our study is based on a secondary dataset with no identifiable information on the survey participants. This dataset is available in the public domain for research use; hence no approval was required from any institutional review board. The data can be downloaded from the website of the United Kingdom Data Archives University of Essex after creating an account (https://www.ukdataservice.ac.uk/). The data for the current study was downloaded from the afore-mentioned website after taking permission (I.D. No. 90978).

Competing Interests

The authors declare that they have no competing interests.

Availability of Data and Materials

The data which support our findings is contained within the manuscript.

Funding

No funding was available for this study.

References

Agadjanian V and Prata N (2003). Civil war and child health: Regional and ethnic dimensionsof child immunization and malnutrition in Angola. S*ocial Science & Medicine*, 56(12): 2515–2527. http://doi.org/10.1016/S0277-9536(02)00286-1.

Avogo WA and Agadjanian V (2010). Forced migration and child health and mortality in Angola. *Social Science & Medicine*, 70(1): 53–60. http://doi.org/10.1016/j.socscimed.2009.09.057.

Barnett I, Ariana P, Petrou S, *et al.* (2012). Cohort profile: The young lives study. *International Journal of Epidemiology*, 42(3): 701–

708. http://doi.org/10.1093/ije/dys082.

Baron RM and Kenny DA (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6): 1173. http://doi.org/10.1037/0022-3514.51.6.1173.

Bennett IM, Schott W and Krutikova S, *et al.* (2015). Maternal mental health, and child growth and development, in four low-income and middle-income countries. *Journal of Epidemiology & Community Health*, 70(2): 168. http://doi.org/10.1136/jech-2014-205311.

Bernal R, Fernández C , Flórez Nieto CE, *et al*. (2009). Evaluation of the early childhood program hogares comunitarios de bienestar in Colombia. Ssrn Electronic Journal. http:// doi.org/10.2139/ssrn.1486209.

Black RE, Victora CG, Walker SP, *et al*. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*, 382(9890): 427–451. https://doi.org/10.1016/S0140-6736(13)60937-X

Cueto S, Leon J, Guerrero G, *et al*. (2012). Psychometric characteristics of cognitive development and achievement instruments in round 3 of young lives.

Dercon S and Outes-Leon I (2008). Survey attrition and attrition bias in Young Lives. *Young Lives Technical Note*, 5. Available from: http://www.younglives.org.uk/files/YL-TN5-OutesLeon-Survey-Attrition.pdf.

Doocy S, Burnham G and Robinson C (2007). Estimating demographic indicators in a conflict-affected population in eastern Sudan. *Prehospital and Disaster Medicine*, 22(02): 112–119. https://doi.org/10.1017/S1049023X00004489.

Dunn LM and Dunn LM (1997). Peabody Picture Vocabulary Test-Third Edition (PPVT-III). Circle Pines, MN: American Guidance Service. *Circle Pines Mn American Guidance Service Inc*, 193(1): 2645–2666.

Esipova N, Pugliese A and Ray J (2013). The demographics of global internal migration. *Migration Policy Practice*, 3(2): 3–5.

Flores E and Escobal D'Angelo J (2009). Maternal migration and child well-being in Peru: Young lives.

Galab M, Gopinath R, Piyush A, *et al*. (2003). Young live preliminary country report: Andhra Pradesh, India 2003

Guha-Sapir D and Panhuis WG (2004). Conflict-related mortality: An analysis of 37 datasets. *Disasters*, 28(4): 418–428. https://doi.org/10.1111/j.0361-3666.2004.00267.x.

Hildebrandt N, McKenzie DJ, Esquivel G, *et al*. (2005). The effects of migration on child health in Mexico [with comments]. *Economia*, 6(1): 257–289.Avaliable from: http://www.jstor.org/stable/20065490.

Ibáñez AM and Moya A (2010). Vulnerability of victims of civil conflicts: empirical evidence for the displaced population in Colombia. *World Development*, 38(4): 647–663. https://doi.org/10.1016/j.worlddev.2009.11.015.

Kawachi I and Berkman LF (2001). Social ties and mental health. *Journal of Urban Health*, 78(3): 458–467. http://doi.org/10.1093/jurban/78.3.458.

Kniveton D, Schmidt-Verkerk K, Smith C, *et al*. (2008). Climate change and migration: Improving methodologies to estimate flows. IOM Migration Series, No.3

Kondylis F (2005). Agricultural production and conflict refugee status: Quasi-experimental evidence from a policy intervention programme in Rwanda. Londres: Economics

Department, Universidad de London, 8. Avaliable from:https://pdfs.semanticscholar.org/c3e1/d71e26f35573283c6c454a3ef7d98c99209a.pdf.

Kumra N (2008). An assessment of the young lives sampling approach in Andhra Pradesh, India. *Oxford: Young Lives*.

Laczko F and Aghazarm C (2009). Migration, environment and climate change: Assessing the evidence: International Organization for Migration Geneva.

Martin P (2013). The global challenge of managing migration. *Population Bulletin*, 68(2): 15. Avaliable from: http://www.prb.org/pdf13/global-migration.pdf.

O'Hare BAM and Southall DP (2007). First do no harm: The impact of recent armed conflict on maternal and child health in Sub-Saharan Africa. *Journal of the Royal Society of Medicine*, 100(12): 564–570. https://doi.org/10.1258/jrsm.100.12.564.

OCHA/IDMC (2009). Monitoring disaster displacement in the context of climate change. Findings of a study by the United Nations Office for the coordination of humanitarian affairs and the internal displacement monitoring centre. Avaliable from: http://www.

refworld.org/docid/4ab9cd4e2.html.

Ortiz BK (2014). Forced displacement and early childhood nutritional development in Colombia: Households in conflict network.

Rossi A (2008). The impact of migration on children in developing countries. Paper presented at the Unpublished manuscript prepared for the Youth Migration Conference. Avaliable from: http://globalnetwork.princeton.edu/bellagio/Rossi.pdf.

Rousseau C, Mekki-Berrada A and Rufagari M (1999). Traumatismes et séparations familiales prolongées chez les réfugiés du Congo-Kinshasa établis à Montréal [in French]. Trauma and separations protests for congo-kinshasa refugees in montreal. *Canadian Journal of African Studies/La Revue canadienne des études africaines*, 33(2–3): 584–592. https://doi.org/10.1080/00083968.1999.10751175.

Senessie C, Gage GN and von Elm E (2007). Delays in childhood immunization in a conflict area: A study from Sierra Leone during civil war. *Conflict and health*, 1(1): 1. https://doi.org/10.1186/1752-1505-1-14.

Stevens GWJM and Vollebergh WAM (2008). Mental health in migrant children. *Journal of Child Psychology and Psychiatry*, 49(3): 276–294. http://doi.org/10.1111/j.1469-7610.2007.01848.x.

UNHCR (2009). 2008 global trends: Refugees, asylum-seekers, returnees, internally displaced and stateless persons

WHO and UNICEF (2012 update). Progress on drinking water and sanitation from New York: World Health Organization & United Nations Children's Fund; 2012

World Bank (2013) Lebanon: Economic and social impact assessment of the Syrian conflict, Report No. 81098-LB, Lebanon, 20 September 2013, 78.

International Journal of Population Studies

RESEARCH ARTICLE

Do young children prohibit mothers from working? A study in the Amhara Region, Ethiopia

Chalachew Getahun Desta

Center for Population Studies, Addis Ababa University, Ethiopia

Abstract: Theoretical work relating economic effect of children suggests that labor market participation decreases for mothers with large number of young children and increases when children are adults. The majority of empirical studies find results consistent with this expectation, but there are some studies which fail to confirm this theoretical prediction for the developing countries. This paper used data from a household survey of rural and urban married women to test the theoretical prediction that labor market participation decreases for mothers with large number of young children and increases when children are adults. Results show that when all households are considered, children seem to have positive effects on the probability of the mother's work participation. However, when household lifecycle and rural-urban location differences are considered, coefficients are negative (but not statistically insignificant) for urban households with large number of young children and positive (and statistically significant) for those households with more adult children; whereas for rural households, these coefficient signs are reversed. Results from the quantitative data combined with qualitative narratives suggest that large numbers of young children do not prohibit rural mothers from working.

Keywords: *maternal work status; Amhara region of Ethiopia; lifecycle fertility; lifecycle maternal labor supply*

ARTICLE INFO
Received: May 31, 2017
Accepted: July 23, 2017
Published Online: July 30, 2017

***CORRESPONDING AUTHOR**
Chalachew Getahun Desta, Center for Population Studies, Addis Ababa University, 150400 Addis Ababa, Ethiopia; chalget@gmail.com

CITATION
Desta CG (2017). Do young children prohibit mothers from working? A study in the Amhara Region, Ethiopia. *International Journal of Population Studies*, 3(2): 29-42. doi: 10.18063/ijps.v3i2.208.

1 Introduction

Theories linking fertility to parental work, in general, view fertility as a response to the parents' demand for children relative to other components in the household's utility maximization decision. In the standard economic literature which considers children as economic goods (Becker, 1960), children are seen to be the result of parental choice relative to other essential household goods, all constrained by financial and time shortages in the household's utility maximization framework (Becker, 1960; Hotz and Miller, 1988). Given budget constraints in the household's utility maximization framework, parents have to make choices between labor market participation or childrearing (Becker, 1960; Hotz and Miller, 1988; Rosenzweig and Wolpin, 1980b). In this line of argument, therefore, the observed fertility is determined by variation in the opportunity cost of parental time in childcare (Becker, 1960; Hotz and Miller, 1988; Rosenzweig and Wolpin, 1980b). However, these static exogenous models of fertility are increasingly criticized since they ignore the inherently sequential nature of fertility and maternal labor supply decision making over the lifecycle (Adda, Dustmann and Stevens, 2011; Fehr and Ujhelyiova, 2012; Francesconi, 2002; Hotz and Miller, 1988;

Moffitt, 1984; Sheran, 2007). In a lifecycle framework, it is argued that since young children are more time intensive than adult children (Hotz and Miller, 1988; Fehr and Ujhelyiova, 2012; Moffitt, 1984), mothers may be less likely to work with an increasing number of young children and more likely to work when children are adults.

Lifecycle frameworks to fertility and labor supply decisions, which are seen to be dynamic and jointly determined, highlight the potential value of focusing on parents' preferred timing of childrearing and labor supply over the lifecycle (Francesconi, 2002; Hotz and Miller, 1988; McNicoll, 1984; Moffitt, 1984; Rosenzweig and Wolpin, 1980b; Sheran, 2007). In this framework, parents' fertility and labor market participation decisions reflect different timing preference. For example, couples may (1) prefer to have children early in their lifecycle and delay participation in the labor market, (2) participate in the labor market first delay childbearing, or (3) prefer to participate in the labor market both early and latter in their lifecycle, allocating their working ages for childrearing (Cho, 2006). The framework is, therefore, considered to be a useful tool in analyzing fertility and maternal labor supply behaviors over the lifecycle (Adda, Dustmann and Stevens, 2011; Cho, 2006; Erosa, Fuster and Restuccia, 2016; Fehr and Ujhelyiova, 2012; Francesconi, 2002; Hotz and Miller, 1988; McNicoll, 1984; Moffitt, 1984; Sheran, 2007).

However, while ample cross-national macro-level empirical evidence for the age dependency hypothesis (population level) is available, the lifecycle hypothesis (household level) is empirically little explored (see *e.g.*, Browning and Crossley, 2001; Chernichovsky, 1978; Davies, 1988; Kelley and Schmidt, 2001). In addition, despite the ample research undertaken to estimate the relationship between fertility and maternal labor participation, empirical research is complicated by the endogeneity of fertility in female labor supply decisions. Several studies have used instrumental variable estimation as a solution to this problem (*e.g.*, Aassve and Arpino, 2007; Angrist and Evans, 1998; Bloom, Canning, Fink, *et al*., 2009; Chun and Oh, 2002; Cruces and Galiani, 2007; Kim and Aassve, 2006; Kim, Engelhardt, Prskawetz, *et al*., 2009; Orbeta, 2005). The vast majority of empirical studies find results consistent with this theoretical prediction relating fertility and maternal labor participation, but there are some studies in the case of developing countries which fail to replicate the expected relationship (*e.g.*, Aghajanian, 1979 for Iran; Angrist and Evans, 1998 for references on a similar evidence for other countries; Cho, 2006 for Korea; Solomon and Kimmel, 2009 for Ethiopia).

When it comes to the context of Ethiopia, the allegedly adverse consequence of rapid population growth on economic development has been acknowledged by the national population policy and the different national development consecutive plans (IMF, 2006; Ministry of Finance and Economic Development, 2002; 2006; 2010; Transitional Government of Ethiopia, 1993; UNDP, 2001; UN Population Division, 2005). The development plans identify, among other things, maternal labor market participation as critical for achieving the planned development. One major strategy suggested by the plan documents to achieve this is ensuring balanced population and economic growth, for example, through reducing fertility.

In Ethiopia, fertility has been one of the highest among the developing countries, but substantial decline has begun in recent years (United Nations, 2017: 33). Modern contraceptive use has recently risen, for example, from 11 percent in 2000 (Central Statistical Agency of Ethiopia and ORC Macro, 2000) and 15 percent in 2005 (Central Statistical Agency of Ethiopia and ORC Macro, 2005) to 29 percent in 2010 (Central Statistical Agency of Ethiopia and ORC Macro, 2011). Correspondingly, total fertility rate is declining substantially though still high, for example, from 5.9 in 2000 (Central Statistical Agency of Ethiopia and ORC Macro, 2000) and 5.4 in 2005 (Central Statistical Agency of Ethiopia and ORC Macro, 2005) to 4.8 in 2010 (Central Statistical Agency of Ethiopia and ORC Macro, 2011). The average annual rate of population growth has also dropped from 2.9 percent during the 1984 – 1994 intercensal period to 2.6 percent during the 1994 – 2007 intercensal period (Hailemariam, Alayu and Teller, 2011; UNFPA, 2010).

However, with fertility declining much more slowly than mortality, the country is yet in the early stage of the demographic transition (Ringheim, Teller and Sines, 2009). The percentage of women working in productive activities is low (Transitional Government of Ethiopia, 1993). For example, the Ethiopian DHS 2005 (Central Statistical Agency of Ethiopia and ORC Macro, 2005) shows female participation rate to be 32 percent by the time of the survey. Apart from its implication for achievability of the planned growth and transformation, this low maternal labor supply amid falling fertility rate is interesting given the forgoing theoretical discussion regarding maternal labor market participation effect of fertility.

The question of interest to the present study, therefore, is whether and the extent to which fertility influences labor market participation of Ethiopian households. Available research regarding maternal labor market participation effect of fertility is spatially polarized, concentrating in Latin American and South and Southeast Asian countries (Aassve and Arpino, 2007).

For Ethiopia, evidence is very limited. The first published work known to the present author is by Solomon and Kimmel (2009), which sought to examine the relationship between fertility and the labor supply of the mother, and the second is by Desta (2013), which, however, examined the effect of children on maternal hours of work for working mothers, instead. As such, the available research is very inadequate to inform policy and that there is no adequate context

specific evidence for the government of Ethiopia to support or not to support policy targeted at reducing fertility rate. Formulation and implementation of sound national population and development policy and programs requires context-relevant research evidence.

Therefore, the present study uses a household sample survey dataset from rural and urban married women with at least two live children to document the effect of the number of children on work participation of Ethiopian women. It seeks to bridge the gap by examining how effect of fertility on maternal work participation varies by the different lifecycles of the household and by rural-urban location.

2 Data and methods

2.1 Data sources

Quantitative data regarding demographic, employment and other socioeconomic characteristics over the last four months prior to the commencement of the survey were collected from a sample of 493 rural and urban married mothers with at least two children living with the household, in two different time periods. First, a sample of 254 households were interviewed in October 2010, and then with the view to increasing the earlier sample size, additional 239 sample households residing in the same place as the previous sample were interviewed in 2013. The urban households were randomly selected from four kebeles (the smallest unit in the administrative structure of the country) out of the total of nine kebeles of the Bahir Dar City, the Amhara Regional State capital. Likewise, the rural sample households were randomly selected from two kebeles drawn from two different rural districts.

While, as is evident from the abundance of published articles, research on the link between fertility and maternal labor supply so far has typically been dominated by a quantitative approach, the use of qualitative approach to supplement and rectify weaknesses of the former is conspicuously missing. There is now an increasing consensus among scholars that the use of qualitative data within a quantitative one offers important value-adding advantages. Some of the value-adding advantages include its ability in improving household survey design, interpreting counter-intuitive or surprising findings from household surveys, explaining the reasons behind observed outcomes, probing motivations underlying observed behavior, suggesting the direction of causality, assessing the validity of quantitative results, understanding conceptual categories, and facilitating analysis of locally meaningful categories of social differentiation (Davis and Baulch, 2010; Hulme, 2007; Kanbur and Shaffer, 2005; Lawson, Hulme and Muwonge, 2007; Shaffer, 2012). This article used qualitative observation and interview conducted with survey households, with the view to discussing results from the quantitative data analysis, and this was found especially helpful where results from the quantitative data alone were found to be inconsistent with the theoretical expectation or with most previous evidence. An observation was made where a member or members of a household were engaged in any type of work activity for the household.

In addition, government policy and program documents were used as data source. Specifically, the national population policy document of Ethiopia and other policy and program documents related to population and development such as the Sustainable Development and Poverty Reduction Program (SDPRP), Plan of Action for Sustained Development to End Poverty (PASDEP), Growth and Transformation Plan (GTP), and other relevant sectoral policies and programs were reviewed.

With the view to assessing if the widely acknowledged rural-urban structural difference also translates into rural-urban differentials in fertility and maternal economic outcomes, data were analyzed first using the full sample and then separately for the urban and the rural sub-samples. As analytical framework, the paper also categorized the women by their children's average age groups to capture lifecycle variations in the effects.

2.2 Measurements

The objective of the present study is to investigate the effect of the number of children on the mother's (participant's) work participation. The dependent variable is the mother's participation in economic activities and the independent variable of interest is the number of children. In addition, a number of other demographic and socio-economic variables were also used as control variables, including average age of children, sex and age of the household head, participant's age at first marriage, participant's years of schooling, contraceptive use (yes = 1; otherwise, 0), loan receipt (yes = 1; otherwise, 0), members in the household other than parents engaged in productive or non-productive work, mean hours of their work, and value of household assets in term of *Ethiopian Birr* (ETB). The choice of these variables is consistent with most previous studies on the similar topic.

2.3 Analytical strategies

The causal effect of fertility on the economic wellbeing of children is complicated by their endogeneity. Although there are a few studies which failed to find endogeneity (Orbeta, 2005), the fact that fertility is endogenous to maternal work participation is widely acknowledged in the economic demographic literature, in the presence of which the use of the ordinary least squares estimator biases the effect of the number of children.

While the econometric literature offers various approaches to account for endogeneity, one of these is the use of an instrumental variable. Using instrumental variable methods yields unbiased estimates even when fertility is or is not exogenous (Schultz, 2007). Different studies used different instrumental variables to generate exogenous variation in fertility. These include, for example, twin first birth (Chun and Oh, 2002; Kim, Engelhardt, Prskawetz, *et al*., 2009; Rosenzweig and Wolpin, 1980a), abortion legislation (Bloom, Canning, Fink, *et al*., 2009), contraceptive choice of couples (Kim and Aassve, 2006), sibling sex composition (Angrist and Evans, 1998; Cruces and Galiani, 2007), sibling sex composition and contraception unavailability (Aassve and Arpino, 2007) and sex of the first birth (Chun and Oh, 2002; Orbeta, 2005).

The present study uses two-step instrumental variable probit (ivprobit hereafter) method, which is one of the most common instrumental variable estimators (Wooldridge, 2009). The instrumental variable used consists of sex composition of the first two siblings born to a mother (same sex = 1; otherwise, 0). This instrument is chosen because sex composition of children is a random assignment and hence the sex of the siblings has no direct significant effect on maternal participation in economic activities while it impacts the number of children.

In this procedure, the first step equation uses ordinary least squares to predict the number of children as a function of the sex mix of the first and the second siblings, controlling for other covariates. Once the number of children is exogenously predicted in this way, the final equation which estimates the mother's work participation can be specified by inserting the predicted number of children as key independent variable of interest, also controlling for the same covariates in the first equation (refer to Appendix A for details).

3 Results

3.1 Characteristics of the study population

The tables below offer some descriptive statistics on the demographic and economic characteristics of sample households. **Table 1** and **Table 2** provide mean values and frequency respectively for sample households on selected demographic and economic variables across the rural-urban economies. As expected, **Table 1** shows that households in the urban sub-sample have higher average age at first marriage/child bearing and years of schooling compared to households in the rural sub-sample. **Table 1** also shows that household members including children for the urban sub-sample work for longer hours (perhaps due to urban children's older average ages) compared to their rural counter parts.

Table 1. Demographic and economic characteristics of sample households (means)

Variables	Mean values		
	Full sample	Urban sub-sample	Rural sub-sample
Age of household head (years)	45.8659	48.5623	41.9875
	(11.9563)	(12.3264)	(9.8497)
Number of children of the participant	4.8911	4.3952	5.5412
	(2.0945)	(2.1098)	(2.4102)
Average age of children of the participants (years)	13.7410	17.0014	11.0108
	(8.0197)	(8.9971)	(5.8945)
Age of the participant at first marriage (years)	15.7618	16.9961	14.5276
	(3.61805)	(3.85991)	(2.88352)
Age of the participant at bearing first child (years)	17.3415	18.7034	16.9856
	(3.0002)	(3.5142)	(2.4315)
Participant's years of schooling*	3.7981	5.9107	2.0098
	(4.8475)	(4.8910)	(2.7458)
Value of household assets (ETB)**	13904.0397	15152.3254	11001.7491
	(15047.4125)	(16049.1152)	(15012.3124)
Mean hours of work per day by household members (excluding parents)	3.02	4.00	2.1
	(0.21)	(0.34)	(0.21)
N	493	248	245

Note: Standard deviations are reported in parenthesis. Source: Survey data (2010 and 2013).
* if no formal education attended, years of schooling is recorded as 0.
** 1 USD=16.636 ETB on December 2010, and 19.1218 ETB on 31 December 2013.

The older average age and the fewer number of urban children is probably due to the relatively higher educational attainment of the urban women leading to the higher rates of contraceptive use (**Table 2**) and the mother's delayed age at first marriage/child bearing (**Table 1**).

Table 2 shows that, compared to the rural households, urban households have a higher proportion of female-headed households, a lower proportion of households who received loan, a lower rate of maternal work participation, a lower proportion of households with more than two children, and a higher proportion of households with members other than parents who work for the household. While the higher proportion of female-headed households and the lower rate of maternal work participation for urban relative to rural households are consistent with previous evidence, the lower rate of loan receipt by urban compared to rural households is unexpected since urban households are expected to have better access to the service given their proximity to credit facilities and the relatively capital-intensive nature of urban jobs.

Table 2. Demographic and economic characteristics of sample households (percentage)

Variables	Percent the event occurred		
	Full sample	Urban sub-sample	Rural sub-sample
Household head is female	19.7	29.3	10.1
Participant used contraceptives	51.5	74.4	28.6
Household received loan	43.5	36.2	51.5
Participant participated in productive work	47.7	38.8	56.5
Members other than parents participate in productive work	58.7	62.7	54.6
Members other than parents participate in non-productive work	65.2	67.3	63.1
First two siblings are the same sex	62.2	57.5	66.8
Households with more than two children	83.3	78.2	88.3
N	493	248	245

Source: Survey data (2010 and 2013).

3.2 Number of children and maternal employment status

Having described the characteristics of the study population, we now turn on to analyzing the effect of the number of children on the maternal productive work participation, using the two step instrumental variable estimator of the ivprobit model. Before that, however, we describe the maternal rate of work participation in relation to the number of children (**Table 3**).

Table 3. Number of children and percentage of participants who participated in productive work prior to the survey

No. of children	Full sample		Urban		Rural	
	N	% women working	N	% women working	N	% women working
2	82	34.3	53	31.2	29	42.6
3-4	148	46.8	86	55.1	62	37.9
5-6	149	42.1	63	21.8	86	55.1
7-8	76	60.7	32	36.2	44	77.5
9-10	30	58.2	14	27.3	16	86.2
≥11	8	100	---		8	100
Total	493	47.7	248	38.8	245	56.5

Source: Survey data (2010 and 2013).

Table 3 depicts that the number of children and the work participation rate of the mother during the specified period differ for the urban and the rural sub-samples. It shows that mothers who participated in productive work increase with increase in the number of children for the rural sub-sample, and, generally, for the full sample. The situation is less consistent for the urban sub-sample. However, the table clearly shows that the mother's work participation is relatively smaller for the urban sub-sample both on average and for each age group.

Table 4 and **Table 5** present results of the ivprobit regressions (and for the exogenous probit model for comparison). **Table 4** shows the amount of variance explained for maternal work participation given the number of children and control covariates. The p-value associated with the Wald χ^2 statistic (ivprobit) is significant at p = 0.022 for the full sample, suggesting that the model is well fit to the data overall, while this is not the case for the sub-samples. For the exogenous probit model, however, the p-values are significant for the sub-samples as well.

Table 4. Proportion of variance explained for maternal work by the number of children and other covariates

Model	Full sample		Urban sub-sample		Rural sub-sample	
	Wald χ^2 (LR χ^2 for OLS)	Prob > χ^2	Wald χ^2 (LR χ^2 for OLS)	Prob > χ^2	Wald χ^2 (LR χ^2 for OLS)	Prob > χ^2
Exogenous probit (%)	28.32	0.005	23.56	0.018	22.60	0.030
Ivprobit (%)	24.51	0.022	16.08	0.237	14.35	0.130
N	493		248		245	

Source: Survey data (2010 and 2013).

Table 5 shows parameter estimates for the ivprobit model (and exogenous probit model). It is worth noting, at this juncture, that this study does not intend to discuss coefficients from control variables (see Appendix B for coefficients from control covariates).

The table consists of three panels. Each panel compares results for the rural and urban locations. The first panel shows results for all households that differ only in their rural-urban location. The second and the third panels show results for households that differ by the age group of their children, in addition.

For the first panel, first, the ivprobit estimate has all positive coefficients for both the rural and urban locations, suggesting that an increase in the number of children is associated with an increase in the probability of the mother's work participation for the households, despite the lack of statistical significance for the ivprobit coefficients. While the lack of statistical significance for the ivprobit coefficients as opposed to those using the exogenous model is consistent also with several other previous research, the lack of difference in coefficient signs by rural-urban location is surprising because the difference in the employment structure between the rural and the urban economies is expected to respond to the effect of the number of children differently for the rural and the urban locations. However, this difference becomes fairly apparent once the lifecycle effect is considered by categorizing households according to age groups of their children (the last two panels of **Table 5**).

Table 5. Parameter estimates for maternal work participation by the number of children (with control variables)

Group of households	Model	Full sample		Urban sub-sample		Rural sub-sample	
		Coef.	p > z	Coef.	p > z	Coef.	p > z
All households	Exogenous probit	0.0918 (0.0321)	0.070	-0.2156 (0.0452)	0.061	0.1568 (0.0425)	0.004
	Ivprobit	0.1671 (0.2031)	0.418	0.0304 (0.1549)	0.315	0.8412 (0.1456)	0.113
	N	493		248		245	
Households with children of ages < 10 years	Exogenous probit	0.1903 (0.1102)	0.021	0.1843 (0.3201)	0.321	0.1497 (0.1222)	0.107
	Ivprobit	0.3349 (0.9742)	0.420	-0.0989 (1.7025)	0.498	0.7008 (0.5079)	0.092
	N	217		99		118	
Households with children of ages ≥ 10 years	Exogenous probit	0.9963 (0.1515)	0.088	-0.0845 (0.0852)	0.541	0.2111 (0.0981)	0.026
	Ivprobit	0.0711 (0.1845)	0.476	0.1252 (0.3602)	0.084	-0.9932 (6.961)	0.566
	N	276		149		127	

Note: Standard errors are reported in parenthesis. Source: Survey data (2010 and 2013).

The second panel of the table shows results for mothers with children of ages less than ten years. The third panel shows results for those mothers with children of ages ten years or older. For the urban sub-sample, the ivprobit coefficient is negative for the second panel, suggesting, as expected, that large number of young children decreases the probability of the mother's work participation, although it is not statistically significant. In the third panel, the ivprobit coefficient is positive and statistically significant at p = 0.084, suggesting that for urban mothers with more adult children, the negative effect of the number of children disappears and contributes positively. For the rural sub-sample, the ivprobit coefficient is positive and statistically significant for the second panel at p=0.092, suggesting that large number of young children increase the mother's probability of work participation. By contrast, for the third panel the ivprobit coefficient is negative, suggesting a

reversal in the positive effect of the number of children, although it is not statistically significant.

4 Discussion

Some key points emerge for discussion from the results section. The first is the relative importance of coefficients' magnitudes for estimates from exogenous probit and ivprobit models. That is, in some cases coefficients from exogenous estimates are larger than those from endogenous estimates, and in some other cases, the reverse is the case. The lack of consistency in coefficient size from the ivprobit estimator compared to the exogenous probit estimator in the present study, however, is in line with the available research evidence for several other countries. Most previous research shows larger coefficients from exogenous models exaggerating the effect compared to those from endogenous models (*e.g.*, Angrist and Evans, 1998 and references therein). However, there is also evidence documenting larger coefficients for estimates from endogenous rather than exogenous models. For example, Rosenzweig and Wolpin (1980a) note that instrumenting endogeneity increases the coefficients compared to the exogenous model. For Korea, Chun and Oh (2002) found larger coefficients using endogenous estimates compared to exogenous estimators when using households with at least one child, but smaller endogenous estimates when using households with at least two children.

Researchers (*e.g.*, Aassve and Arpino, 2007) attribute this inconsistency in exogenous and endogenous coefficients to the instrumental variable used. Whereas the use of sibling sex composition provides a natural experiment whereby households with same sex siblings are treatment groups and those with mixed sex siblings are control groups, a lack of consistency is expected since the two models estimate essentially different things owing to their reference to different samples (Aassve and Arpino, 2007). That is, the exogenous estimator coefficient represents the average effect of the number of children over the entire population in the sample, whereas the endogenous estimator coefficient represents the average effect of the number of children for those households whose first and second siblings have same sex. In such a case, results from the exogenous estimator may be due to variables other than the number of children such as biases from omitted variables, hence making causal inferences problematic.

The second point is regarding the heterogeneity in coefficient signs between the rural and the urban sub-samples. The negative ivprobit coefficient on the probability of maternal work participation effect of young children for the urban mother in the second panel (despite its being statistically insignificant) and the positive and statistically significant ivprobit coefficient in the third panel is consistent with most previous evidence, although most such research is based on rural-urban dummy instead of running separate analysis for rural and urban mothers (see *e.g.*, Angrist and Evans, 1998; Cáceres-Delpiano, 2008; Chun and Oh, 2002; Cruces and Galiani, 2007; Dupta and Dubey, 2003; Kim and Aassve, 2006; Orbeta, 2005; and references therein). This result is also consistent with previous lifecycle evidences. For example, Hotz and Miller (1988) found that children tended to have negative effects during their early ages but not during their adult ages, and that the intensity of time the mother spent tending her children markedly declined as children grew up. Similarly, Assaad and Zouari (2003) for urban Morocco found that the presence of school-age children significantly reduced the participation of women from all types of paid work.

The positive and statistically significant ivprobit coefficient for the rural sub-sample in the second panel, and the negative though statistically insignificant ivprobit coefficient in the third panel, nevertheless, are inconsistent with the theoretical prediction that holds that, other factors held constant, the mother's probability of work participant decreases with an increase in the number of young children and increases when children become more adult. However, consistent with this result, using data from the 2000 Ethiopian DHS and instrumenting the number of children with the husband's desire for children, Solomon and Kimmel (2009) found positive (but statistically insignificant) labor supply effect of children. In this connection, Angrist and Evans (1998:463) also cite a review that found that fertility either has no effect on maternal labor supply, or it has a positive effect when endogeneity is considered.

The question now is why is this so? In the present study, it is argued that, despite the lack of statistical significance for many of the ivprobit coefficients, the quantitative results' inconsistency with theory and most previous research for the rural households is rather due to the rural-urban difference in the employment structure and the effect of the household's lifecycle. Context-specific literature review and qualitative data seem to be revealing in this particular case.

First, the prevalence of household enterprises and traditional nature of farming in rural areas of poor economies including Ethiopia means that more rural women have to work longer compared to urban women (see *e.g.*, Arbache, Kolev and Filipiak, 2010). In Ethiopia, farm plots are fragmented, farming is done manually, and productivity is low. Households have to invest a lot of manual labor per unit area, and, as such, it would be likely for women to work in the farms especially when there are other children to look after very young children at home. In such circumstances, children may not be considered that much prohibitive to the mother's work given the nature of the economy and the mother's need to work for the family, despite the adverse health implications that this is likely to have on the young children.

There is ample evidence on this from maternal work and child care literature (*e.g.*, Samman, Presler-Marshall and Jones, 2016 and the references therein).

Second, work conditions are more flexible for rural economies than for urban economies (Kim and Aassve, 2006). Farms are not that much far from the house and the mother can flexibly use her time taking care of her child at home and working in the nearby farm plot. Even where farm plots are away from home, the mother can still manage to work. It is common, in Ethiopia for example, to see mothers doing the farming activities holding children on their backs or placing them in a tree shade beside the farm with another young child to look after the youngest child. This has also been well confirmed by information from qualitative observation and interview with some sample households.

A twenty-eight year old woman was holding her five-month daughter on her back when I met her cutting fodder from her maize farm. Her responses to questions I raised in the interview illustrate the argument above:

> My husband is sowing shimbra (a local name for chickpeas) in another farm. My daughter is sleeping now on my back, but she will begin crying eventually as she feels hot in the sun. Sometimes, she develops fever at night. I know this is bad for her, but I have to do this because there is no option. My younger sister, whom I brought to help with the household activities following my delivery is there working in the farm with my husband removing the weeds. I could have brought my other two children with me here to look after my daughter in the tree shade, but they are in the house watching for sito [some raw food stuffs such as grains, cereals, *etc*. put to dry in the sun before further processing] from bird pests.

Third, young children in many cases contribute to the family labor by taking care of the domestic chores. The domestic labor contribution of young children is also well documented in the literature (Aghajanian, 1979; Boserup, 1985; Caldwell and Caldwell, 1987; Cho, 2006). According to Cho (2006), children contribute to household economy by replacing the mother's activity at home. Boserup (1985) and Caldwell and Caldwell (1987) argue that in addition to their labor input, children in sub-Saharan Africa demand little child care, allowing the mother to spend her time on work. In the case of Ethiopia, Solomon and Kimmel (2009) also note that young children often contribute to domestic chores, which allows the mother to work away from home. They also note that Ethiopian mothers are unlikely to leave their jobs in the face of high unemployment and underemployment rates in the country. The quantitative data for the present study shows that members engaged in non-productive work for the household increased the work participation rate of the mother (see Appendix B). This result has also been well confirmed by qualitative observation and interview with children and parents. The qualitative data revealed that children were able to work for the household both as non-school children, pre-school children and school children. Some school-age children had their parents refused to send them to school, and so work for them; some children were too young for schooling, but can do some kind of work for their family (as in the young children's taking care of the sito in the house in the quotation above), and some others use their non-school hours for family work, such as the mornings, evenings, and the other half, non-school shift of the school day. In addition, many rural parents make their children be absent from school for several days, especially when there is a sign of untimely rain coming during periods of crop harvest.

Equally inconsistent with theory and most available evidence, despite the lack of statistical significance, is the more adult children's negative contribution to rural maternal work participation. In Ethiopia, it is often common to see older mothers working only a few hours a day or absenting themselves from farm work at all and staying home doing domestic chores compared to younger mothers. Results from the quantitative data show that, contrary to the positive maternal work participation effect of members engaged in non-productive work, members engaged in productive work decreased the mother's work participation. It can be argued that this is probably the result of work substitutability between the mother and other members of the household including children (especially of adult ones). This appears to have been also reinforced by other context-specific circumstances such as land tenure and the subsistence nature of farming coupled with poverty. In Ethiopia, land is owned by the state since 1974 and farmers have utilization rights of the land they have held. There were periodic land redistribution schemes during the Dergue government (1974–1991). The existing government had also redistributed land in 1995/96 and has certified the farmers as a security to their land utilization right. There has not been any redistribution carried out thereafter. As a result, male children were implicitly obliged to continue to work on parents' land, mainly as sharecroppers, even having been married and have own family. Their marriage also increases the household's labor force thus encouraging the mother to stay home doing the household chores. Even when married children might in some cases work on their own farm, or migrate to cities where they engage in non-farm activities, they might have to spend some days helping on their poor parents' farms, or hiring some daily labor for them. The rural to urban migration effect of restrictions in youth access to land is well documented in Bezu and Holden (2014). On the other hand, some better-off parents were able to hire daily labor by their own.

Evidence from the qualitative interview and observation is much more revealing. A forty-one year old woman having five live children out of six demonstrates this as follows:

> .... My eldest son has now passed 24 and has obtained two children. My second child is male and my third child is female. Both are married and have one child each. My fifth child [the fourth was dead] is male and is a grade 11 student living far away from us, and my youngest child is female and is a grade 8 student. .... Despite our land's being not big enough, we gave a small plot to our eldest son, but the kebele (a local equivalent for village administration) said he will not be given ownership certificate since no more land sub-division is allowed. As to our second son, we convinced him to farm with us and share the harvest. Our son and his wife and my husband as well took care of the farm work, and I stayed home preparing food for them and doing other activities. Meanwhile, our son left for ketema (a local equivalent for urban area) and now works there as gimbegna (a local equivalent for construction worker). Initially, his wife was still with us and was of great assistance working in the farms. Later on, however, he took her (along with her child) and I had to come back to farm work. This was difficult for me and my husband. We have grown these children, and are now getting old that we are not as strong to work in the farms as before. We also do not have money to pay for yeqen-serategna (a local equivalent for daily wage laborer) as our rich neighbors do. But, Egziabher yimesgen (a local equivalent for thanks to God), our eldest son agreed to work with us instead since the plot of land we gave him was too small, and so we merged the farm plots. Once again, my activity was confined to housework. Occasionally, especially during periods of peak farm activity, our son from ketema also sends us money to hire yeqen-serategna.

5 Conclusions

This paper used data from a household survey of rural and urban married women to test the theoretical prediction that labor market participation decreases for mothers with large number of young children and increases when children are adults.

Results show that, when all households are considered, children seem to have positive effects on the probability of the mother's work both for the full sample and also when the full sample is split into the rural and urban sub-samples. However, when household lifecycle differences are considered, coefficients are negative for households with large number of young children and positive for those households with more adult children for the urban sub-sample, roughly suggesting consistency with the lifecycle hypothesis and previous research evidence for other countries. Whereas for the rural sub-sample, coefficients signs are reversed, suggesting inconsistency with the theory.

While inconsistency of results among many previous studies has been acknowledged in the literature, the lack of statistical significance of ivprobit results in many of the cases in the present paper, suggests the difficulty of drawing a valid conclusion from the quantitative results alone in this particular case. At this point, at least two limitations of the present paper are noteworthy as potential causes for the weak statistical results. The first is potential problems in data quality including small sample size. The second is the attempt to capture lifecycle variations by categorizing women by their children's average age groups. Lifecycle effects are well studied using panel data which offer a unique opportunity for tracking the changes in the effect for a given household over a period of time. Cross-sectional data used in the present study do not have that capability and are thus inappropriate. The qualitative data and the reviews of context-relevant literature appear to have offered possible explanations for this inconsistent though, in many cases, not statistically significant result.

Conflict of Interest

No conflict of interest has been reported by the authors.

Ethics Approval

No ethics approval was required for this study.

Funding

The author has received financial support from the William and Flora Hewlett Foundation.

Acknowledgements

The author would like to thank the two reviewers for their helpful comments. The author is also grateful to the William and Flora Hewlett Foundation who funded part of the research.

References

Aassve A and Arpino B (2007). Estimation of Causal Effects of Fertility on Economic Wellbeing: Evidence from Rural Vietnam. ISER Working Paper No. 24. University of Essex and University of Florence.

Adda J, Dustmann C and Stevens K (2011). The career costs of children. IZA Discussion paper No. 6201. Institute for the Study of Labor.

Aghajanian A (1979). Family economy and economic contribution of children in Iran: An overview. *Journal of South Asian and Middle Eastern Studies*, 35(1): 21-30.

Angrist JD and Evans W (1998). Children and their parents' labor supply: evidence from exogenous variation in family size. *The American Economic Review*, 88(3): 450-477.

Arbache JS, Kolev A and Filipiak E (2010). Why study gender disparities in Africa's labor market? In JS Arbache, A Kolev and E Filipiak (eds.). Gender Disparities in Africa's Labor Market. (pp. 357-377) Washington D. C.: The World Bank. https://doi.org/10.1596/978-0-8213-8066-6.

Assaad R and Zouari S (2003). Estimating the Impact of Marriage and Fertility on the Female Labor Force Participation when Decisions are Interrelated: Evidence from Urban Morocco. Topics in Middle Eastern and North African Economies, Vol. 5. Middle East Economic Association and Loyola University Chicago.

Becker GS (1960). An economic analysis of fertility. In: Universities-National Bureau (eds.). Demographic and Economic Change in Developed Countries. Columbia: Columbia University Press.

Bezu S and Holden S (2014). Are rural youth in Ethiopia abandoning agriculture? *World Development*, 64: 259-272. http://dx.doi.org/10.1016/j. worlddev.2014.06.013.

Bloom D, Canning D, Fink G, *et al*. (2009). Fertility, female labor force participation, and the demographic dividend. *Journal of Economic Growth*, 14(2): 79-101. https://doi.org/10.1007/s10887-009-9039-9.

Boserup E (1985). Economic and demographic interrelationships in Sub-Saharan Africa. *Population and Development Review*, 11(3): 383-397.

Browning M and Crossley T (2001). The life-cycle model of consumption and saving. *Journal of Economic Perspectives*, 15(3): 3-22. https://doi.org/10.1257/jep.15.3.3.

Cáceres-Delpiano J (2008). Keeping the Best for Last: Impact of Fertility on Mother's Employment: Evidence from developing countries. Working paper 08-68 No. 32. Universidad Carlos De Madrid.

Caldwell J and Caldwell P (1987). The cultural context of high fertility in Sub Saharan Africa. *Population and Development Review*, 13(3): 409-435.

Central Statistical Agency of Ethiopia and ORC Macro (2000). The Ethiopian Demographic Health Survey. Addis Ababa.

Central Statistical Agency of Ethiopia and ORC Macro (2005). The Ethiopian Demographic Health Survey. Addis Ababa.

Central Statistical Agency of Ethiopia and ORC Macro (2011). Ethiopia Demographic and Health Survey 2011: Preliminary report. Addis Ababa.

Chernichovsky D (1978). Personal savings and family size and composition: The unresolved issue. The Proceedings of the Conference of the International Union for the Scientific Study of Population: Economic and demographic change: issues for the 1980s. (pp. Vol 1: 345-360). Helsinki.

Cho, Y (2006). An Analysis of Women's Fertility and Labor Supply: Implications for Family Policies. Paper presented at International Conference on Declining Fertility in East and Southeast Asian Countries, December 14-15, Tokyo.

Chun H and Oh J (2002). An instrumental variable estimate of the effect of fertility on the labor force participation of married women. *Applied Economics Letters*, 9: 631-634. https://doi.org/10.1080/1350485011011785 0.

Cruces G and Galiani S (2007). Fertility and female labor supply in Latin America: New causal evidence. *Labour Economics*, 14(3): 565-573. https://doi.org/10.1016/j.labeco.2005.10.006.

Davies J (1988). Family size, household production, and life cycle saving. *Annales D'economie Et De Statistique*, 9: 141-165. https://doi. org/10.2307/20075685.

Davis P and Baulch B (2011). Parallel realities: Exploring poverty dynamics using mixed methods in rural Bangladesh. *The Journal of Development Studies*, 47(1): 118-142. https://doi.org/10.1080/00220388.2010.492860.

Desta CG (2013). Fertility and maternal hours of work in Ethiopia: A case study in the Amhara region. *African Population Studies*, 27(2): 89-104. https://doi.org/10.11564/0-0-431.

Dupta ND and Dubey A (2003). Poverty and Fertility: An Instrumental Variables Analysis on Indian Micro Data. Working paper series 11. Department of Economics, Aarhus School of Business.

Erosa A, Fuster L and Restuccia D (2016). A quantitative theory of the gender gap in wages. *European Economic Review*, 85: 165-187. http://dx.doi.org/10.1016/j.euroecorev.2015.12.014.

Fehr H and Ujhelyiova (2012). Fertility, female labor supply, and family policy. *German Economic Review*, 14(2): 138-165.

Francesconi M (2002). A joint dynamic model of fertility and work of married women. *Journal of Labor Economics*, 20(2): 336-380.

Hailemariam A, Alayu S and Teller C (2011). The National Population Policy (NPP) of Ethiopia: Achievements, challenges and lessons learned, 1993-2010. In C Teller and A Hailemariam (eds.). The Demographic Transition and Development in Africa: The Unique Case of Ethiopia. London, New York: Springer.https://doi.org/10.1007/978-90-481-8918-2.

Hotz VJ and Miller RA (1988). An empirical analysis of life cycle fertility and female labor supply. *Econometrica*, 56(1): 19-118.

Hulme D (2007). Integrating Quantitative and Qualitative Research for Country Case Studies of Development. GPRG-WPS No. 063. Global Poverty Research Group.

IMF (2006). The Federal Democratic Republic of Ethiopia: Poverty Reduction Strategy Paper - 2003/04 Annual Progress Report. Country report No. 06/27. Washington D. C.

Kanbur R and Shaffer P (2006). Epistemology, normative theory and poverty analysis: Implications for Q-squared in practice. *World Development*, 35(2): 183-196. https://doi.org/10.1016/j.worlddev.2005.10.016.

Kelley A and Schmidt R (2001). Economic and demographic change: A synthesis of models, findings and perspectives. In N Birdsall, AC Kelley and S Sinding (eds.). Population Matters. New York: Oxford University Press. https://doi.org/10.1093/0199244073.001.0001.

Kim J and Aassve A (2006). Fertility and its Consequence on Family Labor Supply. IZA Discussion paper series 2162. Vienna Institute of Demography & IZA Bonn & University of Essex.

Kim J, Engelhardt H, Prskawetz A, *et al.* (2009). Does fertility decrease household consumption? An analysis of poverty dynamics and fertility in Indonesia. *Demographic Research*, 20(26): 623-656.https://doi.org/10.4054/DemRes.2009.20.26

Lawson D, Hulme D and Muwonge J (2007). Methodological Issues Associated with Combining Quantitative and Qualitative Approaches to Understanding Poverty Dynamics: Evidence from Uganda. GPRG-WPS No. 077. Global Poverty Research Group.

McNicoll J (1984). Consequences of rapid population growth: An overview and assessment. *Population and Development Review*, 10(2): 177-233.

Ministry of Finance and Economic Development (2002). Ethiopia: Sustainable Development and Poverty Reduction Program. Addis Ababa. http://planipolis.iiep.unesco.org/sites/planipolis/files/ressources/ethiopia_prsp.pdf.

Ministry of Finance and Economic Development (2006). Ethiopia: Building on Progress: A Plan for Accelerated and Sustained Development to End Poverty (2005/06-2009/10). Volume I. Addis Ababa. https://extranet.who.int/nutrition/gina/sites/default/files/ETH%202005%20Plan%20for%20Accelerated%20and%20Sustained%20Development%20to%20End%20Poverty%20%28PASDEP%29.pdf.

Ministry of Finance and Economic Development (2010). Growth and Transformation Plan (2010/11 - 2014/15). Volume I. Addis Ababa. http://et.one.un.org/content/dam/unct/ethiopia/docs/GTP%20English%20Vol1%20(1).pdf.

Moffitt R (1984). Profiles of fertility, labor supply and wages of married women: A complete life-cycle model. *The Review of Economic Studies*, 51(2): 263-278. https://doi.org/10.2307/2297691.

Orbeta A (2005). Children and the labor force participation and earnings of parents in the Philippines. *Philippine Journal of Development*, 32(1): 19-52.

Ringheim K, Teller C and Sines E (2009). Ethiopia at a crossroads: Demography, gender, and development. Population Reference Bureau Policy brief. Washington D. C.

Rosenzweig MR and Wolpin KI (1980a). Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica*, 48(1): 227-240.

Rosenzweig MR and Wolpin KI (1980b). Lifecycle labor supply and fertility: Causal inferences from household models. *Journal of Political Economy*, 88(2): 328-348.

Samman E, Presler-Marshall E and Jones N (2016). Women's work: Mothers, children and the global childcare crisis. ODI Report. London: Overseas Development Institute.

Schultz TP (2007). Fertility in Developing Countries. Discussion Paper 953. Economic Growth Center, Yale University.

Shaffer P (2012). Ten Years Of 'Q-Squared': Are Two Disciplines Better than One? Q-Squared Working Paper No. 57. Trent University.

Sheran M (2007). The career and family choices of women: A dynamic analysis of labor force participation, schooling, marriage, and fertility decisions. *Review of Economic Dynamics*, 10: 367-399.

Solomon B and Kimmel J (2009). Testing the Inverseness of Fertility and Labor Supply: The Case of Ethiopia. IZA Discussion paper No. 3949. Western Michigan University and IZA.

Transitional Government of Ethiopia (1993). The National Population Policy of Ethiopia. Office of the Prime Minister, Addis Ababa.

UN Population Division (2005). Population Challenges and Development Goals. Report No. ST/ESA/SER.A/248. Department of Economic and Social Affairs, New York.

UNDP (2001). UNDP Review of the Poverty Reduction Strategy Paper (PRSP). New York. http://www.lcgbangladesh.org/PovertyIssues/reports/UNDP%20PRSP%20Assistance.pdf.

UN Population Division (2017). *World Population Prospects: Key Findings and Advance Tables (The 2017 Revision)*. Report No. ESA/P/WP/248. Department of Economic and Social Affairs, New York.

UNFPA (2010). State of the World Population: Monitoring ICPD Goals.

Wooldridge JM (2009). Introductory Econometrics: A Modern Approach. Fourth Edition. Mason OH USA: South-Western SENGAGE Learning.

Appendix A

Model specification

First, the structural form for the probit regression model (Y_{1i}) can be given as:

$$Y_{1i} = \alpha_0 + x_{1i}\alpha_1 + Y_{2i}\beta + U_i \qquad (1)$$

where, Y_{1i} is the probability of mother's work participation for the ith household (1=If she works, 0=Otherwise)

α_1 is parameter coefficient of the vector of an exogenous variable, x_{1i}

β is parameter coefficient of the vector of the number of children, Y_{2i}

U_i is an error term assumed to be normally distributed with mean zero.

However, the literature tells us that the number of children (Y_{2i}) is endogenous. That is,

$$Cov(Y_{2i}, U_i) \neq 0 \qquad (2)$$

If the equation is estimated by OLS, the estimate will be biased. Therefore, Y_{2i} should be itself predicted first in a reduced form as a function of sibling sex composition, Z_i, and other covariates.

$$Y_{3i} = \delta_0 + x_{2i}\delta_2 + Z_i\gamma + e_i \qquad (3)$$

where, Y_{3i} is the predicted number of children for the ith household,

δ_2 is parameter coefficient of the vector of exogenous variables, x_{2i}, for the ith household,

γ is parameter coefficient of the vector of the instrumental variable, Z_i,

e_i is an error term associated to household i.

The instrumental variable, Z_i, is assumed to be uncorrelated with the error term, but partially correlated with the number of children. That is,

$$Cov(Z_i, U_i) = 0 \qquad (4) \text{ and,}$$

$$Cov(Z_i, Y_{2i}) \neq 0. \qquad (5)$$

The instrument is also assumed to be uncorrelated with other exogenous covariates. That is,

$$Cov(Z_i, X_{1i}) = 0. \qquad (6)$$

Because U_i is unobservable, $Cov(Z_i, U_i)$ is untestable, unlike $Cov(Z_i, Y_{2i})$ which can be readily tested using the data.

Once the number of children is exogenously predicted in (3), the final equation which estimates the mother's work participation can be specified by inserting the predicted number of children, Y_{3i}, in place of Y_{2i} as:

$$Y_{1i} = \alpha_0 + x_{3i}\alpha_3 + Y_{3i}\beta + \varepsilon_i \qquad (7)$$

where, Y_{1i} is defined as in (1) above.

α_3 is parameter coefficient of the vector of exogenous variables, x_{3i}

β is parameter coefficient of the estimated number of children, Y_{3i}

ε_i is an error term associated to household i.

The estimated maternal work participation, Y_{1i}, is now assumed to be unbiased.

Appendix B

Parameter estimates for maternal productive work participation by the number of children and control variables[1]

Variables	Exogenous probit						Ivprobit					
	Full sample		Urban sub-sample		Rural sub-sample		Full sample		Urban sub-sample		Rural sub-sample	
	Coef.	p>z	Coef.	p>z	Coef.	p>z	Coef.	p>z	Coef.	p>z	Coef.	p>z
Number of children	0.0918 (0.0321)	0.070	-0.2156 (0.0452)	0.061	0.1568 (0.0425)	0.004	0.1671 (0.2031)	0.418	0.0304 (0.1549)	0.315	0.8412 (0.1456)	0.113
Average age of children	0.0982 (0.0241)	0.113	0.0352 (0.0098)	0.421	0.0934 (0.0401)	0.101	0.0745 (0.0127)	0.211	0.0112 (0.0198)	0.107	-0.0785 (0.0345)	0.113
Sex of household head	0.1845 (0.1987)	0.451	0.5145 (0.1845)	0.054	0.1305 (0.3512)	0.625	0.1562 (0.2189)	0.408	0.3190 (0.3163)	0.301	0.1052 (0.4009)	0.651
Age of household head	-0.0107 (0.0074)	0.201	0.0212 (0.0151)	0.213	-0.826 (0.0321)	0.071	-0.0564 (0.0170)	0.215	-0.1151 (0.0338)	0.412	-0.0777 (0.0307)	0.137
Participant's age at first marriage	0.1342 (0.0361)	0.105	0.2221 (0.0212)	0.265	0.0997 (0.0121)	0.415	0.1121 (0.0415)	0.511	0.2241 (0.0501)	0.671	0.1057 (0.0512)	0.253
Years of schooling of the participant	0.0095 (0.0555)	0.524	0.0886 (0.0346)	0.111	-0.0213 (0.0358)	0.671	0.0652 (0.0398)	0.214	0.0757 (0.0322)	0.201	0.0152 (0.0333)	0.221
Contraceptive use (Yes=1, Otherwise=0)	0.1412 (0.1042)	0.346	0.4141 (0.2112)	0.208	0.1111 (0.5242)	0.741	0.1127 (0.2020)	0.581	0.4025 (0.4240)	0.289	0.0120 (0.4151)	0.888
Loan receipt (Yes=1, Otherwise=0)	0.1919 (0.1701)	0.230	0.7194 (0.1939)	0.031	0.1145 (0.1212)	0.366	0.2191 (0.0881)	0.444	0.4171 (0.2235)	0.111	0.1515 (0.2320)	0.424
Members other than parents engaged in non-productive work	0.3323 (0.1545)	0.012	0.6652 (0.2145)	0.051	0.2002 (0.2525)	0.216	0.4097 (0.1818)	0.143	0.6076 (0.3041)	0.068	0.2451 (0.4041)	0.019
Members other than parents engaged in productive work	-0.0989 (0.1801)	0.601	0.4909 (0.2554)	0.134	-0.5021 (0.2444)	0.129	-0.0666 (0.1965)	0.184	0.4098 (0.2828)	0.113	-0.5142 (0.2099)	0.101
Mean hours of daily work by household members (excluding parents)	-0.1452 (0.1745)	0.521	0.3541 (0.2513)	0.125	-0.0819 (0.2242)	0.210	-0.2535 (0.2004)	0.241	0.2514 (0.2156)	0.121	-0.6852 (0.2002)	0.127
Constant	0.0194 (0.2524)	0.699	-0.6523 (0.5124)	0.214	0.5262 (0.1426)	0.115	0.0098 (0.4251)	0.721	-0.8898 (0.9859)	0.235	0.1104 (9445)	0.546

[1] Covariates controlled. Because of the endogeneity of fertility to economic indicators, employing the ordinary least squares (OLS) estimator in which maternal labor market participation is regressed on the observed number of children becomes misleading. To acknowledge this problem, the two stage instrumental variable was used. In the first stage, the observed number of children were regressed on sex composition of the first two siblings borne to a woman (1=same sex; 0, Otherwise), plus other covariates controlled in the model. In the second stage, maternal labor supply was regressed on the predicted number of children (predicted in the first stage) as the key independent variable of interest, plus the same variables control in the first stage. The idea is that sibling sex mix (the instrumental variable) determines the number of children exogenously (*i.e.*, it has direct effect on the number of children, but no effect on maternal labor supply). For comparison purpose, both exogenous (exogenous probit) and endogenous models (ivprobit) were estimated. Standard errors are reported in parentheses.

International Journal of Population Studies

RESEARCH ARTICLE

Multilevel analysis of infant mortality and its risk factors in South Africa

Samuel Abera Zewdie[1*] and Vissého Adjiwanou[2]

[1] Ethiopian Development Research Institute, Addis Ababa, Ethiopia

[2] Centre for Actuarial Research, University of Cape Town

Abstract: The study analyzed infant mortality and its risk factors in South Africa. It aimed to examine infant mortality in the country by taking into account the hierarchical nature of the problem and investigate the with-in country variation in modeling. In addition to the usual individual level risk factors of infant mortality, living standard, mother's education, and income inequality were defined at municipal level, while HIV prevalence was fixed at province level. A multilevel logistic regression model was then fitted with Bayesian MCMC parameter estimation procedure using the 2011 South African census data. Most of the demographic and socioeconomic variables identified at individual level were found significant. More remarkably, the result indicated that communities with better living standard and women's education were associated with lower infant mortality rates, while higher income inequality and HIV prevalence in the communities were associated higher levels of infant mortality. The changes in infants' odds of death were estimated to be 26%, -21%, 13% and 8% respectively for HIV, women's education, income inequality and level of the living standard. In addition, unobservable municipal and province level random effects significantly affected the level of infant mortality rates.

Keywords: Infant mortality; multilevel; poverty; inequality

ARTICLE INFO
Received: June 21, 2017
Accepted: August 25, 2017
Published Online: September 3, 2017

***CORRESPONDING AUTHOR**
Samuel Abera Zewdie, Ethiopian Development Research Institute, Addis Ababa, Ethiopia; samuel_abera@yahoo.com

CITATION
Zewdie SA and Adjiwanou V (2017). Multilevel analysis of infant mortality and its risk factors in South Africa. *International Journal of Population Studies*, 3(2): 43-56.
doi: 10.18063/ijps.v3.i2.330

1. Introduction

Infant mortality rate is an important indicator of health and development. Biologically, infants have much weaker immune systems than adults and are therefore far more vulnerable to environmental or social complications (Caldwell, 1996). In addition, they are unable to care for themselves and are hence completely dependent on others. As a result, children are generally the group first and most strongly affected by poor living standards. Likewise, advances in health or social conditions are often first observed in improvements in infant or child mortality (Omran, 1971). Studies on infant mortality have accumulated a huge list of determinants or associates, including individual- and community-level factors such as maternal age, race, income, sanitation, water source, electricity, urban/rural residence, region of residence, household composition, occupation, female education, access to health care, and so forth (Caldwell, 1979; Hobcraft, McDonald and Rutstein, 1985; Kembo and Ginneken, 2009; Omariba, Beaujot and Rajulton, 2007; Victora, Wagstaff, Schellenberg *et al*., 2003; Wang, 2003).

Within-country variation in child mortality has been well documented, with rates often varying substantially across different regions and social groups (Moser, Leon and Gwatkin, 2005; Mosley and Chen, 1984). It is said that although global or national level results are vital for assisting policy makers to better prepare for the emerging health needs of populations, they constitute a less effective guide for refocusing health priorities because efforts to reduce health disparities would be more successful if they are based on evidences from lower administrative units (Heuveline, Guillot and Gwatkin, 2002). This research, therefore, studies infant mortality at national, provincial and municipality levels to highlight concentration at lower levels of geography that in turn underscores the ineffectiveness of national level indicators for monitoring progress in health achievement.

The objective of this research is to analyze infant mortality using hierarchical model and to identify important risk factors for infant mortality in South Africa. The study uses the 2011 South African census data and all the risk factors are defined at three levels: Individual, municipal and province level. A three-level logistic regression model is fitted using data on children born twelve months before the census date where child, municipality and province are the first, second, and third levels.

Factors affecting infant mortality are investigated by fitting multilevel logistic regression models in order to quantify the impacts of socioeconomic factors, including poverty and inequality, on infant mortality. The hypothesis is that there are significant spatial variations of child mortality, which are associated with socioeconomic differentials in the country, and hence multilevel modelling helps to measure the impacts of the risk factor at different administrative levels.

Mortality of children has been explained by different theories such as the social and economic explanation, the public health explanation, and the Mosley and Chen analytical framework. Mosley *et al*., (1984) identified a set of proximate determinants that directly have an impact on the morbidity and mortality of children. The factors are then grouped into 5 sets: Maternal factors (age, parity and birth interval), environmental contamination (air, food/water/fingers, skin/soil/inanimate objects, insect vectors), nutrient deficiency (calories, protein, micronutrient, vitamins and minerals), injury (accidental, intentional); the last set is personal illness control (personal preventive measures, medical treatment). It is within this framework that many studies on child mortality and its correlates have been carried out. This study also follows this framework but based on the following classification of the determinants: New-born demographics (Boco, 2010; Hill and Upchurch, 1995; Kembo and Ginneken, 2009; Mustafa and Odimegwu, 2008), Maternal factors (Boco, 2010; Hobcraft, McDonald *et al*., 1985; Kabir, Islam, Ahmed *et al*., 2001; Kembo and Ginneken, 2009; Omariba, Beaujot *et al*., 2007), socioeconomic factors (Bawah and Zuberi, 2005; Cleland, 1990; Hobcraft, 1993; Mustafa and Odimegwu, 2008; Sastry, 1996; Wagstaff, 2000), environmental factors (Bartlett, 2005; Kabir, Islam, Ahmed *et al*., 2001; Kazembe, Clarke and Kandala, 2012; Kembo and Ginneken, 2009) and HIV/AIDS (Dorrington, Johnson, Bradshaw *et al*., 2006; Ng'weshemi, Urassa, Usingo *et al*., 2003; Wang, Liddell, Coates *et al*., 2014; Zaba, Marston and Floyd, 2003). These factors are defined at individual level, which could be child, mother or household, municipal level or province level depending on the nature of the variable as well as the availability of data.

2. Data and Methods

2.1 Data

The study used data from the 10%-unit record of the 2011 de facto population and housing census of South Africa (StatsSA 2014). Children born within 12 months before the census date are considered for this research. The sample data, among other things, consisted of data on children's demographics,general health functioning, income, educational attainment, employment status, fertility, household characteristics and mortality variables. After removing missing, unknown and inconsistent cases, there are 86 877 (un-weighted) children with valid survival status for analysis. These children can be viewed as they are nested in a structure under the 234 municipalities and 9 provinces of South Africa. The mortality status of these children was the outcome measure of the study, and it was assumed that the rate of under-reporting of births in the past 12 months is the same as that of under-reporting of deaths in the past 12 months so that the effect on mortality estimates is negligible.

HIV prevalence rates were taken from the 2012 South African National HIV Prevalence, Incidence and Behaviour Survey conducted by the Human Science Research Council (HSRC, 2014). More details of the survey can be found elsewhere (HSRC, 2014)

2.2 Risk Factors Considered in the Study

The individual level risk factors considered are: sex, age and birth order of the child; age, racial group, marital status and employment status of the mother as well as the living standard of the household where the infant resides. Living standard was computed by constructing an index from different variables which are supposed to be related with the living

Table 1. Summary of variables used for LS index construction

Variable	Category (code)	Mean	SD	Factor loading	Coeff-icient
Dwelling Type	House (1), Other (0)	0.66	0.48	0.384	0.066
Room per person	Greater or equal to 1 (1), less than 1 (0)	0.69	0.46	0.257	0.030
Roof made of	Tiles (3), Concrete/Block (2) Other (1)	1.98	0.66	0.431	0.058
Wall made of	Brick (3), Concrete/Block (2) Other (1)	1.93	0.60	0.388	0.067
Energy used for lighting	Electricity (1), Other (0)	0.85	0.36	0.631	0.128
Energy used for cooking	Electricity/Gas (1), Other (0)	0.77	0.42	0.674	0.123
Piped water on premises	Available (1), Not available (0)	0.73	0.44	0.667	0.106
Flush Toilet	Available (1), Not available (0)	0.60	0.49	0.717	0.179
Television	Available (1), Not available (0)	0.76	0.43	0.595	0.096
Satellite Dish	Available (1), Not available (0)	0.26	0.44	0.554	0.092
Refrigerator	Available (1), Not available (0)	0.70	0.46	0.641	0.118
Washing Machine	Available (1), Not available (0)	0.32	0.47	0.645	0.120
Vacuum Cleaner	Available (1), Not available (0)	0.17	0.38	0.536	0.097
Computer	Available (1), Not available (0)	0.22	0.41	0.555	0.105
Internet access	Available (1), Not available (0)	0.36	0.48	0.436	0.057
Rubbish collected by local authority	Yes (1), No (0)	0.62	0.49	0.625	0.104

Source: Stats SA census 2011

standard (LS) of people. The LS index was constructed based on different indicators of wellbeing from the 2011 census data, specifically on those variables which measure how good the environment is for the infant to live in. Factor analysis (FA) was chosen for constructing the index (Hair, Black, Babin *et al*., 2010). FA mainly involves extracting the factor(s) by partitioning the total variance in each of the variables into variances which are shared and have unique variance. The detail theory and application of FA can be found in any standard multivariate text like Hair, Black, Babin *et al*., (2010). The descriptions of the variables used for constructing the index including some summary statistics of the variables are shown in **Table 1**. As a measure of internal consistency of the scale, Cronbach Alpha–a known measure of reliability – is computed giving a scale reliability coefficient of 0.8597.

The first factor was found to be enough to explain about 80% of the variance in the dataset and hence it was used to construct the index. The factor loadings and the coefficients of each variable used to generate the index are also given on **Table 1**. For ease of understanding, the constructed index was divided into 5 quintiles which can be used as ranking the level of living standard to households. A household lying in the first quintile was categorised as to have the poorest living standard, whereas a household lying in the fifth quintile was categorised to have the best living standard.

Community level variables used in this study were: Poverty and inequality levels of the municipality; education level of the municipality and HIV-prevalence rate of the provinces. The LS index was used to determine whether the municipality was poor or not, whereas Gini-index (GI) was used to determine the level of income inequality. GI is expected to be positively correlated with infant mortality as greater inequality in income within communities reflects unequal access to healthcare, nutrition and other services which is likely to reduce the health of the poor (Waldmann, 1992; Rodgers, 2002). GI is a number between 0 and 1, where 0 corresponds with perfect equality and 1 corresponds with perfect inequality. It is computed from a Lorenz curve (LC) which is literally a plot of the cumulative percentage of population versus the cumulative percentage of wealth/income.

All the independent variables were defined as categorical and hence, the odds ratio of death given in the last column of the tables measures the odds of the category compared to the reference group. Note that the independent variables are listed according to their level such that the variables on proportion of poor mean mothers' years of education and income

inequality as measured by Gini index, are identified at municipal level, while HIV prevalence rate is at province level. All these four variables were classified into two categories: lower and higher magnitude of the respective measures. The lower and higher values dictate that the respective quantity in the area is less than and greater than the national estimate. For instance, about 49% of the children live in municipalities where the level of income poverty is higher than the national poverty head count ratio of 41%. Note also that among the child level variables, age of the child is an indicator variable showing whether the child has age of less than one month (neonatal) or not.

2.3 Multilevel Models

Multilevel analysis is a suitable approach to take into account community level contexts at different levels, like at municipal and province levels, as well as individual subjects. A three-level random intercept logistic regression model was considered where the first level is children born 12 months before the census, whereas the municipalities and provinces in which the children live are the second and third levels respectively. Let π_{ijk}be the probability that child i living in municipality j and province k died before reaching age one. Then, the three-level random intercept logistic regression model in question with the predicator variables described above can, therefore, be expressed as

$$\ln[\pi_{ijk}/(1-\pi_{ijk})] = \beta_{0jk} + \sum_{l=1}^{16}\beta_l X_{lijk} \qquad \text{[Level 1]}$$

$$\beta_{0jk} = \beta_{00k} + \sum_{l=17}^{22}\beta_l X_{ljk} + u_{0jk} \qquad \text{[Level 2]}$$

$$\beta_{00k} = \beta_{000} + \sum_{l=23}^{24}\beta_l X_{lk} + v_{00k} \qquad \text{[Level 3]}$$

where $v_{00k} \sim N(0, \sigma^2_{v0})$, $u_{0jk} \sim N(0, \sigma^2_{u0})$, and the notations of the independent variables are as given in **Table 2**. The coefficients $\beta_1, \beta_2, \ldots, \beta_{24}$, called fixed effects, measure the impact of the corresponding predicator variable on the log of odds of death, whereas β_{0jk}, the random intercept, measures the combination of municipal and provincial level effects as defined in the second and third level of the model. Unlike ordinary logistic regression, there are two types of residual terms, u_{0jk} and v_{00k}, defined at level 2 and level 3 respectively and assumed to be normally distributed with mean zero and constant variance. Bayesian approach with Markov Chain Monte Carlo (MCMC) was implemented to the parameters of the above model. Further information regarding methods of parameter estimation is given in the appendix.

3. Results

3.1 Descriptive Statistics of Variables

The descriptive statistics of all individual, municipal and province level variables chosen for the analysis including the bivariate odds of infant death are shown in **Table 2**. It shows that some of the variables, such as race and education of the mother, living standard, birth order and HIV prevalence contribute to greater odds of death of the infant than others.

3.2 Multilevel Model Outputs

The final results of the regression are shown in **Table 3**. All parameter estimates were measured on the log-odds (logit) scale. In order to make more specific and meaningful inference about the effect of the risk factors on the infant mortality, the odds ratios (ORs) were given corresponding to each coefficient estimate in the same table. Note that among the independent variables, proportion of poor people, income inequality and mean years of mother's education were measured at municipality level, whereas HIV prevalence rate was computed at province level. All these four variables were dichotomised as higher and lower values of the respective quantities.

All coefficients of the living standard dummy variables are negative and their 95% confidence intervals exclude zero. Compared to infants who were in the first quintile of living standard, those who were in the second to fifth quintiles had 6%, 7%, 14 % and 24% lower odds to die, respectively. Likewise, the income poverty has a positive and significant coefficient, entailing that children living in a household whose members earned a per capita income of less than the South African poverty line were more likely to die than those who were above the poverty line.

Most of the municipal level indicator variables are significant, which implies that the level of poverty, women education and inequality of the municipality affected the survival status of infants. An infant was more likely to die in a highly poor and more unequal municipality compared to municipalities where the levels of poverty and inequality were lower after controlling for other risk factors. Considering the magnitude of the impact, it seems that the income inequality mattered more for infant mortality risk than the size of poverty in that more unequal municipalities were associated with 13% higher odds ratio of infant death than less unequal municipalities, whereas municipalities where poverty was high were

Table 2. Summary statistics of the variables in the regression model

Label	Variable		Mean	Std. Dev.	Odds Ratios of Dying
Y	Child died		0.0187	0.1353	0.0190
	Individual level				
X1	Sex of child		0.5066	0.5000	1.2108
X2	Age less than 1 month		0.1139	0.3177	0.8152
	Mother's age at birth	<20 yrs	0.1852	0.3885	1.0551
X3		20–34 years	0.1524	0.3594	0.8552
X4		>34 years	0.4000	0.4899	1.2212
	Birth order	1	0.2984	0.4575	0.8636
X5		2	0.1623	0.3687	0.8342
X6		3	0.1393	0.3463	1.0893
X7		4+	0.1312	0.3377	1.5342
	Mother's educ	No/primary educ	0.7722	0.4194	1.5116
X8		Secondary educ	0.0966	0.2954	0.9552
X9		Higher educ	0.5250	0.4994	0.5006
X10		Mother never married	0.2028	0.4021	1.1803
X11		Mother works	0.8592	0.3478	0.9151
X12		Mother is Black African	0.2317	0.4219	2.0703
	Living Standard Quintiles	Q1	0.2040	0.4030	1.4821
X13		Q2	0.1900	0.3923	1.1295
X14		Q3	0.1320	0.3385	1.0532
X15		Q4	0.3371	0.4727	0.7591
X16		Q5	0.3865	0.4869	0.4331
	Municipal level variables				
	Municipal poverty level	Low	0.3371	0.4727	1.2865
X17		Medium	0.3865	0.4869	1.0610
X18		High	0.2764	0.4472	0.6718
	Municipal Inequality	Low	0.3335	0.4715	0.8963
X19		Medium	0.3421	0.4744	1.0924
X20		High	0.3244	0.4682	1.0173
	Municipal women educ	Low	0.3362	0.4724	1.4668
X21		Medium	0.3507	0.4772	0.7895
X22		High	0.3131	0.4638	0.8325
	Province level variables				
	HIV prevalence rate	Low	0.3710	0.4831	0.7494
X23		Medium	0.3047	0.4603	0.9941
X24		High	0.3243	0.4681	1.3334

Note: Odds ratios were bivariate.

Table 3. Outputs from the multilevel logistic regression model

Variables		Coefficient	p	Odds Ratios
Fixed Effect parameters				
Cons		-2.2274	0.000	0.11
Sex of child (Male)		0.0790	0.000	1.08
Age of child (months)		-0.0774	0.010	0.93
Birth order	1 (Reference)			
	2	-0.0019	0.4760	1.00
	3	0.0736	0.0150	1.08
	4+	0.1440	0.0000	1.15
Mother's age at birth	20-34 years (Reference)			
	< 20 years	0.0297	0.125	1.03
	> 34 years	0.0225	0.261	1.02
Mother's years of education		-0.0708	0.0080	0.93
		-0.1922	0.0000	0.83
Mother never married		0.0518	0.0130	1.05
Mother works		0.0891	0.0010	1.09
Mother is Black African		0.1252	0.0020	1.13
Living Standard Quintile	Q1(Reference group)			
	Q2	-0.0572	0.0270	0.94
	Q3	-0.0716	0.0130	0.93
	Q4	-0.1471	0.0001	0.86
	Q5	-0.2808	0.0001	0.76
Municipal-level variables				
Level of poverty	Low Poverty (Reference group)			
	Medium	0.0466	0.1160	1.05
	High	0.0264	0.0450	1.03
Inequality	Low Inequality (Reference group)			
	Medium	0.0737	0.0210	1.08
	High	0.0530	0.0690	1.05
Proportion of educated women	Low Proportion (Reference group)			
	Medium	-0.1023	0.0010	0.90
	High	-0.1167	0.0200	0.89
Province level variable				
HIV prevalence rate	Low Prevalence (Reference group)			
	Medium	0.1352	0.0420	1.14
	High	0.1152	0.0420	1.12
Random effect parameters				
Province effect (level 3)		0.0048		
Municipality effect (level 2)		0.0085		

only associated with 3.5% higher odds of infant death than those municipalities where poverty was low. Similarly, infants living in municipalities where the average years of education of women was higher had a better likelihood of survival irrespective of the education level of their own mother. They were about 16% less likely to die than infants living in areas where there was less education of mothers.

Considering the effect of HIV on mortality of infants, it is apparent that its coefficient is positive and significant. Infants in provinces with high HIV prevalence were 12% more likely to die than infants in other provinces after controlling other factors in the model.

The sex effect is also positive and significant at $p<0.001$ level. More specifically, boys were 8% more likely to die during their first year of life than girls, all else being equal. Similarly, age of the child significantly affected the mortality of the child in that children who survived the first month after birth were less likely to die than those who were less than one month old by about 7%.

In terms of mother characteristics, the results in **Table 3** show that mothers' age at the birth of the child was not statistically significant. All the other independent variables associated with mothers were significant, which include years of education, employment status, birth order, marital status and population group.

In accordance with the literature, the more education a mother gets, the less likely the infant dies. Infant whose mothers completed secondary and higher education were 7% and 17% respectively less likely to die than those children whose mothers had no education or only had primary education. In addition, infants of higher birth orders, infants born to mother who were never married or single mothers, and infant born to mothers from African population group had a greater chance of dying than their counterparts of lower birth order, born to mothers who were not single and born to mothers from non-African population group. Specifically, an infant of the fourth or higher birth orders was associated with 15% higher odds of dying than first born infants, whereas an infant of the third birth order was associated with 8% of higher odds of dying. No difference was found in infants mortality between the second birth order and the first birth order. Similarly, children born to Black women or born to not-married women were associated with 13% and 5% higher odds of dying before age 1 in comparison with children born to non-Black or born to married women.

It is also worthy to see the distribution of the predicted log-odds of infant deaths in the country to see how the multilevel model performs in estimating infant mortality. The predicted log-odds of death of the municipalities and provinces are shown in **Figure 1** and **2** below. It can be observed that municipalities such as City of Cape Town, Liansburg, Mossel Bay and Nama Khoi had the lowest level of infant odds of death, while Disobotla, Mamusa, Msukaligwa and Ingwe scored

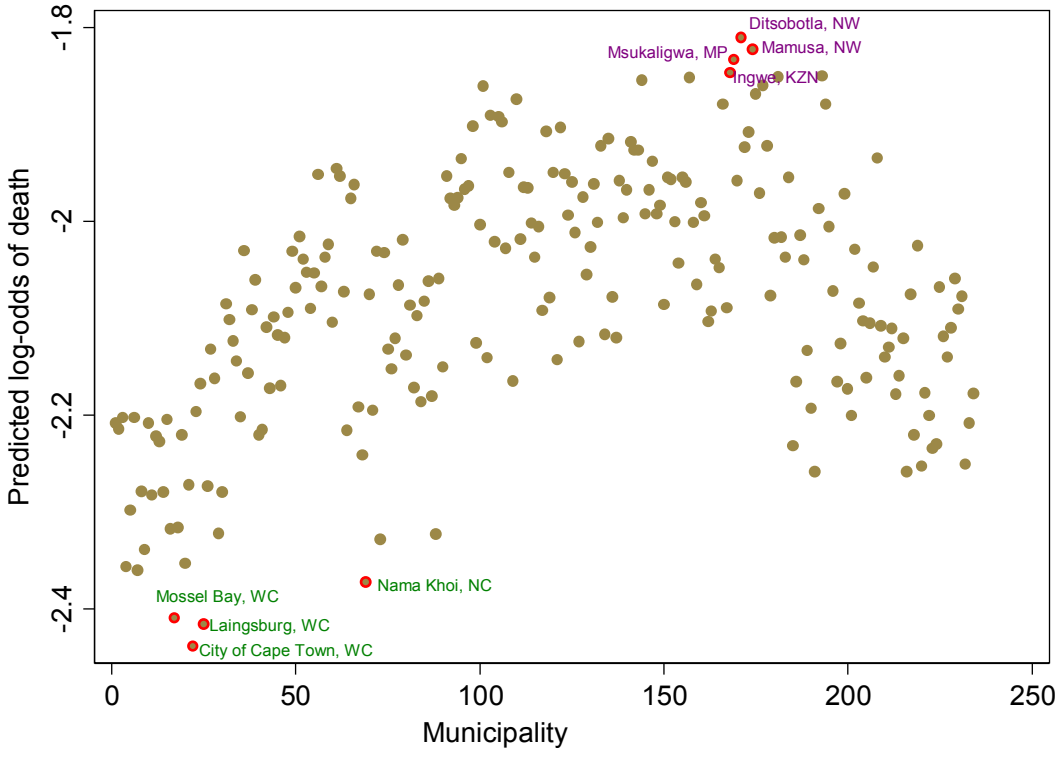


Figure 1. Predicated log-odds of infant death across the municipalities of South Africa

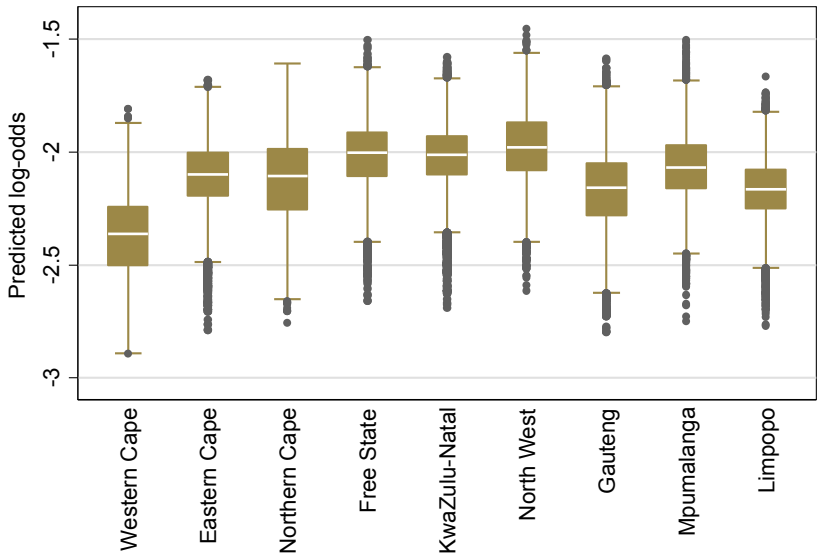


Figure 2. Predicated log-odds of infant death across the provinces South Africa

the highest mortality. Among the province, Western Cape showed the lowest odds of death followed by Gauteng and Limpopo.

The random effects terms included in the model were also significant. Hence, there was a unique effect for each province (level 3) and for each municipality (level 2) in addition to the fixed effects discussed above. The reported mean values in **Table 3** are the average random effect estimates of all provinces and all municipalities. The addition of the municipality specific effects as well as province specific effects made the model more accurate than the fixed only model. Specifically, controlling for the municipal and province level variations had on average increased the odds of death of infants by 1% and 0.5% respectively.

4. Discussion

In order to study the determinants or associates of infant mortality in South Africa, a multilevel logit model was fitted using data on the survival status of children born twelve months before the census and its risk factors which include several demographic, socioeconomic and environmental variables. The hierarchical nature of the data was taken into account in the process by considering provinces, municipalities and children as third, second and first levels respectively– hence, fitting a three-level logistic regression model employing the Bayesian MCMC procedure for estimating the parameters of the model. The results obtained from the regression model suggest that child mortality in South Africa is jointly determined by the observed individual demographic and socioeconomic characteristics of the child and the mother, and by municipal and province level covariates, as well as unobserved municipal and province level effects.

Among the demographic factors incorporated in the model, only age of mother at birth is not significant. Specifically, the odds of death for infants of four or more birth order is 15% higher than first born infants, whereas third born infants are 8% more likely to die. It shows that mortality risk is increasing with the addition of successive births, which implies that last-born (fourth and higher order) infants are at significantly greater risk of dying compared with earlier-born children. These findings are highly consistent with the Resource Depletion Hypothesis (Ashiabi and O'Neal, 2007), according to which, with the addition of successive children in the family, parental resources – both material and emotional – become diminished and consequently last-borns suffer the most. This is particularly true in the post-neonatal period since resource depletion is so heavily influenced by socioeconomic and environmental conditions that come into play in the later phases of infancy (Kembo and Ginneken, 2009). Similarly, male infants are 8% more likely to die than female, where as each month of life of infants contributes to 7% reduction on the odds of death of the infant. Genetic components result in newborn females being biologically advantaged when it comes to surviving their first birthday. Biologically, males, have lower chances of surviving infancy in comparison to female babies (Boko, 2010; Hill and Upchurch, 1995).

All the socioeconomic variables considered at level-one of the model are significant at $p<0.05$ levels and are in agreement with the results from other researches (Mosley and Chen, 1984; Hobcraft, 1993; Sastry, 1996; Kabir, Islam *et al*., 2001) as well as our child mortality estimation results presented above. For instance, children of black African mothers have a higher risk of death as compared to other population groups, while those who are from better educated mother have much lower risk of death. As expected, the living standard (LS) index, which is used as a measure of poverty, is highly significant too in that the higher the LS index of the household where the child lives the less likely the risk of dying. Note that because of multicollinearity issue the household income poverty indicator variable is not included in the model–the LS index and income poverty have a strong correlation coefficient of 0.85. Hence, in relation to poverty the result could be interpreted as, for example, children in the least poor and the second least poor household have more than 24% and 14% chance of dying respectively as compared to those living in the poorest households. Children living in municipalities where there is higher level of income poverty and inequality have greater likelihoods of death. Similarly, average years of schooling of women at municipal-level significantly affects child survival positively, whereas higher women HIV prevalence rate of provinces is highly related with higher risk of death of children as one expects. Poverty, inequality and women's education of the municipalities affected the odds of infant death by 8% 13% and -21% respectively, while the impact of province level HIV/AIDS is estimated to be 26. The results of the regression model also indicate small but statistically significant residuals which can convey province-level and municipal-level random effects on the risk of dying, even after controlling for a range of child-level, municipal-level, and province-level variables.

The study examined a comprehensive array of multilevel risk factors for infant mortality in South Africa by giving more attention to provincial and municipal variation. The modelling and estimation strategies utilized are appropriate and supported by large number of observation from the latest available census data. The use of multilevel analysis helped to understand the impact of some community level infant mortality risk factors of infant. Developmental indicators are likely to still vary significantly across the municipalities and also their effects on infant mortality. Hence, for further improvements, there is a need to focus on municipal level planning as well instead of only at the national or province level. Taking into account the findings under the present study, for a data involving hierarchical structure, there is a need to emphasize the use of the possible highest levels in hierarchical models. To further emphasize, such optimal considerations may provide additional important clues to policy planners leading to optimal use of available resources regarding public health programs.

Although the study achieved its objectives, there were some unavoidable limitations. First, our analyses were primarily based on the 2011 South African census data. Therefore, the significance and reliability of the results depends on the quality of the census data which includes the quality of enumeration and data processing. Any defect in the census data might seriously impact the results in the research. Second, the cross-sectional nature of this study design cannot us determine the causal relationship between independent variables and the outcome variable. The use of longitudinal data would have been much better. Third, the assumption that under reports of birth and deaths in the census data are the same. Fourth, the unavailability of municipal level HIV prevalence rates, and hence the assumption that these prevalence rates are the same as the rates at the respective provinces. Ignoring the HIV prevalence variation within provinces might especially impact the results of the regression model to some extent.

5. Conclusions

The main objective of the research was to investigate infant mortality risk factors with special emphasis on the impact of poverty and inequality. The results from the multilevel logistic regression model suggest that most of the demographic and socioeconomic factors as well as the province and municipal level random effects are significant. The significant predictors at individual-level include birth order and sex of the child, education, employment status, race, and marital status of the mother, and living standard of the family. These factors can bring from -24% to 15% change on the odds of death of infants. The changes in odds of infant death due to municipal and province level effects are estimated to be 12%, -11%, 7%, and 3% respectively for the HIV prevalence women's education, inequality and LS poverty. This implies that communities with better living standard and women education are associated with lower infant mortality rates, while higher income inequality and HIV prevalence in the communities depict higher levels of infant mortality. In addition, unobservable municipal and province level random effects significantly affect the level of infant mortality rates. The study helps to analyze infant mortality in the country by taking the hierarchical nature of the theme into account and to investigate the with-in country variation of infant mortality.

Authors' Contribution

SA Zewdie designed the study, prepared the data, performed the analysis, drafted and revised the manuscript, and

interpreted the results. V Adjiwanou supervised the study design and analysis, revised the manuscript and interpreted the results.

Conflict of Interest

This paper is part of an MPhil degree thesis at the university of Cape Town in 2014. The authors declare that there is no conflict of interest.

Ethics

Consent was obtained from all participating persons in the study.

Appendix

Methods of Parameter Estimation

There are two commonly used estimation methods for multilevel logistic regression models: quasi-likelihood (QL) approach and Bayesian approach with Markov Chain Monte Carlo (MCMC) methods (Goldstein, 2011). In QL approach, the non-linear logistic regression equation is estimated first using a Taylor series expansion which approximates a nonlinear function by an infinite series of terms (Breslow *et al*., 1993). If the Taylor series is expanded about the fixed and the random parameters, then the estimation is known as penalised quasi-likelihood (PQL) (Breslow *et al*.). Once the quasi-likelihood has been formed, unbiased estimates of the random parameters can be found by applying either iterative generalised least squares (IGLS) or restricted generalised least square (RGLS) which are estimation procedures in the case of continuous response variables (Goldstein, 2011). On the other hand, the Bayesian approach using MCMC estimation methods can be used by first specifying starting values prior distributions for each of the model parameters and then sequentially sampling subsets of parameters from their conditional posterior distributions using Markov chain. A discussion and technical details of MCMC estimation methods for multilevel models can be found from in Browne (2003) and Goldstein (2011)The MCMC procedure followed by MLwiN–software dedicated for multilevel modelling and used by this research–by default assigns flat prior distributions to the parameters of the model. That is, for fixed terms $p(\beta) \propto 1$ and for random terms, $p(1/\sigma^2) \sim Gamma\ \ (\varepsilon\varepsilon, \varepsilon\varepsilon)$ where $\varepsilon\varepsilon$ is a very small number. After assigning initial values, usually estimates from QL methods, the MCMC procedure in MLwiN then performs the simulations in two phases. In the initial burn-in period it runs until the chain converges to its stationary distribution; and in the next stage (monitoring period) it runs so that the means and standard errors of the parameters are estimated. The 2.5th and 97.5th quantiles of the chains provide Bayesian 95% credible intervals in order to make inferences concerning the estimated parameters, serving the same purpose as 95% confidence intervals. For fitting the aforementioned model, the number of iterations run is 1000 in the burn-in period and 10 000 for the monitoring period.

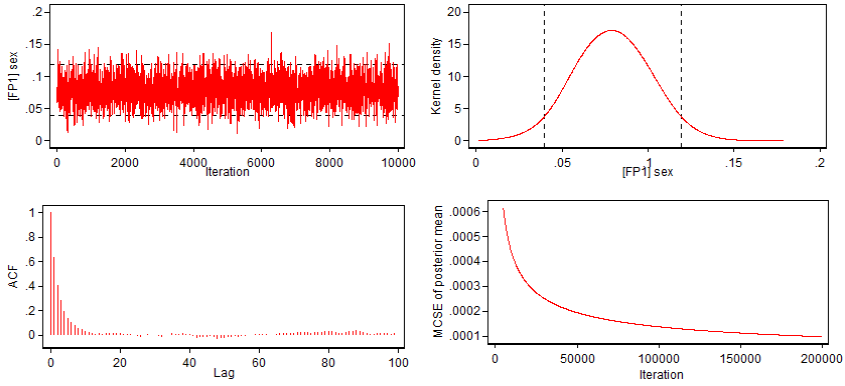
After running the model, residuals at municipal and province level (estimates of u_{0jk} and v_{00k}) are calculated so that the underlying assumptions, such as normality and constant variance of residuals, be investigated with the help diagnostic plots. Furthermore, as part of model diagnostics, the trace of the chains, autocorrelations (AC) and partial autocorrelations (PAC) functions at iteration t and t-k having accounted for iterations $t-1,\ldots,t-(k-1)$, and Monte Carlo standard errors (MCSE) are investigated for each of the posterior distributions of the parameter in the model. For the model to be good it is expected that the traces be not skewed, the AC and PAC functions be less correlated and the MCSE be close to zero. Increasing the number of iterations produces better results in all these dimensions. A comprehensive detail of parameter estimation and model diagnostics using MCMC simulations methods can be found from MLwiN manual (Rasbash, Charlton, Browne et al., 2012).

Model Diagnostics

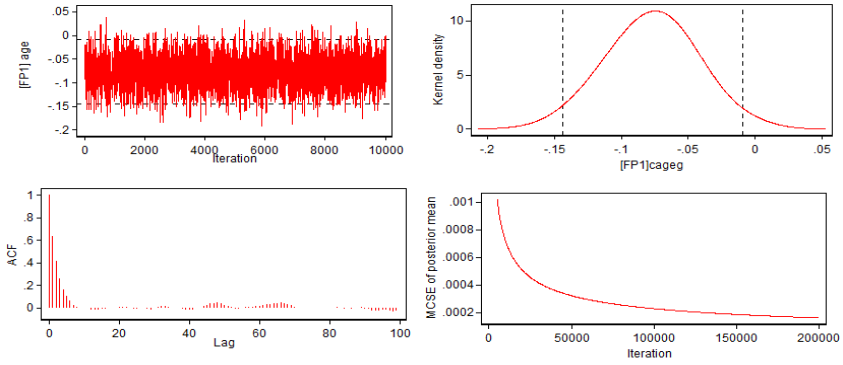
A three-level logistic regression model is fitted on the survival status of children born twelve months before the census. The parameters of the model are estimated using the Bayesian MCMC procedure by running the simulation for 1000 burn-in and 10000 monitoring period. After fitting the model, the reasonableness of the parameter estimates are assessed by looking at some diagnostics plots including the autocorrelation plots of successive iterations of the chains and Monte Carlo standard error plots for checking convergence of the posterior distributions. These are done for each of the fixed and random terms in the model. Some of these plots are given in the annex from **Figure (1)** to **Figure (f)**. The assumptions of normality of the residual terms at municipality and province level are approximately maintained.

MCMC Diagnostic Plots

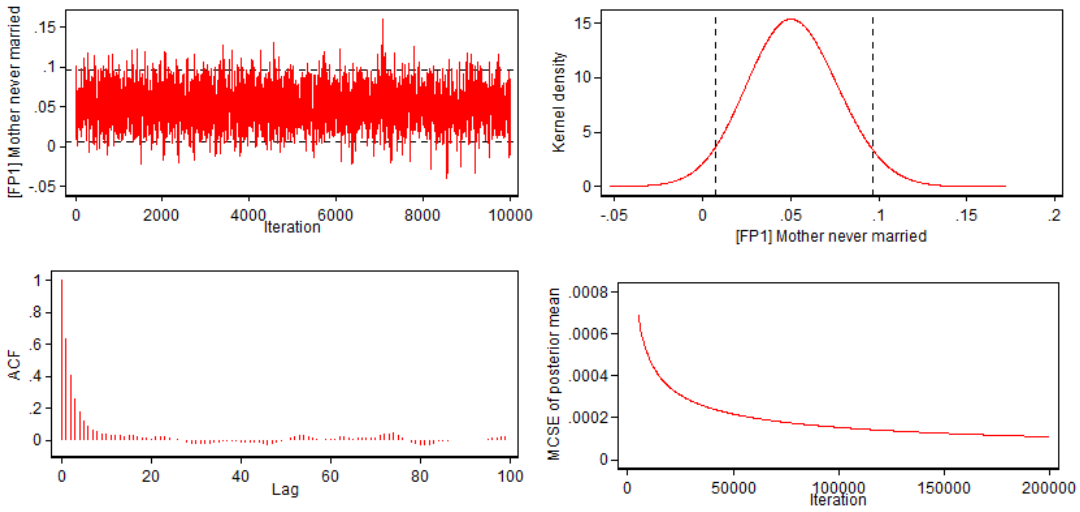
a)Fixed effect - Child sex



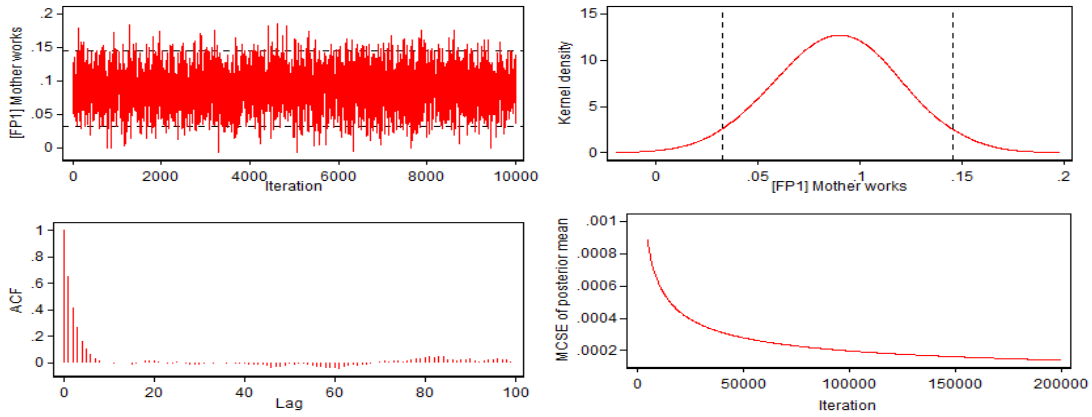
b)Fixed effect - Child age



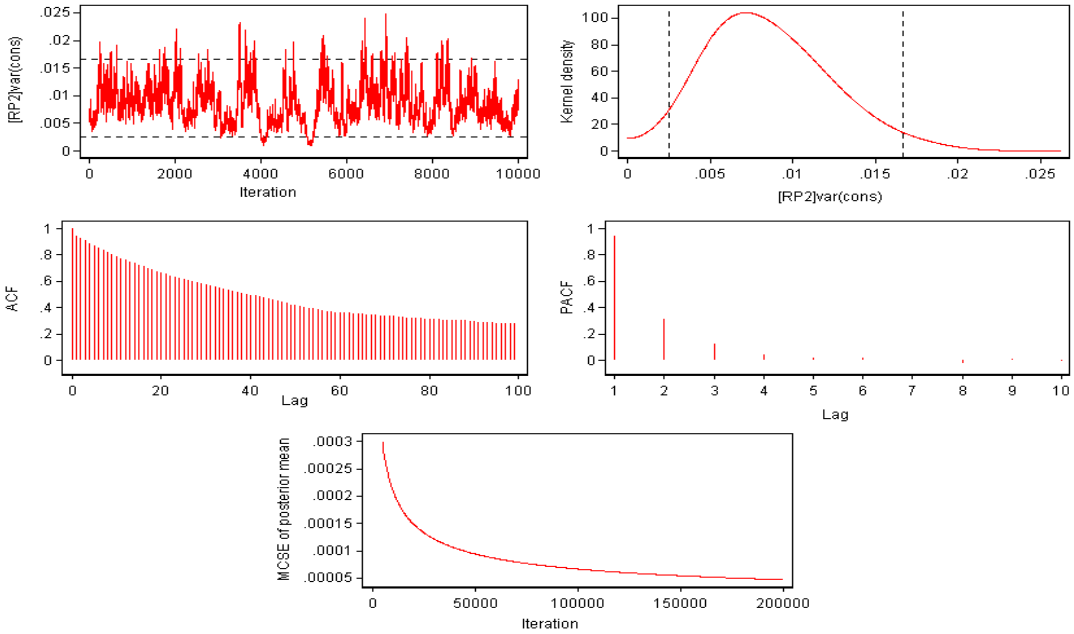
c) Fixed effect - Mother never married



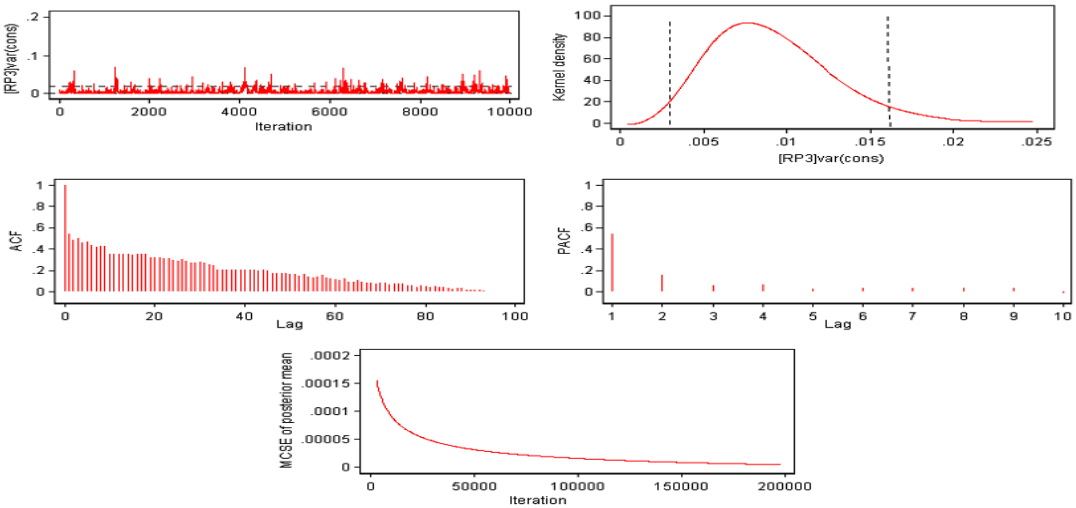
d) Fixed effect - Mother works



e) Level 2 random effect



f) Level 3 random effect



References

Ashiabi GS, O'Neal KK (2007). Children's health status: Examining the associations among income poverty, material hardship, and parental factors. *PLoS ONE*, e940; 2: 1–9. http://doi.org/10.1371/journal.pone.0000940

Bartlett S (2005). Water, sanitation and urban children: The need to go beyond "improved" provision. *Children, Youth and Environments*, 15(1): 115–137. http://doi.org/10.7721/chilyoutenvi.15.1.0115.

Bawah AA and Zuberi T (2005). Socioeconomic status and child survival in southern Africa. *Genus*, 61(2): 55–83. http://doi.org/29788852.

Boco AG (2010). Individual and community level effects on child mortality: An analysis of 28 demographic and health surveys in sub-saharan Africa. Calverton, Maryland, USA, ICF Macro.

Breslow NE and Clayton DG (1993). Approximate inference in generalised linear mixed models. *Journal of the American Statistical Association*, 88(421): 9–25.http://doi.org/ 10.1080/01621459.1993.10594284.

Browne WJ (2003). MCMC estimation in MLwiN. Avaliable from: http://www.bristol.ac.uk/cmm/media/migrated/2-31/mcmc-print.pdf.

Caldwell JC (1979). Education as a factor in mortality decline an examination of Nigerian data. *Population Studies*, 33(3): 395–413. http://doi.org/stable/2173888.

Caldwell P (1996). Child survival: Physical vulnerability and resilience in adversity in the European past and the contemporary Third World. *Social Science and Medicine*, 43(5): 609. https://doi.org/10.1016/0277-9536(96)00109-8.

Cleland JC (1990). Maternal education and child survival: Further evidence and explanations. What we know about health transition: The cultural, social and behavioural determinants of health. Caldwell JC *et al*. (editorss), Health Transition Center, The Australian National University.

Dorrington R, Johnson L, Bradshaw D, *et al*. (2006). The demographic impact of HIV/AIDS in south Africa. National and Provincial Indicators for 2006. Cape Town, Centre for Actuarial Research, South African Medical Research Council, Actuarial Society of South Africa.

Goldstein H (2011). Multilevel Statistical Models, John Wiley & Sons Ltd.

Hair J, Black C, Babin J, *et al*. (2010). Multivariate data analysis, Pearson Prentice Hall.

Heuveline P, Guillot M, Gwatkin DR (2002). The uneven tides of the health transition. *Social Science and Medicine*, 55(2): 313–322. https://doi.org/10.1016/S0277-9536(01)00172-1.

Hill K, Upchurch DM (1995). Gender differences in child health: Evidence from the demographic and health surveys. *Population and Development Review*, 21(1): 127–151. https://doi.org/10.2307/2137416.

Hobcraft J (1993). Women's education, child welfare and child survival: A review of the evidence. *Health Transition Review*, 3(2): 159–175.

Hobcraft JN, McDonald JW and Rutstein SO (1985). Demographic determinants of infant and child mortality: A comaparative analysis. *Population Studies*, 39(3): 363–385. Avaliable from: http://www.jstor.org/stable/2174100.

HSRC (2014). South African national HIV prevalence, incidence and behaviour survey. Human Science Research Coucil.

Kabir A, Islam MS, Ahmed MS, *et al*. (2001). Factors influencing infant and child mortality in Bangladesh: Research paper. *The Sciences*, 1(5): 292–295. Avaliable from: http://docsdrive.com/pdfs/ansinet/jms/2001/292-295.pdf.

Kazembe L, Clarke A and Kandala NB(2012). Childhood mortality in sub-Saharan Africa: cross-sectional insight into small-scale geographical inequalities from Census data. *BMJ Open*, 2(5): e001421. https://doi.org/10.1136/bmjopen-2012-001421.

Kembo J and Ginneken JKV (2009). Determinants of infant and child mortality in Zimbabwe: Results of multivariate hazard analysis. *Demographic Research*, 21(3): 1255–1267. https://doi.org/10.4054/DemRes.2009.21.13.

Moser KA, Leon DA and Gwatkin DR (2005). How does progress towards the child mortality millennium development goal affect inequalities between the poorest and least poor? Analysis of Demographic and Health Survey data. *BMJ*, 331(7526): 1180–1182. https://doi.org/10.1136/bmj.38659.588125.79.

Mosley WH and Chen L (1984). An analytical framework for the study of child survival in developing countries. *Population and Development Review*, 10: 25–45.

Mustafa EH and Odimegwu C (2008). Socioeconomic determinants of infant mortality in Kenya: Analysis of Kenya DHS 2003.

Journal of Humanities and Social Sciences, 2(2): 1–16. Avaliable from: https://www.scientificjournals.org/journals2008/articles/1409.pdf.

Ng'weshemi J, Urassa M, Usingo R, et al. (2003). HIV impact on mother and child mortality in rural Tanzania. *Journal of Acquired Immune Deficiency Syndromes*, 33: 393–404. https://doi.org/10.1097/00126334-200307010-00015.

Omariba DWR, Beaujot R and Rajulton F (2007). Determinants of infant and child mortality in Kenya: An analysis controlling for frailty effects. *Population Research and Policy Review*, 26(3): 299–321. https://doi.org/10.1097/00126334-200307010-00015.

Omran AR (1971). The epidemiologic transition: A theory of the epidemiology of population change. *Milbank Memorial Fund Quarterly*, 49(4): 509–538. http:// doi.org/10.2307/3349375.

Rasbash J, Charlton C, Browne WJ, *et al*. (2012). MLwiN Version 2.26, Centre for Multilevel Modelling, University of Bristol.

Rodgers G B (2002). Income and inequality as determinants of mortality: An international cross-section analysis. *International Journal of Epidemiology*, 31: 533–538.

Sastry N (1996). Community characteristics, individual and household attributes, and child survival in Brazil. *Demography*, 33(2): 211–229. http://doi.org/10.2307/2061873.

Victora CG, Wagstaff A, Schellenberg JA, *et al*. (2003). Applying an equity lens to child health and mortality: More of the same is not enough. *The Lancet*, 362: 233–241. http://doi.org/10.1016/S0140-6736(03)13917-7.

Wagstaff A (2000). Socio-economic inequalities in child mortality: Comparisons among nine developing countries. Bulletin of World Health Organization, WHO. 78.

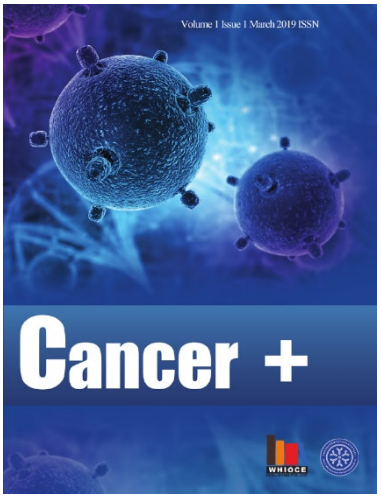
Waldmann RJ (1992). Income distribution and infant mortality. *The Quarterly Journal of Economics*, 107(4): 1283–1302. http://doi.org/10.2307/2118389.

Wang H, Liddell CA, Coates MM, *et al*. (2014). Global, regional, and national levels of neonatal, infant, and under-5 mortality during 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, 384(9947): 957–979. http://doi.org/10.1016/S0140-6736(14)60497-9.

Wang L (2003). Determinants of child mortality in LDCs. Empirical findings from demographic and health surveys. *Health Policy*, 65(3): 227–299. http://doi.org/10.1016/S0168-8510(03)00039-3.

Zaba B, Marston M and Floyd S (2003). The effect of HIV on child mortality trends in sub-Saharan Africa. Training Workshop on HIV/AIDS and Adult Mortality in Developing Countries, New York, Population Division Department of Economic and Social Affairs United Nations Secretariat.

OUR JOURNALS



Cancer Plus is a research journal that aims to provide a peer-reviewed, open access platform for researchers and practitioners to present novel and high-impact articles in the field of cancer research and clinical trials. This journal is a collaboration between the Shanxi Academy of Medical Sciences, Shanxi Bethune Hospital (China), and the Chinese Academy of Sciences. Areas of interest for this journal include, but are not limited to, the following topics:

- Mechanisms of tumorigenesis
- Genetic alterations of cancer
- Epigenetics of cancer
- Cell signalling
- Cancer stem cells
- Cancer prevention
- Diagnosis
- Cancer drugs/chemotherapy/therapeutic methods
- Clinical trials

The Journal of Clinical and Translational Research (JCTR) is an open-access, peer-reviewed, multidisciplinary scientific journal that publishes studies with at least an ex vivo, in vivo, or clinical component. The published research centers on any clearly defined clinical problem, which may comprise a disease or the basis of disease, a form of therapy or intervention, and clinical diagnostics or prognostics. Articles (original research, reviews, technical reports, medical hypotheses, commissioned articles, special issue articles, and editorials) are published bi-monthly.

Studies performed in cells only will generally not be accepted unless they contain critical data that are in line with the scope of the journal. Some examples of such studies include molecular pathways that lie at the basis of a disease, novel biotechnological approaches, e.g., the production of drugs, or new techniques that improve clinical diagnostics and prognostics. Articles that combine preclinical and clinical data are given priority. Contributions from academic institutions and industry are welcome.

Start a new journal

Write to us via email if you are interested to start a new journal with Whioce Publishing. Please attach your CV, professional profile page and a brief pitch proposal in your email. We shall inform you of our decision whether we are interested to collaborate in starting a new journal.

Contact: info@whioce.com

Whioce Publishing Pte. Ltd.

Publishing Office:

7030 Ang Mo Kio Avenue 5, #04-15 Northstar@AMK, Singapore 569880
Editorial Office: editor@whioce.com
General inquires: info@whioce.com
Whioce Publishing Pte. Ltd. Is a company registered in Singapore (No. 20147293E)