UNIVERSITY OF CALIFORNIA
RIVERSIDE

# Essays on Trade, Infrastructure, and Human Capital Outcomes in Developing Countries

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Economics

by

Pallavi Panda

June 2015

Dissertation Committee:
Dr. Anil B. Deolalikar, Chairperson
Dr. Aman Ullah
Dr. Mindy Marks

The Dissertation of Pallavi Panda is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

I would like to thank my advisor, Prof. Anil B. Deolalikar, for giving me the time and support to complete this dissertation. He has supported me in my ideas and provided critiques to help me become better and have a stimulating and productive PhD experience. He has been a mentor and guide and steered me through the tough times of this PhD process and I am very grateful for that. I would also like to thank Prof. Mindy Marks and Prof. Aman Ullah, who have been on my committee and have provided continued guidance. Prof. Mindy Marks has been instrumental in hearing out my ideas, giving extremely useful feedback, and providing ample time and energy in honing me as a researcher. I am thankful to her for teaching me perseverance and shaping an individual who can think critically as well as provide valuable inputs. Being a successful female economist in this department, she has been a role model for me since the very beginning.

I also acknowledge the continuous support of the faculty, friends, and administrative staff in the Economics Department at UCR in supporting me emotionally, financially, and intellectually in completing this dissertation. Teaching opportunities at the department has helped me develop a holistic view of economics and prepared me as an academician. I sincerely thank Prof. Aman Ullah, the chair of Department of Economics, for always looking out for me. I am grateful to the administrative staff, especially Gary Kuzas, for being always ready to help and being most patient with me during some of the hard times in the PhD process. I am indebted to the graduate division at UCR for providing numerous fellowship and travel grants to cover my tuition and other expenses for me to

be able to be productive in my research, without a need to divert my attention to financial considerations.

I gratefully acknowledge the funding from Hewlett Foundation/IIE and awarding the 2013-2015 Hewlett Foundation Dissertation Fellowship in Population, Reproductive Health and Economic Development for the research on effects of trade on child health outcomes in sub-Saharan Africa. Apart from providing me the funding, this fellowship introduced me to a network of economists with similar interests and research focus. The presentation of my paper at the annual conference organized by the foundation also helped me get feedback from various renowned economists in the field, which was extremely helpful. I also thank Paul Novosad and Sam Asher for providing me the administrative data and useful discussions on the paper on road construction and health outcomes.

At the inception of this thesis, Prof. Jorge M. Agüero, who was a part of my oral defense committee, provided invaluable advice and helped me shape the research ideas. He pushed me and challenged me to come up with research ideas that I am excited about every morning of every day. I have imbibed inquisitiveness and intellectual drive from him, which laid the foundations of a successful researcher. I also thank him for providing suggestions and comments on the chapter of trade and child health to help it develop into a better paper. I am grateful to Prof. Joseph Cummins who provided his time and inputs to improve this thesis. I also give thanks to Prof. Catalina Herrera and Prof. Urmee Khan for reading my paper and giving useful comments. I would also like to thank Prof. Mary

Gauvain, a member of my oral defense committee, for giving me her time and asking insightful questions.

My time at UCR would not have been as enjoyable without the company of wonderful friends, who became like a family, when I left India to come here to pursue my PhD. I have immensely enjoyed living here and I am grateful for the numerous trips and fond memories. I am thankful to Indian Students' Association at UCR for introducing me to some of my best friends and providing experiences of a lifetime. Being the president of Indian Students' Association at UCR was a rewarding experience, for which I am grateful for. I also thank my friends back in India who supported me in taking the big decision to pursue my PhD in US and supported me emotionally throughout the process. I, especially, would like to thank Amber Qureshi Urrutia for being with me at every step of the way in the PhD process. We started this journey together and I am grateful to her for being my support both personally and professionally in most difficult times. She provided sanity and stability in this crazy Economics PhD world.

Last, but not the least, I would like to thank my family for unconditional love, encouragement, and support. I would like to thank my mother, Nalini Panda, who raised me in a way that helped me have the courage and determination to pursue a PhD and supported me in realizing my dreams. I thank my father, Prof. Santosh C. Panda, for being a role model in more ways than one. This dissertation would not have had been possible without them looking out for me at every step of the way. I would also like to thank my cousins, Narayan Dash and Sahasranshu Panda and their family, for comforting me in times of distress and being my support system in US especially in the initial period

of my PhD. And most of all, I thank my fiancé Nikhil Nagane who faithfully stood by me and supported me through long working nights and days of my PhD. I thank him for being the most dependable person, with whom I could share my “eureka” moments as well as rely on him in times of distress. I have deeply benefitted from discussions of new ideas with him and his feedback from a different point of view, which helped the thesis evolve. I thank God for giving me the strength, health, opportunity, and ability to complete this PhD.

ABSTRACT OF THE DISSERTATION

Essays on Trade, Infrastructure, and Human Capital Outcomes in Developing Countries

by

Pallavi Panda

Doctor of Philosophy, Graduate Program in Economics
University of California, Riverside, June 2015
Dr. Anil B. Deolalikar, Chairperson

This dissertation presents three chapters on child health outcomes in India and sub-Saharan Africa. The goal of this research is to analyze how large-scale policies have intended or unintended impacts on the lives of people living in underdeveloped regions. In the first chapter, I study the effects of a change in trade policy, the African Growth and Opportunity Act (AGOA), on infant and neonatal mortality in sub-Saharan Africa. Increased exports and increasing opportunities for employment of mothers may contribute towards improving health of the child, due to rising incomes (income effect) or may deteriorate health of the child as the mother stays away from home (substitution effect). Empirically, I find that the increase in exports from 30 sub-Saharan countries help in reducing infant mortality by 9% of the sample mean or around 7 deaths per 1000 using the most comprehensive health data available for these countries, Demographic and Health Surveys (DHS). This decrease in infant deaths operates through increasing health seeking behavior of mother, increased possession of assets, and increased maternal labor

supply in non-agricultural sectors. The second chapter analyzes the paradox of decreasing infant mortality but not a corresponding improvement in malnutrition in India over time. We look at the effects of mortality selection on anthropometric scores in India using three rounds of National Family and Health Surveys (NFHS) and find evidence of significant negative mortality selection. Specifically, children with sample average characteristics that survive, with controls for unobservable characteristics of groups of women, have lower HAZ scores than a child randomly drawn from the population. In the third chapter, I delve into the impact of rural road creation (Pradhan Mantri Gram Sadak Yojana, PMGSY) on infant and neonatal mortality and sex-ratio in India. To study this effect at the village-level a unique micro dataset using nationwide health data surveys and online road construction data is created. Using two different empirical approaches, I find no significant changes in these statistics after the road creation in the short term. These findings will help in conceptualizing future policy actions to effectively improve child health for these developing economies.

# Table of Contents

# List of Figures

## List of Tables

# Chapter 1

## Does Trade reduce Infant Mortality? Evidence from Sub-Saharan Africa[*]


### Abstract

This paper investigates the impact of a trade policy, the African Growth and Opportunity Act (AGOA), on infant mortality and fleshes out the likely mechanisms. This paper derives a causal estimate by developing a micro panel dataset across countries and exploiting within-mother variation in survival of infants. I find that the policy reduces infant mortality by about 9% of the sample mean. I also identify channels through which the mortality reduction operates including increasing health seeking behavior, increased possession of assets, and increased maternal labor supply in non-agricultural sectors. Disaggregation reveals heterogeneity of treatment at the country and individual level.



***Keywords:*** Infant Mortality, Child Health, Trade Openness, sub-Saharan Africa.

***JEL Codes***: F13, I15, J21, J82, O15, O18, O19, O24.



[*] I would like to thank Hewlett foundation/IIE for awarding the 2013-2015 Hewlett Foundation Dissertation Fellowship in Population, Reproductive Health and Economic Development for this research. I am grateful to Anil B. Deolalikar, Aman Ullah, Mindy Marks, Jorge M. Agüero, and Joseph Cummins for their comments and guidance. This paper has benefitted from presentation at the PopPov conference in Kenya (Jan 2014), Population Association of America Conference (May 2014), International Society for New Institutional Economics Annual Conference (June 2014) and UC Global Health Day (April 2014). All remaining errors are my own.

## 1.1 Introduction

Historically, trade routes have played a major role in increasing the prosperity of nations.[1] Free trade can create access to a better variety of goods, increase women labor force participation, increase incomes and often leads to improvements in infrastructure investment (Dollar and Kraay, 2001; Broda and Weinstein, 2006; Wood, 1991; Storeygard, 2013). There are few empirical studies estimating the effect of trade on development, especially child health.[2] This paper estimates the effect of being exposed to a trade policy on infant and neonatal mortality, analyzes heterogeneous effects both at the macro and micro level, and examines possible pathways in the context of sub-Saharan Africa.[3]

Empirically, it is difficult to identify causal effects, as trade policy is likely to be endogenous to other socio-economic factors that also affect development. To overcome this, identification in previous literature has come from using instrumental variables like predicting trade volumes as a ratio of GDP using geographic factors.[4] However, these approaches might have potential threats to validity as geographical trade share may be correlated with other factors that affect children's welfare like presence of strong

[1] The infrastructure created to boost trade becomes the main arteries of countries and turn cities into "engines of growth". Railroads had a great impact on the counties in American economy due to a change in that county's "market access" (Donaldson and Hornbeck, 2013).
[2] Edmonds, Pavcnik, and Topalova (2010) analyze effect of trade policy on schooling and child labor decisions in India; Topalova (2007), Porto (2004) focus on poverty and inequality in India after trade liberalization, Levine and Rothman (2006) focus on effect of trade on infant and child mortality and stunting in a cross-country cross-sectional setting.
[3] It is developed by collating household surveys across 30 sub-Saharan African countries. Details are discussed in the Data section.
[4] Levine and Rothman (2006), Frankel and Romer (1999)

institutions.[5] This paper uses a novel way of combining micro datasets across countries to study the effect of macroeconomic trade policy on development outcomes. The effect of trade policy on infant mortality will be gauged by studying the varying exposure between the children born to same mothers but exposed to the trade policy or not in both policy-affected and non-affected countries.

Recently, sub-Saharan African countries signed a huge non-reciprocal trade agreement – the African Growth and Opportunity Act (AGOA) - which conferred on these sub-Saharan African countries largely duty-free and quota-free access to US markets. The head of the countries signing these agreements hoped the agreement would increase export volumes and spur economic growth in these economies. Frazer and Van Biesebroeck (2010) found that AGOA had a large and robust impact on exports to US without decreasing the country's export to Europe. Some countries like Kenya experienced an almost 700% increase in exports to US (from $36 million in 2000 to $284 million in 2010) (Onyago and Ikiara, 2011). This agreement took effect in 2000 with 34 sub-Saharan African countries eligible for the trade benefits included in the AGOA.

Identification in this analysis is based on each country's exposure to the trade policy at different points of time. Using retrospective birth histories from Demographic and Health Survey (DHS), I develop a micro panel dataset that spans 30 sub-Saharan African countries, and about 686,000 children born to 212,000 mothers. By observing different

[5] For example, Mauritius is surrounded by sea and has experienced an export boom in garments. But, this boom has been attributed to a sound economic strategy by the government underpinned by social and political arrangements. Mauritius also has very low infant mortality rates which would have been brought about by the safety nets provided by the government. Hence, role of institution has been playing an instrumental role in decreasing infant mortality as well as increasing trade, which may not be properly captured by an instrumental variable approach. For more details on institutions, see Rodrik (2001)

children of the same mother conceived before and after the trade policy change between AGOA affected and non-AGOA affected countries, a within-mother variation in survival of infant is carried out rather than cross-country or within-country variation. This analysis ensures that it is able to separate the effect of trade policy from other country level confounding factors since it is able to control for country specific time trends and unobservable time invariant characteristics of mothers and countries. Moreover, since this developed dataset are collated micro-level surveys across countries, it also overcomes the problem of small samples endemic to cross-country analysis.

The results of this analysis suggest that infant mortality falls by about 0.7 to 0.8 percentage points, or about 9%-10% of the sample mean, even after controlling for country-time linear trends as well as mother's time invariant characteristics. The results are also robust to controlling for some time variant country characteristics. A large portion of this fall comes from a decrease in neonatal mortality. The drop in infant mortality is evident immediately after AGOA takes place. Dynamics of mortality reveal that there exists no effect prior to AGOA being implemented, corroborating that the decrease in infant mortality is due to AGOA. The benefits of AGOA are not equally distributed across countries and across households. The effect of AGOA on infant survival is stronger for countries that export large amounts of agricultural goods and mineral ores as compared to oil exporting countries. Results suggest that infant mortality falls more for socially disadvantaged women. Infant deaths fall more for employed women than unemployed women, hinting towards spurring employment in export sectors.

The theoretical effect of trade on household is ambiguous.[6] Trade increases employment opportunities and, especially in a developing country context, opportunities for low-skilled labor. Increasing opportunities for employment of mothers may contribute towards improving health of the child, due to increasing incomes (income effect) or may even deteriorate health of child as the mother stays away from home (substitution effect).[7] By collating multiple rounds of household surveys for each country, this paper is able to identify three potential mechanisms at the household level: (a) change in household income/assets; (b) change in female employment in labor force; and (c) changing health seeking behavior of mothers.[8] Since AGOA mainly boosted exports out of the country, access to a variety of goods for consumers does not seem to be a major pathway in infant reduction. This study finds that AGOA led to a significant decline in the number of households possessing no assets, a realignment of employment of female labor from agriculture to manual labor and increased delivery care and tetanus toxoid injections to kids.

Evidence of macroeconomic trade policies on microeconomic outcomes like child health has been understudied. This paper provides new evidence of a causal effect of trade on development. AGOA has been studied mainly to look at the effectiveness of the policy in

[6] Various routes through which trade can affect infant health, both at the macro and micro level, are discussed in Section 2.2.

[7] Kishor and Parasuraman (1998), find using NFHS Data for India in 1992-93 that mothers who are employed have a 10 percent higher IMR for their children and 36 percent higher child-mortality than mothers who are not employed. Many studies find strong relationship between increased female employment and increased exports (Wood, 1991, Standing, 1999).

[8] Multiple rounds of DHS datasets are available for 22 of 30 countries. See Appendix for more details.

increasing exports from these countries.[9] This study provides first estimates of the effects of AGOA on infant mortality, to the best of my knowledge. Results point towards dominance of an income effect in reducing infant mortality in this setting. Income shocks may lead to realignment of preferences with respect to health care investment in children.[10] The results are consistent with an increase and realignment of employment and increasing incomes being the driving force of the observed decrease in infant mortality. Mothers are also choosing to receive additional health care for their children after AGOA, either because of easier availability and better infrastructure or a change in investment preference in their child's health when times are better.

## 1.2 Background

### 1.2.1 African Growth and Opportunity Act (AGOA)

The African Growth and Opportunity Act (AGOA) provide for preferential treatment of exports from sub-Saharan Africa in the form of duty-free and largely quota-free access to US markets. It entails a series of incentives provided to African countries by the US opening its market for exports originating from these countries. AGOA has been part of the US international cooperation efforts for Africa since 2000. At the onset of the legislation, 34 countries were eligible for AGOA benefits.[11] AGOA was initially set to

[9] Frazer and Biesebroeck (2007), Condon and Stern (2011), Collier and Venables (2007), Paulos et al. (2010)
[10] Case (2001), Paxson and Schady (2005), Bhalotra (2007)
[11] More countries were added for the benefits later and some were removed due to failures regarding political or democratic freedom. Cote D'Ivoire was removed from the list in January, 2005. Effective December 23, 2009, the President removed Guinea, Madagascar and Niger from the list of AGOA eligible countries.

expire in 2008 but was eventually extended to 2015. Under AGOA provisions, four main sectors account for over 90% of the exports - energy-related products, textiles and apparel, transportation equipment, and minerals and metals. Figure 1.2 plots the total US imports and exports from sub-Saharan Africa. There is a significant increase in exports from sub-Saharan Africa to US after 2001 when the AGOA took effect. Overall, total US imports from AGOA countries have increased from $5B in 2000 to over $25B in 2005 (Paulos et al., 2010).

The AGOA was implemented not only to boost exports but also improve and foster economic growth. Country eligibility for AGOA is determined by the US President, and takes into account whether countries have made efforts to improve human rights, follow open market economic policies, protect worker rights and remove child labor, combat corruption, and establish rule of law among others.[12] The eligibility criteria for the Generalized System of Preferences (GSP), a US trade preference program that applies to more than 120 developing countries, and AGOA substantially overlap, and countries must be GSP eligible in order to be eligible for AGOA but the AGOA covers more product lines and includes additional criteria. Under initial AGOA legislation, 1800 additional items were allowed to be exported duty-free in addition to the 4,600 under GSP. These newly added lines included items such as footwear, handbags, many agricultural products, chemicals, steel etc. These constituted the non-apparel exports

[12] Country eligibility is listed in Section 104 of African Growth and Opportunity Act. It states that countries need to "have established, or are making continual progress toward establishing the following: market-based economies; the rule of law and political pluralism; elimination of barriers to U.S. trade and investment; protection of intellectual property; efforts to combat corruption; policies to reduce poverty, increasing availability of health care and educational opportunities; protection of human rights and worker rights; and elimination of certain child labor practices".

under AGOA and could be exported at zero import duties as soon as the countries were declared AGOA beneficiary.

AGOA also places heavy emphasis on Africa's emerging textile and apparel industry as the primary sector for trade benefits. While AGOA removes import duties on eligible African imports, preferential market access is granted only upon compliance with the relevant Rules of Origin. These rules prescribe the percentage value added that must take place locally in an AGOA-beneficiary country, while special provisions relating to apparel outline what processing must take place locally. However, the lesser-developed countries were eligible for a Special Rule and could source raw materials from all over the world until 2004 while still receiving AGOA benefits.[13] AGOA also benefits these signatory countries as the exports under AGOA are not subject to a maximum volume ceiling as under GSP. However, with the ending of Multi-Fiber Agreements (MFA) in 2005, the apparel exports from African countries have decreased in the face of competition from China, Bangladesh, and India.

Many studies have been conducted to study the effectiveness of AGOA in increasing exports to US. Frazer and Biesebroeck (2007), Condon and Stern (2011) and Collier and Venables (2007) find a positive and significant impact of AGOA on exports.[14] Thompson (2004) and Mattoo et al. (2006) show that the largest share of US imports from Africa remain to be the oil and energy sectors. These studies provide evidence on heterogeneous

[13] The lesser-developed beneficiary countries are countries with a per capita income of less than $1,500 in 1998. By the end of 2002, 33 countries were beneficiaries of Special Rule provision.

[14] There have been few studies questioning the distribution of benefits of AGOA inside the country. Paulos et al. (2010) review the progress of a decade of AGOA and find that even though exports may be increasing, it may not be benefitting the countries internally. Kimenyi (2009) argues that only a few countries from the whole of Africa actually reap the maximum benefits.

effects of AGOA based on country's main item of export and volume of exports from these countries. In line with this argument, along with looking at the effect of AGOA on the people living in these economies, this study also explores the heterogeneity and inherent differences between the 30 sub-Saharan African beneficiaries to capture the differences in effects on those countries.

### 1.2.2 Trade Linkages

Trade can affect the development process of a country via various direct and indirect mechanisms, both at the macro and micro level. Trade affects the overall aggregate or macro state of the economy by affecting economic growth, government health expenditures, urbanization and increased job creation which in turn affects the socio-economic indicators. Trade may improve health conditions by increasing tax revenues of the government allowing it to possibly increase health expenditures.[15] Economic growth also results in higher household incomes, which in turn could improve health outcomes via mechanisms like improved nutrition, improved access to sanitation and health care etc. Trade also spurs employment in labor-intensive sectors in a developing economy. Increasing employment can benefit child health due to increasing incomes or worsen health due to increased time away from home by the mothers.

[15] (Adam, Bevan, and Chambas 2001) find that openness raises trade tax revenues in CFA franc countries while it has little effect in non-CFA franc countries in sub-Saharan Africa. (Agbeyegbe, Stotsky, and WoldeMariam 2006) find that trade liberalization is not strongly linked to aggregate tax revenue, but with one measure, is linked to higher income tax revenue in sub-Saharan Africa.

Most of studies concerning trade liberalization look at outcomes like growth rates, income inequality, productivity and wages.[16] There are many studies estimating the effects of trade policy on income growth rates, showing a positive effect (Dollar, 1992; Frankel and Romer, 1999; Dollar and Kraay, 2001). Dollar and Kraay (2001) argue that the increase in growth rates with trade leads to proportionate increases in the income of the poor and therefore poverty reduction in poor countries. They also find that there is no systematic relationship between changes in trade and changes in inequality. The studies estimating the effect of trade on income growth rates are cross country studies over many developing nations using different indicators of openness like decade-over-decade changes in the volume of trade, index of real exchange rate distortion or geographically determined amount of trade to counter the issue of endogeneity. But, this large literature on trade remains inconclusive about a clear-cut effect of trade on both growth and development. Though they provide insights into the effect of trade on macroeconomic outcomes, drawing a causal conclusion between the two is difficult.

Trade may also affect the individuals and households directly at the micro-level through multiple channels. International trade can affect economic outcomes by decreasing transportation costs (Clark, Dollar, and Micco, 2004; Hummels and Skiba, 2004; Storeygard, 2013). As transportation costs decrease, the price of traded goods decreases. In the scenario where trade promotes better infrastructure and better market access, it has been argued that access to transportation networks may not have large impact on the relative economic performance of those areas affected by improvements in transportation

[16] Goldberg & Pavnick (2007a), Frankel and Romer (1999), Winters, McCulloch, and McKay (2004), Hanson (2007)

infrastructure in a developing country (Banerjee, et al., 2012). Mobility of factors and development of institutions play a critical role in realizing the gains from trade. Institutional differences are a large factor in explaining the differences in economic prosperity between nations (Acemoglu, et al., 2001).

Trade can also bring variety gains by making a bigger set of consumption bundle available for people in the country (Broda and Weinstein, 2006). Increased variety in food may also result in increased diversity of micronutrient consumption, which can be important for maternal and child health in countries where micronutrient deficiency is endemic.[17] Topalova (2010) finds evidence that Indian districts with greater exposure to trade do not have a significant gain in consumption levels. In the Indian context, Topalaova (2007) establishes that rural areas experienced slower progress in poverty reduction but with no significant impact on inequality.

Effect of trade directly on child health has been under-studied. Levine and Rothman (2006) use Frankel and Romer's approach in predicting how much a country will trade based on exogenous geographical characteristics and then use this predicted trade share to obtain a cross-sectional effect of trade on children's health. They find that for an average country, a 15-percentage point increase in predicted trade as a share of GDP (an increase of about 1 standard deviation) corresponds to approximately 4 fewer infant deaths per 1000 births. However, they do not use a panel data set and hence are unable to capture how the change in trade affects change in infant mortality. The country specific effects are taken care by this latter channel of estimation.

[17] (Cutler, Deaton, and Lleras-Muney 2006), (Porto 2004)

## 1.3 Data

The micro level health data for the sub-Saharan African countries comes from the Demographic and Health Surveys (DHS). DHS are nationally representative household surveys that provide data for population, health and nutrition The DHS questionnaire is (mostly) standardized across countries and rounds, and so allows for comparisons across countries. The Standard DHS Surveys have large sample sizes (about 5,000 households) and are typically conducted about every 5 years. Information regarding child health, immunizations, antenatal care, etc. is found in the surveys, along with mother and household characteristics.

DHS collects data using three types of questionnaires – household, women's and men's questionnaires. Household questionnaire is used to collect data on household dwelling units, nutritional status, and anemia; while women's questionnaire is used to collect data from women about the characteristics, reproductive behavior, contraception, children's health etc. Women of reproductive age (15-49 years) are interviewed about the date of birth and death (if applicable) for up to 20 children they have had. This kind of retrospective survey gives an opportunity to build a panel dataset of mothers, with the time dimension being the child birth year.

One problem with the recall data is measurement error. To be robust to measurement error and to capture the maximum effect of carrying out the siblings analysis, all children

born before 1990 were dropped from the sample.[18] This ensures that the siblings are not very far apart in age and hence are broadly comparable. This also reduces the recall bias. Moreover, some sub-Saharan African countries in the sample gained their independence between 1975 and 1990[19] and also experienced higher rates of civil wars, which may muddle with the effects of trade on infant mortality. Since the most recent year in the surveys have few observations, they have been merged with the previous year to prevent sharp spikes in infant death due to limited observations in the last year of survey.[20] Even though, Mozambique, Liberia, Ethiopia and Cote D'Ivoire were given AGOA rights, in the sample they effectively behave as not treated since the law came into effect in the last year of the survey. These are in the non-treated group, along with Zimbabwe, which is not AGOA eligible.

There are 36 DHS Surveys publically available for the sub-Saharan countries where DHS survey has been carried out at least once.[21] The surveys for Central African Republic, Comoros, Gabon, South Africa, Sudan and Togo were all carried out before AGOA was implemented in these countries. The remaining 30 surveys are included in this analysis.[22] A dummy variable indicating if the child has died before reaching the age of 1 year is constructed based on mother's birth history. This will be the indicator for *individual-level*

[18] The results are not dependent on the year of birth cut off. Other models with different year of birth (1994 or 1995) cutoffs gave similarly significant results. See Table 1.13 for results.
[19] Mozambique (1975), Cape Verde (1975), Comoros (1975), Sao Tome and Principe (1975), Angola (1975), Zimbabwe (1980), Namibia (1989)
[20] Results are robust to not doing this.
[21] These are most recent surveys at the time of analysis. Newer DHS has been carried out in past 2-3 years.
[22] The list of DHS used and respective sample periods are listed in Table 1.A1 in Appendix.

infant mortality. As long as at least one round of survey has been conducted in a particular country, a panel dataset of mothers for that country can be built.[23]

After dropping data for children born within twelve months of the survey, to ensure full exposure for every child in the sample and reduce measurement error, the sample includes 686,093 children born to 212,738 mothers. Infant (Neonatal) mortality rate is the number of deaths of children before reaching the age of one year (month) per 1000 live births. In this sample, it is calculated by multiplying the sample mean child deaths (in the appropriate age group) by thousand. The sample average infant mortality rate is 8.15% of live births while the sample neonatal mortality rate is 3.8% of live births. Since determinants of neonatal mortality may differ from infant mortality, an indicator for child dying before the age of 1 month is also constructed and effect of AGOA on it is studied.

In Table 1.1, I show average infant deaths for the whole population, as well as infant mortality based on different characteristics of mother like education, place of residence and wealth levels. Infant mortality varies significantly (based on t-statistic) between AGOA and non-AGOA countries. A mother is labeled as educated if she has attended any type of school and uneducated is defined as mother did not attend any school. Since DHS does not have income data, definition of poor is based on possession of assets. The wealth index is calculated using easy-to-collect data on a household's ownership of selected assets, such as televisions and bicycles; materials used for housing construction; and types of water access and sanitation facilities using principal components analysis

[23] DHS dataset has been used in this manner to study the effect of income fluctuations on infant mortality (Bhalotra, 2007, Paxson and Schady,2005) and effect of democracy on infant mortality (Kudamatsu, 2012). This paper follows that methodology.

and is reported in DHS. A mother is categorized as poor if the wealth index is marked as poor or poorer, while mothers with wealth index being middle, richer or richest are categorized as non-poor. Rural or urban are defined by the place of residence of mother during the time of interview. Infant and neonatal deaths also vary between AGOA and non-AGOA countries for women of different socio-economic status. Among the child characteristics, birth order differs between these countries and is also included in controls.[24] It is observed that these countries are similar in terms of sex composition, but the composition of mother's age at birth is different across these countries. I include controls for maternal age at birth.

Figure 1.1 plots the sample mean infant mortality rates by year for countries affected by AGOA by 2001, affected after 2001 and never affected by AGOA. It shows that AGOA affected countries have higher infant mortality than non-AGOA countries at the time of first implementation of AGOA in 2001. All three groups of countries exhibit a declining trend in infant mortality over the years and the difference seems to be decreasing after AGOA has been implemented. The differential trends in infant mortality will be accounted for by using country time trends in this analysis. The mean infant mortality rates for the 25 AGOA affected sub-Saharan African countries in the sample by year of birth of child, 1990 onwards is shown in Appendix Figure 1.A2. The data shows a declining trend in infant mortality over time for all countries. A sharp fall in infant deaths in some of the countries after the year AGOA is implemented is observed, more prominently than others.

[24] Country specific birth order is also controlled for in robustness check.

Since the estimation strategy includes maternal fixed effects, it is the mothers giving birth both before and after AGOA that contribute to the identification of effect of AGOA. Thus, in Table 1.2, I compare the characteristics of mothers having two or more births before and after AGOA (column (3)) with mothers in the entire sample (column (1)) as well as mothers in AGOA countries (column (2)). Mothers who give birth at least twice and both before and after AGOA are less well educated, live in rural areas and poor. Since mother fixed effects estimation derives the effect of AGOA on infant mortality using those mothers giving birth both before and after AGOA, Table 1.3 shows the sample mean infant and neonatal mortality rates for mothers giving birth both before and after AGOA and for mothers with more than two births only before or after AGOA. The sample mean infant and neonatal mortality rates fall for mothers giving birth both before and after AGOA, after AGOA is implemented. Column (2) shows that the mean infant mortality is higher for the mothers giving birth only after AGOA than mothers in Column (1), but it is lower than the mean infant mortality for mothers giving birth only before AGOA.

## 1.4 Empirical Strategy

To analyze the effect of trade on infant mortality, I estimate the following equation using a linear probability model[25]:

[25] I also check for Logit estimates. They are significant. But due to easier interpretation of LPM estimates, I present those in the results. Also, LPM Model allows me to use a general form of mother fixed effects and since the estimations are deviations from the trend (after controlling for country-time control), the coefficient is likely to be small and should lie in the range of 0 to 1.

$$IMR_{imct} = \alpha_m + \beta_t + \gamma T_{ct} + X_{imct}\delta + \mu_c.t + \varepsilon_{imct} \quad (1)$$

for child i, born to mother m in country c in year t. IMR is a dummy which takes the value 1 if child i dies before reaching the age of 1 year, $\alpha_m$ is mother fixed effect, $\beta_t$ is birth-year fixed effect and $\mu_c.t$ captures the country-time specific trend. $T_{ct}$ takes the value 1 if the specific country was under AGOA throughout time t. $X_{imct}$ is a vector of control characteristics including sex of the child, whether or not they are born in multiple births (i.e. twins, triplets, etc.), dummies for their birth order, mother's age at birth and birth month of the child. It may also be argued that birth order trends differ between countries. As a robustness check, country specific birth orders are also controlled for. $\gamma$ provides the estimate of the effect of AGOA on infant mortality. The standard errors are clustered at the country level to take into account any correlation of the error across space and time within each country.

Treatment in this paper is defined as a child's exposure to AGOA.[26] This is a dummy variable which takes the value 1 if the child is born *after* AGOA is implemented. This ensures that the child has been fully exposed to AGOA through their lifespan. For example, if AGOA was signed and passed on 1st October, 2001 for country $C_1$, then AGOA takes the value of 2002 for $C_1$. If instead, AGOA is passed on January 1, 2003 for country $C_2$, then AGOA takes the value 2003 in $C_2$. A child is then said exposed to AGOA if in $C_1$, they are born in 2003 or later while in $C_2$, they are born in 2004 or later.

[26] There have been concerns on using trade volumes as a measure of trade policy (Rodrik, 2001). Trade volumes generally reflect the outcomes of many different things like economy's overall performance as well as productive capacity of the economy. Hence, trade volumes are not entirely controlled by the government, while trade policies are. Keeping this in mind, this study abstracts away from using trade volumes as an indicator of trade policy.

However, the results are not sensitive to this definition. Even if AGOA variable is defined to be the year it was announced, the results still remain highly significant.

In the next specification, an interaction between mother's birth cohort by child's birth year ($\beta_{bt}$)[27] fixed effects are included:

$$IMR_{imbct} = \alpha_m + \beta_{bt} + \gamma T_{ct} + X_{imbct}\delta + \mu_c.t + \varepsilon_{imbct} \quad (2)$$

This specification accounts for the possibility that women may be delaying their fertility based on improvements in survival of their kids overtime. $\gamma$ captures the average difference in changes in probability of death of infants born to the same mother between those countries that have been affected by AGOA vis-à-vis those that are not. Since AGOA implementation varies by countries as well as time, fixed effects estimation can be carried out for more robust estimates.

For the estimates to be unbiased, the error cannot be correlated with any of the covariates and outcomes, not only contemporaneously but also in leads and lags as the same mother gives birth. Specifically:

$$E(\varepsilon_{imbct} | T_{ct}, \beta_{bt}, \alpha_m, \mu_c.t, X_{imbct}) = 0 \quad (3)$$

This specification also assumes that treatment selection can be based on unobserved heterogeneity at country level, but within country which mothers and children got the treatment is unrelated to the gain from the program. Another concern is that mothers affected by AGOA in AGOA affected countries do not behave differently than mothers in

[27] Subscript b denotes the mother's birth cohort.

non-AGOA countries, if they had been AGOA affected. Thus, mothers cannot be timing their fertility in response to AGOA. I test to see later if fertility selection bias is a major concern in the data.

The main concern in studying the effect of such a policy is the difficulty of disentangling the effect of this policy from the prerequisites for being a signatory on the AGOA. In terms of disentangling this effect, this study does better than cross-country studies. Time-invariant heterogeneity regarding geography, history, culture, politics and attitudes etc. are taken care of by the mother fixed effects ($\alpha_m$) since this is implicitly a country fixed effect – mothers of the children belong to a certain country of residence and hence controlling for mother's characteristics implies controlling for the country characteristics.

The year fixed effects ($\beta_t$) control for an aggregate time variation involving improvement of health technology and year shocks. $\beta_{bt}$ controls for changing time of mother's age at birth. The mother cohort by year fixed effects controls for fertility changes overtime in that region due to improvements in health technology. The country specific trends ($\mu_c.t$), in fact, also allow country specific improvement in infant and maternal health i.e. differential states of development of the countries.

But, there may be time variant heterogeneity which may affect both trade and infant mortality rates. Implementation of AGOA or how well the country does after its implementation may depend on the country's political situation, GDP per capita, average female education of the country etc. Countries with a higher GDP per capita or in a democratic regime may experience a lower IMR too (Kudamatsu, 2012). Hence these

may bias the estimates. As a robustness check, at the country level there is a control for additional characteristics ($Z_{ct}$) like GDP per capita, political regime of the country, whether it is a democracy, degree of openness overtime, average level of female education etc. which may help control some of the time variant heterogeneity at the country level. To capture these effects, I estimate the following equation:

$$IMR_{imbct} = \alpha_m + \beta_{bt} + \gamma T_{ct} + X_{imbct}\delta + \mu_c.t + \lambda Z_{ct} + \varepsilon_{imbct} \quad (4)$$

To check for heterogeneity based on mother's level of education, place of residence and possession of assets, the mother-FE regression is run with interactions to tease out the effects:

$$IMR_{imbct} = \alpha_m + \beta_{bt} + \gamma(T*MC)_{ct} + X_{imbct}\delta + \mu_c.t + \varepsilon_{imbct} \quad (5)$$

Where, MC defines the mother's characteristics. The interaction term $(T*MC)_{ct}$ provides an estimate of treatment effect of AGOA on probability of infant death for a specific subsection of the population based on assets, education, employment and place of residence in comparison to the reference population.

Heterogeneity at the country level is also necessary to observe given the difference in exports variety and volumes across AGOA beneficiaries. To capture these effects, I estimate the following regression:

$$IMR_{imbct} = \alpha_m + \beta_{bt} + \gamma(T*CC)_{ct} + X_{imbct}\delta + \mu_c.t + \varepsilon_{imbct} \quad (6)$$

CC captures differences in country characteristics like whether a country is a predominant petroleum exporter, apparel exporter, agricultural exporter, low income country, region etc. The

interaction term will indicate which types of countries are actually accruing the most benefits in reducing infant deaths via AGOA.

## 1.5 Results

### 1.5.1 Event-Study Graph

I create an event-study graph for the treated countries to show the effect of AGOA on infant and neonatal mortality. Figure 1.3 graphs the likelihood of child death by year, with respect to the treatment, for the treated countries. The plotted estimates depict the differential trends in infant and neonatal mortality over four years before and after the AGOA announcement (with the year of announcement being the omitted year). The estimates $\theta_j$ are derived from the following regression:

$$Death_{imct} = \alpha_c + \beta_t + \sum_{j=-4}^{4} \theta_j T_{c,t+j} + X_i'\delta + \varepsilon_{imct} \quad (7)$$

where $T_{c,t+j}$ is 1 for j years of announcement of AGOA in country c. The specification controls for country and year fixed effects. The covariates included are mother's age at birth, mother's socio-economic characteristics, mother's education, sex of child, birth order, birth month, and whether born in multiple birth. The standard errors are clustered at the country level.

In both neonatal and infant mortality, there are no noticeable trends in the pre-treatment period. Consequently, F-test rejects the null hypothesis of joint significance of pre-treatment years. After AGOA is announced, there is a sharp fall in both infant and

neonatal mortality. The child death drops are significant at the conventional levels for the 1st, 3rd and 4th years after AGOA implementation. The point estimates are presented in Appendix Table 1.A8. The infant and neonatal mortality is below the pre-treatment level even after four years of implementation.

### 1.5.2 Main Results

Table 1.4 provides the main regression results of the effects of treatment on infant mortality Column (1) shows the results for country fixed effects, without controlling for linear country time trend. The resulting coefficient on AGOA is negative, but not statistically significant. Since there are country specific trends in infant mortality, not controlling for those trends is confounding with the effect of AGOA. Controlling for country time trends along with country fixed effects in (2) makes the coefficient statistically significant. The coefficient now indicates that AGOA reduces the probability of infant dying by 0.8 percentage points. Controlling for maternal fixed effects in (3) decreases the coefficient to about 0.7 percentage points.[28]

Table 1.4, specification 2 (columns (4)-(7)) additionally controls for cohort-year fixed effects. In this specification, the changing time of mother's age at birth due to improvements in survival of babies over time in Africa is accounted. The fixed effect controls for fertility changes over the years due to improvement in health technology.

[28] To be robust to the possibility of small number of clusters in the sample, I also test the significance of coefficients using method outlined in Donald and Lang (2007) and Cameron and Miller (2013). Since N within each group (country) is large, the resulting t-statistic will have a T(G-2) distribution rather than standard normal, where G is the number of groups. The t-statistic is computed using the estimated coefficient and clustered standard errors and is tested using the T(G-2) distribution critical levels. The critical values for a two-tailed test using T distribution with 28 degrees of freedom are 1.70 at 10%, 2.048 at 5% and 2.763 at 1%. The coefficients still remain significant at these levels.

Even after controlling for these with an interaction of dummies for mother's birth year (cohort) with child's year of birth, the magnitude of the coefficient remains around 0.7 percentage points statistically significant.[29]

The absolute value of coefficient remains between 7.9 to 6.9 reductions in infant deaths per 1000 live births, as we move across specifications. Thus, the results are robust to various specifications. Mother fixed effects controls for factors like maternal ability to raise kids, genetic traits, household environment, parental education, place of residence etc. On carrying out mother fixed effects analysis of AGOA on neonatal mortality in column (7), a significant negative effect is found. Neonatal deaths reduce by 4.5 deaths per 1000, which is about 12% of the sample mean. Hence, about half of the reduction in infant deaths is coming via a decrease in neonatal deaths.

The magnitude of the estimated effect is economically significant effects as well. Exposure to AGOA reduces infant mortality by 0.7 percentage points which is 9% of the sample mean and decreases deaths before age of one month by 0.4 percentage points, which is 12% of the sample mean. For comparison, the effect of Progresa, a conditional cash transfer program in Mexico, is an 8% reduction in rural IMR (Barham, 2011). Comparing this with previous literature[30], the effect is higher in absolute magnitude using mother FE than in the cross-country setting, with trade openness contributing to a reduction of around 7 infant deaths per 1000 births.

[29] On testing sex selection at birth, I do not find differences in probability of infant dying based on gender. Results are presented in Appendix Table 1.A7.

[30] Levine and Rothman (2006) find that for an average country, a 15-percentage point increase in predicted trade as a share of GDP results in 4 fewer infant deaths per 1000 births.

It is crucial that it is AGOA which brings about the change in infant mortality and we are not wrongly attributing the effects of some other change to AGOA. For the estimates to be unbiased, the error should not be correlated with any of the covariates and outcomes, not only contemporaneously but also in leads and lags as the same mother gives birth. To corroborate this, a regression involving lags and lead periods for AGOA has been estimated.[31] Figure 1.2 graphs the dynamics of infant mortality from 2 years before AGOA implementation to 4 years after it. It can be seen that change in infant mortality had been almost constant, not significantly different from zero in the two years before AGOA was implemented. There is a significant drop in infant mortality as compared to 3 years before implementation of AGOA in year 1, year 3 and year 4 of the AGOA being implemented.

### 1.5.3 Robustness to Time-variant Factors

Table 1.5 controls for country level variables like log GDP per capita, Democratic regime, Openness, female education etc. in the mother FE specification with cohort-year FE. Since some benefits of AGOA were based on income threshold, especially for Apparel exports, it is imperative to control for changing GDP per capita levels for the countries since higher income countries may also display better health of children. GDP per capita data is obtained from PWT 7.0 and log of GDP per capita is used to run the regression with cohort year fixed effects in Table 1.5 (1). Infant mortality was observed to decrease with an increase in the GDP per capita (significant at 10% level), but even

[31] Point estimates are shown in Appendix Table 1.A2.

controlling for GDP per capita did not reduce the magnitude of the AGOA coefficient much nor remove statistical significance.

Some studies find that democracy and political regime may affect child health (Kudamatsu, 2012). Since AGOA emphasized political stability, it was the politically stable countries which acquired and retained AGOA rights.[32] It may also have served as an incentive to turn into a democratic country to acquire AGOA rights. Hence, democracy may have served as a pre-condition for getting AGOA benefits. The effect of democracy and political regime is controlled for by using the democracy-dictator data from Cheibub et al. (2010) which is an updated dataset based on Przeworski et al. (2000). They define democracy as: the executive is directly elected or indirectly elected via the legislature; the legislature is directly elected; there is more than one party; and the executive power alternates between different parties under the same electoral rule. If a country satisfies these conditions, the democracy indicator takes the value 1. In Table 1.5 (2) controlling for democracy status of the country, does not change the magnitude of the coefficient much from the results in Table 1.4. Democracy tends to reduce infant mortality but the coefficient is not significantly different from zero at the conventional level.

Sub-Saharan African countries have increasingly received Official Developmental Assistance (ODA) from developed countries to promote economic development. It may be the case that at the same time AGOA was introduced, the trade-related or other parts

[32] For example, Cote D'Ivoire was removed as an AGOA beneficiary due to not implementing a peace plan and Eritrea was removed after failing to bring about democratic reforms.

of ODA also increased. ODA is intended to provide assistance in development and hence will aid in infant mortality reduction. To ensure that the actual effects of AGOA are observed, ODA is included as a control variable in Table 1.4 (3). It may be argued that a country which already had trade routes open under GSP would have benefitted more from AGOA and hence its coefficient maybe capturing the effects of already increased trade flows. But controlling for openness from PWT 7.0 in Table 1.5 (4), it is observed that the coefficient is not significantly different from zero and the original coefficient on AGOA does not decrease in absolute value or significance.

Countries with higher growth of human capital such as average years of education of females in a country may be benefitting more than others in attracting trade flows as well as decreasing infant mortality. Thus, data for the average years of schooling of females 15 years or older is collated from Barro and Lee (2010) and there is a control for average years of female education of the country in Table 1.5 (5). The number of countries for which this data is available falls to 21. It is seen that the coefficient is not significantly different from zero and also the coefficient on treatment to AGOA does not change much and stays statistically significant.

Commodity price fluctuations have contributed to improved incomes and growth in Africa over the last decade (Deaton, 1999). Changes in international commodity prices can work through changes in consumption and government expenditure, which results in changes in national output. Since sub-Saharan Africa relies a lot on primary exports and these are subject to volatility in commodity prices, it is necessary to separate the effects

of a commodity price boom from the effects of AGOA. Considering this finding, the commodity price index derived from PWT 8.0 in Table 1.5 (6) is controlled for but this does not decrease the magnitude of the coefficient on AGOA much. The coefficient on commodity price index is itself significant and tends to increase infant mortality. In Table 1.5 (7), all the macro variables are controlled for and that also does not reduce the magnitude or significance of the variable in question. It confirms that the coefficient on AGOA is robust to controlling for some of the important country level time variant factors.[33]

### 1.5.4 Heterogeneity in Effects

AGOA may affect the recipient countries differentially based on their composition of exports at the country level. At the individual level, heterogeneity may exist based on characteristic of the mother and the household. I explore these in the following section.

Table 1.6 columns (1)-(4), check for heterogeneity in effects based on mother's place of residence, education, possession of assets and employment.[34] The AGOA differentially decreases infant deaths for uneducated mothers more than educated mothers.[35] AGOA also has a significant effect in reducing infant mortality for mothers living in rural areas, but not for those living in urban areas. AGOA seems to be effective in significantly

[33] I also check for robustness to controlling for ODA since sub-Saharan African countries were highly dependent on foreign aid and these changed with time, the coefficient still remain significant and similar in magnitude.
[34] The definitions of variables are elaborated in Data Section.
[35] The coefficients are not significantly different from each other using an F-test for educated and uneducated mothers. Though, individual coefficients point towards more reduction in infant deaths for uneducated mothers. F-statistic for difference in coefficients is significant for rural and poor mothers vis-à-vis urban and non-poor mothers.

reducing infant deaths for poor more than non-poor; negating the widely held notion that trade increases inequality. While interpreting these results, it should be kept in mind that data on asset variables is available at the time of survey. So, as long as mothers have not moved across wealth categories (moving within wealth categories does not pose a problem), these results are informative. AGOA seems to be affecting the more backward sections of the society, where there is a larger scope of reducing infant mortality. This is consistent with standard economic theory (Hecksher-Ohlin model) stating that gains of trade should flow to abundant factors, and in this developing country setting, unskilled labor (uneducated rural poor mothers) should benefit the most.

One of the main channels through which AGOA may affect women is by affecting their employment. The women who are working should benefit more than women who have no gainful employment. The results in Table 1.6 column (4) show this. Moreover, among the employed, women employed in agriculture and manual labor benefit the most since many of these countries were apparel, mineral or agriculture exporting countries. This can be seen in Table 1.7. The infant mortality for mothers employed in the agriculture and manual labor sector, falls significantly by around 1.2 percentage points while those involved in household and services do not show a significant decline with respect to the fall in infant mortality for unemployed women.

AGOA may affect the recipient countries differentially based on their predominant commodity of export. I use the trade volumes data from the Office of the United States Trade Representative, by commodity classification, to determine the main commodity of

export at the 3-digit level. In Table 1.8, it is seen that countries producing and exporting agricultural products as well as mineral ores and petrol and metals gain the most from AGOA. These countries see a higher relative decline in infant mortality than others. On further disaggregation, the five countries with highest declines in infant mortality are Rwanda, Kenya, Lesotho, Tanzania and Zambia.[36] These are also majorly agriculture, apparel or mineral exporting countries, which benefitted the most under AGOA. It has been argued that with Oil and Gas being the most valuable export to US with AGOA, it does not create long term benefits for the economy. The results are in line with this reasoning. In fact Table 1.A7 shows that oil exporting countries like Angola and Nigeria see an increase in infant deaths (although statistically insignificant in case of Angola). It may well then be that resource rich countries are not able to effectively harness the developmental gains from trade.

Table 1.8 column (2) shows that low-income countries in sub-Saharan Africa experience a significant decline in infant mortality due to AGOA vis-à-vis the middle income countries. There are heterogeneous effects by predominant religion of the country. Predominantly Islamic countries experience a larger fall in infant mortality than predominantly Christian countries. Many of the Islamic countries are in West Africa, dealing with mineral and ore exports.[37] In terms of regional heterogeneity, East Africa experience larger gains in infant deaths. Four out of the five nations in East Africa are

[36] The full set of results for country effects are shown in Appendix Table 1.A5.
[37] Guinea, Niger, Sierra Leone are mineral exporters which saw a significant decline (refer Appendix Table 1.A5)

predominantly agricultural exporters, which corroborates the previous finding. Even at the macro level, AGOA helps in levelling the disparities.

The benefits and heterogeneity we see at the micro level are reflected in the country level heterogeneity analysis. The benefits to the poor and uneducated are consistent with low income countries and countries with agricultural product exports benefitting more than countries with predominantly oil exports. This is because countries with agricultural exports or mining must be employing unskilled labor which is depicted by greater benefits accruing to that section of the society. This is also reflected in the results showing larger falls in infant mortality accrue to women employed in the agriculture or manual labor. On the other hand, countries with predominantly oil and energy exports are not able to reap the benefits despite generating new revenues.

## 1.6 Pathways

Table 1.9 delves into finding the possible pathways through which the effects are operating at the macro level. Country level macro data from World Development Indicators and Penn World Table has been used for this analysis. Controlling for GDP per capita and health expenditure per capita at country level does not change the magnitude of the effect of AGOA on infant mortality much. The magnitude of the coefficient on AGOA does not change may be due to the use of aggregate country-level data. The effects of AGOA on the country-level variables are hard to decompose since these variables are highly correlated with aggregate trends. Results suggest that increasing GDP per capita, health expenditure per capita and access to paved roads are important in

themselves in reducing infant mortality. Apart from income boost, AGOA possibly brings about changes in public and private health expenditures that are benefiting the individuals in those countries. AGOA does not seem to operate via increase in labor force participation rates as well as changing fertility.[38] The pathway findings are similar to Levine and Rothman (2006).[39]

Due to availability of micro level data, I can further verify at the household level if increase in income and increasing health expenditures after AGOA has been implemented contribute to this reduction in infant mortality. Families can increase investments in children's health if income increases due to AGOA, thereby reducing infant mortality. Moreover, AGOA may directly affect the availability of health care interventions that are also known to affect the probability of infant survival, for instance tetanus toxoid injections to pregnant mothers, skilled delivery assistance, and access to piped water and toilets.[40] AGOA may also affect the employment and distribution of employment in various sectors, thereby having an impact on infant mortality. Women working more as manual labor or managerial services may earn more or have better bargaining power affecting health of children than women staying at home.[41]

[38] Since the data from these data sources (World Bank and PWT) are sparse for some of the indicators, these results should be interpreted as more of a correlation than causation.
[39] The authors find that trade predicts higher income, higher immunization rate, and larger public health expenditures.
[40] Black et al (2003), Jones et al (2003)
[41] Aguayo-Tellez et al (2010) show in the Mexican case of trade liberalization that employment of women increased and household bargaining power shifted in favor of women

DHS does not collect data retrospectively for other variables like possession of assets by the household and employment.[42] These are collected only at the time of survey. Information on health care variables is available for live births for last five years. Therefore, mother fixed effects on a retrospective panel dataset created previously, can no longer be employed for this analysis. To gain more variation, a repeated cross-section sample of infants at each survey is created by collating data for various rounds of survey for each country. I have data on assets, employment and health care variables for 22 countries, where DHS survey has been carried out more than once.[43] Since mother fixed effects cannot be controlled for, I instead create 'mother-cohorts' defined by their year of birth, place of residence (country and urban/rural), and level of education (attended primary school or not). In the estimation of the effect of AGOA on change in health care services, assets of household and maternal labor force supply, I control for fixed effects by these mother cohort categories.

Since tetanus toxoid injections are given to pregnant mothers, I estimate the effect of AGOA on probability of getting more tetanus toxoid injections if AGOA was present in the year when the mother was pregnant. Hence, the coefficient of interest is of $T_{c,t-1}$. For other healthcare variables like skilled delivery assistance, access to piped water and access to flush toilet, the treatment is the same as before as these help in reducing infant mortality after a child is born. The regressions control for birth year fixed effects,

[42] DHS does not have data on total income of the household. Instead it asks if the household possess certain assets at the time of survey. Accordingly I define a household as poor if it possesses no assets at the time of survey.

[43] I drop Angola MIS surveys, which do not have data on the required variables and Burundi and Guinea do not have data on employment at time of survey. For more details of surveys used, see Appendix A4.

multiple births, gender, birth order and mother cohort characteristics fixed effects. The standard errors are clustered at the country level.

Table 1.10 shows the results for healthcare and sanitation. The evidence suggests that mothers are more likely to get immunization and get a skilled delivery birth attendant for their deliveries after AGOA. This may be brought about by increased availability of infrastructure or by behavioral changes in mothers due to perception of giving birth in better times. In this sense, if trade shock is considered a positive income boost (or an upturn), health seeking behavior of mother changes according to the environment.[44] The significant increase in health care access also points to the fact that we observe a big decline in neonatal mortality. On the other hand, the access to flush toilets and piped drinking water decreases after AGOA. The increase in health care expenditure per capita at the macro level indicates better access to health care facilities rather than improving infrastructure at homes.

Maternal labor supply is likely to change after AGOA as the policy changes the opportunities available for women. After AGOA, more women are likely to be working in apparel or mineral and ores sector as manual labor since AGOA increased labor opportunities in those export sectors. This is reflected in the results presented in Table 1.11. After AGOA, the probability of a woman working in agriculture decreases while probability of work in manual labor and managerial services increases. It is documented that children of mothers in agricultural work are more likely to contract diseases and less likely to be treated in downturns (Bhalotra, 2010). This shift away from agricultural work

[44] Paxson and Schady (2005) point out cyclicality in health seeking due to income cycle.

into manual labor and managerial services, points towards a benefit in terms of reduced infant deaths.

AGOA also affects possession of assets by women. Since data on wealth index is not available for all surveys, I instead create an indicator for being poor as mother not owning any assets. Though it is not clear why some assets increase while others decrease after AGOA, results suggest a decrease in probability of being "poor" by 6.3 percentage points in Table 1.12. AGOA seems to be increasing affluence of mothers and hence help in decreasing infant deaths. Due to non-availability of data on consumption and food intake, I cannot estimate the effect of AGOA on that channel. But, since AGOA mainly increased exports, a greater increase in availability of food or decrease in consumer prices across the country via increased imports do not seem to be plausible channels in this case.

## 1.7 Robustness Checks

One concern in this analysis may be that AGOA may have affected fertility differently so that mothers who gave birth in those years in AGOA countries are not the same as the mothers who gave birth in the same year in non-AGOA countries. This may also differ by socio-economic status of women. To test this, I calculate the percentage of women of socio-economic status z giving birth in each country-cohort in year t by creating a panel for women aged 15 to 49 years of age.[45] The regression estimates the effect of AGOA on this fertility indicator. To account for changes in maternal age by year of child birth,

[45] Fertility(Z) = number of births(Z)/total number of women(Z)*100 where Z refers to socio-economic type of woman being uneducated, poor or rural. The types are segregated as there is evidence of these types of women systematically differing in their fertility/mortality behavior (Paxson and Schady, 2005).

differing by 'type' of woman, the regression controls for cohort-year-type fixed effects. These regressions also control for effect of birth cohort of women on fertility differing by country and 'type' using cohort-country-type fixed effects. The results, presented in Table 1.13, show that the coefficients are not significantly different from zero. Fertility selection bias does not seem to be a major concern in the estimations.

Placebo tests by allocating fake timing of AGOA is also carried out to rule out spurious effects of policy change.[46] I re-estimate the effect of AGOA on infant mortality by assuming that AGOA has been implemented 1 to 5 years before the actual implementation. The t-statistics for each of the regressions is plotted in Figure 1.6. I expect to observe no significant effect of the fake treatment. None of the t-statistics for the previous periods are significant at the standard levels.[47]

To alleviate the concern that the results may be driven by one outlier country, Figure 1.5 shows that the result is robust to dropping one country at a time implying that these are not driven by changes due to an outlier country.[48] Also, since there are 30 countries and there may be country specific differences in birth order or mother's age trend, Table 1.14 controls additionally for country specific birth order dummy and country specific mother's age quadratic trend. The magnitude and significance of the coefficient derived in the main specification is unchanged, implying that the result is robust to differing trends and decline among countries in birth order and mother's age.

[46] I also carry out another test where I allocate AGOA treatment and year of the treatment to the countries in the sample, randomly from a uniform distribution. I do not find any significant results. Results are available on request.
[47] Point estimates are in Appendix Table 1.A6.
[48] Point estimates are shown in Appendix Table 1.A3.

There may be concerns about whether it is the pre-conditions that are required for a sub-Saharan African country to become AGOA eligible that is bringing about the change or it is implementation of AGOA and the changes in government policies thereafter which is helping in reducing infant mortality. This has been addressed in Figure 1.4. If it were the pre-conditions that were making infant mortality fall then we would have seen the drop even before AGOA was implemented; which is not the case. As an additional check I divide the group of countries into two – one who got AGOA status in 2001 and in the other group those who got later. It may be argued that countries who got the AGOA status in 2001, were already "ready" while those which got later, needed to work on pre-requisites to get themselves an AGOA Beneficiary status. Hence, the estimates for the latter group should be bigger and significant if in fact it is the preconditions which lead to a fall in infant deaths. Table 1.15 (1) shows the results. It is observed that the group of countries which got AGOA status later does not significantly do better than those countries which got their AGOA status earlier in 2001 and in fact the group of countries which got the status earlier is more effective in decreasing infant mortality.

I run additional robustness tests to include different cut offs for dropping the sample based on various year of birth of kids in Table 1.15 (2). The result is robust to using different years as cut offs on both ends. There may exist bunching of deaths at 12 months in the data in DHS. To account for this, I redefine the infant mortality variable to include children who died at 12 months as well. The results are presented in Table 1.15 (4). This does not change the effect of AGOA on infant mortality and the effect is still statistically significant with a fall in 0.8 percentage points in infant mortality.

Lastly, I re-estimate the models using different definitions for treatment to AGOA. In my first specification in Table 1.15 (3), a model is estimated where in instead of choosing an indicator variable to indicate the presence of AGOA policy, percentage change in trade volumes interacted with an indicator of the country becoming AGOA eligible, is used as the independent variable. Table 1.15 (3) indicates that the coefficient is still negative and significant. A 1% increase in percentage in trade volumes in AGOA affected countries decreases the probability of an infant dying by 0.0014 percentage points. With year-to-year increases about 100% for some countries, this means it decreases the probability of infant dying by 0.14 percentage points. But, as pointed out earlier, it should be kept in mind that trade volumes in fact embody different aspects of an economy which may correlate with infant mortality and hence may not provide the best estimate.

It could be argued that countries which got treated with AGOA earlier in 2001 are different from countries which got AGOA later and hence a treatment variable is defined such that it takes the intensity of treatment into account. In Table 1.15 (3), I redefine the treatment variable such that it differs by number of years exposed to AGOA. The countries which got AGOA earlier in 2001 get a value 2 for treatment while those which got AGOA after 2001, get a value of 1. The never treated countries in the sample get a value of zero. The resulting coefficient remains statistically significant at about 0.4 percentage points fall in infant deaths. This fall in magnitude could be due to the heterogeneity in fall in infant mortality for those countries which got AGOA in 2001 vis-à-vis later.

## 1.8 Conclusion

The empirical study of the effect of trade on development has been limited, and in many cases confounding. By creating a micro panel dataset for 30 sub-Saharan African countries in DHS, this study is able to better control for confounding factors at the country and mother level like poor institutions, macroeconomic instability, geography etc. and hence is able to derive a causal estimate of trade on infant mortality. The reduced-form results indicate trade policy has a positive developmental effect on the population in terms of reducing probability of infant and neonatal deaths. It should be noted that this does not imply that the trade policy leads to overall decrease in infant mortality rates; rather it changes the household experience of child death.

I also find evidence of differential benefits to different sections of the population. AGOA reduces infant death significantly for the uneducated and rural mothers. This may be happening because uneducated rural mothers provide cheap labor which is employed with the job creation that comes with trade openness. In this sense, trade closes the gap between the groups. At the country level, countries with predominantly agricultural exports benefit more than others and so do low-income countries. Some oil and gas exporting countries see an increase in infant deaths after the policy.

The fall in infant deaths accrues via a change in response to AGOA by mothers by increasing maternal labor supply in non-agricultural sectors and increased health seeking behavior. The results also indicate the presence of income channel in decreasing infant deaths as AGOA decreases the probability of being “poor”. The decrease in infant and

neonatal deaths is observed immediately after AGOA is implemented. The increased health seeking behavior of mother could bring about immediate changes in neonatal mortality. The improved health seeking behavior of the women could be attributed to increase in relative bargaining power for women due to increased income (Aguayo-Tellez et al., 2010) or realignment of preferences due to changing environment (Paxson and Schady, 2005).

This study emphasizes that macroeconomic policies could have a positive causal effect on microeconomic development outcomes like health. With such a large policy change, which affects so many countries, it is difficult to simulate this experience through a randomized control trial. Even in a non-randomized setting, this analysis derives a clean estimate of the effect of trade policy on health outcomes and behavior using the variation of implementation in AGOA. It is important to develop a macro-micro synthesis and study the relationship between health and macroeconomic outcomes as this underdeveloped route will open up new channels of effective policy intervention which would help harness all the benefits that any macroeconomic policy may have on society's welfare.

## References


Acemoglu, Daron, Simon Johnson, and James A Robinson. 2009. "The Colonial Origins of Comparative Development : An Empirical Investigation." *American Economic Review* 91 (5): 1369–1401.

Adam, Christopher S., David L. Bevan, and Gerard Chambas. 2001. "Exchange Rate Regimes and Revenue Performance in Sub-Saharan Africa." *Journal of Development Economics* 64 (1). Elsevier: 173–213.

Agbeyegbe, Terence D., Janet Stotsky, and Asegedech WoldeMariam. 2006. "Trade Liberalization, Exchange Rate Changes, and Tax Revenue in Sub-Saharan Africa." *Journal of Asian Economics* 17 (2): 261–84. doi:10.1016/j.asieco.2005.09.003.

Aguayo-Tellez, Ernesto, Jim Airola, and Chinhui Juhn. 2010. "Did Trade Liberalization Help Women? The Case of Mexico in the 1990s." *NBER Working Paper No. 16195*, July.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian. 2012. "On the Road: Access to Transportation Infrastructure and Economic Growth in China." *NBER Working Paper No. 17897*, March.

Barham, Tania. 2011. "A Healthier Start: The Effect of Conditional Cash Transfers on Neonatal and Infant Mortality in Rural Mexico." *Journal of Development Economics* 94 (1): 74–85. doi:10.1016/j.jdeveco.2010.01.003.

Barro, Robert J., and Jong Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950–2010." *Journal of Development Economics* 104 (September): 184–98. doi:10.1016/j.jdeveco.2012.10.001.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics* 119 (1). MIT Press: 249–75.

Besley, Timothy, and Masayuki Kudamatsu. 2009. "Health and Democracy." *The American Economic Review* 96 (2). American Economic Association: 313–18.

Bhalotra, Sonia. 2010. "Fatal Fluctuations? Cyclicality in Infant Mortality in India." *Journal of Development Economics* 93 (1): 7–19. doi:10.1016/j.jdeveco.2009.03.006.

Black, Robert E, Saul S Morris, and Jennifer Bryce. 2003. “Where and Why Are 10 Million Children Dying Every Year?” *Lancet* 361 (9376): 2226–34. doi:10.1016/S0140-6736(03)13779-8.

Broda, Christian, and David E. Weinstein. 2006. “Globalization and the Gains from Variety.” *The Quarterly Journal of Economics* 121 (2). MIT Press: 541–85.

Cameron, Colin A., and Douglas L. Miller. 2013. “A Practitioner’s Guide to Cluster-Robust Inference.” *Journal of Human Resources, Forthcoming*.

Case, Anne. 2002. “Health, Income, and Economic Development.” In *Annual World Bank Conference on Development Economics*, 221–41. World Bank Publications.

Cheibub, José Antonio, Jennifer Gandhi, and James Raymond Vreeland. 2009. “Democracy and Dictatorship Revisited.” *Public Choice* 143 (1-2): 67–101. doi:10.1007/s11127-009-9491-2.

Chou, Shin-Yi, Jin-Tan Liu, Michael Grossman, and Theodore J Joyce. 2007. “Parental Education and Child Health: Evidence from a Natural Experiment in Taiwan.” *NBER Working Paper No. 13466*, October.

Clark, Ximena, David Dollar, and Alejandro Micco. 2004. “Port Efficiency, Maritime Transport Costs, and Bilateral Trade.” *Journal of Development Economics* 75 (2): 417–50. doi:10.1016/j.jdeveco.2004.06.005.

Cochrane, Susan H., Joanne Leslie, and Donald J. O’Hara. 1982. “Parental Education and Child Health: Intracountry Evidence.” *Health Policy and Education* 2 (3-4): 213–50. doi:10.1016/0165-2281(82)90011-X.

Collier, Paul, and Anthony J. Venables. 2007. “Rethinking Trade Preferences: How Africa Can Diversify Its Exports.” *The World Economy* 30 (8): 1326–45. doi:10.1111/j.1467-9701.2007.01042.x.

Condon, Niall, and Matthew Stern. 2011. “The Effectiveness of African Growth and Opportunity Act (AGOA) in Increasing Trade from Least Developed Countries: A Systematic Review.” *EPPI- Centre, Social Science Research Unit, Institute of Education, University of London*.

Cutler, David, Angus Deaton, and Adriana Lleras-Muney. 2006. “The Determinants of Mortality.” *Journal of Economic Perspectives* 20 (3). American Economic Association: 97–120.

Deaton, Angus. 1999. “Commodity Prices and Growth in Africa.” *Journal of Economic Perspectives* 13 (3). American Economic Association: 23–40.

Desai, Sonalde, and Soumya Alva. 1998. “Maternal Education and Child Health: Is There a Strong Causal Relationship?” *Demography* 35 (1): 71. doi:10.2307/3004028.

Dollar, David. 1992. “Outward-Oriented Developing Countries Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976-85.” *Economic Development and Cultural Change* 40 (3): 523–44.

Dollar, David, and Aart Kraay. 2002. “Growth Is Good for the Poor.” *Journal of Economic Growth* 7 (3). Kluwer Academic Publishers: 195–225. doi:10.1023/A:1020139631000.

Donald, Stephen G., and Kevin Lang. 2007. “Inference with Difference-in-Differences and Other Panel Data.” *The Review of Economics and Statistics* 89 (2). MIT Press: 221–33.

Donaldson, Dave, and Richard Hornbeck. 2013. “Railroads and American Economic Growth: A ‘Market Access’ Approach.” *NBER Working Paper No. 19213*, July.

Edmonds, Eric V., Nina Pavcnik, and Petia Topalova. 2010. “Trade Adjustment and Human Capital Investments: Evidence from Indian Tariff Reform.” *American Economic Journal: Applied Economics* 2 (4). American Economic Association: 42–75.

Feenstra, Robert C, Robert Inklaar, and Marcel P Timmer. 2013. *The Next Generation of the Penn World Table, PWT 8.0*.

Frankel, Jeffrey A, and David Romer. 1999. “Does Trade Cause Growth?” *The American Economic Review* 89 (3). American Economic Association: 379–99. doi:10.2307/117025.

Frazer, Garth, and Johannes Van Biesebroeck. 2010. “Trade Growth under the African Growth and Opportunity Act.” *Review of Economics and Statistics* 92 (1). The MIT Press: 128–44. doi:10.1162/rest.2009.12111.

Hanson, Gordon H. 2007. “Globalization, Labor Income, and Poverty in Mexico.” *University of Chicago Press and NBER*, March. University of Chicago Press, 417–56.

Heston, Alan, Robert Summers, and Bettina Aten. 2011. *Penn World Table Version 7.0*.

Hummels, David, and Alexandre Skiba. 2004. “Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture.” *Journal of Political Economy* 112 (6). University of Chicago Press: 1384–1402.

Jones, Gareth, Richard W Steketee, Robert E Black, Zulfiqar A Bhutta, and Saul S Morris. 2003. "How Many Child Deaths Can We Prevent This Year?" *Lancet* 362 (9377): 65–71. doi:10.1016/S0140-6736(03)13811-1.

Kimenyi, Mwangi S. 2009. "African Growth and Opportunity Act: A Case of Vanishing Benefits." *Africa Growth Initiative, Opinion Piece, The Brookings Institution*.

Kishor, Sunita, and Sulabha Parasuraman. 1998. "Mother's Employment and Infant and Child Mortality in India". Mumbai, India: International Institute for Population Sciences and Calverton, Maryland: Macro International Inc.

Kudamatsu, Masayuki. 2012. "Has Democratization Reduced Infant Mortality in Sub-Saharan Africa? Evidence from Micro Data." *Journal of the European Economic Association* 10 (6): 1294–1317. doi:10.1111/j.1542-4774.2012.01092.x.

Levine, David I, and Dov Rothman. 2006. "Does Trade Affect Child Health?" *Journal of Health Economics* 25 (3): 538–54. doi:10.1016/j.jhealeco.2005.10.001.

Levine, Ross, and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions." *The American Economic Review* 82 (4). American Economic Association: 942–63. doi:10.2307/2117352.

Lindahl, Mikael, and Alan B. Krueger. 2001. "Education for Growth: Why and for Whom?" *Journal of Economic Literature* 39 (4). American Economic Association: 1101–36.

Lofgren, Hans. 2000. "Trade Reform and the Poor in Morocco: A Rural – Urban General Equilibrium Analysis of Reduced Protection." In *Earnings Inequality, Unemployment, and Poverty in the Middle East and North Africa*. Greenwood Publishing Group.

Mattoo, Aaditya, Devesh Roy, and Arvind Subramanian. 2003. "The Africa Growth and Opportunity Act and Its Rules of Origin: Generosity Undermined?" *The World Economy* 26 (6): 829–51. doi:10.1111/1467-9701.00550.

Nicita, Alessandro. 2004. "Who Benefited from Trade Liberalization in Mexico?: Measuring the Effects on Household Welfare." *World Bank Policy Research Working Paper 3265*. World Bank Publications, 55.

Onyango, Christopher, and Moses Ikiara. 2011. "Reflections on Kenya's Experience Under AGOA: Opportunities and Challenges." In *Improving AGOA: Toward a New Framework for US-Africa Commercial Engagement*. African Growth Initiative at Brookings. The Brookings Institution, Washington D.C.

Páez, Laura, Stephen Karingi, Mwangi Kimenyi, and Mekalia Paulos. 2010. "A Decade (2000-2010) of African-U.S. Trade under the African Growth Opportunities Act (AGOA): Challenges, Opportunities and a Framework for Post AGOA Engagement." *African Economic Conference*.

Paxson, Christina, and Norbert Schady. 2005. "Child Health and Economic Crisis in Peru." *The World Bank Economic Review* 19 (2): 203–23. doi:10.1093/wber/lhi011.

Porto, Guido. 2004. "Trade Reforms, Market Access and Poverty in Argentina." *World Bank Policy Research Working Paper No. 3135*.

Przeworski, Adam, Michael E. Alvarez, Jose A. Cheibub, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*. Cambridge University Press.

Rodriguez, Francisco, and Dani Rodrik. 2001. "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." In *NBER Macroeconomics Annual 2000*, 261–338. MIT Press.

Rodrik, Dani. 2001. "The Global Governance of Trade: As If Development Really Mattered." *United Nations Development Programme Report*, no. July.

Schultz, T Paul. 1993. "Mortality Decline in the Low-Income World: Causes and Consequences." *The American Economic Review* 83 (2). American Economic Association: 337–42. doi:10.2307/2117687.

Seaman, L. C. B. 2002. *Victorian England: Aspects of English and Imperial History 1837-1901*. Vol. 1. Routledge.

Standing, Guy. 1999. "Globalization through Flexible Labor: A Theme Revisited." *World Development1* 27 (3).

Storeygard, Adam. 2013. "Farther on Down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa." *World Bank Working Paper 6444*, May.

The World Bank. 2012. *World Development Indicators*.

Thompson, Carol B. 2004. "US Trade with Africa: African Growth & Opportunity?" *Review of African Political Economy* 31 (101). Routledge: 457–74. doi:10.1080/0305624042000295549.

Topalova, Petia. 2007. "Trade Liberalization, Poverty and Inequality: Evidence from Indian Districts." *NBER Working Paper w11614*, March. University of Chicago Press, 291–336.

———. 2010. “Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India.” *American Economic Journal: Applied Economics* 2 (4). American Economic Association: 1–41. doi:10.1257/app.2.4.1.

Winters, L Alan, Neil McCulloch, and Andrew McKay. 2004. “Trade Liberalization and Poverty: The Evidence so Far.” *Journal of Economic Literature* 42 (1). American Economic Association: 72–115. doi:10.2307/3217037.

Wood, A. 2001. “Africa’s Export Structure in a Comparative Perspective.” *Cambridge Journal of Economics* 25 (3): 369–94. doi:10.1093/cje/25.3.369.

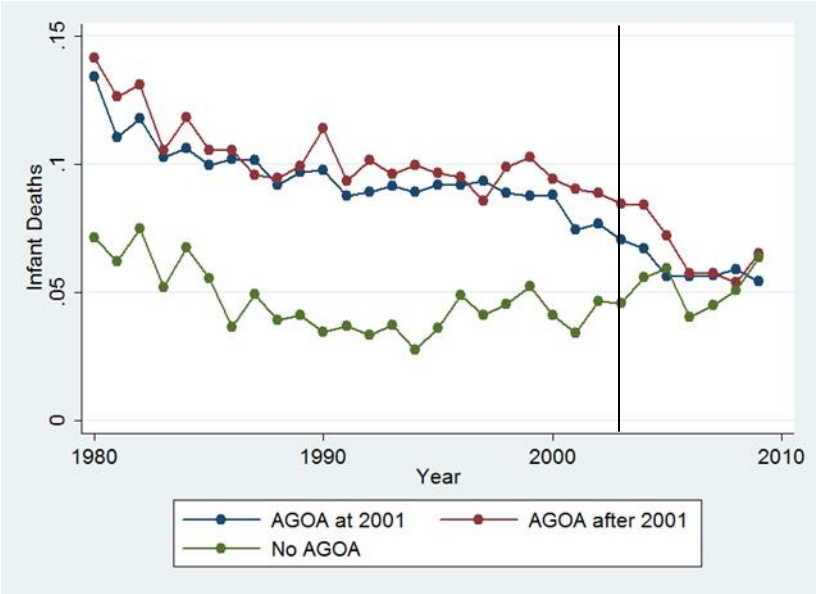
Wood, Adrian. 1991. “North-South Trade and Female Labour in Manufacturing: An Asymmetry.” *Journal of Development Studies* 27 (2). Routledge: 168–89. doi:10.1080/00220389108422191.

Younger, Stephen D. 2001. “Cross-Country Determinants of Declines in Infant Mortality: A Growth Regression Approach.” *Cornell University, Food and Nutrition Policy Program*.

Zakir, Mohammed, and Phanindra V. Wunnava. 1999. “Factors Affecting Infant Mortality Rates: Evidence from Cross–sectional Data.” *Applied Economics Letters* 6 (5). Routledge: 271–73. doi:10.1080/135048599353203.

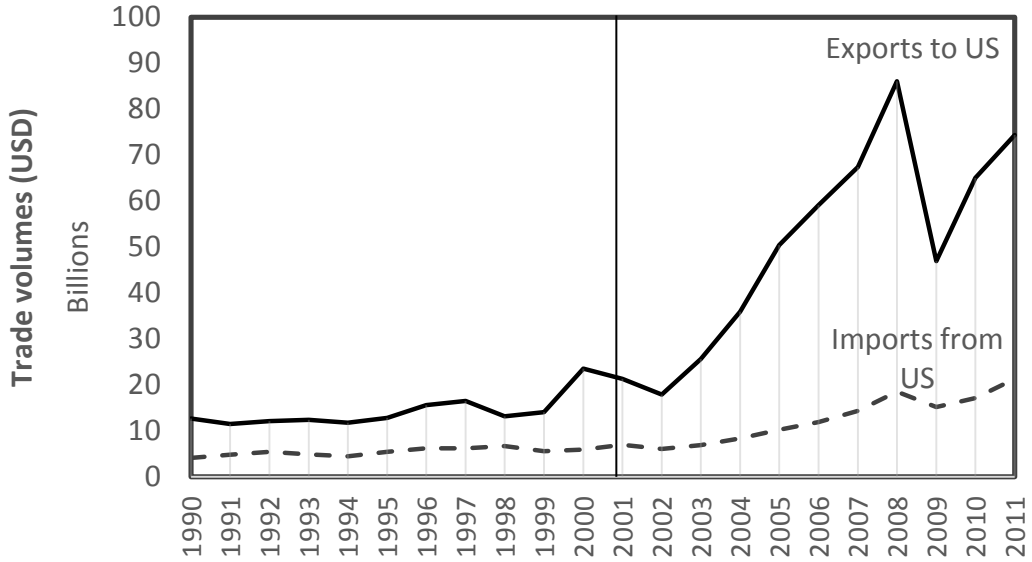
# Figures

## Figure 1.1: Infant Deaths for AGOA and not-AGOA eligible countries



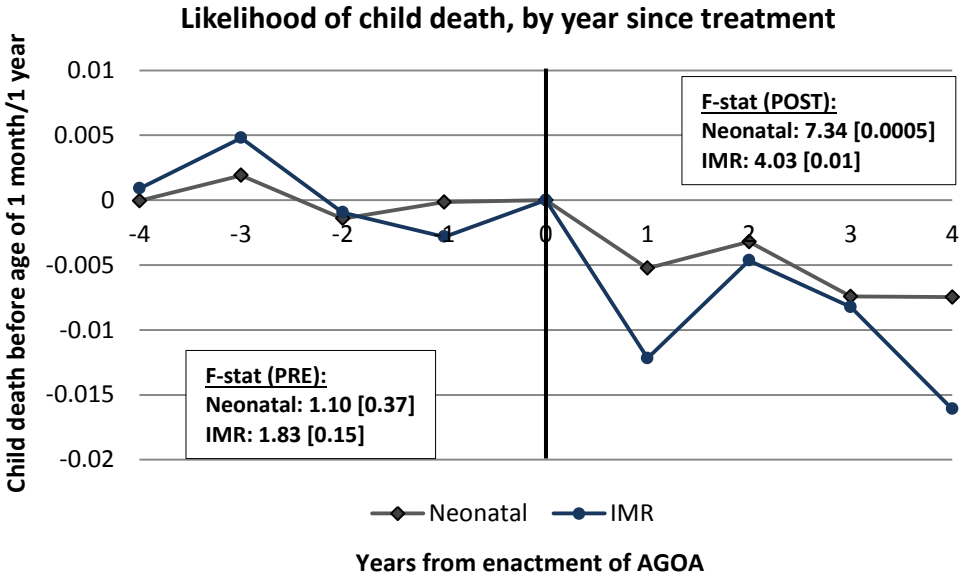

Note: This graph plots the sample infant mortality rates for countries affected by AGOA by 2001, countries affected by AGOA after 2001 and never affected by AGOA countries, by year of child birth. The countries affected by AGOA by 2001 are – Benin, Cameroon, Chad, Congo, Ethiopia, Ghana, Guinea, Kenya, Lesotho, Madagascar, Malawi, Mali, Mozambique, Namibia, Niger, Nigeria, Rwanda, Sao Tome and Principe, Senegal, Swaziland, Tanzania, and Zambia. Countries which received AGOA benefits after 2001 are – Angola, Burkina Faso, Burundi, Cote D'Ivoire, Democratic Republic of Congo, Liberia and Sierra Leone. Zimbabwe has never been an AGOA beneficiary.

**Figure 1.2: Trade Volumes between US and sub-Saharan Africa**



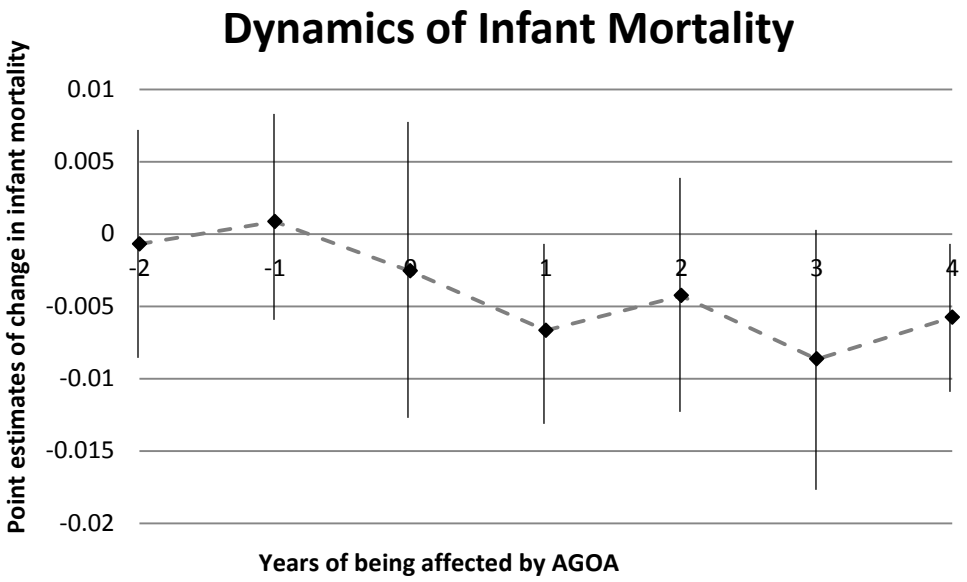

Note: This graph has been plotted using the data from International Trade Administration, U.S. Department of Commerce. It depicts the total exports and imports between US and all the sub-Saharan African countries from 1990-2011. The solid black line represents the imports into US from sub-Saharan Africa while the dotted line represents the exports from US to sub-Saharan Africa. It is observed that both exports and imports from sub-Saharan Africa increase dramatically after 2001. A more distinct increase in exports from sub-Saharan Africa to US is observed.

**Figure 1.3: Event-Time Study**



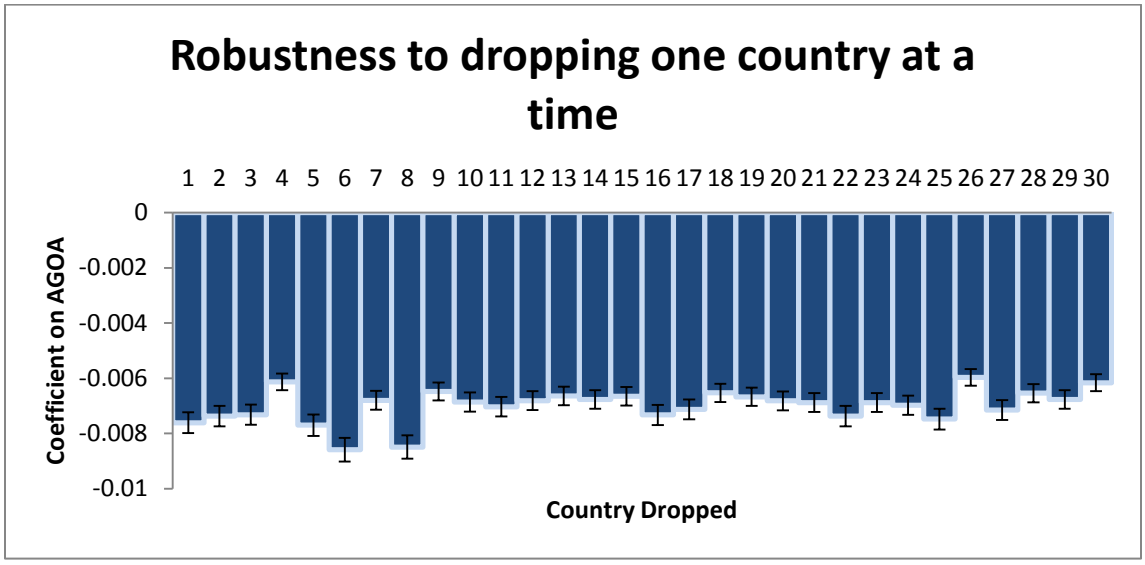

Note: These are the θj estimates plotted from estimating this equation:
$Death_{imct} = \alpha_c + \beta_t + \sum_{j=-4}^{4} \theta_j T_{c,t+j} + X_i'\delta + \varepsilon_{imct}$
Death takes the value of infant mortality of neonatal mortality. The sample is restricted to treated countries. The solid line at zero indicates the year of announcement of AGOA. The control variables are whether born in multiple birth, birth order, birth month, mother's age at birth, mother's education, place of residence, asset index. Both the specifications control for year and country fixed effects.

**Figure 1.4: Dynamics of Infant Mortality**



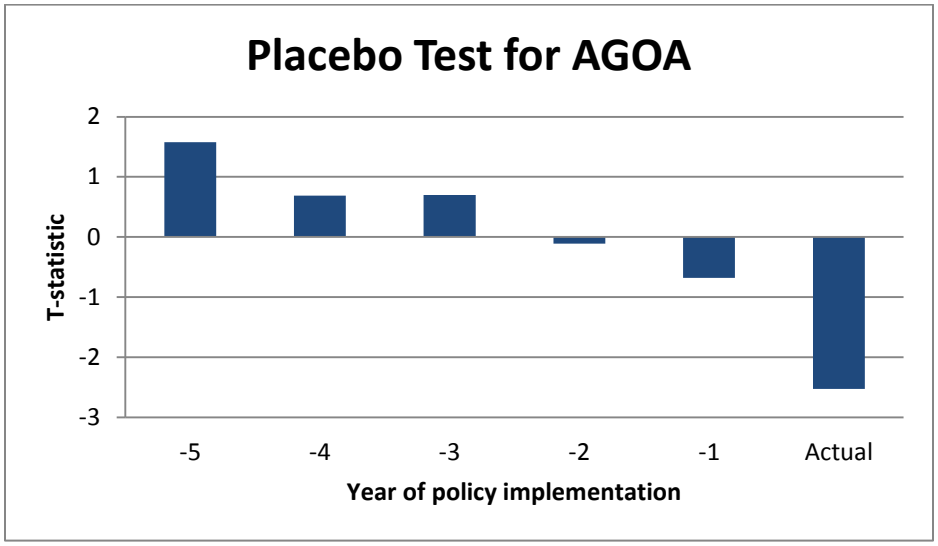

Note: The solid black line depicts the change in infant mortality compared to 3 years before implementation of AGOA controlling for mother fixed effects, cohort-year fixed effects, country specific linear trends, sex of child, whether born in multiple birth, birth order and birth month. Year 0 is the year of implementation of AGOA, such that the countries have been at least partially affected by AGOA in that year. The dotted lines represent the 95% confidence interval with standard errors clustered at the country level. The point estimates for 1, 3 and 4 years after being affected by AGOA are significant at least at 90% significance level. The point estimates are provided in Appendix Table 1.A2.

**Figure 1.5: Robustness Check – Dropping one country at a time**




Note: In each of the separate regressions, one of the countries is dropped at a time in alphabetical order. The graph depicts the point estimates of the effect of treatment on infant mortality controlling for sex of child, whether born in multiple birth, birth order, birth month, mother fixed effects, country specific linear trends, mother's cohort by child birth year FE. Standard errors are clustered at country level. All the estimates are significant at least at 5% significance level. The point estimates are provided in Appendix Table 1.A3.

**Figure 1.6: Placebo Test – t-statistics for effect of policy change**




Note: In each of the separate regressions, the effect of AGOA is estimated at false policy timings. The graph depicts the t-statistic of the effect of "policy change" on infant mortality controlling for sex of child, whether born in multiple birth, birth order, birth month, mother's age, year of birth fixed effects, mother fixed effects, and country specific linear trends. None of the years before the actual policy change give statistically significant results. The point estimates are provided in Appendix Table 1.A6.

# Tables

## Table 1.1: Summary Statistics – Child Variables

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | **All** | **Before AGOA** | **After AGOA** | **Non-AGOA** | **T-test** |
| *Child Variables* | | | | | |
| **Infant Mortality** | 0.0815 | 0.089 | 0.065 | 0.089 | 9.57 |
| *Uneducated* | 0.0939 | 0.1016 | 0.0719 | 0.104 | 8.34 |
| *N* | *342382* | *201754* | *93566* | *47062* | |
| *Educated* | 0.0691 | 0.075 | 0.059 | 0.074 | 4.26 |
| *N* | *343693* | *174889* | *124351* | *44453* | |
| *Poor* | 0.0902 | 0.101 | 0.069 | 0.098 | 6.42 |
| *N* | *300418* | *161300* | *97069* | *42049* | |
| *Non-Poor* | 0.0747 | 0.081 | 0.061 | 0.082 | 6.53 |
| *N* | *385675* | *215353* | *120856* | *49466* | |
| *Rural* | 0.0866 | 0.096 | 0.067 | 0.095 | 8.27 |
| *N* | *501284* | *272892* | *163277* | *65115* | |
| *Urban* | 0.0677 | 0.071 | 0.057 | 0.076 | 5.73 |
| *N* | *184809* | *103761* | *54648* | *26400* | |
| **Neonatal Mortality** | 0.038 | 0.041 | 0.032 | 0.040 | 3.40 |
| **Female** | 0.492 | 0.492 | 0.493 | 0.492 | -0.06 |
| **Multiple Births** | 0.035 | 0.034 | 0.037 | 0.033 | -2.45 |
| **Birth Order** | 3.47 | 3.38 | 3.64 | 3.45 | -1.96 |
| **Month of birth** | 6.15 | 6.10 | 6.23 | 6.07 | -7.03 |
| **Mother's age at birth(20-29yrs)** | 0.50 | 0.499 | 0.483 | 0.49 | -6.43 |
| **Mother's age at birth(30-39yrs)** | 0.24 | 0.23 | 0.26 | 0.24 | -1.16 |
| **Mother's age at birth(40-49yrs)** | 0.02 | 0.015 | 0.045 | 0.027 | 5.88 |
| **N** | 686093 | 468168 | 217925 | 91515 | |

Note: Sample means of all child level variables are reported. Column (1) is for the whole sample with AGOA affected and non-affected countries. Columns (2) and (3) report the sample mean infant mortality before and after the implementation of AGOA in AGOA affected countries. Column (4) reports the sample mean in non-AGOA countries. Column (5) gives the t-statistic testing if the means are significantly different between AGOA and non-AGOA countries. N refers to the number of observations in each sample. Educated implies having attended any type of school and uneducated is defined as mother did not attend any school. Poor is defined by a wealth index as defined as poor or poorer vis-à-vis with mothers who are non-poor based on the wealth index being middle, richer or richest. Rural and Urban are defined by the place of residence of mother during the time of interview. Female is 1 if sex of child is female. Multiple birth is a dummy variable indicating if the child is born in a multiple birth. It is 0 for a single birth and 1 for twins, triplets or quadruplets.

**Table 1.2: Mother Characteristics – Full Sample and 2+ Mothers**

| | (1) | | (2) | | (3) | | (4) |
|---|---|---|---|---|---|---|---|
| | **Full Sample** | | **AGOA Countries All mothers** | | **2+ Sample both before and after AGOA** | | **T-test** |
| | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** | |
| **Mother's age at birth** | 25.72 | 6.44 | 25.72 | 6.42 | 26.03 | 6.39 | -23.7 |
| **Mother's Education** | 0.501 | 0.499 | 0.503 | 0.499 | 0.478 | 0.499 | 24.78 |
| **Mother's wealth index** | 2.86 | 1.402 | 2.86 | 1.39 | 2.77 | 1.38 | 35.92 |
| **Mother's Residence (Rural)** | 0.73 | 0.444 | 0.73 | 0.442 | 0.77 | 0.422 | -37.29 |
| **N** | 686093 | | 594578 | | 391425 | | |
| **N (M.Educ)** | 686075 | | 594560 | | 391414 | | |

Note: Sample means and standard deviations are reported for different samples of mothers. N refers to the number of observations in each sample. Column (1) gives the mean and standard deviation for different mother characteristics for the whole sample with AGOA affected and non-affected countries. Column (2) reports the same for all mothers in AGOA affected countries. Column (3) reports the sample mean and standard deviation for mothers with two or more children giving birth before and after AGOA. All variables are categorical variables except mother's age at birth. Column (4) provides a difference in means t-test between (2) and (3).

**Table 1.3: Mean Infant and Neonatal Mortality for Sample of 2+ Mothers in AGOA Countries**

| | (1) | | (2) | |
|---|---|---|---|---|
| | **Both before and after AGOA** | | **Only before or after AGOA** | |
| | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** |
| **Infant Mortality** | | | | |
| Before AGOA | 0.090 | 0.286 | 0.091 | 0.286 |
| After AGOA | 0.063 | 0.243 | 0.077 | 0.267 |
| **Neonatal Mortality** | | | | |
| Before AGOA | 0.041 | 0.198 | 0.0436 | 0.202 |
| After AGOA | 0.029 | 0.168 | 0.0425 | 0.187 |
| **N** | 391425 | | 165098 | |
| Before AGOA | 247784 | | 117811 | |
| After AGOA | 143641 | | 47287 | |

Note: Sample mean is reported in the top row and number of live birth observations for AGOA affected countries in the bottom row. Column (1) gives the sample mean and standard deviation for infant and neonatal mortality for the sample of mothers giving birth both before and after AGOA. Column (2) reports the sample mean and standard deviation for mothers with two or more children either only before AGOA or after AGOA. N represents the number of live births.

## Table 1.4: Effect of AGOA treatment on infant and neonatal mortality

| | Specification 1 | | | Specification 2 | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| **Dependent Variable** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Neonatal Mortality** |
| **Treatment** | -0.0071 | -0.0081*** | -0.0071** | -0.0079*** | -0.0079*** | -0.0069** | -0.0046*** |
| | (0.0028) | (0.0028) | (0.0028) | (0.0019) | (0.0028) | (0.0027) | (0.0011) |
| **Explanatory Variables** | YES | YES | YES | YES | YES | YES | YES |
| **Country time trend** | NO | YES | YES | YES | YES | YES | YES |
| **Country FE** | YES | YES | NO | NO | YES | NO | NO |
| **Mother FE** | NO | NO | YES | NO | NO | YES | YES |
| **Cohort-year FE** | NO | NO | NO | YES | YES | YES | YES |
| **Number of countries** | 30 | 30 | 30 | 30 | 30 | 30 | 30 |
| **Number of mothers** | 212738 | 212738 | 212738 | 212738 | 212738 | 212738 | 212738 |
| **Observations** | 686093 | 686093 | 686093 | 686093 | 686093 | 686093 | 686093 |

Note: Treatment is defined as 1 for a child born after AGOA has been implemented in an AGOA affected country. The other control variables included in the specifications are sex of child, whether born in multiple birth, year of birth, mother's age at birth, birth order and birth month. Standard errors clustered at the country level are reported in brackets. Specification 2 allows for changing mother's age at birth for different year of birth of child. Hence, controls for mother's age and year of birth of child are subsumed in these specifications.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 1.5: Country-level time varying variables

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| **Dependent Variable** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** |
| **Treatment** | -0.0068** | -0.0076*** | -0.0082*** | -0.0069** | -0.0067* | -0.009*** | -0.0066** |
| | (0.0025) | (0.0026) | (0.0025) | (0.0026) | (0.0032) | (0.0025) | (0.0028) |
| **Log GDP per capita** | -0.0099* (0.0054) | | | | | | -0.0175* (0.0094) |
| **Democracy** | | -0.0041 (0.0029) | | | | | -0.0043 (0.0028) |
| **ODA** | | | 0.00009 (0.0001) | | | | -0.00003 (0.00007) |
| **Openness** | | | | -0.00002 (0.00007) | | | 0.00009 (0.00005) |
| **Female Education** | | | | | 0.0029 (0.0053) | | -0.001 (0.0048) |
| **Commodity Price Index** | | | | | | 0.0327*** (0.0067) | 0.0311*** (0.0066) |
| **Number of countries** | 30 | 30 | 29 | 30 | 21 | 29 | 21 |
| **Number of mothers** | 212738 | 209721 | 205420 | 212738 | 134952 | 206137 | 131959 |
| **Observations** | 686093 | 673646 | 655443 | 686093 | 410833 | 663838 | 394715 |

Note: The regressions control for sex of child, whether born in multiple birth, birth order, birth month, mother fixed effects, country specific linear trends, mother's cohort by child birth year FE. Standard errors clustered at country level are reported in brackets. Data for (1) and (4) taken from PWT 7.0, (2) is taken from Democracy-Dictatorship (DD) Data by Cheibub et al (2010), (3) Net Official Development Assistance received as a % of GNI is taken from World Bank Indicators, (5) from Barro and Lee (2010), and (6) from PWT 8.0. Number of observations and number of mothers varies depending on availability of country level control variable from different data sources.
*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

**Table 1.6: Heterogeneity across different types of mothers**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Dependent Variable** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** |
| **Educated** | -0.0054*<br>(0.0031) | | | |
| **Uneducated** | -0.0082***<br>(0.0029) | | | |
| **Rural** | | -0.0085***<br>(0.0028) | | |
| **Urban** | | -0.0018<br>(0.0031) | | |
| **Poor** | | | -0.0102***<br>(0.0028) | |
| **Non-Poor** | | | -0.0044<br>(0.0029) | |
| **Employed** | | | | -0.0095***<br>(0.0028) |
| **Unemployed** | | | | -0.0057<br>(0.0038) |
| **F-Stat** | 0.83<br>(0.371) | 5.71<br>(0.021) | 7.82<br>(0.009) | 2.25<br>(0.145) |
| **Number of Countries** | 30 | 30 | 30 | 28 |
| **Number of mothers** | 212732 | 212738 | 212738 | 197632 |
| **Observations** | 686075 | 686093 | 686093 | 632951 |

Note: The control variables are sex of child, whether born in multiple birth, birth order, birth month, mother fixed effects, country specific linear trends, mother's cohort by child birth year FE. Standard errors clustered at country level are reported in brackets. The treatment is interacted with the type/characteristic of mothers to get the treatment effect on those types of mothers vis-à-vis all mothers of control group. Column (1) includes the effect on infant mortality for educated mothers where educated implies having attended any type of school and uneducated mothers, where uneducated is defined as mother did not attend any school. Column (2) assesses this heterogeneity between women living in rural areas and urban areas at the time of interview. Column (3) has effect on infant mortality for mothers having a wealth index as defined as poor or poorer vis-à-vis with mothers who are non-poor based on the wealth index being middle, richer or richest. Column (4) specifies the effect of infant mortality for mothers who are employed, where employment has been categorized into 9 categories – Professional and managerial, clerical, sales, Agricultural self-employed, Agricultural employee, household and domestic, services, skilled manual and unskilled manual. Unemployed is defined for a mother who is not working. Data for employment status is not available for mothers in Angola and Nigeria in the DHS survey used. F-stat and the corresponding p-values for equality of coefficients are also reported.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

**Table 1.7: Heterogeneity across different employment groupings for mothers**

| | Treat | Agriculture | Manual Labor | Managerial Services | Household and Services |
|---|---|---|---|---|---|
| **Infant Mortality** | 0.0063<br>(0.0040) | -0.0185***<br>(0.0035) | -0.0155***<br>(0.0043) | -0.0081***<br>(0.0026) | -0.0022<br>(0.0061) |
| **F-Stat** | 3.16<br>(0.041) | | | | |
| **Number of Countries** | 28 | | | | |
| **Number of mothers** | 148006 | | | | |
| **Observations** | 484754 | | | | |

Note: The control variables are sex of child, whether born in multiple birth, birth order, birth month, mother fixed effects, country specific linear trends, mother's cohort by child birth year FE. All the coefficients are derived from the same regression. Robust standard errors clustered at country level are reported in brackets. Employment is categorized into four major sectors: (1) Agriculture - if the mother is working either as Agricultural self-employed or Agricultural employee, (2) Manual Labor - if the mother is employed as skilled manual or unskilled manual, (3) Managerial - if the mother is employed as Professional and managerial, clerical or sales, and (4) Household and services - if the mother is working in household or domestic services or the services sector. F-stat and corresponding p-value for equality of coefficients on employment categories is reported. Omitted category is the unemployed mothers. Data for employment status is not available for mothers in Angola and Nigeria in the DHS survey used.
*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 1.8: Heterogeneity across different country groupings

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Dependent Variable** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** |
| **Apparel** | -0.00023<br>(0.0046) | | | |
| **Oil** | 0.00142<br>(0.0034) | | | |
| **Agricultural Products** | -0.0132***<br>(0.0031) | | | |
| **Mineral and ore** | -0.0109**<br>(0.0048) | | | |
| **Others** | -0.00764<br>(0.0057) | | | |
| **Non-Islamic** | | -0.0038<br>(0.0026) | | |
| **Islamic** | | -0.0107***<br>(0.0034) | | |
| **Low income countries** | | | -0.0094***<br>(0.0031) | |
| **Middle income countries** | | | 0.00012<br>(0.0038) | |
| **East** | | | | -0.0181***<br>(0.0031) |
| **West** | | | | -0.0064<br>(0.0038) |
| **Central** | | | | -0.0055<br>(0.0043) |
| **South** | | | | 0.0006<br>(0.0009) |
| **F-Stat** | 4.40<br>(0.0066) | 3.30<br>(0.079) | 6.02<br>(0.0204) | 20.21<br>(0.00) |
| **Number of Countries** | 30 | 30 | 30 | 30 |
| **Number of mothers** | 212738 | 212738 | 212738 | 212738 |
| **Observations** | 686093 | 686093 | 686093 | 686093 |

Note: The regressions control for sex of child, whether born in multiple birth, birth order, birth month, mother fixed effects, country specific linear trends, mother's age and child birth year FE. Standard errors clustered at country level are reported in brackets. The treatment is interacted with the different country groupings to get the treatment effect on those groups of countries. Column (1) includes separate effect of being affected by AGOA based on their predominant commodity of export. Countries with high volume of apparel exports are Kenya, Lesotho, Madagascar, Namibia, Malawi and Swaziland. Countries having majorly oil and gas exports are Angola, Congo, Cameroon, Nigeria and Democratic Republic of Congo. Countries which had highest share of agricultural products exports are Burkina Faso, Burundi, Rwanda, and Tanzania. Countries with major mineral and ore exports (includes petrol, coal, minerals and ores) were Guinea, Ghana, Niger, Sierra Leone, and Zambia. Products not being classified under these above categories have been labeled as "other exports". These include Forestry, animal and wood products, electronics chemicals etc. The countries exporting these types of products are Benin, Chad, Mali, Sao Tome and Principe, and Senegal. The data for predominant commodity of export from these countries into US has been collected from Office of the United States Trade Representative. Column (2) assesses heterogeneity in reduction in infant mortality for countries based on their predominant religion. Data for predominant religion of each country has been collected from CIA World Factbook. Column (3) divides the 30 countries based on World Bank's ranking of incomes into low and middle income countries. Column (4) analyzes the impact on the countries based on their geographic location in sub-Saharan Africa. F-stat and the corresponding p-values for equality of coefficients are also reported.
*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

**Table 1.9: Possible Macro Pathways**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| **Dependent Variable** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** |
| **Treatment** | -0.0067** | -0.0071** | -0.0086*** | -0.0087*** | -0.0066** | -0.0086*** |
| | (0.0025) | (0.0025) | (0.0026) | (0.0026) | (0.0027) | (0.0029) |
| **Log GDP per capita** | -0.0099* (0.0054) | | | | | |
| **Health expenditure per capita** | | 0.00010** (0.00004) | | | | |
| **Paved Roads Access** | | | -0.00043** (0.00016) | | | |
| **Female LFPR** | | | | -0.00032 (0.0007) | | |
| **Inequality** | | | | | 0.000028 (0.00044) | |
| **Fertility** | | | | | | -0.00017 (0.016) |
| **Number of countries** | 30 | 29 | 29 | 29 | 27 | 29 |
| **Number of mothers** | 212738 | 194638 | 190014 | 206137 | 163946 | 206137 |
| **Observations** | 686093 | 519738 | 593076 | 663838 | 526782 | 663838 |

Note: The control variables are sex of child, whether born in multiple birth, birth order, birth month, mother fixed effects, country specific linear trends, mother's cohort by child birth year FE. Standard errors clustered at country level are reported in brackets. Data for (1) is taken from PWT 7.0, (2), (3), (4) and (5) are taken from World Bank Development Indicators. Inequality is measured by Gini Index and varies from 0 to 100 with 0 being perfect equality. Even though Gini Index numbers are available from World Bank, the data is sparse between years and therefore an interpolated Gini Index has been constructed as a measure of inequality. Number of observations and number of mothers varies depending on availability of country level control variable from different data sources.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

**Table 1.10: Micro Pathways – Health Care and Sanitation**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Dependent Variable** | Tetanus Toxoid | Delivery Assistance | Piped Water | Flush Toilets |
| **Treatment** | 0.132*** | 0.102*** | -0.069** | -0.008* |
| | (0.044) | (0.032) | (0.025) | (0.0048) |
| **Number of countries** | 22 | 22 | 22 | 22 |
| **Observations** | 118784 | 121797 | 119705 | 119657 |

**Table 1.11: Micro Pathways – Maternal Labor Force**

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Dependent Variable** | Agriculture | Manual Labor | Managerial Services | Household and Services | Not Working |
| **Treatment** | -0.149*** | 0.095** | 0.061* | -0.009 | -0.044 |
| | (0.015) | (0.037) | (0.034) | (0.019) | (0.039) |
| **Number of countries** | 22 | 22 | 22 | 22 | 22 |
| **Observations** | 74478 | 74478 | 74478 | 74478 | 122053 |

**Table 1.12: Micro Pathways –Ownership of Assets**

| | (1) | (2) | (3) | (4) | (5) | (6) | (6) | (7) |
|---|---|---|---|---|---|---|---|---|
| **Dependent Variable** | Electricity | Radio | Refrigerator | Bike | TV | Scooter | Car | Poor |
| **Treatment** | -0.055*** | 0.078*** | -0.024*** | 0.041* | -0.033** | 0.051*** | 0.001 | -0.063*** |
| | (0.014) | (0.017) | (0.007) | (0.019) | (0.013) | (0.009) | (0.005) | (0.014) |
| **Number of countries** | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 |
| **Observations** | 115771 | 119206 | 113511 | 119149 | 117053 | 117921 | 117869 | 119148 |

Note: These estimates in Table 1.10, 1.11 and 1.12 are derived from a pooled sample of mothers in multiple surveys across 22 countries. The sample includes all babies, both living and dead, born within twelve months of survey date. For details on the surveys included, refer to Appendix Table 1.A4. The models control for sex of child, whether born in multiple births, birth order, country specific linear trends, mother's age at birth, dummy for year and mother group fixed effects. Standard errors clustered at country level are reported in brackets. For definition on employment categories, see notes in Table 1.7. Description of how health care and sanitation variables are created and definition of mother group is described in Section 6.
*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 1.13: Fertility Selection Effect

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | **Fertility (All)** | **Fertility (Uneducated)** | **Fertility (Poor)** | **Fertility (Rural)** |
| **AGOA** | -1.242<br>(1.07) | -0.567<br>(0.981) | -1.261<br>(1.07) | -1.252<br>(1.135) |
| **AGOA*Woman's type** | | -0.265<br>(0.712) | 0.338<br>(0.839) | -0.101<br>(1.04) |
| **F-stat** | | 0.39<br>[0.54] | 0.5<br>[0.48] | 1.18<br>[0.29] |
| **Number of Countries** | 30 | 30 | 30 | 30 |
| **Observations** | 19250 | 38199 | 38290 | 38325 |

Note: The dependent variable is percentage of ('type' of) women giving birth. Woman's type is a dummy variable referring to if the woman is uneducated, poor or rural. For definitions of these, check notes in Table 1.1. (1) refers to all types of women, (2) to uneducated women, (3) to poor women and (4) to rural women. Standard errors clustered at the country level are reported in brackets. F-test reports F-statistics and its associated p-values in brackets for the null that the sum of coefficients on AGOA and on its interaction term with Woman's type is zero. All regressions control for country by woman's birth cohort fixed effects and year of giving birth by woman's birth cohort fixed effects which are also allowed to differ by woman's type.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

**Table 1.14: Robustness Checks**

| | (1) | (2) |
|---|---|---|
| **Dependent Variable** | **Infant Mortality** | **Infant Mortality** |
| **Specification** | Country Specific Birth Order | Country Specific mother's age quadratic trends |
| **Treatment** | -0.00688** | -0.00681** |
| | (0.0025) | (0.0026) |
| **Explanatory Variables** | YES | YES |
| **Country time trend** | YES | YES |
| **Country Specific Birth Order Dummy** | YES | YES |
| **Country specific mother's age quadratic trend** | NO | YES |
| **Mother FE** | YES | YES |
| **Cohort-year FE** | YES | YES |
| **Number of countries** | 30 | 30 |
| **Number of mothers** | 212738 | 212738 |
| **Observations** | 686093 | 686093 |

Note: The other control variables included in the specifications are sex of child, whether born in multiple birth, year of birth, mother's age at birth, birth order and birth month. Standard errors clustered at the country level are reported in brackets.
*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 1.15: Effect of AGOA on different specifications

| | (1) | | (2) | | (3) | | (4) |
|---|---|---|---|---|---|---|---|
| **Dependent Variable** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** |
| **Sample** | AGOA in 2001 | AGOA after 2001 | YOB cutoff = 1991 | 1993< YOB< 2008 | % change in trade volumes | Intensity of treatment | Including death at 12 months |
| **Treatment** | -0.0163*** | -0.0029 | -0.0070** | -0.0082*** | -0.000014** | -0.0041** | -0.0085*** |
| | (0.0039) | (0.0065) | (0.0028) | (0.0028) | (0.0027) | (0.0019) | (0.0030) |
| **Explanatory Variables** | YES | YES | YES | YES | YES | YES | YES |
| **Country time trend** | YES | YES | YES | YES | YES | YES | YES |
| **Mother FE** | YES | YES | YES | YES | YES | YES | YES |
| **Cohort Year FE** | NO | NO | NO | NO | NO | NO | NO |
| **Number of countries** | 25 | 10 | 30 | 30 | 30 | 30 | 30 |
| **Number of mothers** | 176295 | 69667 | 209970 | 197072 | 209970 | 212738 | 212738 |
| **Observations** | 559498 | 218110 | 635844 | 536137 | 635844 | 686093 | 686093 |

Note: YOB stands for year of birth. The covariates are sex of child, whether born in multiple birth, birth order, birth month, mother's age, birth year. Standard errors clustered at country level are reported in brackets. (1) includes effect of trade on infant mortality in AGOA affected countries vis-à-vis no-AGOA countries, where AGOA was implemented in 2001 in column 1 vis-à-vis those nations where AGOA was implemented after 2001. (2) includes robustness check for birth year cut off for children in sample. (3) redefines the independent variable to percentage change in trade volumes from previous year for AGOA vs non-AGOA countries and another definition of treatment, with treatment taking the value 2 for AGOA affected countries in 2001, 1 for countries getting AGOA after 2001 and 0 for not being AGOA affected in the sample. (4) redefines the dependent variable to include deaths at 12 months as well.
*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

# Appendix

**Figure 1.A1: Map of AGOA eligible and not AGOA-eligible countries**

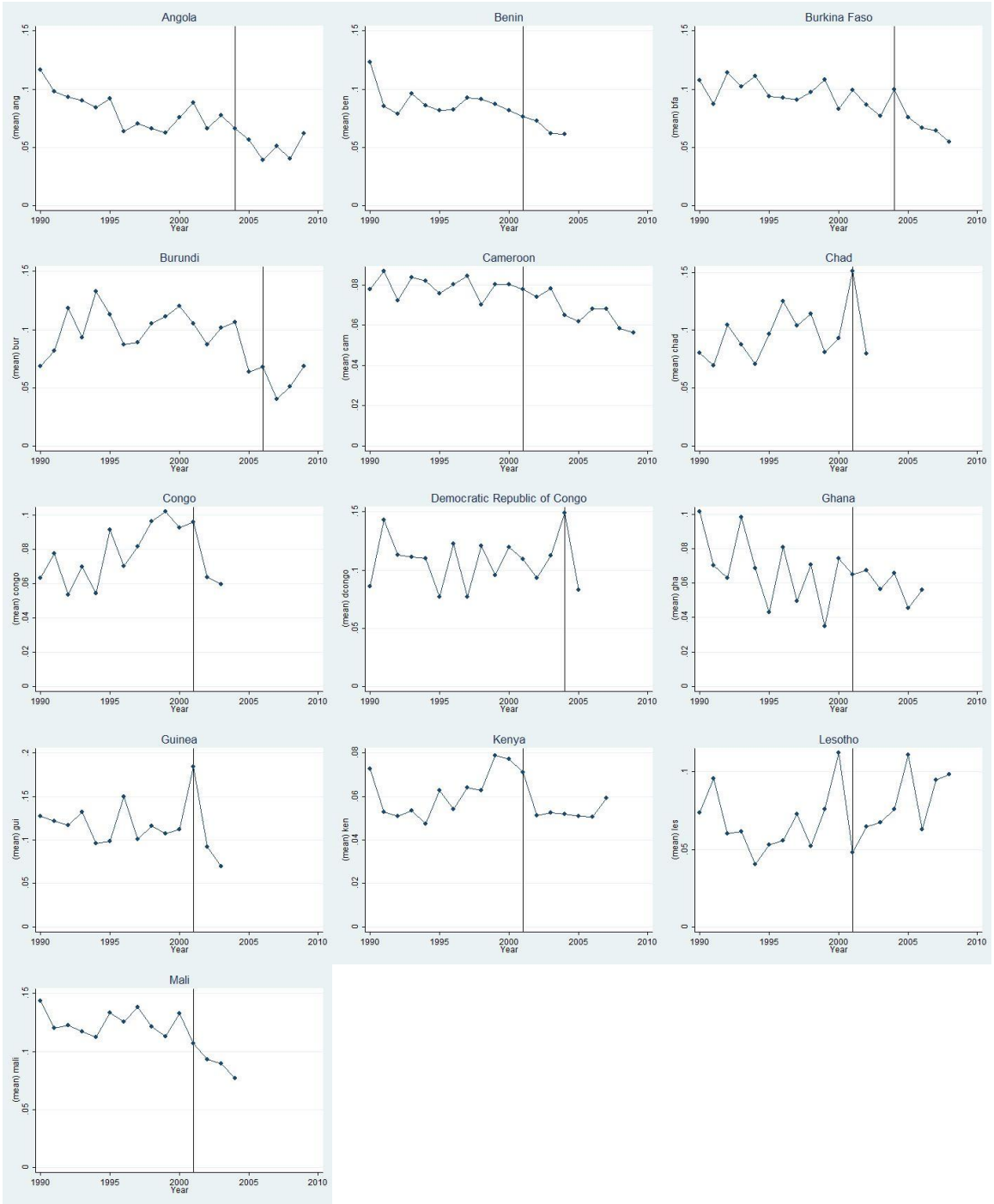



**Figure 1.A2(a): Sample mean infant mortality rates by country for AGOA affected countries overtime, 1990 onwards**

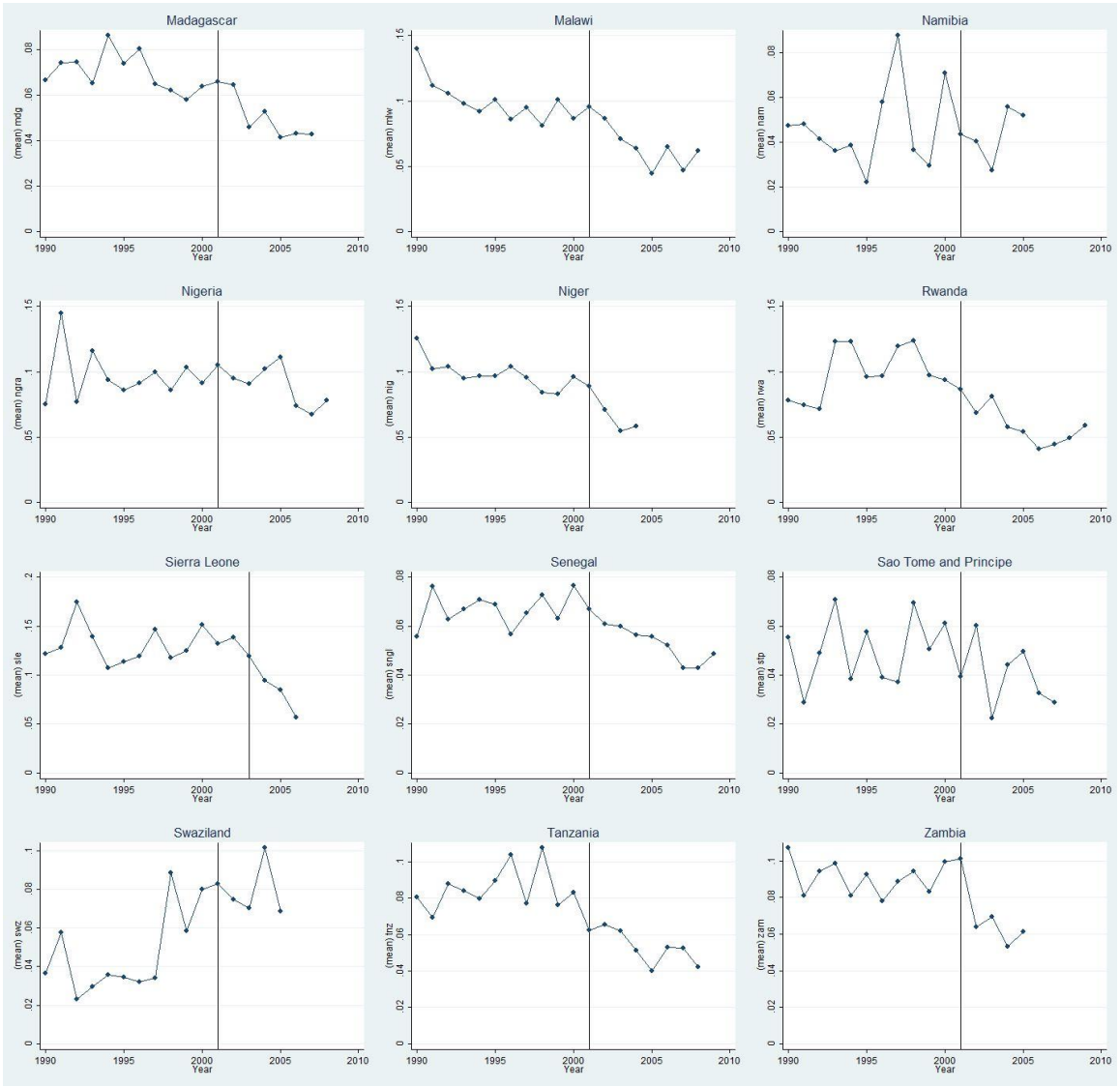



**Figure 1.A2(b): Sample mean infant mortality rates by country for AGOA affected countries overtime, 1990 onwards**

**Table 1.A1: List of 30 countries in sub-Saharan Africa used in the study, categorized by AGOA Eligibility, year made AGOA eligible, DHS survey used and sample period of births**

| Sub-Saharan Africa | AGOA Eligible | Year made AGOA Eligible | DHS used | Sample period |
|---|---|---|---|---|
| Angola | Y | December 30, 2003 | 2011 | 1990-2010 |
| Benin | Y | October 2, 2000 | 2006 | 1990-2005 |
| Burkina Faso | Y | December 10, 2004 | 2010 | 1990-2009 |
| Burundi | Y | January 1, 2006 | 2010 | 1990-2010 |
| Cameroon | Y | October 2, 2000 | 2011 | 1990-2010 |
| Chad | Y | October 2, 2000 | 2004 | 1990-2003 |
| Republic of the Congo | Y | October 2, 2000 | 2005 | 1990-2004 |
| Democratic Republic of the Congo | Y | October 31, 2003 – Suspended 2011 | 2007 | 1990-2006 |
| Cote d'Ivoire | Y | 2003 – Suspended 2005; restored 2011 | 2005 | 1990-2003 |
| Ethiopia | Y | October 2, 2000 | 2011 | 1990-2002 |
| Ghana | Y | October 2, 2000 | 2008 | 1990-2007 |
| Guinea | Y | 2000- Suspended 2009; restored 2011 | 2005 | 1990-2004 |
| Kenya | Y | October 2, 2000 | 2008-09 | 1990-2008 |
| Lesotho | Y | October 2, 2000 | 2009 | 1990-2009 |
| Liberia | Y | December 29, 2006 | 2007 | 1990-2006 |
| Madagascar | Y | 2000-Suspended 2009; restored 2014 | 2008-09 | 1990-2008 |
| Malawi | Y | October 2, 2000 | 2010 | 1990-2009 |
| Mali | Y | 2000 – Suspended 2012; restored 2014 | 2006 | 1990-2005 |
| Mozambique | Y | October 2, 2000 | 2003 | 1990-2002 |
| Namibia | Y | October 2, 2000 | 2006-07 | 1990-2006 |
| Niger | Y | 2000-Suspended 2009; restored 2011 | 2006 | 1990-2005 |
| Nigeria | Y | October 2, 2000 | 2010 | 1990-2009 |
| Rwanda | Y | October 2, 2000 | 2010 | 1990-2009 |
| Sao Tome and Principe | Y | October 2, 2000 | 2008-09 | 1990-2008 |
| Senegal | Y | October 2, 2000 | 2010-11 | 1990-2010 |
| Sierra Leone | Y | October 23, 2002 | 2008 | 1990-2007 |
| Swaziland | Y | October 2, 2000 | 2006-07 | 1990-2006 |
| Tanzania | Y | October 2, 2000 | 2010 | 1990-2009 |
| Zambia | Y | October 2, 2000 | 2007 | 1990-2006 |
| Zimbabwe | N | Non-AGOA | 2010-11 | 1990-2009 |

Note: Since Liberia has sample size till 2006 and AGOA was implemented in 2006 for the country, it effectively in the sample behaves as not being AGOA affected. Similarly for Cote d'Ivoire, Ethiopia and Mozambique, since I merge the last year data with previous year due to few data points in the final year, these countries effectively behave as not affected by AGOA.

**Table 1.A2: Point estimates for Dynamics of infant mortality**

| | Pre 2 | Pre 1 | Pre 0 | Post1 | Post2 | Post3 | Post 4 |
|---|---|---|---|---|---|---|---|
| **Infant** | -0.0007 | 0.0008 | -0.0026 | -0.0067** | -0.0043 | -0.0087* | -0.006** |
| **Mortality** | (0.0037) | (0.0036) | (0.0050) | (0.0032) | (0.0039) | (0.0043) | (0.0024) |
| **F-test** | 0.15 [0.924] | | | | | | |
| **Number of countries** | 30 | | | | | | |
| **Number of mothers** | 212738 | | | | | | |
| **Observations** | 686093 | | | | | | |

Note: The explanatory variables included in the specifications are sex of child, whether born in multiple birth, birth order, birth month, mother's age, country specific linear trends, year fixed effects and mother fixed effects. Standard errors clustered at country level are reported in brackets. F-test is for the coefficients on years before AGOA implementation are all zero.
*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 1.A3: Robustness Check – Dropping one country at a time

| Dependent Variable | Infant Mortality | Infant Mortality | Infant Mortality | Infant Mortality | Infant Mortality | Infant Mortality | Infant Mortality |
|---|---|---|---|---|---|---|---|
| | **(1)Angola** | **(2)Benin** | **(3) Burkina Faso** | **(4)Burundi** | **(5)Came-roon** | **(6)Chad** | **(7)Congo** |
| **Treatment** | -0.00761** | -0.00737** | -0.00732** | -0.00613* | -0.0077** | -0.00859*** | -0.0068** |
| | (0.0029) | (0.0026) | (0.0025) | (0.0032) | (0.0031) | (0.0027) | (0.0026) |
| | **(8)Congo, Dem.** | **(9)Cote d'Ivoire** | **(10)Ethiopia** | **(11)Ghana** | **(12)Guinea** | **(13)Kenya** | **(14)Les-otho** |
| **Treatment** | -0.00849*** (0.0026) | -0.00648** (0.0026) | -0.00686** (0.0026) | -0.00703** (0.0026) | -0.00681** (0.0026) | -0.00664** (0.0026) | -0.00677** (0.0026) |
| | **(15)Liberia** | **(16)Madag-ascar** | **(17)Malawi** | **(18) Mali** | **(19)Mozam-bique** | **(20)Namibia** | **(21)Niger** |
| **Treatment** | -0.00665** (0.0027) | -0.00733** (0.0027) | -0.00713** (0.0028) | -0.00653** (0.0026) | -0.00667** (0.0026) | -0.00682** (0.0026) | -0.00688** (0.0026) |
| | **(22)Nigeria** | **(23)Rwanda** | **(24)Sao Tome & Principe** | **(25)Senegal** | **(26)Sierra Leone** | **(27)Swazi-land** | **(28)Tanz-ania** |
| **Treatment** | -0.00737*** (0.0026) | -0.00688** (0.0026) | -0.00698** (0.0026) | -0.00748*** (0.0026) | -0.00597** (0.0025) | -0.00715** (0.0026) | -0.00654** (0.0026) |
| | **(29)Zambia** | **(30)Zimba-bwe** | | | | | |
| **Treatment** | -0.00677** (0.0026) | -0.00616** (0.0025) | | | | | |

Note: The explanatory variables included in the specifications are sex of child, whether born in multiple birth, birth order, birth month, mother fixed effects, country specific linear trends, mother's cohort by child birth year FE. Standard errors clustered at country level are reported in brackets. In each of the separate regressions, one of the countries is dropped at a time in alphabetical order.

## Table 1.A4: List of countries and surveys used in Pathway Analysis

| Sub-Saharan Africa | Year made AGOA Eligible | Infant mortality DHS used | Pathway Analysis DHS Used |
|---|---|---|---|
| Angola | 2003 | 2011 | No data for employment - dropped |
| Benin | 2000 | 2006 | 1996, 2001, 2006 |
| Burkina Faso | 2004 | 2010 | 1993, 1998-99, 2003, 2010 |
| Burundi | 2006 | 2010 | No employment data - dropped |
| Cameroon | 2000 | 2011 | 1991, 1998, 2004, 2011 |
| Chad | 2000 | 2004 | 1996-97, 2004 |
| Republic of the Congo | 2000 | 2005 | 2005, 2011-12 |
| Democratic Republic of the Congo | 2003 | 2007 | Two surveys not available - dropped |
| Cote d'Ivoire | 2003 – Suspended 2005; restored 2011 | 2005 | 1994, 2011-12 (1998/2005 do not have data on toxoid injections) |
| Ethiopia | 2000 | 2011 | 2000, 2005, 2011 |
| Ghana | 2000 | 2008 | 1993, 1998, 2003, 2008 |
| Guinea | 2000- Suspended 2009; restored 2011 | 2005 | No employment data - dropped |
| Kenya | 2000 | 2008-09 | 1993, 1998, 2003, 2008-09 |
| Lesotho | 2000 | 2009 | 2004, 2009 |
| Liberia | 2006 | 2007 | No employment data for 2 rounds of survey - dropped |
| Madagascar | 2000-Suspended 2009; restored 2014 | 2008-09 | 1992, 1997, 2003-04, 2008-09 |
| Malawi | 2000 | 2010 | 1992, 2000, 2004, 2010 |
| Mali | 2000 – Suspended 2012; restored 2014 | 2006 | 1995-96, 2001, 2006 |
| Mozambique | 2000 | 2003 | 1997, 2003, 2011 |
| Namibia | 2000 | 2006-07 | 1992, 2000, 2006-07 |
| Niger | 2000-Suspended 2009; restored 2011 | 2006 | 1992, 1998, 2006 |
| Nigeria | 2000 | 2010 | 1990, 1999, 2008 (2010 is MIS Data) |
| Rwanda | 2000 | 2010 | 1992, 2000, 2005, 2010 |
| Sao Tome and Principe | 2000 | 2008-09 | Two surveys not available - dropped |
| Senegal | 2000 | 2010-11 | 1992-93, 1997, 2005, 2010-11 |
| Sierra Leone | 2002 | 2008 | Two surveys not available - dropped |
| Swaziland | 2000 | 2006-07 | Two surveys not available - dropped |
| Tanzania | 2000 | 2010 | 1991-92, 1996, 1999, 2010 |
| Zambia | 2000 | 2007 | 1992, 1996, 2001-02, 2007 |
| Zimbabwe | Non-AGOA | 2010-11 | 1994, 1999, 2005-06, 2010-11 |

Note: Later survey has been included in the pathway analysis, if a newer survey is available at the time of the analysis, than the infant mortality analysis. The surveys needed to have information on health care variables, maternal employment and possession of assets to be included in the analysis.

## Table 1.A5: Effect of AGOA on infant mortality by year and country

### Time effects:

| | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 |
|---|---|---|---|---|---|---|---|---|
| **Infant** | -0.0059 | -0.0069** | -0.0118 | -0.0028 | 0.0028 | -0.0084** | -0.014*** | -0.0104* |
| **Mortality** | (0.0042) | (0.0030) | (0.0073) | (0.0102) | (0.0082) | (0.0034) | (0.0048) | (0.0061) |
| **Number of countries** | 30 | | | | | | | |
| **Number of mothers** | 212738 | | | | | | | |
| **Observa-tions** | 686093 | | | | | | | |

Note: These are estimates of AGOA interacted with the year dummies for each AGOA year. The explanatory variables included in the specifications are sex of child, whether born in multiple birth, birth order, birth month, mother's age at birth, country time trend and mother fixed effects. Standard errors clustered at country level are reported in brackets.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

### Country effects:

| **Dependent Variable** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** |
|---|---|---|---|---|---|---|
| | **(1)Angola** | **(2)Benin** | **(3)Burkina Faso** | **(4)Burundi** | **(5)Cameroon** | **(6)Chad** |
| **Treatment** | 0.0043 | -0.0022 | -0.0068** | -0.0148*** | 0.0029 | -0.0089* |
| | (0.0026) | (0.0043) | (0.0027) | (0.0024) | (0.0044) | (0.0044) |
| | **(7)Congo** | **(8)Congo, Dem.** | **(11)Ghana** | **(12)Guinea** | **(13)Kenya** | **(14)Lesotho** |
| **Treatment** | -0.0154*** | -0.0019 | 0.0175*** | -0.0157*** | -0.0217*** | 0.0179*** |
| | (0.0037) | (0.0033) | (0.0045) | (0.0039) | (0.0043) | (0.0044) |
| | **(16)Mada-gascar** | **(17)Malawi** | **(18) Mali** | **(20)Namibia** | **(21)Niger** | **(22)Nigeria** |
| **Treatment** | 0.0026 | -0.00073 | -0.0209*** | -0.00668 | -0.0104** | 0.00858* |
| | (0.0044) | (0.0043) | (0.0042) | (0.0044) | (0.0042) | (0.0044) |
| | **(23)Rwanda** | **(24)Sao Tome & Principe** | **(25)Senegal** | **(26)Sierra Leone** | **(27)Swazi-land** | **(28)Tanzania** |
| **Treatment** | -0.0223*** | 0.0057 | -0.0037 | -0.0175*** | 0.0072 | -0.0176*** |
| | (0.0043) | (0.0045) | (0.0044) | (0.0030) | (0.0043) | (0.0044) |
| | **(29)Zambia** | | | | | |
| **Treatment** | -0.0186*** | | | | | |
| | (0.0044) | | | | | |

Note: These are estimates of AGOA interacted with the country dummies for each 25 AGOA affected country in the sample. The explanatory variables included in the specifications are sex of child, whether born in multiple birth, birth order, birth month, mother's age at birth, country time trend and mother fixed effects. Standard errors clustered at country level are reported in brackets.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 1.A6: Placebo test – False timing of AGOA

| | 5 years before | 4 years before | 3 years before | 2 years before | 1 year before | Actual |
|---|---|---|---|---|---|---|
| **Infant** | 0.006 | 0.0030 | 0.0032 | -0.0004 | -0.0023 | -0.0071** |
| **Mortality** | (1.58) | (0.69) | (0.70) | (-0.11) | (-0.68) | (-2.53) |
| **Number of countries** | 30 | | | | | |
| **Number of mothers** | 212738 | | | | | |
| **Observations** | 686093 | | | | | |

Note: Each cell represents a different regression. The explanatory variables included in the specifications are sex of child, whether born in multiple birth, birth order, birth month, mother's age, country specific linear trends, year fixed effects and mother fixed effects. Standard errors are clustered at country level. Resulting t-statistics are reported in brackets. These are placebo test run to test if there are false effects of AGOA on infant mortality before AGOA has been actually implemented.
*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 1.A7: Heterogeneity in effects by child's gender

| | (1) | (2) | (3) |
|---|---|---|---|
| **Dependent Variable** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** |
| **AGOA** | -0.0072**<br>(0.0028) | -0.0062**<br>(0.0026) | -0.0062**<br>(0.0025) |
| **AGOA*Son** | -0.0017<br>(0.0012) | -0.0017<br>(0.0017) | -0.015<br>(0.0017) |
| **Son** | 0.0138***<br>(0.0011) | 0.014***<br>(0.0011) | 0.014***<br>(0.0011) |
| **Explanatory Variables** | YES | YES | YES |
| **Country time trend** | YES | YES | YES |
| **Country FE** | YES | NO | NO |
| **Mother FE** | NO | YES | YES |
| **Cohort-year FE** | NO | NO | YES |
| **Number of Countries** | 30 | 30 | 30 |
| **Number of mothers** | 212738 | 212738 | 212738 |
| **Observations** | 686093 | 686093 | 686093 |

Note: The control variables are whether born in multiple birth, birth order, birth month, and country specific linear trends. Mother's age at birth and year of birth of child fixed effects is included in specifications (1) and (2), and it is subsumed in Cohort-year fixed effects in (3). AGOA is interacted with the gender of child (if the child is a male). Standard errors clustered at country level are reported in brackets.
*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

**Table 1.A8: Effect on the likelihood of child death**

| | (1) | (2) |
|---|---|---|
| **Dependent Variable** | **Infant Mortality** | **Neonatal Mortality** |
| **T-4** | 0.0009<br>(0.0024) | -0.00005<br>(0.0021) |
| **T-3** | 0.0047*<br>(0.0027) | 0.0019<br>(0.0013) |
| **T-2** | -0.0009<br>(0.0029) | -0.0014<br>(0.0019) |
| **T-1** | 0.0028<br>(0.0048) | -0.0015<br>(0.0027) |
| **F-stat**<br>**PRE** | 1.83<br>(0.1569) | 1.10<br>(0.378) |
| **T+1** | -0.0122**<br>(0.0053) | -0.0052***<br>(0.0016) |
| **T+2** | -0.0046<br>(0.0043) | -0.0032<br>(0.0022) |
| **T+3** | -0.0082***<br>(0.0028) | -0.0074***<br>(0.0018) |
| **T+4** | -0.0161**<br>(0.0059) | -0.0075**<br>(0.0034) |
| **F-stat**<br>**POST** | 4.03<br>(0.0122) | 7.34<br>(0.0005) |
| **Number of countries** | 25 | 25 |
| **Observations** | 594560 | 594560 |

Note: These are the θj estimates derived from estimating this equation:

$Death_{imct} = \alpha_c + \beta_t + \sum_{j=-4}^{4} \theta_j T_{c,t+j} + X_i'\delta + \varepsilon_{imct}$ . All coefficients are from the same regression. The sample is restricted to treated countries. The standard errors are clustered at the country level and reported in brackets. The control variables are whether born in multiple birth, birth order, birth month, mother's age at birth, mother's education, place of residence, asset index. Both the specifications control for year and country fixed effects. F-statistics and the corresponding p-values are reported for the joint significance of pre-treatment years and post-treatment years.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

# Chapter 2

## Are Children at Higher Risk of Malnutrition Surviving? Evidence from India*


### Abstract

India presents itself as a paradox with low infant mortality and high malnutrition. Using pooled health surveys from 1993-2005 and a pseudo-panel selection model, this study finds that the change in HAZ scores can be explained by mortality selection. Specifically, children with sample average characteristics that survive with controls for unobservable characteristics of groups of women, have 14.7% less HAZ scores than a child randomly drawn from the population. This is consistent with the hypothesis of weaker children surviving due to skilled delivery which pulls down the overall HAZ scores. Heterogeneity in selection is observed for children of different gender and birth order with males and males of lower birth order having higher negative selection.



***Keywords:*** Infant Mortality, Child Health, Nutrition, India.

***JEL Codes***: J11, J13, I15.



* I am grateful to Anil B. Deolalikar, Mindy Marks, Joseph Cummins, and Aman Ullah for their comments. All remaining errors are my own.

## 2.1 Introduction

Malnutrition has presented itself as a serious burden, especially in India. Malnutrition has been shown to decrease labor productivity and intelligence and lower ability to earn future income, perpetuating a vicious cycle of poverty (Belli, 1971). Due to the high socio-economic impacts of malnutrition, reduction of malnutrition and eradication of poverty and hunger is one of the UN Millennium Development Goals. India presents itself as a paradox with low infant mortality and high malnutrition. Many sub-Saharan African countries have lower rates of malnutrition than India, despite having much higher infant and child mortality rates and lower income per capita.[1] This paper provides evidence of mortality selection as an explanation to the paradox of low infant mortality and high malnutrition in India – due to improved health infrastructure and neonatal care, India is able to save weaker children from dying and who by surviving, lower the average anthropometric scores leading to higher rates of malnutrition prevalence in the data.

Even though this paradox is widely discussed, there are not many empirical studies analyzing the paradox. Panagariya (2013) compares the malnutrition rates across India and sub-Saharan Africa and concludes that the way malnutrition is measured could be the problem as it does not take into account micro-nutrient deficiencies and other aspects of malnutrition. This paper looks at a different phenomenon and takes a different approach. We use the widely used anthropometric measure, height-for-age z-score (HAZ Scores) as a measure of malnutrition to explain the paradox of low infant mortality and high

[1] For example, Chad had IMR of 124 vis-a-vis 50 for India in 2009. While, Chad had 44.8% children below the age 5 as stunted while 47.9% in India for the same period (World Health Statistics Report,2011).

malnutrition *within* India.[2] Bias in this literature can work in two ways. First, weaker malnourished children are more probable to die in the sample resulting in a sample selection such that it pushes up the anthropometric measures. On the other hand, due to improved technology and skilled birthing assistance, survival of weaker kids increases which instead pushes the HAZ scores down. The second source of bias is consistent with the paradox of low infant mortality and high malnutrition in India in recent years. To study this selection effect, we use three rounds of National Family and Health Surveys (NFHS) to examine the probability of survival of a child and link it to nutritional outcomes of the surviving children.

Using pooled national health surveys across years allows us to study how a change in infant mortality affects change in rates of HAZ scores. This is an improvement over any cross-sectional analysis where it concentrates on levels rather than changes and is unable to explain the relationship between changes in the outcome and explanatory variable. Moreover, since we can observe different cohorts of women overtime, we can control for cohort specific time-invariant unobserved heterogeneity which can produce biased estimates in both mortality and malnutrition regression. As long as women belonging to certain cohorts based on their age group, location and socio-economic characteristics are predicted to have similar unobserved characteristics, this study does better than others. In our empirical analysis, we estimate a sample selection model where in stage one, we estimate a pooled probit model of whether or not the child is alive for all children. For

[2] HAZ Scores are the most common anthropometric measure to track malnutrition and is used by UN and WHO. We focus on height instead of weight because the height is considered as a long-run measure of an individual's health (Behrman and Deolalikar, 1988). But, we do also check for robustness of results with WAZ scores.

those children that are currently alive, we then estimate a least squares model with state and cohort fixed effects to examine the factors influencing child malnutrition. If infant mortality is ignored and only data from those children who are currently alive is used to study nutritional outcomes, there is the possibility of sample selection bias.

The paper's findings suggest that infant mortality selection is significant and leads to an underestimation of HAZ scores in the sample, after controlling for child characteristics and mother cohort's time invariant characteristics. This result is also robust to controlling for state-time trends, flexible specification to control for survey timing and mother's age at birth. We observe negative mortality selection in the estimates with respect to both infant and neonatal mortality. A child with sample average characteristics who survives with controls for unobservable characteristics of groups of women, has 14.7% less HAZ scores than a child randomly drawn from the population. The negative mortality selection points towards the survival of weaker kids in the sample which would have otherwise died if not saved by superior birthing technology and skills.

The mortality selection differs by the part of HAZ or WAZ distribution the child belongs to. For stunting and being underweight which are measures of severe malnourishment, we in fact find a positive selection indicating that severely malnourished kids die leaving the sample with higher anthropometric scores than the population. The negative selection is observed at the upper end of the distribution.

The micro level health dataset also helps in identifying the heterogeneous effects based on child characteristics like birth order and gender. We observe the familiar gender and

low birth order preference pattern in India. With better technology, parents put more effort in saving a male child and even a lower birth order male child. This in turn reflects in a negative selection for these groups and the average HAZ scores for these groups are lower than the overall population. Results suggest there is spatial heterogeneity in selection as well. States with historically better child sex-ratios like Kerala display no mortality selection while states like Punjab and Haryana display high negative mortality selection.

Change in HAZ can be brought about by improvement in nutrition which is a causal channel. But, the changes in HAZ could be brought about by changes in probability of survival which changes latent health. Another way in which the increasing trends could be explained is by fertility change where better mothers give birth or certain kind of mothers stop giving birth. This paper provides evidence of mortality selection being a significant channel in change in HAZ scores after controlling for demographic change of mothers and types of households.

## 2.2 Literature Review

Mortality and fertility selection has been acknowledged as a potential source of bias in the literature. A large number of studies show the effect of anthropometric indicators on child mortality indicating malnourished children are more likely to die (Caulfield et al., 2004; Pelletier, 1994; Rice et al., 2000). Boerma et al. (1991) study 17 cross-sectional surveys and other longitudinal data in developing countries to find that malnutrition is more prevalent for deceased children but the survivor bias is small. Pitt (1997) estimates

factors affecting child mortality and child health allowing for past selective fertility and mortality behavior in the context of sub-Saharan Africa. He finds fertility selection to be a significant determinant of mortality in 14 sub-Saharan African countries but with very little change in parameters when selection is accounted for. Dancer, Rammohan, and Smith (2008) study the differences in in survival probabilities by gender and consequent differences in gender based child nutrition in Bangladesh and find that after correcting for selection, female children were more likely to have lower HAZ and WHZ scores. Even though these studies concentrate on selective mortality, the context and approach to the problem is different in this paper.

Another strand of literature looks at the effect of selective mortality and nutrition on the heights of the adult population with the premise that childhood nutrition and disease environment has a significant effect on adult height (Akachi and Canning, 2010; Bozzoli et al., 2009; Moradi, 2010). Deaton (2007) offers a framework of scarring and selection where he explains that positive selection effect in terms of removing shorter individuals by mortality outweighs the negative scarring effect in high mortality environments, resulting in taller adults in the case of sub-Saharan Africa. Bozzoli et al (2009) examine adult height of 31 cohorts in England, US and 10 European countries and show that the postneonatal mortality rate of the country predicts the average adult height of the birth cohort. Unlike developed countries where fall in infant mortality is accompanied by increasing heights, Akachi and Canning (2010) do not find evidence of the same in sub-Saharan Africa. If adult height is regarded as a measure of well-being, falling infant mortality may not be contributing to increased health in different settings.

The paradox between falling infant mortality but small improvements in health status has been documented where the reductions in child mortality are brought about by directed interventions. Epidemiology literature has documented that vaccinations and interventions aimed at reducing mortality does not reduce morbidity. Pinchinat et al. (2004) find that in southern Senegal, no nutritional improvement was found in children in 1962-1992, despite a big drop in infant mortality. With medical improvements and directed efforts at reducing infant mortality, without complementary increase in protein and other micronutrient intake by kids, India may be facing a similar situation.

The paper closest to this research is by Alderman et al. (2011), where they construct simulated nutritional status of the child population for all the children in India who die before the age of 3 years, assuming they were alive using a proportional hazard model. This is based on matching of individuals based on observable characteristics to impute HAZ scores of children who had died. They use the three rounds of NFHS to construct the simulations and find a 5 percent difference between the counterfactual and the actual height-for-age z-scores. The problem with matching on observables is that if there are unmeasured confounding factors, the analysis could be biased. In terms of controlling for unobserved mother cohort characteristics which may affect survival and malnutrition, this study does better.

## 2.3 Model

Following Pitt (1997), we can estimate the effect of selective mortality in anthropometric measures of child health by reduced form equations of infant and neonatal mortality and health:

$$CH = X_c\beta_c + \delta_{mh}\mu_m + \nu_c = X_c\beta_c + \varepsilon_c \quad (1)$$

$$M^* = X_m\beta_m + \mu_m + \nu_m = X_m\beta_m + \varepsilon_m \quad (2)$$

Where $X_c$ and $X_m$ are exogenous regressors, the error in both equations contain a heterogeneous error term $\mu_m$ which determines parental preferences over unwanted births (especially in the case of India, son preference), woman or cohort specific effects of women who observe higher mortality of kids tend to also have kids which are malnourished, general medical interventions and shocks which may affect mortality etc. Child health data is observed only for surviving children and hence it needs to be corrected for selective mortality.

If the errors have zero means, and $\nu_c$ and $\nu_m$ are uncorrelated, then covariance between the error terms can be written as:

$$Cov(\varepsilon_m, \varepsilon_c) = \delta_{mh}Var(\mu_m) \quad (3)$$

Therefore, selection bias results if $\delta_{mh}$ is not equal to 0 that is there is a feedback between mortality and child health. If medical technology is saving weaker infants, which then become malnourished due to improper care, then $\delta_{mh}>0$.

In the analysis, we define survival (s) as a kid not experiencing neonatal or infant mortality. If the population errors are jointly normally distributed, the health of children conditioned on survival is:

$$E(CH|X_c, S^*>0) = X_c\beta_c + Cov(\varepsilon_m, \varepsilon_c)\lambda \quad (4)$$

where $\lambda$ is the Inverse Mill's Ratio. Omitting this $\lambda$ creates a bias in the estimates of $\beta_c$.

Considering the HAZ scores for the whole population,

$$Z_{pop} = Z_s(1-P_d) + Z_dP_d \quad (5)$$

Where $Z_s$ and $Z_d$ are average Z-scores of surviving and deceased children and $P_d$ is the proportion of deceased children. The population Z score will change if deceased children survived:

$$\theta = Z_s - Z_{pop} = (Z_s-Z_d)P_d \quad (6)$$

According to (6), $Z_s>Z_d$ and $\theta>0$ if more malnourished children die, leading to higher existing HAZ scores of the sample than the population. Instead, if weaker children survive due to improvements in technology and other interventions such that $Z_s<Z_d$ and $\theta<0$, then it will result in an underestimation of HAZ-score distribution in the sample.

## 2.4 Data

This analysis uses data from three waves of India's National Family Health Survey (NFHS) – 1992/93, 1998/99 and 2005/06. The NFHS follows the pattern of a standard

Demographic and Health Survey and is a large scale survey covering a representative sample of households throughout India. For the 1992/93 survey, interviews were conducted with a nationally representative sample of 88,562 households and 89,777 ever-married women in the age group 13-49, from 24 states and the then National Capital Territory of Delhi. The 1998/99 survey covered a representative sample of about 91,000 ever-married women age 15-49 from 26 states in India. NFHS-3 conducted interviews with over 230,000 women age 15-49 and men age 15-54 throughout India.[3]

The survey is administered to ever-married females and contains detailed information about their reproductive history, asset ownership, vaccinations and preventive care, reproductive health, and educational characteristics etc. Women of reproductive age are interviewed about the date of birth and death (if applicable) of their pregnancy histories. This kind of retrospective survey gives an opportunity to build an indicator of infant and neonatal mortality and in turn the survival probabilities.[4]

The NFHS provides height and weight data for children under age of 48 months in 1992/93, under age of 36 months in 1998/99, and under age of 60 months in 2005/06.[5] The NFHS contains no anthropometric information for deceased children at the time of their death. The NFHS collects information on weight at birth in addition to weight at the time of the survey and asks mothers to categorize the weight of their children at birth as large, average, or small. It also collects information on height/length of the child and age

[3] http://www.rchiips.org/nfhs

[4] One problem that can be raised with the recall data is the measurement error problem. Since the birth histories do not go too much into the past, it is lesser of a problem in this case. Moreover, since deaths of a kid are important in a mother's life, this variable should be recorded without much measurement error.

[5] We also check for the mortality selection effect if we constrain the results to kids born under 36 months across different survey years in Table 2.7. The results are unchanged.

for children up to five years of age. In the sample, we have 766364 children born to 234548 mothers. Out of these, 72116 children have non missing data on HAZ scores and 81018 children have data on WAZ scores.

In Table 2.1, we show the summary statistics. The average infant mortality rate is 82 kids per 1000 while the average neonatal mortality rate in the sample is 51 per 1000 children. The mean HAZ score is -2.13 and mean WAZ score is -1.84. About 55% of the children in the sample are stunted (HAZ<-2) and 44% are underweight (WAZ<-2). Over the survey years, HAZ has been increasing and infant mortality has been falling (Appendix Figure 2.1). The infant mortality was about 92 deaths per 1000 in 1993 and has fallen to 70 deaths per thousand live births. Mean HAZ scores did not change a lot between 1993 and 1998, with it being around -2.3 but it increased by 2005 with the score at -1.89. This is shown in Fig. 3f. The average mother's age at birth is 23 years. Around 64% mothers have delivery at home and 44% mothers have their delivery by a skilled birth attendant.

Figure 2.1 plots the relationship between HAZ and WAZ and decline in infant mortality. All panels show a decline in HAZ or WAZ score with higher rates of declines in infant mortality between 1993-1998, 1998-2005 and 1993-2005. The scatterplots show the states. As expected, Gujarat, Bihar, Uttar Pradesh and Madhya Pradesh lie below the fitted line with lower HAZ for the mortality decline. Figure 2.2 plots the sample mean infant mortality rates by year in different regions of India. Most states exhibit a declining trend in infant mortality over the years. Kerala stands out in the southern region with a distinctly lower level of mortality rate.

A mother is labeled as educated if she has attended any type of school and uneducated if mother did not attend any school. Since DHS does not have income data, definition of poor is based on possession of assets. A mother is categorized as poor if she does not possess any asset. Rural or urban are defined by the place of residence of mother during the time of interview. Around 59% of women interviewed are uneducated, 68% live in rural areas and 29% hold no assets in their possession. Figure 2.3 plots the difference in HAZ densities by mother characteristics. HAZ distributions for rural, uneducated, and poor women are skewed to the left. Mothers of different age groups also showcase different HAZ distributions. Keeping this in mind the mother cohorts are constructed later.

## 2.5 Empirical Strategy

We have pooled cross-section data overtime where same individuals are not observed in different time periods and interview years. We expect individual heterogeneity to be present in the error term. Moreover the selection process may be varying with time and individual.[6] Deaton (1985) emphasizes on using cohort to obtain consistent estimates in pseudo-panels even in the case of correlation between individual effects and explanatory variables. The individual heterogeneity can be written as cohort effect plus individual

[6] If the selection process is identical over time, then the fixed effects estimator will remove the selection bias in a panel data.

deviation. The child health equation in our case can be written with HAZ score as the dependent variable:[7]

$$HAZ_{i(t)t} = X_{i(t)t}\beta + \alpha_{i(t)} + \mu_{it} \quad (7)$$

$$\text{Where, } \alpha_{i(t)} = \Sigma\delta_c\alpha_c + \sigma_{i(t)} \quad (8)$$

And the survival equation, indicating selective mortality is described by:

$$S_{i(t)t}^* = Z_{i(t)t}\gamma + \eta_{i(t)} + u_{i(t)t} \; ; \; S_{i(t)t} = 1\,[S_{it}^* > 0] \quad (9)$$

Infant mortality is defined as a child dying before the age of 1 year and neonatal mortality is defined as child dying before the age of 1 month. Survival accordingly is the observation of child not dying in the data till the age of 1 year and in case of neonatal survival, till the age of 1 month. The child health equation estimated by pooled OLS will be biased if correlated individual heterogeneity is present or selection process is nonrandom and $\alpha_{i(t)}$ is not a random component of the error. To eliminate individual effects, fixed effects method can be used where cohort-specific dummies control for the cohort effect. Since deviation from cohort is independent from the selection process itself, the correlation between them should be zero which would lead to efficient estimates.

Rodriguez and Muro (2014) develop a selection bias estimation for pseudo-panel data and show that using probit model with cohorts as instruments, a cohort level inverse

[7] We use the 2006 WHO standards for HAZ, computed using the Stata package "haz06" which requires the data for age, height, and gender from NFHS. I have also run this using the HAZ scores reported in the NFHS. But, it does not change the results qualitatively.

Mills ratio can be derived and which can be replaced in the child health equation to derive estimates of selection bias. Accordingly, the selection equation is estimated by a probit model with controls for cohort dummies. A cohort is defined by a group of mothers belonging to the same age group (based on their year of births), residence, education and economic condition. Controlling for mother cohorts also controls for changing demography and fertility which may be bringing about a change in HAZ, which we may erroneously attribute to improvement in nutrition. Since there are considerable differences in the distribution of HAZ scores by these characteristics, the mother-cohort groups are constructed keeping that in mind.[8] Since NFHS does not have data on income or wages, the economic condition of woman or household is determined by the assets that she possesses. In this analysis, if the woman has no assets, she is defined as poor. Along with cohort dummies, the selection equation controls for gender, multiple births, month of birth of the child, survey year, birth order, interaction between being a female and birth order, caste of the household, religion of the household, mother's age at birth, whether the residence has access to piped water, has a toilet and has electricity.

For us to be able to identify the parameters of reduced form mortality selection and determinants of health, we need at least one exogenous variable that affects mortality or survival but does not affect child health and anthropometric scores. It has been noted in the literature that to avoid weak identification, there should be some variables in the selection equation which are not there in the child health equation so that the effect is not

[8] Including different characteristics with missing values, increases the number of cohorts but decreases the observations within cohorts; which is not desirable computationally and otherwise for consistency (Borjas and Sueyoshi, 1993)

identified solely off of nonlinearity in the inverse Mills ratio (Little, 1985; Vella, 1998). Not having any excluded variable may inflate the standard errors and give unreliable estimate of $\beta$ in the second stage. In this analysis, this excluded variable is if the woman has her delivery by a professional doctor or nurse.[9] Having a delivery done by a professional affects the probability of survival of the kid but does not directly affect the HAZ score of the child later on.

In the parametric estimation, after the estimation of selection equation over all observations, the inverse Mills ratio is constructed, which is the ratio of probability density function and the cumulative distribution function of the standard normal distribution. This ratio then is used as an additional regressor in the child health equation to consistently estimate the parameters:

$$HAZ_{i(t)t} = X_{i(t)t}\beta + \theta\widehat{\lambda_{i(t)t}} + \mu_{it} \qquad (10)$$

The equation controls for all the independent variables in the selection equation, except for the excluded variables (delivery by skilled birth attendant). If $\theta$ is statistically significant, it points towards a selection effect operating through selective mortality on health measures. If $\theta>0$, there is a positive selection meaning those who survive have a higher HAZ than a random drawing from the population with the same characteristics. If $\theta<0$, there is a negative selection meaning those who survive have a lower HAZ than a random drawing from the population with the same characteristics. Moreover since this

[9] Skilled delivery assistance is measured by delivery being assisted by doctors, nurse/midwife, auxiliary midwife, ayurvedic doctor, and any other India-specific health professional. It is not considered as skilled assistance if the baby is born with the help of trained birth attendant, traditional birth attendant, relatives, other persons or no one.

equation is able to control for mother-cohorts, any cohort specific time invariant heterogeneity is taken care of. Any health specific shocks and survey-year differences are captured through the year dummies. Standard errors are clustered at the mother-cohort level to improve inference.

In India, especially in rural areas, studies have shown that differential health outcomes can be expected by birth order of the child and also by gender (Behrman and Taubman, 1986; Horton, 1988; Savage, Derraik, Miles, et al., 2013). To test if there is a differential selection effect, we run the child health specification separately for children with first and second birth order versus children of later birth order. The negative selection effect should be prominent in later birth children as they are more probable to be saved by neonatal care units but get lower nutrition inputs in a budget constrained household. Similarly, we postulate that females with higher birth order should be more affected than males or females in lower birth order. Heterogeneity is also expected by state and time.

## 2.6 Results

### 2.6.1 Selection Effects

The mortality selection effect is captured by the coefficient on inverse Mills ratio and is presented in Table 2.2. Columns (1)-(3) provide the effect of inverse Mills ratio on HAZ and WAZ scores which conditions for survival of kids till 1 year of age. Probit estimates are provided in Appendix Table 2.A1.[10] Skilled birth assistance is significant at the 1%

[10] The results include the coefficients of important variables. The full table is available on request.

level and more skilled birth delivery increases the survival and therefore decreases both infant and neonatal deaths. Table 2.2, column (1) finds evidence of negative selection, even in the absence of cohort fixed effects. Columns (2) and (3) control for mother cohorts and see the effect on HAZ and WAZ respectively. In both the results, the coefficient is similar to (1) and highly statistically significant at 1% significance level. With mean inverse Mills ratio in the sample being 0.116, the coefficient of -1.18 implies an average truncation effect of (0.116*1.18) = 0.137. Thus the HAZ scores are shifted down due to the selection effect. A child with sample average characteristics who survives with controls for demographic change in fertility of women, has -(exp(0.137)-1)*100 = 14.7% less HAZ scores than a child randomly drawn from the population.

In Table 2.2, (4)-(6), first stage probit estimates the probability of survival till 1 month of age. The resulting inverse Mills ratio is calculated and included in these regressions. With only the state fixed effects, the ratio is highly significant and similar in magnitude as before. With the inclusion of cohort effects, the coefficients are now significant at 10% level. But, the sign of the coefficients is still negative and magnitude similar, albeit a little bit smaller. The negative inverse Mills ratio points towards the direction of a negative selection, implying that the HAZ and WAZ scores of the sample are lower than a population taken at random. This supports our hypothesis that weaker kids are surviving due to skilled delivery which pulls down the sample anthropometric scores.

It is interesting to note the coefficients on some other important variables that are historically deemed to be important in determination of malnutrition in children. We run

the OLS regression of HAZ score on all the control variables in Table 2.2, (7) without accounting for selection, where being a female or born in a multiple birth significantly affects HAZ. However, after controlling for selection, being born in a multiple birth is not statistically significant in almost all specifications and the coefficient of female dummy reduces. Birth order, access to clean drinking water, having a toilet and access to electricity all remain significant and similar in magnitude, even after accounting for selective mortality. With skilled delivery and medical inputs, the complications arising due to multiple births are taken care of, resulting in no differentiation in HAZ between multiple birth children and single birth children conditional on surviving.

According to WHO Standards, both HAZ and WAZ score use the cutoff of -2 to measure moderate and severe under nutrition.[11] Having a low HAZ score is termed as stunted growth while a low WAZ score leads to the child being underweight. Stunted growth refers to a child below 5 years being short for his/her age and is an indicator of chronic malnutrition. Low WAZ can be either due to the child being thin or short for his/her age. This is a combination of chronic and acute malnutrition. Table 2.3 checks for the presence of mortality selection at various points of the HAZ and WAZ distribution. For stunting and underweight children, we find effects of positive selection whereas for HAZ and WAZ scores greater than -2 negative mortality selection we observe negative selection.

[11] A z-score of zero indicates the median of gender and age specific reference population, -1 is 1 standard deviation below and +1 is 1 standard deviation above the reference median population.

This is consistent with our hypothesis. At the lower end of the distribution, the kids are severely under-nourished. This would mean that they are at a higher risk of dying of less nourishment leaving the sample HAZ scores higher than the whole population. On the other hand, with HAZ and WAZ scores greater than -2, if medical intervention is able to save the children and they do not get enough nourishment later they survive but the presence of these children lowers the HAZ score than the anthropometric score for the population.

Table 2.4 approaches the problem of selection by another method. We can calculate the predicted HAZ score if dead children were included in the sample. In NFHS, we have fertility history of mother and many women give birth to more than one child. Assuming that mothers have similar abilities in raising all their children and siblings who died will be similar to the ones who survived, after controlling for child characteristics, we develop a predicted HAZ score for the whole sample, including the kids who died. This predicted HAZ is better in terms of being able to control for mother unobserved characteristics than random matching based on observable covariates. Since over the years medical technology has improved with better neonatal care, HAZ scores in later survey years will be an underestimate in the sample. We expect that in the predicted sample, the coefficient on 2005 survey year is less positive and statistically significant in increasing HAZ scores.

HAZ has been increasing over the survey years.[12] But some of this increase is explained by changing child demographics and women demographics. Table 2.4, column (1) provides the results of regression of HAZ on child covariates and survey years. The

[12] Graph in Appendix Figure 2.A1.

coefficient on 2005 survey year is positive and statistically significant. Further, we control for mother characteristics in (2). The coefficient on 2005 dummy is still positive, statistically significant but a little lesser in magnitude. Column (3) now uses the predicted HAZ as the outcome variable. Predicted HAZ is obtained by regressing HAZ on child covariates and family fixed effects and getting a linear prediction. This now includes the HAZ scores for kids who have died, if instead they would have survived. The coefficient on the 2005 interview year dummy is significant but almost similar in magnitude. Similarly, in (4)-(6), the WAZ scores are positive and significant, but not very different in magnitudes.

### 2.6.2 Heterogeneity in Selection

India presents a case of strong gender preferences and birth order differences in health outcomes (Behrman, 1988). As plotted in Appendix Figure 2.A2, the changes in infant mortality also differ by birth order and gender.[13] Hence, we expect to see differential selection outcomes based on gender, birth order and an interaction between the two. This hypothesis is tested and results are presented in Table 2.5. For both infant and neonatal survival selection, we see that while male HAZ scores have a significant negative effect of selection, no such effect is found for the female group. This would be the case if parents especially put more effort in getting even a weaker male kid to survive but do not put such an effort in the case of a girl child. Given the patriarchy structure in India, with a lot of emphasis on male kids being the driver of the family heritage, this seems plausible.

[13] Similarly, percentage change in HAZ also differs by gender and birth order, as shown in Appendix Figure 2.A4.

Parents and families face different financial constraints over time and hence may devote less or more resources to latter-born kids. If the family is more constrained, the outcome for health of latter born kid will be negatively affected than their older counterparts. Mothers giving birth at an older age way back in time would face higher probability of a latter-born kid dying than mother giving late birth in a new era when technology is able to save her kids, without changing family's health behavior much. This is represented in the data in Appendix Figure 2.A2. States with higher proportion of high birth order kids (>2) display a greater decline in infant mortality over the survey years than states with low birth order (<=2).

Results in Table 2.5, columns (2) and (4), reflect this finding. High birth order kids display a negative selection bias implying that technology may have been able to save the inherently weak kids lowering the HAZ and WAZ scores. On the other hand, low birth order kids have positive selection bias implying those who are surviving are in fact healthier than the population in general.

The distinction between birth orders is also apparent when females are higher birth order kids. Being a female at higher birth order acts as double disadvantage for female kids as in a resource constrained environment, females receive fewer resources than their male counterparts and if family size is bigger and they are later born, this distinction should be sharper. Table 2.5, columns (3) and (6), test for the differences in selection by birth order and gender. Both male and female kids with high birth order display negative selection, consistent with columns (2) and (4). Male kids see a higher negative selection (-3.29)

than females (-2.64). It reinforces the fact that higher order male kids are more likely to be survived by advanced technology than the female kids at higher birth orders.

We expect to see space and time variation in selection effect. Later years should display a negative relationship between survival and child health as later years will have improved technology vis-a-vis earlier years making it possible to save weaker kids from dying. The results are presented in Table 2.6, columns (2) and (5). Even though the coefficients on inverse Mills ratio are negative and later survey years are more negative than the first survey year, they are not statistically significant.

Based on regions displaying different culture of son preference, heterogeneity by states is also expected. Kerala has distinctively better male-to-female child sex ratio (1.04) than the average in India (1.08), according to 2001 census. Since most of the negative mortality selection is observed in male child, with parity, we should observe no selection effect for this state. Table 2.6 column (1) and (4) displays the results. There is no statistically significant evidence of mortality selection in Kerala. The magnitude of the coefficient is also small and positive. According to Census 2001, while Kerala has a high child sex ratio, there are six states which perform poorly and below the national average – Gujarat, Punjab, Haryana, New Delhi, and Himachal Pradesh. With a strong preference for boys documented for these states, the effect of negative mortality selection should be more pronounced. Table 2.6 columns (3) and (6) show that the coefficient for mortality selection is negative, high in magnitude, and statistically significant.

### 2.6.3 Robustness Checks

It could be argued that the excluded variable of skilled birth delivery may not satisfy the exclusion restriction since it can be taken as a proxy for accessibility to health center which also affects HAZ. Even though this exclusion conditional on other covariates, cannot be tested econometrically, we provide evidence that it does not belong in the second stage HAZ regression. To account for accessibility to health inputs, another variable detailing the number of antenatal visits made by mother to the hospital is included in both the first and second stage regressions. For infant mortality selection, as presented in Table 2.7 (5), the inverse Mills ratio is negative and statistically significant.

The results are also robust to changing the outcome variable from HAZ and WAZ to height in centimeters and weight in kilograms. As seen in columns (3) and (4) of Table 2.7, we still find effects of negative mortality selection and it is significant at the conventional levels. The model is also robust to inclusion of state time trends, which take into account differential levels of development of the states implying that this selection is observable within mothers of similar characteristics within state as well. Lastly, since there are three pooled surveys which have been carried out in different years and at different times of the years, to control more flexibly for survey time; a linear, quadratic and cubic term of the month-year survey time is added to the specification along with age of child and year of birth. This does not significantly affect the results and the coefficient is similar as before.

Since the height and weight data for children in different rounds of NFHS differs by age of child, we restrict the sample to children within 36 months of interview year across all survey years to maintain comparability. The estimates are given in Table 2.7, column (6). The coefficient on inverse Mills ratio is virtually unchanged from the previous regressions.

## 2.7 Conclusion

This paper analyzes the paradox of decreasing infant mortality but not a corresponding increase in health anthropometric scores in India over time. Using three rounds of health surveys in India, we find evidence of negative mortality selection. This negative selection is observed when the anthropometric scores are above -2 standard deviations. With improved technology, weaker kids are surviving pulling down the sample HAZ and WAZ scores than what the scores would have been otherwise. With a lot of emphasis on reduction in mortality and improving child nutrition as part of Millennium Development Goals, this result should be taken into consideration. The improved efforts to decrease infant and neonatal deaths by providing one shot interventions should be followed by health and nutrition interventions to keep malnourishment away. If weaker kids survive due to successful neonatal interventions, without provision of appropriate care afterwards, it would worsen the case of malnutrition.

We also find evidence of heterogeneous effects based on gender and birth order of the child. Given the strong son preference in India this is not surprising. Male children are more likely to experience negative selection in mean HAZ and WAZ scores with no such

effects for the female child. With a patrilineal structure in India, all efforts to save a male child are expected, which makes the probability of finding a weaker male child in the distribution higher; pulling down the overall anthropometric scores. In terms of birth order, we observe negative selection effect in higher birth order children, for both males and females. Spatial heterogeneity is also observed with states historically better off in child sex ratio composition like Kerala displaying no evidence of selection while states with skewed sex ratios like Punjab, Haryana, Gujarat, New Delhi and Himachal Pradesh, display a strong negative selection indicating the presence of son preference. These patterns of mortality selection may be depressing the child anthropometry scores, hiding the fact that India maybe doing better in terms of number of malnourished children in the population overtime than the raw data suggests.

## References


Akachi, Yoko, and David Canning. 2010. “Health Trends in Sub-Saharan Africa: Conflicting Evidence from Infant Mortality Rates and Adult Heights.” Economics and Human Biology 8 (2). Elsevier B.V.: 273–88. doi:10.1016/j.ehb.2010.05.015.

Alderman, Harold, Michael Lokshin, and Sergiy Radyakin. 2011. “Tall Claims: Mortality Selection and the Height of Children in India.” Economics and Human Biology 9 (4): 393–406. doi:10.1016/j.ehb.2011.04.007.

Behrman, Jere R. 1988. “Nutrition, Health, Birth Order and Seasonality: Intrahousehold Allocation among Children in Rural India.” Journal of Development Economics 28 (1): 43–62. doi:10.1016/0304-3878(88)90013-2.

Behrman, Jere R, and Anil Deolalikar. 1988. “Health and Nutrition.” In Handbook of Development Economics, edited by Hollis Chenery and Srinivasan T.N., 1st ed., 1:631–711. Elsevier.

Behrman, Jere R, and Paul Taubman. 1986. “Birth Order , Schooling , and Earnings.” Journal of Labor Economics 4 (3): S121–45.

Belli, Pedro. 1971. “The Economic Implications of Malnutrition: The Dismal Science Revisited.” Economic Development and Cultural Change 20 (1): 1. doi:10.1086/450527.

Boerma, J Ties, A Elisabeth Sommerfelt, and George T Bicego. 1992. “Child Anthropometry in Cross-Sectional Surveys in Developing Countries: An Assessment of the Survivor Bias.” American Journal of Epidemiology 135 (4): 438–49.

Borjas, George J, and Glenn T Sueyoshi. 1994. “A Two-Stage Estimator for Probit Models with Structural Group Effects.” Journal of Econometrics 64 (1-2): 165–82. doi:10.1016/0304-4076(94)90062-0.

Bozzoli, Carlos, Angus Deaton, and Climent Quintana-domeque. 2009. “Adult Height and Childhood Disease.” Demography 46 (4): 647–69. doi:10.1353/dem.0.0079.

Caulfield, Laura E, Mercedes de Onis, Monika Blössner, and Robert E Black. 2004. “Undernutrition as an Underlying Cause of Child Deaths Associated with Diarrhea, Pneumonia, Malaria, and Measles.” The American Journal of Clinical Nutrition 80 (1): 193–98.

Cole, Shawn, Andrew Healy, and Eric Werker. 2012. “Do Voters Demand Responsive Governments? Evidence from Indian Disaster Relief.” Journal of Development Economics 97 (2): 167–81. doi:10.1016/j.jdeveco.2011.05.005.

Dancer, Diane, a N U Rammohan, and Murray D Smith. 2008. “Infant Mortality and Child Nutrition” 1035 (July): 1015–35. doi:10.1002/hec1379.

Deaton, Angus. 2007. “Height, Health, and Development.” Proceedings of the National Academy of Sciences of the United States of America 104 (33): 13232–37. doi:10.1073/pnas.0611500104.

Derosas, Renzo. 2009. “The Joint Effect of Maternal Malnutrition and Cold Weather on Neonatal Mortality in Nineteenth-Century Venice: An Assessment of the Hypothermia Hypothesis.” Population Studies 63 (3): 233–51. doi:10.1080/00324720903165449.

Horton, Sue. 1988. “Birth Order and Child Nutritional Status: Evidence from the Philippines.” Economic Development and Cultural Change 36 (2). University of Chicago Press: 341–54.

Leroy, Jef. 2011. “ZSCORE06: Stata Module to Calculate Anthropometric Z-Scores Using the 2006 WHO Child Growth Standards.” Statistical Software Components. Boston College Department of Economics. http://ideas.repec.org/c/boc/bocode/s457279.html.

Little, Roderick J A. 1985. “A Note about Models for Selectivity Bias.” Econometrica 53 (6). Econometric Society: 1469–74.

Moradi, Alexander. 2010. “Selective Mortality or Growth after Childhood? What Really Is Key to Understand the Puzzlingly Tall Adult Heights in Sub-Saharan Africa.” Working Paper, 1–23.

Panagariya, Arvind. 2013. “Does India Really Suffer from Worse Child Malnutrition Than Sub-Saharan Africa?” Economic and Political Weekly 48 (18): 98–111.

Pelletier, D L, E A Frongillo, D G Schroeder, and J P Habicht. 1995. “The Effects of Malnutrition on Child Mortality in Developing Countries.” Bulletin of the World Health Organization 73 (4): 443–48.

Pinchinat, Sybil, Catherine Enel, Gilles Pison, Géraldine Duthé, Emmanuel Lagarde, François Simondon, and Kirsten B. Simondon. 2004. “No Improvement in Weight-for-Age of Young Children in Southern Senegal, 1969-1992, despite a Drastic Reduction in Mortality. Evidence from a Growth Monitoring Programme.” International Journal of Epidemiology 33 (6): 1202–8. doi:10.1093/ije/dyh174.

Pitt, Mark M. 1997. “Estimating the Determinants of Child Health When Fertility and Mortality Are Selecetive.” The Journal of Human Resources 32 (1 (Winter)): 129–58.

Rice, Amy L, Lisa Sacco, Adnan Hyder, and Robert E. Black. 2000. “Malnutrition as an Underlying Cause of Childhood Deaths Associated with Infectious Diseases in Developing Countries.” Bulletin of the World Health Organization 78 (10). World Health Organization: 1207–21. doi:10.1590/S0042-96862000001000007.

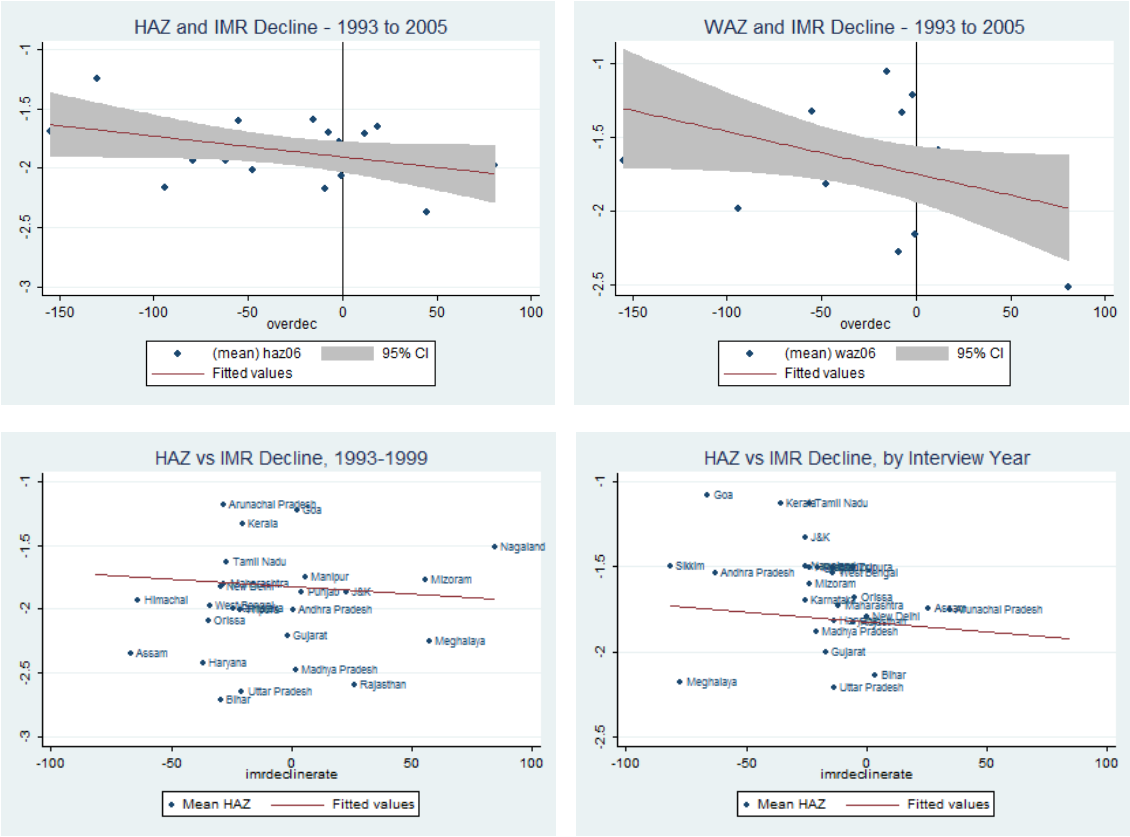
Savage, Tim, José G B Derraik, Harriet L Miles, Fran Mouat, Paul L Hofman, and Wayne S Cutfield. 2013. “Increasing Maternal Age Is Associated with Taller Stature and Reduced Abdominal Fat in Their Children.” PloS One 8 (3): e58869. doi:10.1371/journal.pone.0058869.

Vella, Francis. 1998. “Estimating Models with Sample Selection Bias: A Survey.” The Journal of Human Resources 33 (1): 127–69.

World Health Statistics Report, 2011. 2011. World Health Organization. http://www.who.int/whosis/whostat/2011/en/.

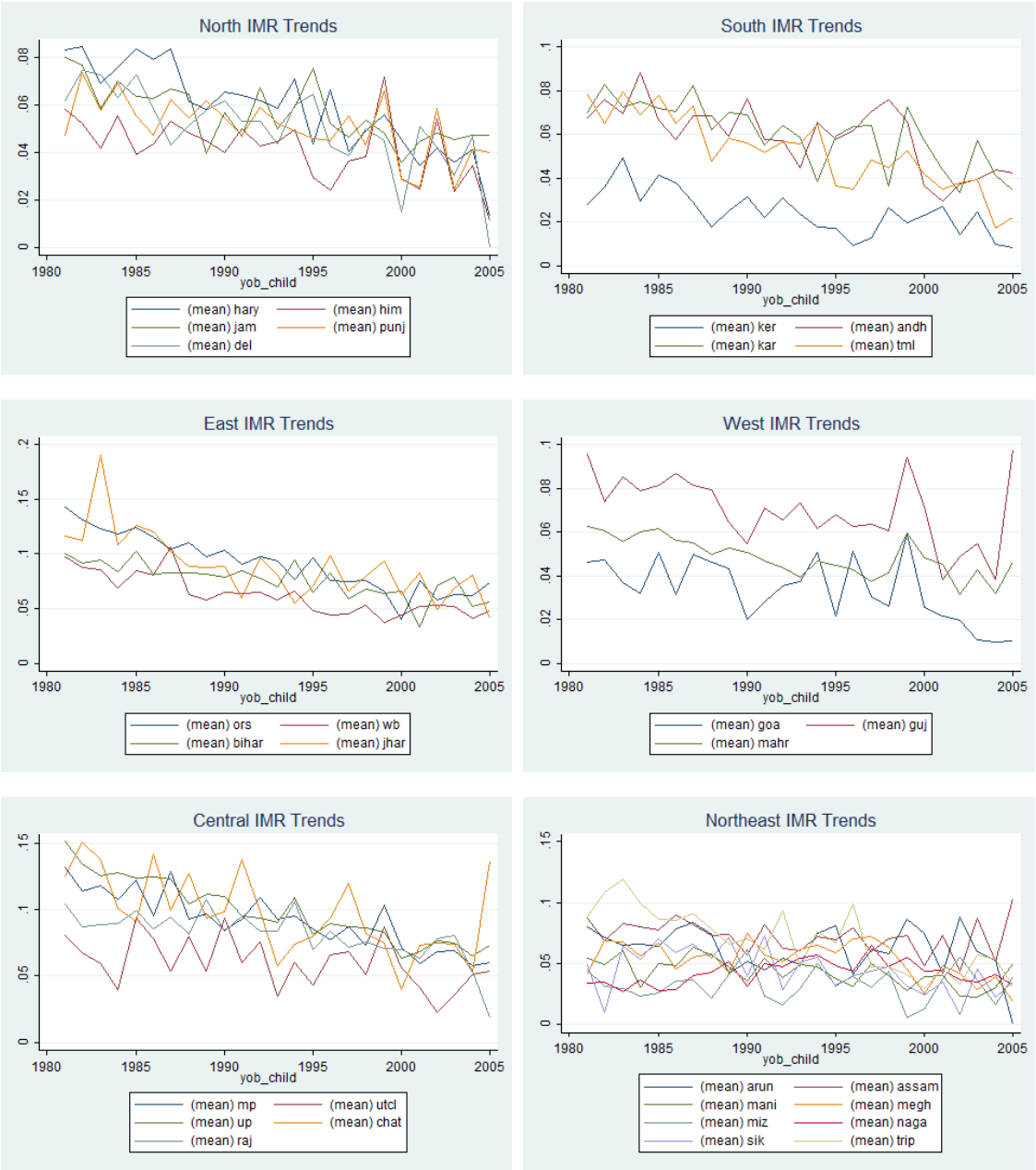
# Figures

## Figure 2.1: Infant Deaths and Malnutrition - By Survey Years




Note: In the first row, these graphs plot the decline in sample infant mortality deaths for India between 1993 and 2005. First panel plots the fitted line and the confidence interval of the relationship between IMR Decline Rate and HAZ scores. Second panel plots the relationship between WAZ scores and IMR Decline rate. In the second row, these graphs plot the decline in sample infant mortality deaths for various states between 1993 and 1999 (first and second NFHS Surveys) and 1999 and 2005 (second and third NFHS surveys). For both the panels, a fitted line is drawn to indicate the relationship between IMR Decline Rate and HAZ scores.

**Figure 2.2: Infant Mortality Rates - By Region**




Note: These graphs plot the mean infant mortality deaths over time for different geographic regions of India. North region consists of Haryana, Punjab, Jammu and Kashmir, Himachal Pradesh, and New Delhi. South consists of Kerala, Karnataka, Andhra Pradesh, and Tamil Nadu. East region has the states of West Bengal, Orissa, Bihar, and Jharkhand. West has Goa, Gujarat, and Maharashtra. Central is divided into Madhya Pradesh, Uttar Pradesh, Uttaranchal, Chhattisgarh, and Rajasthan. North-east consists of Assam, Arunachal Pradesh, Meghalaya, Mizoram, Nagaland, Sikkim, and Tripura.

**Figure 2.3: Variation in HAZ Scores Density**

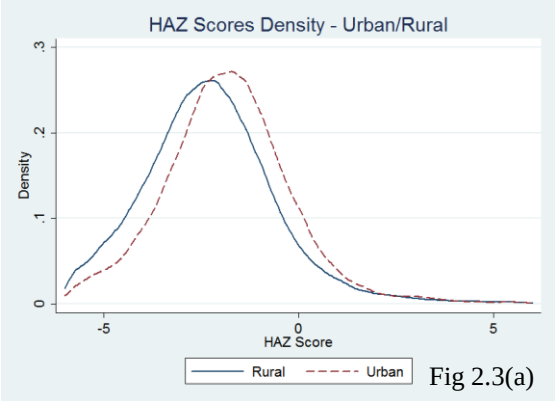



Fig 2.3(a)

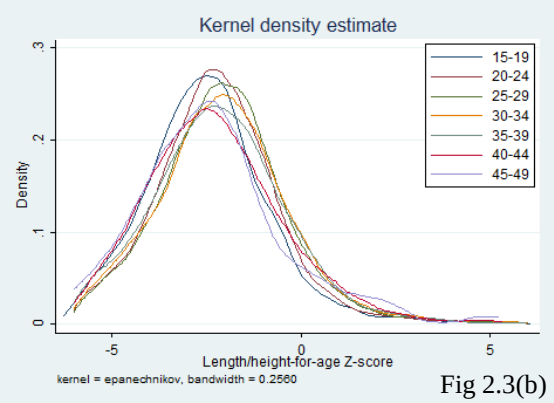



Fig 2.3(b)

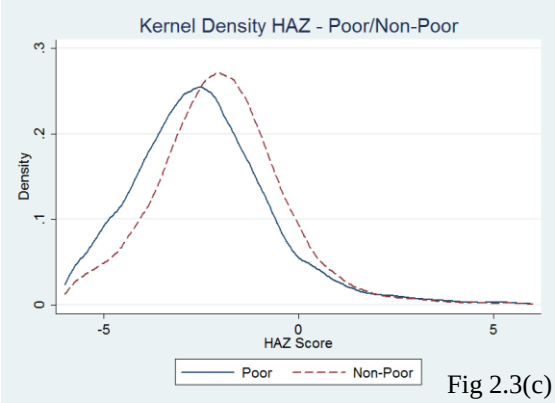



Fig 2.3(c)

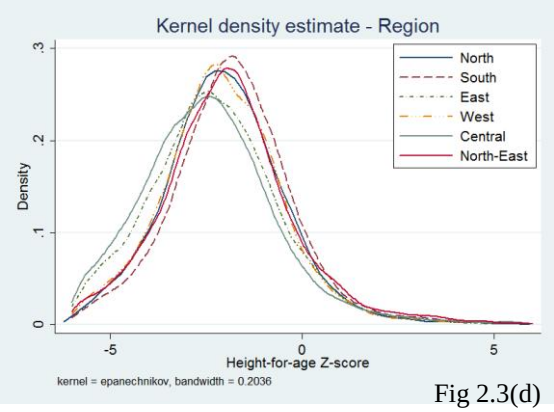



Fig 2.3(d)

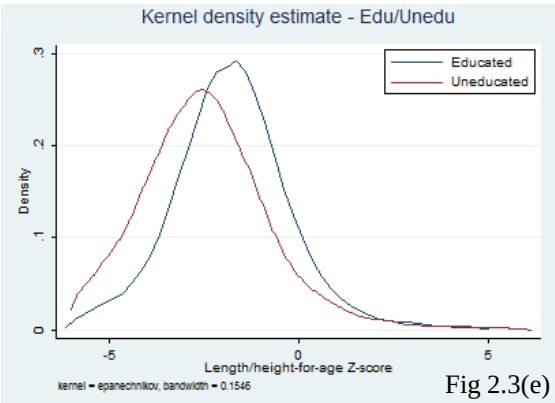



Fig 2.3(e)

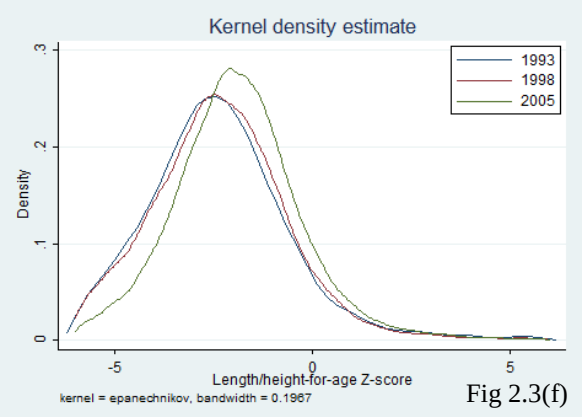



Fig 2.3(f)

Note: These graphs plot the kernel density estimates of HAZ Scores by various mother, year and region characteristics. Fig 3(a), 3(b), 3(c), and 3(e) plot the kernel density by mother characteristics like place of residence, age group, possession of assets and education respectively. Fig 3(d) plots the HAZ scores by region and 3(f) plots it by year of survey.

# Tables

## Table 2.1: Summary Statistics

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | **Obs** | **Mean** | **Std. Dev.** | **Min** | **Max** |
| **Infant Mortality** | 766364 | 0.0818 | 0.274 | 0 | 1 |
| **Neonatal Mortality** | 766364 | 0.0511 | 0.22 | 0 | 1 |
| **HAZ Score** | 72116 | -2.13 | 1.7 | -6 | 6 |
| **Stunted (HAZ<-2)** | 72116 | 0.546 | 0.497 | 0 | 1 |
| **Normal (HAZ>-1)** | 72116 | 0.22 | 0.414 | 0 | 1 |
| **WAZ Score** | 81018 | -1.84 | 1.3 | -5.99 | 4.93 |
| **Underweight (WAZ<-2)** | 81018 | 0.444 | 0.496 | 0 | 1 |
| **Normal (WAZ>-1)** | 81018 | 0.252 | 0.434 | 0 | 1 |
| **Age in Months** | 90818 | 30.5 | 12.4 | 10 | 59 |
| **Height (cm)** | 74795 | 82.9 | 12.17 | 0 | 922.2 |
| **Weight (Kg with decimal)** | 81499 | 10.48 | 2.91 | 0.5 | 97 |
| **Female** | 766364 | 0.48 | 0.499 | 0 | 1 |
| **Multiple Births** | 766364 | 0.013 | 0.113 | 0 | 1 |
| **Birth Order** | 766364 | 2.71 | 1.79 | 1 | 18 |
| **Mother's age** | 766364 | 35.4 | 7.74 | 14 | 49 |
| **Mother's age at birth** | 766364 | 23.3 | 5.12 | 13 | 48 |
| **Uneducated** | 766364 | 0.59 | 0.49 | 0 | 1 |
| **Rural** | 766364 | 0.68 | 0.47 | 0 | 1 |
| **Poor** | 759987 | 0.29 | 0.45 | 0 | 1 |
| **Non-Working** | 765768 | 0.612 | 0.48 | 0 | 1 |
| **Father Uneducated** | 761687 | 0.325 | 0.47 | 0 | 1 |
| **Hindu** | 766364 | 0.749 | 0.43 | 0 | 1 |
| **Muslim** | 766364 | 0.138 | 0.35 | 0 | 1 |
| **SC/ST** | 766364 | 0.294 | 0.45 | 0 | 1 |
| **Home Delivery** | 99187 | 0.639 | 0.48 | 0 | 1 |
| **Public Delivery** | 99187 | 0.198 | 0.40 | 0 | 1 |
| **Private Delivery** | 99187 | 0.156 | 0.37 | 0 | 1 |
| **Delivery, Skilled Personnel** | 99282 | 0.444 | 0.49 | 0 | 1 |
| **Access to piped water** | 766364 | 0.192 | 0.39 | 0 | 1 |
| **No access to toilet** | 766364 | 0.567 | 0.49 | 0 | 1 |
| **Electricity** | 760183 | 0.642 | 0.47 | 0 | 1 |

Note: Sample statistics of the variables are reported. Educated implies having attended any type of school and uneducated is defined as mother did not attend any school. Poor is defined by a mother not owning any asset as collected in NFHS. Rural and Urban are defined by the place of residence of mother during the time of interview. Female is 1 if sex of child is female. Multiple birth is a dummy variable indicating if the child is born in a multiple birth. It is 0 for a single birth and 1 for twins, triplets or quadruplets.

**Table 2.2: Selective Mortality and correction for selection**

| | Infant Mortality | | | Neonatal Mortality | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| **Dependent Variable** | **HAZ** | **HAZ** | **WAZ** | **HAZ** | **HAZ** | **WAZ** | **HAZ** |
| **Inverse Mills Ratio** | -1.42** | -1.38*** | -1.18*** | -1.99*** | -1.21* | -0.82* | |
| | (0.52) | (0.45) | (0.37) | (0.33) | (0.62) | (0.47) | |
| **Female** | 0.048** | 0.045* | 0.028 | 0.04* | 0.05** | 0.039* | 0.077*** |
| | (0.021) | (0.02) | (0.02) | (0.021) | (0.025) | (0.019) | (0.022) |
| **Multiple** | 0.221 | 0.21 | 0.121 | 0.34** | 0.091 | -0.05 | -0.27*** |
| | (0.22) | (0.17) | (0.127) | (0.15) | (0.19) | (0.14) | (0.06) |
| **Birth order 2** | -0.22*** | -0.24*** | -0.19*** | -0.24*** | -0.24*** | -0.18*** | -0.18*** |
| | (0.025) | (0.027) | (0.022) | (0.024) | (0.026) | (0.02) | (0.022) |
| **Birth order 3** | -0.37*** | -0.40*** | -0.29*** | -0.44*** | -0.40*** | -0.29*** | -0.33*** |
| | (0.028) | (0.032) | (0.028) | (0.032) | (0.034) | (0.03) | (0.026) |
| **Birth order 4** | -0.39*** | -0.41*** | -0.35*** | -0.49*** | -0.42*** | -0.35*** | -0.37*** |
| | (0.033) | (0.040) | (0.030) | (0.037) | (0.042) | (0.031) | (0.033) |
| **Birth order 5** | -0.50*** | -0.50*** | -0.42*** | -0.64*** | -0.53*** | -0.43*** | -0.49*** |
| | (0.057) | (0.047) | (0.032) | (0.058) | (0.048) | (0.034) | (0.046) |
| **Birth order 6** | -0.62*** | -0.60*** | -0.49*** | -0.74*** | -0.62*** | -0.52*** | -0.65*** |
| | (0.072) | (0.053) | (0.042) | (0.068) | (0.054) | (0.04) | (0.052) |
| **Birth order 7** | -0.72*** | -0.67*** | -0.48*** | -0.88*** | -0.71*** | -0.51*** | -0.76*** |
| | (0.086) | (0.066) | (0.049) | (0.078) | (0.065) | (0.05) | (0.065) |
| **Birth order 8** | -0.74*** | -0.67*** | -0.52*** | -0.88*** | -0.68*** | -0.55*** | -0.82*** |
| | (0.091) | (0.089) | (0.066) | (0.093) | (0.09) | (0.06) | (0.08) |
| **Birth order 9** | -0.86*** | -0.73*** | -0.62*** | -1.04*** | -0.79*** | -0.68*** | -0.95*** |
| | (0.145) | (0.095) | (0.08) | (0.153) | (0.92) | (0.077) | (0.10) |
| **Birth order>10** | -0.87*** | -0.70*** | -0.56*** | -1.03*** | -0.77*** | -0.61*** | -0.99*** |
| | (0.166) | (0.126) | (0.102) | (0.162) | (0.126) | (0.10) | (0.116) |
| **Piped water** | 0.15*** | 0.13*** | 0.13*** | 0.19*** | 0.14*** | 0.14*** | 0.17*** |
| | (0.018) | (0.021) | (0.014) | (0.017) | (0.021) | (0.014) | (0.018) |
| **No toilet** | -0.21*** | -0.21*** | -0.25*** | -0.29*** | -0.22*** | -0.26*** | -0.24*** |
| | (0.032) | (0.02) | (0.014) | (0.031) | (0.021) | (0.014) | (0.017) |
| **Electricity** | 0.117*** | 0.13*** | 0.13*** | 0.19*** | 0.14*** | 0.14*** | 0.15*** |
| | (0.029) | (0.019) | (0.015) | (0.027) | (0.02) | (0.016) | (0.02) |
| **State FE** | YES | YES | YES | YES | YES | YES | NO |
| **Cohort FE** | NO | YES | YES | NO | YES | YES | NO |
| **Number of groups** | 27 | 233 | 233 | 27 | 223 | 223 | |
| **Observations** | 68345 | 67886 | 76662 | 67532 | 67532 | 76274 | 68422 |

Note: Mortality selection effect is captured by the Inverse Mills Ratio. The procedure of deriving the ratio is described in the text. Mother cohort is defined by mothers grouped by rural or urban residence, education, poverty and age. The other control variables included in the specifications are age of child, mother's age at birth, female-specific birth order, birth order, birth month, year of survey, caste, religion, and state dummies. Standard errors clustered at the group level (cohort/state) are reported in brackets. Column (7) provides OLS estimates without any selection correction.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

**Table 2.3: Selection by HAZ and WAZ Profile**

| | HAZ | | | WAZ | | |
|---|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| **Dependent Variable** | **HAZ<-2** | **-2<HAZ<-1** | **HAZ>-1** | **WAZ<-2** | **-2<WAZ<-1** | **WAZ>-1** |
| **Inverse Mills Ratio** | 1.05*** | -0.493*** | -0.562*** | 0.578*** | -0.424*** | -0.432*** |
| | (0.087) | (0.073) | (0.071) | (0.085) | (0.079) | (0.092) |
| **Explanatory Variables** | YES | YES | YES | YES | YES | YES |
| **Cohort FE** | YES | YES | YES | YES | YES | YES |
| **Interview Dummy** | YES | YES | YES | YES | YES | YES |
| **Number of groups** | 233 | 233 | 233 | 233 | 233 | 223 |
| **Observations** | 67886 | 67886 | 67886 | 76662 | 76662 | 76274 |

Note: Mortality selection effect on HAZ and WAZ scores are evaluated at different cutoffs. The other control variables included in the specifications are sex of child, whether born in a multiple birth, mother's age at birth, birth order, female interacted with birth order, birth month, year of survey, whether the household has access to piped water, electricity and toilet, caste, religion, and state dummies. Mother cohort is defined by mothers grouped by rural or urban residence, education, being poor and age. Standard errors clustered at the mother cohort level are reported in brackets.
*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 2.4: Including dead kids in HAZ and WAZ

| | HAZ | | | WAZ | | |
|---|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| **Dependent Variable** | **HAZ** | **HAZ** | **HAZ-hat** | **WAZ** | **WAZ** | **WAZ-hat** |
| **Child's Age** | -0.004*** | -0.009*** | -0.005*** | -0.002*** | -0.006*** | -0.0035*** |
| | (0.0007) | (0.0015) | (0.006) | (0.0005) | (0.001) | (0.0003) |
| **Female** | 0.294*** | 0.30*** | -0.0065 | 0.27*** | 0.28*** | -0.28 |
| | (0.037) | (0.037) | (0.033) | (0.028) | (0.028) | (0.019) |
| **Multiple** | -0.149** | -0.216** | -0.342*** | -0.186*** | -0.25*** | -0.502*** |
| | (0.073) | (0.070) | (0.074) | (0.061) | (0.058) | (0.044) |
| **Interview Year** | | | | | | |
| **1998** | -0.021 | -0.041** | 0.180 | 0.103*** | 0.064*** | 0.0107 |
| | (0.018) | (0.018) | (0.016) | (0.013) | (0.013) | (0.009) |
| **2005** | 0.496*** | 0.374*** | 0.340*** | 0.279*** | 0.134*** | 0.203*** |
| | (0.016) | (0.016) | (0.014) | (0.012) | (0.012) | (0.008) |
| **Mother and House Controls** | NO | YES | NO | NO | YES | NO |
| **Observations** | 72116 | 70271 | 90818 | 81018 | 79141 | 90818 |

Note: HAZ and WAZ represent the available anthropometric scores of all the living children. HAZ-hat and WAZ-hat are the predicted HAZ and WAZ scores for the full sample including the dead children. The predicted values are calculated by a regression of HAZ and WAZ on child characteristics and controlling for mother fixed effects. The other child control variables included in the specifications are birth order, female specific with birth order, and birth month. Mother and house characteristics include mother's age at birth, caste, religion, rural or urban residence, education, being poor and mother's age. Standard errors clustered at the mother level are reported in brackets.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 2.5: Heterogeneity by gender and birth order

| | Infant Mortality | | | Neonatal Mortality | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Dependent Variable** | **HAZ** | **HAZ** | **HAZ** | **HAZ** | **HAZ** | **HAZ** |
| **Female** | -0.266 | | | 1.64 | | |
| | (0.739) | | | (1.13) | | |
| ***N*** | 32586 | | | 32417 | | |
| ***Number of Cohorts*** | 231 | | | 221 | | |
| **Male** | -2.43*** | | | -1.44** | | |
| | (0.73) | | | (0.657) | | |
| ***N*** | 35300 | | | 35300 | | |
| ***Number of Cohorts*** | 233 | | | 233 | | |
| **Birth order<=2** | | 1.96*** | | | 5.62*** | |
| | | (0.56) | | | (0.853) | |
| ***N*** | | 38190 | | | 38138 | |
| ***Number of Cohorts*** | | 217 | | | 212 | |
| **Birth Order>2** | | -2.94*** | | | -3.03*** | |
| | | (0.438) | | | (0.673) | |
| ***N*** | | 29696 | | | 29394 | |
| ***Number of Cohorts*** | | 219 | | | 209 | |
| **Birth order<=2 & Male** | | | 2.02*** | | | 4.52*** |
| | | | (0.691) | | | (0.95) |
| ***N*** | | | 19726 | | | 19698 |
| ***Number of Cohorts*** | | | 205 | | | 200 |
| **Birth order>2 & Male** | | | -3.29*** | | | -3.29*** |
| | | | (0.663) | | | (0.88) |
| ***N*** | | | 15574 | | | 15417 |
| ***Number of Cohorts*** | | | 213 | | | 203 |
| **Birth order<=2 & Female** | | | 2.13** | | | 7.61*** |
| | | | (1.06) | | | (1.40) |
| ***N*** | | | 18464 | | | 18440 |
| ***Number of Cohorts*** | | | 207 | | | 203 |
| **Birth order>2 & Female** | | | -2.64*** | | | -2.87*** |
| | | | (0.59) | | | (1.00) |
| ***N*** | | | 14122 | | | 13977 |
| ***Number of Cohorts*** | | | 212 | | | 202 |

Note: All the cells represent different regressions on a pooled sample of mothers in multiple surveys in India, according to different criterions. The sample is restricted to females or males in (1) and (4), lower birth order (<=2) or higher birth order (>2) in (2) and (5), and interaction between gender and birth order in (3) and (6). The control variables are sex of child, whether born in multiple birth, birth order, birth month, mother cohort fixed effects, interview-time dummies, mother's age at birth, caste, religion, piped water availability, access to toilet, and whether house has electricity. Standard errors clustered at mother cohort level are reported in brackets.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 2.6: Heterogeneity by place and time

| | Infant Mortality | | | Neonatal Mortality | | |
|---|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| **Dependent Variable** | **HAZ** | **HAZ** | **HAZ** | **HAZ** | **HAZ** | **HAZ** |
| **Kerala** | 0.649 | | | 2.86 | | |
| | (7.79) | | | (4.36) | | |
| ***N*** | 596 | | | 592 | | |
| ***Number of Cohorts*** | 81 | | | 79 | | |
| **Interview year=1993** | | -0.92 | | | -0.54 | |
| | | (0.81) | | | (1.15) | |
| ***N*** | | 20459 | | | 20337 | |
| ***Number of Cohorts*** | | 232 | | | 222 | |
| **Interview year=1998** | | -2.12** | | | -1.87 | |
| | | (0.93) | | | (1.18) | |
| ***N*** | | 16705 | | | 16639 | |
| ***Number of Cohorts*** | | 225 | | | 218 | |
| **Interview year=2005** | | -1.13 | | | -0.95 | |
| | | (0.712) | | | (0.928) | |
| ***N*** | | 30722 | | | 30556 | |
| ***Number of Cohorts*** | | 230 | | | 220 | |
| **Worse sex ratio states** | | | -7.2*** | | | -9.63*** |
| | | | (0.82) | | | (1.20) |
| ***N*** | | | 15133 | | | 15060 |
| ***Number of Cohorts*** | | | 227 | | | 217 |

Note: All the cells represent different regressions on a pooled sample of mothers in multiple surveys in India, according to different criterions. The sample is restricted to state of Kerala in (1) and (4), the three survey years in (2) and (5), and worse sex ratio states – Gujarat, Punjab, Haryana, New Delhi and Himachal Pradesh in (3) and (6). The control variables are sex of child, whether born in multiple birth, birth order, birth month, mother cohort fixed effects, interview-time dummies, mother's age at birth, caste, religion, piped water availability, access to toilet, and whether house has electricity. Standard errors clustered at mother cohort level are reported in brackets.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

**Table 2.7: Robustness - Different specifications for mortality selection**

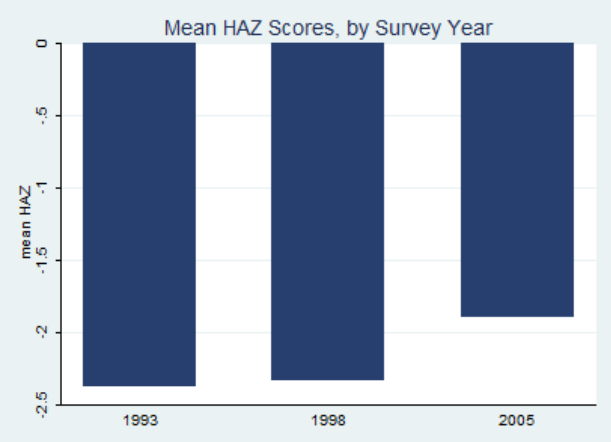
| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| **Dependent Variable** | **HAZ** | **HAZ** | **Height in CM** | **Weight in KG** | **HAZ** | **HAZ** |
| **Sample** | Including state time trends | Including year of birth, square and cubic terms of survey year | Replace the outcome variable, HAZ | Replace the outcome variable, WAZ | Include Antenatal Visits (infant) | <3years from interview date |
| **Inverse Mills Ratio** | -3.3*** | -1.51*** | -15.55*** | -1.27*** | -2.29*** | -1.87*** |
| | (0.37) | (0.46) | (1.33) | (0.37) | (0.324) | (0.53) |
| **Explanatory Variables** | YES | YES | YES | YES | YES | YES |
| **Cohort FE** | YES | YES | YES | YES | YES | YES |
| **Interview Dummy** | YES | YES | YES | YES | YES | YES |
| **Number of groups** | 233 | 233 | 233 | 233 | 229 | 226 |
| **Observations** | 67886 | 67886 | 70480 | 77125 | 57360 | 57025 |

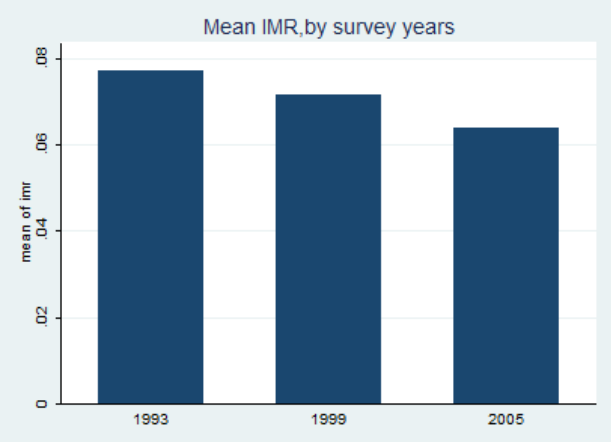
Note: All the cells represent different regressions on a pooled sample of mothers in multiple surveys in India, according to different criterions. The regression controls for state-time trends in addition to state and year of survey dummies in (1), (2) includes the year of birth along with age at birth and linear, quadratic and cubic controls for survey month-year, (3) replaces the HAZ outcome variable to height measured in centimeters and (4) measures the outcome in weight measured in kilograms instead of WAZ, (5) includes antenatal visits as a control in both first and second stage regressions, and (6) includes restricting sample to births within 36 months. The control variables are sex of child, whether born in multiple birth, birth order, birth month, mother cohort fixed effects, interview-time dummies, mother's age at birth, caste, religion, piped water availability, access to toilet, and house has electricity. Standard errors clustered at mother cohort level are reported in brackets.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

# Appendix

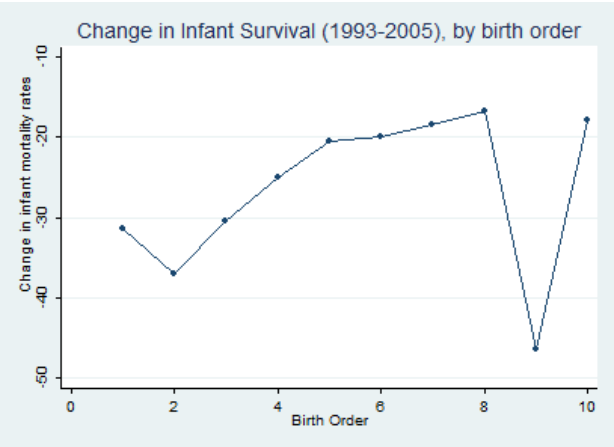
## Figure 2.A1: Mean HAZ Scores and IMR, by Survey Years

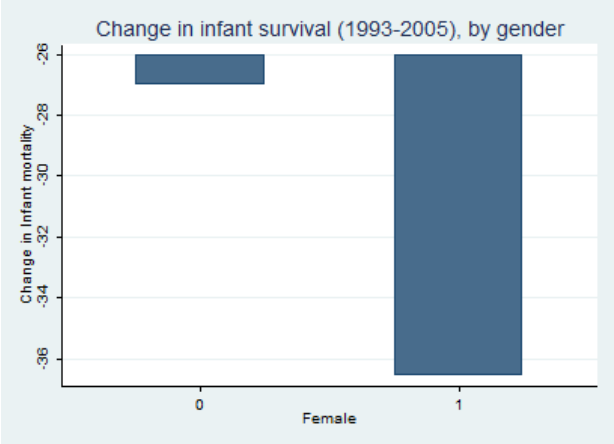



Note: These bar graphs plot the mean HAZ Scores and mean infant deaths by the three survey years – 1993, 1998 and 2005 in our sample.

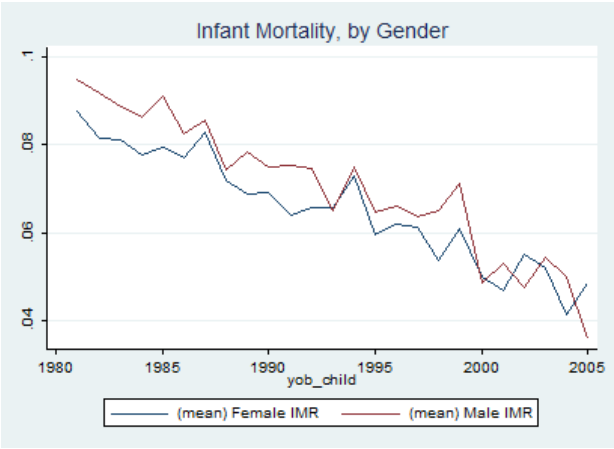
**Figure 2.A2: Infant Mortality Change, by Child Characteristics**

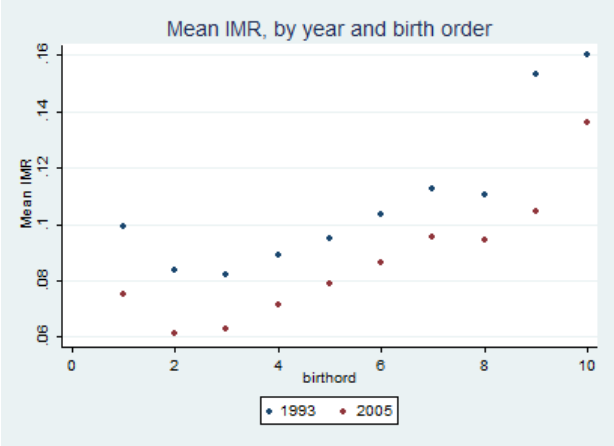



Note: The first graph shows how change in infant mortality varies by birth order. The changes in infant mortality are calculated by changes in mean infant deaths by birth order over 1993 to 2005, divided by mean infant mortality over that cell in 1993 and multiplied by 100. The second graph plots a bar graph of infant mortality changes for males and females where male is denoted by female=0.

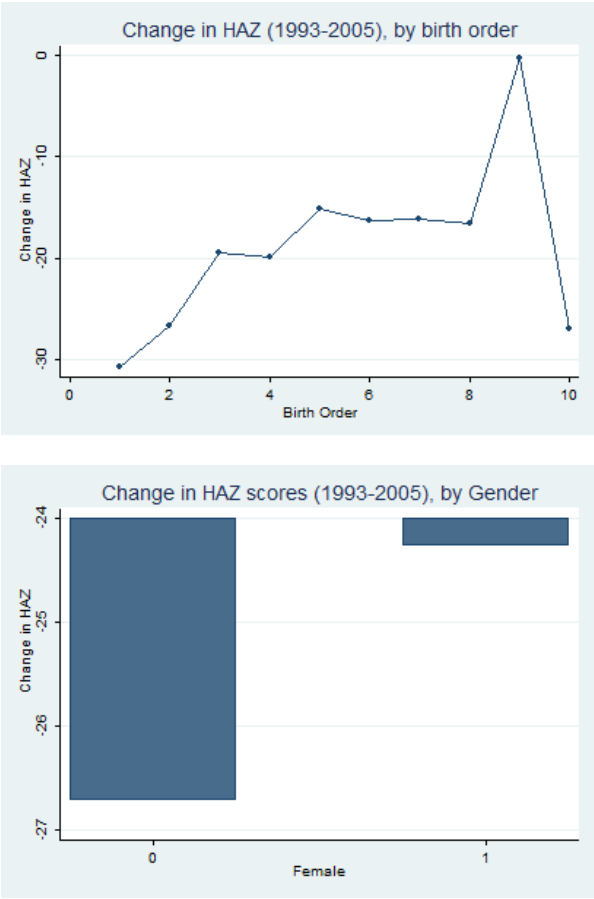
**Figure 2.A3: Infant Mortality Trends, by Gender and Birth Order**




Note: This graph shows the decline in infant deaths overtime (1980-2005) by gender. Panel 2 plots the mean infant deaths by birth order for the two survey years, 1993 and 2005.

**Figure 2.A4: Changes in HAZ (1993-2005), by Gender and Birth Order**




Note: The first graph shows how change in HAZ varies by birth order, where change is defined as the difference between mean HAZ by birth order between 1993 and 2005, divided by HAZ in 1993 and multiplied by 100. The second graph plots a bar graph of HAZ changes for males and females where male is denoted by female=0.

## Table 2.A1: First-Stage Probit Estimates (Marginal Effects)

| | (1) | (2) | (3) |
|---|---|---|---|
| **Dependent Variable** | **Survival till 1 year** | **Survival till 1 month** | **Survival till 1 year** |
| **Delivery** | 0.0076***<br>(0.002) | 0.0038**<br>(0.0019) | 0.0104***<br>(0.002) |
| **Female** | 0.010***<br>(0.003) | 0.007***<br>(0.002) | 0.010***<br>(0.003) |
| **Multiple** | -0.24***<br>(0.012) | -0.195***<br>(0.013) | -0.24***<br>(0.012) |
| **Piped water** | 0.007<br>(0.004) | 0.004<br>(0.023) | 0.009**<br>(0.004) |
| **No toilet** | -0.006***<br>(0.002) | -0.006***<br>(0.016) | -0.008***<br>(0.003) |
| **Electricity** | 0.007***<br>(0.002) | 0.007***<br>(0.002) | 0.01***<br>(0.002) |
| **Mother's Age at Birth** | 0.0014**<br>(0.0005) | 0.0011***<br>(0.0002) | 0.0022**<br>(0.0002) |
| **State FE** | YES | YES | YES |
| **Cohort FE** | YES | NO | NO |
| **Observations** | 93999 | 94614 | 94614 |

Note: These are the first stage probit estimates of deriving the inverse Mills ratio. Other included variables are birth dummy, female interacted with birth dummy, survey year, month of birth, caste, religion, and state dummies. Mother cohorts are also included in (1). Standard errors are clustered at the state level.

# Chapter 3

## Roads to Better Health: Linkages between Infrastructure and Child Health Outcomes in India*


### Abstract

Many people in rural areas in developing countries live at a distance from an all-weather road. This constrains their ability to access health care facilities and markets. This paper studies the impact of road construction on health outcomes of those affected by a large scale rural road construction program in India. This study creates a new dataset involving health and road construction information at the village level for India and estimates the causal outcome using two different estimation techniques to overcome any issue in identification and endogeneity of road placement. In the short run, I do not find any statistically significant effect of road construction on child health outcomes like infant and neonatal mortality and sex ratio at birth.



***Keywords:*** Infant Mortality, Child Health, Infrastructure, India.

***JEL Codes***: I15, O18, J13.



* I am grateful to Anil B. Deolalikar, Aman Ullah, Mindy Marks, and Joseph Cummins for their comments and guidance. I am extremely grateful to Paul Novosad and Sam Asher for sharing their PMGSY matched data. All remaining errors are my own.

## 3.1 Introduction

Lack of access to the nearest market or health center significantly constrains people's economic and social opportunities. About 900 million rural dwellers worldwide live more than two kilometers—a 20 to 25 minute walk— from the nearest all-weather road (World Bank 2006). Henceforth, significant budgetary allocations are made towards infrastructure development in developing countries. In 2000, 40 percent of villages in India lacked an all-weather road. Access to transportation infrastructure can increase labor force participation, increase incomes, create access to a better variety of goods, increase domestic and international trade but on the other it may increase pollution or increase inequality (Akee, 2006; Qian et al., 2012; Limão and Venables, 2001; Bohemen and Lak, 2003). There are few empirical studies estimating the effect of rural road construction on child health.[1] Empirically, it is difficult to establish causality as road placement is likely to be endogenous to other socio-economic factors which also affect development. This paper estimates the effect of being exposed to road construction on infant and neonatal mortality by developing a unique matched micro panel dataset on both health outcomes and road construction in India.

Recently, India experienced a massive rural road building program – Pradhan Mantri Gram Sadak Yojana (PMGSY) - which was launched in 2000 with the aim of providing all-weather access roads to all villages with a population of 500 or more. By March 2011,

[1] There have been studies detailing the effects of road construction on economic growth, consumption, schooling, employment (Mu and Walle, 2007; Khandker et al., 2009; Aggarwal, 2014) but the effect on health is understudied.

over 420,000 km of roads had been sanctioned to connect nearly 110,000 habitations at a total cost of 1.19 trillion INR (about $27 billion) (Ministry of Rural Development, 2012). PMGSY is a large scale India-wide program to provide rural roads with a priority to provide access to villages with a population of 1000 or above and consequently provide roads to villages with a population of 500 or above. This phase-wise allocation of roads to different villages benefitting from the road construction at different points of time, based on their population, permits me to carry out two different empirical strategies to derive a causal estimate.

Estimating the causal impact of infrastructure on any outcome is complicated by the fact that infrastructure is not allocated randomly. These placements are generally endogenous. For example, places which need the infrastructure most and are deemed the most backward would get it first. Alternatively, endogeneity can also arise if program placement is a function of time-varying factors. For example, the cities with maximum growth potential get the better road which makes it difficult to disentangle the effect of roads versus the growth trajectory. Endogeneity is still a concern in the impact evaluation of road construction at the micro level since there might be unobserved but geographically determined individual characteristics which are also correlated with the program placement.[2] In the economic literature, identification has come from a use of instruments based on historical routes or from calculating straight line distance. Apart from the fact that valid instruments are hard to find, it is also difficult to use these

[2] For more details on problems in impact evaluation of road construction, refer to van de Walle (2009).

measures for identification when infrastructure creation leads to path dependencies.[3] The ideal scenario would be to be able to exploit if road projects are randomized but given the intensive investment, random assignment of roads to communities is not feasible. Instead, in this paper, I use the quasi-random experiment of the roll-out of roads under PMGSY and differential access to these roads by the villages at different points of time.

India does not have any nationwide survey individual-level health data which has village identifiers. But, PMGSY was implemented at the village level. To get a cleaner identification of treatment, individual-level data from District Level Household Survey (DLHS-3) is matched with the 2001 Census Data to derive village identifiers and then is matched with the road construction data. This paper then develops a causal estimate by using two different empirical methods. First, using retrospective birth histories from District Level Household Survey (DLHS-3), a within-mother variation in survival of infant is carried out before and after the road construction; which in turn develops a causal analysis of effect of road construction on health. Second, given the roll-out of the program concerning different thresholds, a fuzzy regression discontinuity design is invoked. The paper's findings suggest that road construction under PMGSY does not bring about a statistically significant fall in infant and neonatal deaths in the short run as may be expected. In all the specifications the effect is greater in magnitude for villages with population of 1000 or more. I also do not find any evidence of change in sex ratio at birth, as measured by the probability of a male child being born after the village is affected by PMGSY.

[3] Berger and Enflo (2013); Bleakley and Lin (2012)

The rest of the paper is organized as follows. Section 2 elaborates on the program specification and implementation of PMGSY and the linkages through which road building is expected to affect health outcomes. Section 3 summarizes and describes the data used and section 4 provides details on the different empirical specifications used. Section 5 presents the empirical results and section 6 concludes.

## 3.2 Background

### 3.2.1 Pradhan Mantri Gram Sadak Yojana (PMGSY)

Government of India launched the Pradhan Mantri Gram Sadak Yojana (PMGSY) on 25th December, 2000 to provide all-weather access to unconnected habitations. The goal of the PMGSY was to provide connectivity, by an all-weather road, such that all unconnected Habitations with a population of 1000 persons and above were covered in three years (2000-2003) and all unconnected habitations with a population of 500 persons and above by the end of the Tenth Plan Period (2007).[4] The 2001 Census population was the basis for determining the population size of the Habitation as well as eligibility for the program. Precedence was given to new connectivity with priority to connect villages with higher population. The PMGSY also permitted upgradation of the existing roads in those districts where all the eligible habitations of the designated population size had already been provided all-weather road connectivity. In upgradation work, priority was given to

[4] The threshold for hill states, desert and tribal areas was a population of 250.

through routes of the rural core network.[5] The District Panchayat approved the core network. The program had a huge scope with an investment of USD 26 billion and connecting 170,000 eligible habitations across India. The program has recently been extended to include all habitations with populations above 100.

The unit for this program was a habitation, which is different from a revenue village or a panchayat. A habitation is a cluster of population, living in an area, the location of which does not change over time.[6] An unconnected habitation was defined as a habitation located at a distance of at least 500 meters or more from an all-weather road or a connected habitation.[7] The eligible Unconnected Habitations were intended to be connected to nearby Habitations already connected by an all-weather road or to another existing all-weather road so that services like educational and health facilities and markets, which were not available currently, become available to the residents. The core states that benefitted most under PMGSY were Madhya Pradesh, Uttar Pradesh, Chhattisgarh, Bihar, Rajasthan, Orissa, West Bengal, Jharkhand, Assam and Himachal Pradesh.

[5] A Core Network is a minimal network of roads that is essential to provide basic access to essential social and economic services to all eligible habitations in the selected areas through at least single all-weather road connectivity. A Core Network comprises of Through Routes and Link Routes. Through Routes are the ones which collect traffic from several link roads or a long chain of Habitations and lead it to marketing centers. Link Routes are the roads connecting a single Habitation or a group of Habitations to Through Routes or District Roads leading to market centers. (Source: PMGSY Guidelines)
[6] PMGSY Guidelines (2004)
[7] 1.5 Kilometers in case of hills

### 3.2.2 Linkages

Providing all-weather rural roads to unconnected villages can affect child and maternal health by direct and indirect mechanisms. Roads provide an easier access to health centers by reducing the travel time and transportation cost. Studies have shown that a reduction in travel time and cost leads to increase in use of formal health care (Wong et al., 1987; Adhvaryu and Nyshadham, 2014). Brenneman (2002) surveys the literature on transportation and health and finds that better transport leads to easier access to health care, better staffing in hospitals and easier operation of clinics. Improved access to hospitals and utilization of formal health care will lead to a reduction in child mortality.

Building of roads has shown to increase household per capita income and consumption (Khandker et al., 2009; Gibson and Rozelle, 2003). Gibson and Rozelle (2003) analyze how infrastructure could affect in reducing poverty in Papua New Guinea. They analyze the consumption and price effects of access to roads and find that a 1-hour increase in travelling time decreases real consumption by almost 10%. And hence, an improvement in transportation could be pivotal in poverty reduction. Access to rural roads, specifically in the context of India, leads to variety gains in consumption as well (Aggarwal, 2014). Increased variety in food may also result in increased diversity of micronutrient consumption, which can be important for maternal and child health in countries where micronutrient deficiency is endemic.[8]

[8] Cutler, Deaton, & Lleras-Muney (2006); Porto (2004)

More broadly, rural roads create access to markets and local market development (Mu and Walle, 2007). Mu and Walle (2007) find a sizable increase in perishable and non-perishable food items and also find heterogeneous impacts with poorer communes benefitting more by a rural road construction project in Vietnam. Not only does roads create access to farther off markets, it also can increase agricultural produce by local farm households and decrease cost of living, with a fall in consumer prices (Jacoby, 2000). A study in India using data from the International Crops Research Institute for the Semi-Arid Tropics from 1970 through 1974, with 5,450 observations of irrigated and non-irrigated districts showed that for irrigated districts, road density (km roads/1000 sq km cropped land) was strongly and positively correlated with agricultural output in irrigated zones. With a decrease in cost of living and a resource-constrained rural household setting, an increase in disposable income could be expected which could be applied to accessing health care and preventing child deaths.

At a macroeconomic level, building of transport infrastructure boosts trade (Buys et al. 2006; Donaldson, 2013; Duranton et al., 2013). The production and trade in goods and services, relies heavily on an efficient distribution network. Therefore, road construction will aid trade and improve production. Datta (2011) studies the impact of highway construction in India on the firms' inventories and finds that building of the highway leads to a reduction in inventory of affected firms. Reduced trade barriers also increase the demand for skilled labor. Michaels (2008) uses the U.S. Interstate Highway System as an experiment to identify the labor market effects of reduced trade barriers and finds that in rural counties that had a highly educated workforce, highways increased the relative

wage-bill of nonproduction workers, and where the workforce was less educated, highways decreased the relative wage-bill of nonproduction workers pointing to a heterogeneous effect. Trade policy has been shown to have a causal effect on child health outcomes. Panda (2014) shows that improved exports from sub-Saharan African countries under African Growth and Opportunity Act helps in decreasing infant and neonatal mortality by change in household assets, change in female employment in labor force and changing health seeking behavior of mothers. This indirect channel of increasing trade by access to all-weather roads will improve child health.

Change in transportation costs play an important role in growth of cities and economic activity (Storeygard, 2012; Limão and Venables, 2001; Hummels and Skiba, 2004; Qian et al. 2012; Chandra and Thompson, 2000). Qian et al. (2012) show that proximity to transportation networks has a large positive causal effect on per capita GDP growth rates across sectors. Construction of rural roads helps in decreasing transportation cost and therefore boosting economic activity. Calderon and Serven (2004) evaluate the impact of infrastructure development on economic growth and income distribution using a panel data set for 121 countries for 1960-2000 and show that infrastructure development can be effective in combating poverty. Labor is an important factor in income of rural households in developing countries. Imperfections in rural labor market have been viewed as an impediment to income and employment of rural households (Bardhan, 1979). With improved infrastructure, decreased search costs and access to faraway labor markets, even part-time agricultural workers will have the opportunity to improve their living conditions (Michaels, 2008). Improved economic activity, increased opportunities

of employment and increasing incomes will improve health outcomes by mechanisms of improved nutrition and improved access to sanitation and health care.

While there has been no work on studying the effects of road on child health, there has been some work on the effect of community infrastructure like access to electricity, improved sanitation and public health expenditure on child mortality (Wang, 2003; Galiani et al., 2005; Fay et al., 2005, Benefo and Schultz, 1996). Fay et al. (2005) use Demographic and Health Surveys for 39 developing countries to generate an index of infrastructure access involving household with a sanded floor, access to water, without sanitation and with access to electricity as a proxy for exposure to pathogens and find that community infrastructure is significant in determination of stunting and child mortality. Aggarwal (2014) studies the effect of PMGSY on outcomes like school enrollment, variety gains in consumption, price dispersion, women's labor force participation, and adoption of new technologies. She finds a significant impact on the consumption basket as well as increased women's labor force participation. Since these are some of the mechanisms through which road construction would affect child health, a priori I expect child health to improve in the districts and villages affected by PMGSY.

This paper contributes to the literature in analyzing the effect of rural road construction on child health outcomes in rural areas. To the best of my knowledge, these are the first estimates of PMGSY on child health outcomes.

## 3.3 Data

To study the effect of road construction on child health, a micro panel dataset is developed by matching the household survey data for health information (DLHS-3), the data on road construction (OMMS) and Census 2001 Data. Moreover, the reproductive birth history from DLHS-3 is used to create a panel dataset from the survey.

### 3.3.1 PMGSY Data

Effective monitoring of the program was carried out with the officials are required to furnish, online, all the data and information, as may be prescribed by the NRRDA from time to time, in the relevant module of the Online Management & Monitoring System (OMMS). Data for PMGSY is obtained from the OMMS. This includes data on roads built, habitations affected each year, baseline habitation connectivity status, road sanctioning dates, and population. This data is reported either at the habitation level or the road level.

### 3.3.2 Health Data

The micro level health data comes from the third round of the District Level Household Survey (DLHS-3) of India. DLHS-3 is a district level survey conducted between December 2007 and December 2008 and provides estimates on maternal and child health, family planning and other reproductive health services. The DLHS Surveys is one of the largest ever demographic and health surveys carried out in India, with a sample size of about 720,320 households covering all districts of the country. Information regarding

child health, immunizations, antenatal care, etc. is found in the surveys, along with mother and household characteristics.

DLHS-3 collected data using four types of questionnaires – household, village, individual "ever married" women (15-49 years of age), and individual "never married" women (15-24 years of age). Household questionnaire is used to collect data on household dwelling units, religion, caste, head of the household etc.; while ever-married women's questionnaire is used to collect data from eligible women about the characteristics, reproductive behavior, antenatal and prenatal care, contraception, reproductive health etc. The village questionnaire collects information regarding the availability and accessibility of various services. Women of reproductive age (15-49 years) are interviewed about the date of birth and death (if applicable) during the last three years preceding the survey date.[9] This kind of retrospective survey gives an opportunity to build a panel dataset of mothers, with the time dimension being the year of childbirth given by each mother.[10]

Even though DLHS collects data at village level, no village identifiers are available in the dataset. But, for each village DLHS-3 records the population according to 2001 Census. Therefore, by using village population I match the 2001 Census Data to DLHS-3 Individual file to get village names and codes for each observation. The matching is done at the sub-district level to avoid any duplicates as well as all villages are dropped within a

[9] This gives a complete list of all the pregnancies the woman has had since January 1, 2004.

[10] One problem that can be raised with the recall data is the measurement error problem. Since the birth histories do not go too much into the past, it is lesser of a problem in this case. This ensures that the siblings are not very far apart in age and hence are broadly comparable. Moreover, since deaths of a kid are important in a mother's life, this variable should be recorded without much measurement error. Even if measurement error exists, it is improbable that it is systematically correlated with implementation of PMGSY.

sub-district if they have the same population. This process of matching enables me to uniquely match 18,513 of the 22,825 (81%) villages interviewed in DLHS-3. This individual-level data is also matched to village questionnaire data to obtain village level characteristics.

The PMGSY data is recorded at the habitation level, but it does not record the corresponding village names or codes.[11] Therefore, a matching of this data is carried out with the 2001 census data by matching names of villages and habitations, where village codes are missing, by using a fuzzy matching algorithm.[12] After matching village names to each habitation at the sub-district level, village codes are obtained and are subsequently matched with the merged DLHS-3 data now having the village codes.

Treatment in this paper, for the household fixed effects specification, is defined as a child's exposure to PMGSY. This is a dummy variable which takes the value 1 if the child is born in a village *after* PMGSY is implemented in the village. This ensures that the child has been fully exposed to the road construction through their lifespan. For example, if PMGSY mandated building of a road in village $V_1$ in 2004-2005, then PMGSY takes the value of 2005 for $V_1$. A child is then said exposed to AGOA if in $V_1$; they are born in 2005 or later. A dummy variable indicating if the child has died before reaching the age of 1 year is constructed, based on mother's birth history. This will be the indicator for *individual-level* infant mortality.

[11] This data is taken from the state reports at http://pmgsy.nic.in/

[12] This data has been compiled by Paul Novosad and Sam Asher by developing a hindi-name matching algorithm. I'm thankful to them for providing me with this data to develop the new dataset by matching with DLHS-3.

Since PMGSY was applicable to only rural areas, I restrict the data to the rural sector. After the matching and dropping data for children born within twelve months of the survey, to ensure full exposure for every child in the sample, the sample includes 151,370 children born to 122,526 mothers. Infant (Neonatal) mortality rate is the number of deaths of children before reaching the age of one year (month) per 1000 live births. In this sample, it is calculated by multiplying the sample mean child deaths (in the appropriate age group) by thousand. The sample average infant mortality rate is 3.9% of live births while the sample neonatal mortality rate is 2.3% of live births. Since determinants of neonatal mortality may differ from infant mortality, an indicator for child dying before the age of 1 month is also constructed and effect of AGOA on it is studied.

In Table 3.1, I show the summary statistics. In the sample, 46% children are females. In terms of mother characteristics, 51% of women did not attend any type of school, 76% are hindu, 41% belong to Other Backward Classes (OBC) and 48% have low wealth index. Since DLHS does not collect income or wage data, the wealth index is constructed by using data on asset possession by the individuals and is reported in DLHS. 50% of the surveyed individuals live in villages which has access to electricity for more than 6 hours a day and 78% villages are connected to a health sub-center. The average distance to the nearest town is about 15 kilometers and villages on average are about 40 kilometers away from a railway station.

## 3.4 Empirical Strategy

### 3.4.1 OLS with household fixed effects

To analyze the effect of road construction under PMGSY on infant mortality, I estimate the following equation using a linear probability model[13]:

$$IMR_{imvt} = \alpha_m + \beta_t + \gamma T_{vt} + X_{imvt}\delta + \mu_d.t + \varepsilon_{imvt} \quad (1)$$

for child i, born to mother m in village v in year t. IMR is a dummy which takes the value 1 if child i dies before reaching the age of 1 year, $\alpha_m$ is mother fixed effect, $\beta_t$ is birth-year fixed effect and $\mu_d.t$ captures the district-time specific trend.[14] $T_{ct}$ takes the value 1 if the specific village had a road constructed under PMGSY at time t. $X_{imct}$ is a vector of control characteristics including sex of the child, whether or not they are born in multiple births (i.e. twins, triplets, etc.), dummies for their birth order, mother's age at birth and birth month of the child. $\gamma$ provides the estimate of the effect of PMGSY on infant mortality. The standard errors are clustered at the district level to take into account any correlation of the error across space and time within each district.

It may be of concern that the villages may have selected into getting the roads or the program guidelines were not followed. I plot the probability of getting a road by village population and it is observed that there are jumps at both 500 and 1000 population levels

[13] Due to easier interpretation of LPM estimates, I present those in the results. Also, LPM Model allows me to use a general form of mother fixed effects and since the estimations are deviations from the trend (after controlling for country-time control), the coefficient is likely to be small and should lie in the range of 0 to 1.

[14] Since the plans were drawn at the district levels, I control for district-time trends to control for unobserved factors at the district level that may change with time and could be correlated with road construction.

as well as probability of getting a road for villages with a population of thousand or higher is more than that of villages with a population of 500. Another concern would be corruption involved in such a wide-scale program in India. It could be that the money allocated to road building was spent elsewhere, without any road being built. But, given the monitoring of the project and involvement of state and central governments, it seems unlikely.

In addition, to alleviate any concerns about program placement based on prerequisites, I control for household (and in turn district or village) unobserved characteristics. Time-invariant heterogeneity regarding geography, history, culture, politics and attitudes etc. are taken care of by the mother fixed effects ($\alpha_m$). The year fixed effects ($\beta_t$) control for an aggregate time variation involving improvement of health technology and year shocks. The district specific trends ($\mu_d.t$), in fact, also allow district specific improvement in infant and maternal health i.e. differential states of development of the districts in different states.

Another outcome that this study is interested in is the sex ratio at birth. India has historically seen skewed sex ratios with lower female births than male births and it is especially acute in certain states. To measure the effect of road construction on sex ratio, I estimate:

$$SR_{imvt} = \alpha_m + \beta_t + \gamma T_{vt} + X_{imvt}\delta + \mu_d.t + \varepsilon_{imvt} \quad (2)$$

Here, the outcome variable is an indicator variable taking the value of 1 if the child is male and 0 if child is female. The time of treatment for the sex-ratio regression is 1 year prior to when the road was built and implies treatment at conception.

### 3.4.2 Fuzzy Regression Discontinuity Design

Since the program rollout was according to village population cut-offs with cut-offs at 1000 and 500, I can also exploit a regression discontinuity design framework. Even if selection into PMGSY is based on political or economic factors, these factors should not change discontinuously at the cut-off population levels and this will provide a causal evidence of the effect of road construction on outcome variables. Since the program was a large scale program, some deviations from the plans are expected. For example, if a village lies in the route of construction of connecting a bigger village to link road, even though it has lower population and lesser priority. The likelihood of getting a road in the village should be a function of the population size of the village, as per 2001 Census, and it should discontinuously jump at the population thresholds. Keeping this in mind I use the Fuzzy Regression Discontinuity Design framework.

In Fuzzy RD, the discontinuity becomes an instrumental variable for the treatment status (Angrist and Pischke, 2009). I therefore use the two-stage least squares estimation strategy:

First Stage equation:

$$T_{ivd} = \alpha + f(X_{vd}) + \lambda_4 * C_1 + \lambda_5 * C_2 + \mu_d + \delta_{ivd} \quad (3)$$

Second Stage equation:

$$IMR_{ivd} = \beta o + \beta_1 * X_{vd} + \beta_2 * X^2_{vd} + \beta_3 * X^3_{vd} + \beta_4 * C_1 + \beta_5 * T_{ivd} + \mu_d + \varepsilon_{ivd} \quad (4)$$

Where, C1 = first cutoff, at 1000 and,

$$IMR_{ivd} = \beta 0 + \beta_1 * X_{vd} + \beta_2 * X^2_{vd} + \beta_3 * X^3_{vd} + \beta_4 * C_2 + \beta_5 * T_{ivd} + \mu_d + \varepsilon_{ivd} \quad (5)$$

Where, C2 is the second cut-off at 500

Here, X is the running variable i.e. the population of village v in district d, T is treatment of child and his/her village i.e. the village gets a road by the time of DLHS-3 survey (2008), C1 is an indicator variable indicating the village population is greater than 1000 and similarly C2 is an indicator variable for population greater than 500. $f(X_{vd})$ represents a flexible control for village population.[15] The outcome is measured by infant or neonatal deaths of children living in village v in district d. Though village controls, mother and child controls, and district fixed effects are not required for identification, these are included to improve efficiency in the second stage regressions. $\beta_5$ in the second stage equations are the estimated effect of treatment on infant mortality at cutoffs C1 and C2 respectively. Having two separate second stage equations lets me estimate separate local average treatment effects for the two cutoffs.

[15] I include first, second and third degree polynomial of the population variable in my estimation.

## 3.5 Results

### 3.5.1 Regression Discontinuity Graphs

For the Fuzzy RD design to be valid there should be a discontinuous jump in probability of assignment of roads at the cutoffs. Figure 3.1 plots the probability of road construction/sanction by village size and we observe discontinuity and a jump at cutoffs of 500 and 1000, validating the use of a Fuzzy RD design. Moreover, the average values of outcome variables infant and neonatal mortality display a discontinuity at the same population levels of 500 and 1000, as is shown in Figure 3.2. The outcome variables do not look discontinuous at other population levels as well.

To ensure that it is only the assignment of road that brings about the changes in outcome variable at the population thresholds, and other things are not changing, I graph the covariates in Figure 3.3. The y-axis plots the proportion of villages having access to health sub-center, primary health clinic, access to more than 6 hours of electricity in a day, and villages having drainage facility. The scatterplots show no apparent change in density at the cutoff level of 1000.

Lastly, to check that there is no "gaming of the cutoff" that is villages in order to get the road report their populations to be just above the cutoff, I plot the population density graphs at 500 and 1000 village population cutoffs. Given that the plans were drawn based on Census 2001 population level, it is unlikely to be a problem. If in fact we see a huge bump in density around the cutoffs, it is a cause of concern and invalidates the design.

Figure 3.4 plots the density functions. We do not observe any special increase in density of villages with population of 500 or 1000.

### 3.5.2 Regression Estimates

Table 3.2 presents the results of the OLS specification. Column (1) – (3) present the effect of road construction on infant mortality. Column (1) controls for district fixed effects and year fixed effects along with controls for mother characteristics like religion, caste, education of mother, and being poor and village characteristics like number of households in the village, distance from nearest town, distance to district headquarter, distance from nearest railway station, source of drinking water in village, village connected to health sub center, primary health clinic or dispensary. The coefficient on the treatment is negative, small in magnitude and not statistically significant at the conventional levels. Column (2) adds the household fixed effects instead of district fixed effects, and even though the absolute magnitude of the coefficient increases, significance is unchanged. Column (3) includes district time trends as well without much change in the coefficients. Columns (4) – (6) show the results for neonatal mortality. Again the coefficients are small in magnitude and not statistically significant.

Next, I present the results from the Fuzzy RD analysis. First, to check if the cutoffs are valid instruments, I present the first stage results in Table 3.3. The two cutoffs at 500 and 1000 are highly significant with F-statistics much greater than 10 for both of them. This regression controls for linear, squared and cubic term for population, and village controls for number of households in the village, distance from nearest town, distance to district

headquarter, distance from nearest railway station, source of drinking water in village, village connected to health sub center, primary health clinic or dispensary along with district fixed effects.

Table 3.4 presents the results of effect of road construction on infant mortality using the cutoffs at 500 and 1000 as instruments in a two-stage least squares method following a fuzzy RD design. These regressions include controls for village and mother characteristics to improve efficiency. Since the conditional distribution of covariates is continuous at cutoffs, inclusion of these in the regression should not change the effect of treatment on the outcome. Moreover, the presence of covariates includes the precision if these covariates are in fact linked to the outcome of interest. The effect of road construction as measured at the cutoff of 500, with and without village and mother and child level controls, are similar at 18 deaths per thousand reduction in infant mortality albeit not statistically significant. The effect is higher in magnitude at around reduction of 24 deaths per thousand at cutoff of 1000 but is again not statistically significant at conventional levels. The coefficient in column (5) is significant at 10% level, when no mother-child characteristics are included in the regression. With a mean sample infant mortality of about 39 per thousand, this implies a reduction by around 38% of sample mean.

The Fuzzy RD estimates of effect of PMGSY on neonatal mortality are presented in Table 3.5. The magnitudes are similar across specifications with road construction having a greater effect on villages with population nearing 1000 than at 500. Road construction

reduces neonatal mortality by 5-7 deaths per thousand in villages with population around 500 while it decreases neonatal deaths by 13 deaths per thousand in villages with population of 1000. In both cases though, the effects are not statistically significant.

I also present the estimates on sex-ratio as an outcome variable. This is defined as probability of male child being born if the mother is affected by PMGSY during conception in the OLS specification. In the Fuzzy RD framework, since this is a cross-section with no time dimension, treatment is defined as any mother being affected by PMGSY till the survey year. Table 3.6 presents both the OLS and the Fuzzy RD results. It is important to keep in mind that the OLS results are more generic in terms of the effect on overall population while the Fuzzy RD estimates are more localized and true for villages with similar characteristics across the cutoffs. The OLS estimates with controls for household fixed effects imply the probability of being born male increases by 1.5 percentage points, though it is not statistically significant. The Fuzzy RD estimates show that the probability of being male decreases by about 4.9 percentage points at the cutoff of 500 (significant at 10% level) while it decreases by about 1 percentage point when the cutoff is 1000.

## 3.6 Conclusion

Road infrastructure can affect the health of a population by giving them access to health care, variety of goods and services, and markets. This paper studies the impact of road construction on infant and neonatal mortality of those affected by a large scale rural road construction program in India. To study this effect a unique micro dataset involving

individual level nationwide health data and road construction data is created. The analysis suggests that road construction does not have a statistically significant impact on infant and neonatal mortality.

This could be due to various reasons. First, since this is a short term estimation of effect the gains of road construction may not have been realized yet. It is required that for roads to have an immediate effect on infant and neonatal mortality, the mechanisms should have a dramatic change almost immediately. This may not be the case. The growth of cities and marketplace as well as increased trade activity may show some lag in responding to this change, which a short term study like this is not able to capture. Second, it is possible that even though the road has been built, it is not specifically used to access better health care due to household financial constraints. Access to healthcare or markets does not imply a necessary change in healthcare behavior especially by financially constrained households. Third, even though roads have been provided, individuals may not have proper access to means of transport to utilize the roads. For example, households need a cycle or car or public transport like buses to visit a hospital which is farther off. Construction of roads reduces time cost to access those places but does not provide means of access.

A longer horizon study in studying the effect of road construction on health outcomes is therefore called for. Unfortunately, nationwide health data available in India till now do not provide us with enough detailed village level information with a wide time range to carry on this type of analysis at this point in time. It will also be interesting to analyze the

mechanisms at work to observe if there is an actual change in the medium through which this program would be able to bring a significant change in the health outcome variables.

## References


Adhvaryu, Achyuta, and Anant Nyshadham. 2014. “Returns to Treatment in the Formal Health Care Sector: Evidence from Tanzania.” Accepted, AEJ: Economic Policy.

Aggarwal, Shilpa. 2014. “Do Rural Roads Create Pathways out of Poverty ? Evidence from India.” Working Paper.

Ahmed, R, and M Hossain. 1990. “Development Impact of Rural Infrastructure in Bangladesh.” International Food Policy Research Institute.

Akee, Randall. 2006. “The Babeldaob Road : The Impact of Road Construction on Rural Labor Force Outcomes in the Republic of Palau.” IZA Working Paper, no. 2452.

Asher, Sam, and Paul Novosad. n.d. “The Road to Riches : The Employment Effects of Road Construction in Rural India.” Working Paper.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian. 2012. “On the Road: Access to Transportation Infrastructure and Economic Growth in China.” NBER Working Paper No. 17897. doi:10.2139/ssrn.2018637.

Bardhan, Pranab K. 1979. “Wages and Unemployment in a Poor Agrarian Economy: A Theoretical and Empirical Analysis.” Journal of Political Economy 87 (3): 479. doi:10.1086/260774.

Benefo, Kofi, and T Paul Schultz. 1996. “Fertility and Child Mortality in Cote d’Ivoire and Ghana.” The World Bank Economic Review 10 (1): 123–58.

Berger, Thor, and Kerstin Enflo. 2013. “Locomotives of Local Growth: The Short- and Long-Term Impact of Railroads in Sweden.” EHES Working Paper 42.

Bleakley, Hoyt, and Jeffrey Lin. 2012. “Portage and Path Dependence.” Quarterly Journal of Economics 127 (2): 587–644. doi:10.1093/qje/qjs011.

Brenneman, Adam, and Michel Kerf. 2002. “Infrastructure & Poverty Linkages - A Literature Review.” World Bank Report, 1–122.

Bryceson, Deborah Fahy, and John Howe. 1993. “Rural Household Transport in Africa: Reducing the Burden on Women?” World Development 21 (11): 1715–28. doi:10.1016/0305-750X(93)90079-O.

Buys, P., U. Deichmann, and D. Wheeler. 2010. “Road Network Upgrading and Overland Trade Expansion in Sub-Saharan Africa.” Journal of African Economies 19 (3): 399–432. doi:10.1093/jae/ejq006.

Calderón, César, and Luis Servén. 2004. “The Effects of Infrastructure Development on Growth and Income Distribution.” Central Bank of Chile Working Paper, no. 3400: 1–43.

Chandra, Amitabh, and Eric Thompson. 2000. “Does Public Infrastructure Affect Economic Activity?” Regional Science and Urban Economics 30 (4): 457–90. doi:10.1016/S0166-0462(00)00040-5.

Cutler, David, Angus Deaton, and Adriana Lleras-Muney. 2006. “The Determinants of Mortality.” Journal of Economic Perspectives 20 (3). American Economic Association: 97–120.

Datta, Saugato. 2012. “The Impact of Improved Highways on Indian Firms.” Journal of Development Economics 99 (1). Elsevier B.V.: 46–57. doi:10.1016/j.jdeveco.2011.08.005.

Dinkelman, By Taryn. 2011. “The Effects of Rural Electrification on Employment : New Evidence from South Africa.” American Economic Review 101 (December): 3078–3108. doi:10.1257/aer.101.7.3078.

Donaldson, Dave, and Richard Hornbeck. 2013. “Railroads and American Economic Growth: A ‘Market Access’ Approach.” NBER Working Paper No. 19213, July.

Duranton, G., P. M. Morrow, and M. A. Turner. 2013. “Roads and Trade: Evidence from the US.” The Review of Economic Studies 81 (2): 681–724. doi:10.1093/restud/rdt039.

Fay, Marianne, Danny Leipziger, Quentin Wodon, and Tito Yepes. 2005. “Child Health Related Millenium Development Goals: The Role of Infrstructure.” World Development 33 (8): 1267–84.

Galiani, Sebastian, Paul J. Gertler, and Ernesto Schargrodsky. 2005. “Water for Life: The Impact of the Privatization of Water Services on Child Mortality.” Journal of Political Economy 113 (1): 83–119. http://papers.ssrn.com/abstract=648048.

Gibson, John, and Scott Rozelle. 2003. “Poverty and Access to Roads in Papua New Guinea.” Economic Development and Cultural Change. doi:10.1086/380424.

Guideliness of PMGSY. 2004.

Hummels, David, and Alexandre Skiba. 2004. “Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture.” Journal of Political Economy 112 (6). University of Chicago Press: 1384–1402.

Jacoby, Hanan G. 2000. “Access to Markets and the Benefits of Rural Roads.” The Economic Journal 110 (465): 713–37.

Khandker, Shahidur R., Zaid Bakht, and Gayatri B. Koolwal. 2009. “The Poverty Impact of Rural Roads : Evidence from Bangladesh.” Economic Development and Cultural Change 57 (4): 685–722.

Limão, Nuno M., and Anthony J. Venables. 1999. “Infrastructure, Geographical Disadvantage, and Transport Costs.” World Bank Policy Research Working Paper 2257, December. The World Bank.

Michaels, Guy. 2008. “The Effect of Trade on the Demand for Skill - Evidence from the Interstate Highway System.” The Review of Economics and Statistics 90 (November): 683–701. doi:10.1162/rest.90.4.683.

Mu, Ren, and Dominique Van De Walle. 2007. “Rural Roads and Poor Area Development in Vietnam.” World Bank Policy Research Working Paper 4340, no. August.

Panda, Pallavi. 2014. “Does Trade Reduce Infant Mortality ? Evidence from Sub-Saharan.” Working Paper.

Porto, Guido. 2004. “Trade Reforms, Market Access and Poverty in Argentina.” World Bank Policy Research Working Paper No. 3135.

Schaffer, Mark E. 2015. “XTIVREG2: Stata Module to Perform Extended IV/2SLS, GMM and AC/HAC, LIML and K-Class Regression for Panel Data Models.” Boston College Department of Economics. http://econpapers.repec.org/RePEc:boc:bocode:s456501.

Storeygard, Adam. 2012. “Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa.” Working Paper, Tufts University, no. April.

Van Bohemen, H D, and W H Janssen van de Laak. 2003. “The Influence of Road Infrastructure and Traffic on Soil, Water, and Air Quality.” Environmental Management 31 (1): 50–68. doi:10.1007/s00267-002-2802-8.

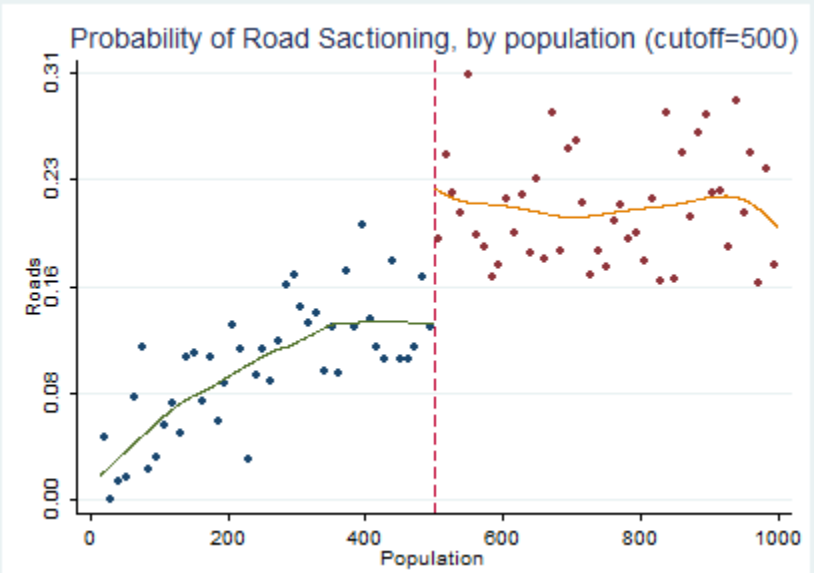
Van de Walle, Dominique. 2009. “Impact Evaluation of Rural Road Projects.” Journal of Development Effectiveness. doi:10.1080/19439340902727701.

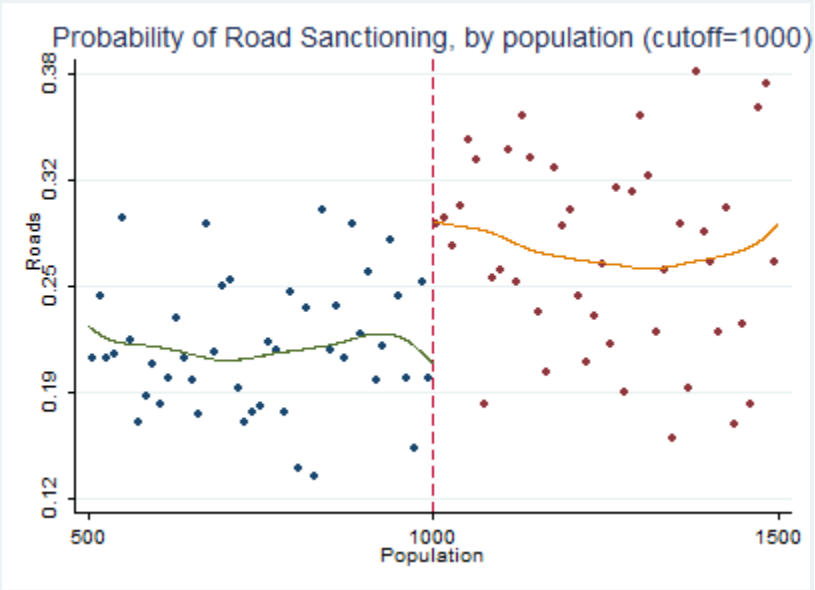
Wang, Limin. 2003. “Determinants of Child Mortality in LDCs: Empirical Findings from Demographic and Health Surveys.” Health Policy 65 (3): 277–99.

Wong, Emelita L., Barry M. Popkin, David K. Guilkey, and John S. Akin. 1987. “Accessibility, Quality of Care and Prenatal Care Use in the Philippines.” Social Science & Medicine 24 (11): 927–44. doi:10.1016/0277-9536(87)90286-3.

## Figures

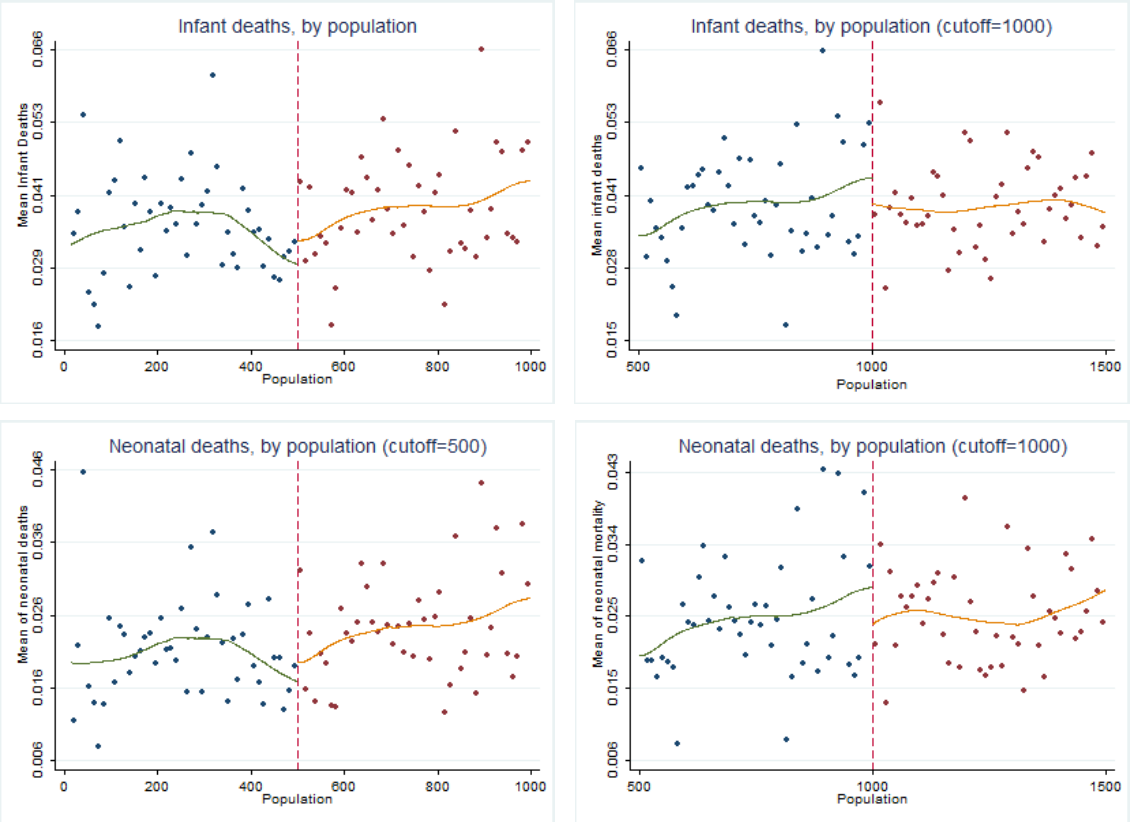
### Figure 3.1: Probability of road sanctioning at population cutoffs



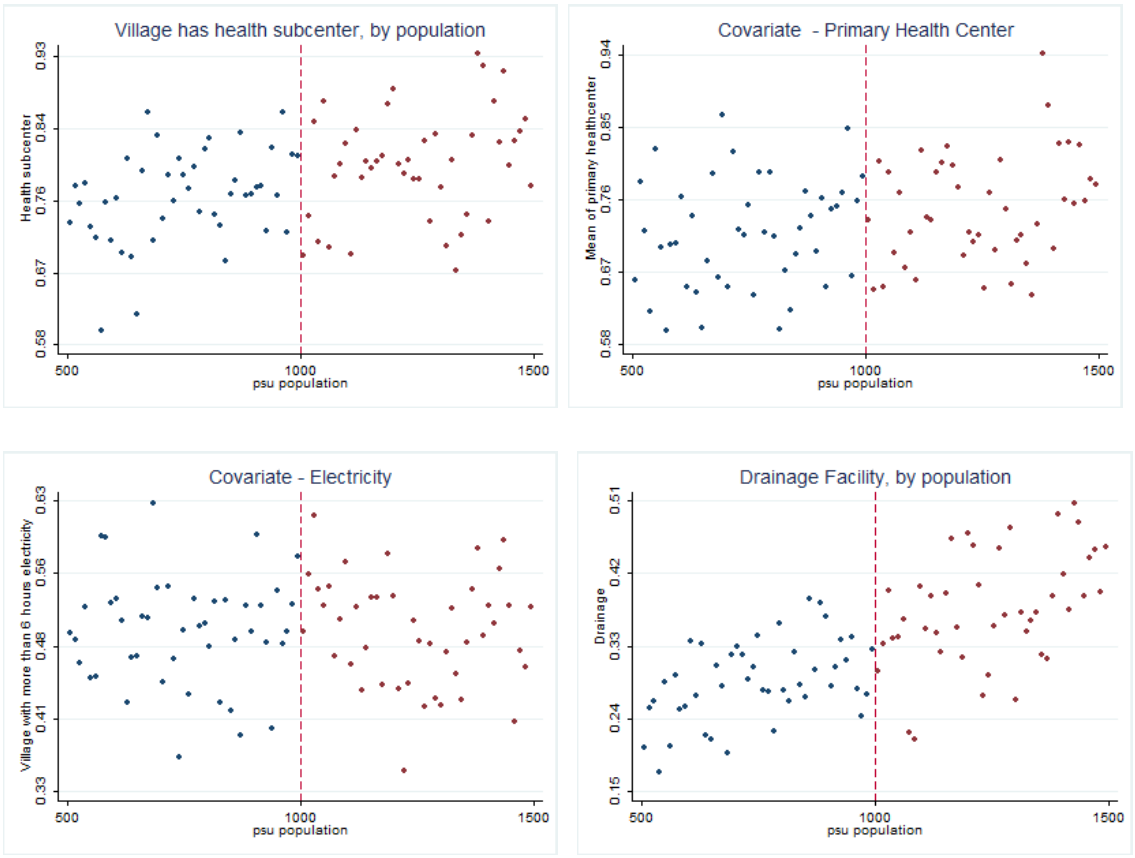

Note: This graph plots probability of assignment of road at the two village population cutoffs, 500 and 1000. Separate local linear smoothed fits are plotted to the left and the right of the cutoff points.

**Figure 3.2: Mean infant and neonatal deaths at population cutoffs**



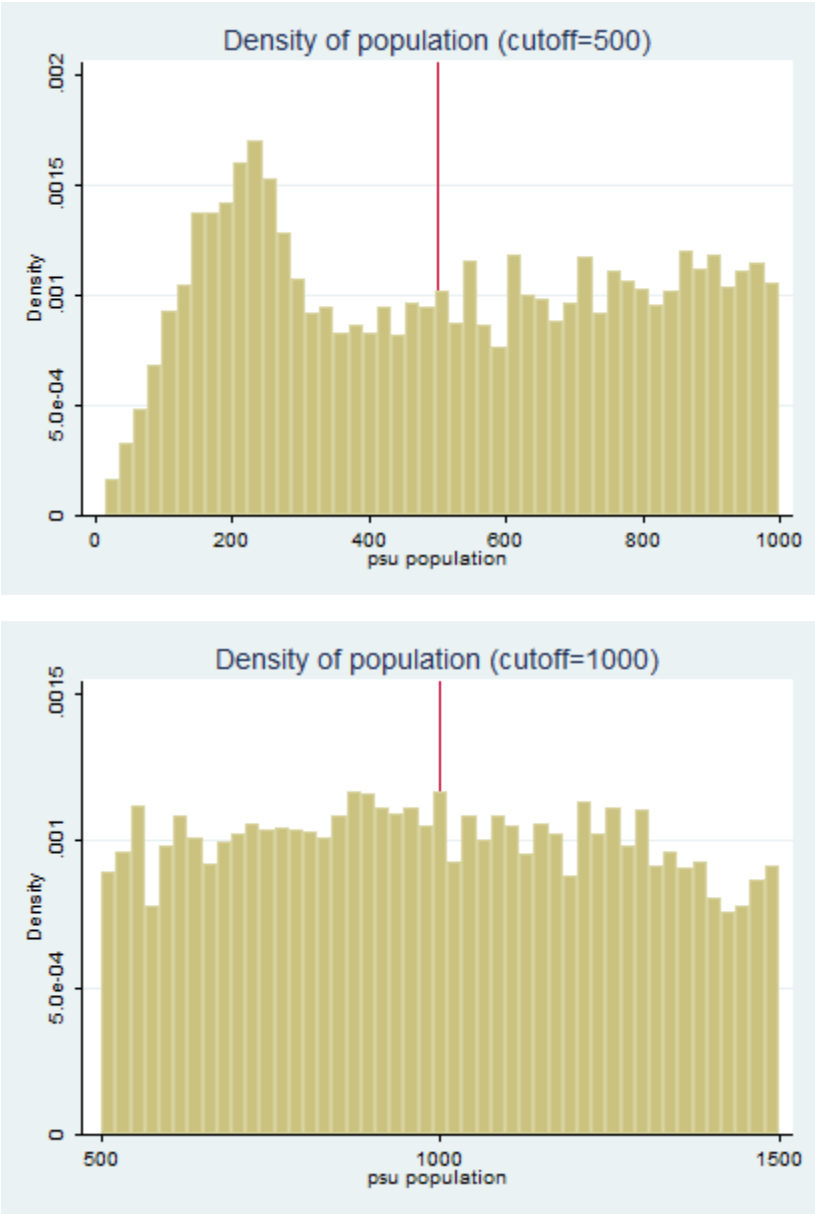

Note: The first row of graphs plot the discontinuity in average infant deaths as the outcome variable based on the village population size. The second row of graphs plot the discontinuity in average infant deaths as the outcome variable based on the village population size. The cutoffs are at village population of 500 and 1000. Separate local linear smoothed fits are plotted to the left and the right of the cutoff points.

**Figure 3.3: Covariates at population cutoffs**




Note: These scatterplots show the distribution of mean village covariates by the village population size. The covariates included are access to health sub-center, access to primary health center, access to at least 6 hours of electricity and availability of drainage facility in the village.

**Figure 3.4: Density of population at population cutoffs**




Note: These graphs plot the density of population at the two village population cutoffs, 500 and 1000, in the micro health and road merged data.

# Tables

## Table 3.1: Summary Statistics

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | **Obs** | **Mean** | **Std. Dev.** | **Min** | **Max** |
| **Infant Mortality** | 151370 | 0.039 | 0.19 | 0 | 1 |
| **Neonatal Mortality** | 151370 | 0.023 | 0.15 | 0 | 1 |
| **Female** | 149838 | 0.46 | 0.49 | 0 | 1 |
| **Multiple Births** | 151370 | 0.015 | 0.12 | 0 | 1 |
| **Birth Order** | 142611 | 2.69 | 1.89 | 1 | 25 |
| **Village Population (500-1000)** | 32114 | 757.8 | 143.5 | 501 | 999 |
| **Village Population (>1000)** | 88754 | 3733 | 5016 | 1000 | 99506 |
| ***Mother Characteristics*** | | | | | |
| **Uneducated** | 151350 | 0.51 | 0.49 | 0 | 1 |
| **Hindu** | 151368 | 0.76 | 0.42 | 0 | 1 |
| **Muslim** | 151368 | 0.14 | 0.34 | 0 | 1 |
| **SC** | 148337 | 0.20 | 0.40 | 0 | 1 |
| **ST** | 148337 | 0.18 | 0.39 | 0 | 1 |
| **OBC** | 148337 | 0.41 | 0.49 | 0 | 1 |
| **Low Wealth Index (Poor)** | 151336 | 0.48 | 0.49 | 0 | 1 |
| **Medium Wealth Index** | 151336 | 0.23 | 0.42 | 0 | 1 |
| **High Wealth Index** | 151336 | 0.28 | 0.45 | 0 | 1 |
| **Mother's age at birth** | 151369 | 24.5 | 5.37 | 0 | 1 |
| ***Village Characteristics*** | | | | | |
| **No electricity** | 151361 | 0.20 | 0.40 | 0 | 1 |
| **More than 6 hours electricity** | 151361 | 0.50 | 0.49 | 0 | 1 |
| **Drainage Facility** | 151361 | 0.36 | 0.48 | 0 | 1 |
| **Access to Dispensary** | 151361 | 0.71 | 0.45 | 0 | 1 |
| **Access to Primary care** | 151361 | 0.74 | 0.44 | 0 | 1 |
| **Access to Health sub center** | 151361 | 0.78 | 0.41 | 0 | 1 |
| **Distance to nearest town (KM)** | 151361 | 14.9 | 14.5 | 0 | 99 |
| **Distance to district headquarter (KM)** | 151248 | 44.9 | 35.7 | 0 | 865 |
| **Distance to nearest railway station (KM)** | 150390 | 40.7 | 64.7 | 0 | 900 |
| **Number of households in village** | 150818 | 390.5 | 765.4 | 0 | 9905 |

Note: Sample statistics of the variables for the rural sample are reported. Educated implies having attended any type of school and uneducated is defined as mother did not attend any school. Female is 1 if sex of child is female. Multiple birth is a dummy variable indicating if the child is born in a multiple birth. It is 0 for a single birth and 1 for twins, triplets or quadruplets.

## Table 3.2: Effect of PMGSY on infant and neonatal mortality - OLS

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| **Dependent Variable** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Neonatal Mortality** | **Neonatal Mortality** | **Neonatal Mortality** |
| **Road** | -0.0008 | -0.011 | -0.003 | -0.0002 | -0.0013 | 0.003 |
| | (0.0017) | (0.0087) | (0.009) | (0.0014) | (0.0074) | (0.0083) |
| **Explanatory Variables** | YES | YES | YES | YES | YES | YES |
| **District FE** | YES | NO | NO | YES | NO | NO |
| **Mother FE** | NO | YES | YES | NO | YES | YES |
| **Year FE** | YES | YES | YES | YES | YES | YES |
| **District-time trends** | NO | NO | YES | NO | NO | YES |
| **Number of groups** | 561 | 120765 | 120765 | 561 | 120765 | 120765 |
| **Observations** | 138368 | 142606 | 142606 | 138368 | 142606 | 142606 |

Note: Treatment (Road) is defined as 1 for a child born after PMGSY has been implemented in PMGSY affected district. The other control variables included in the specifications are sex of child, whether born in multiple birth, year of birth, mother's age at birth, birth order and birth month. The regressions with only district FE also control for mother characteristics like religion, caste, education of mother, and being poor and village characteristics like number of households in the village, distance from nearest town, distance to district headquarter, distance from nearest railway station, source of drinking water in village, village connected to health sub center, primary health clinic or dispensary are also included. Standard errors clustered at the district level are reported in brackets.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

**Table 3.3: First Stage – Fuzzy Regression Discontinuity Design**

| Road/Treatment | Coefficient | Robust Standard Error | t-statistics | p-value |
|---|---|---|---|---|
| **Cutoff 1 (500)** | 0.104*** | 0.011 | 9.28 | 0.000 |
| **Cutoff 2 (1000)** | 0.058*** | 0.012 | 4.95 | 0.000 |
| **District FE** | YES | | | |
| **Flexible population controls** | YES | | | |
| **Village controls** | YES | | | |
| **F-Stat (Cutoff1)** | 86.16 | | | |
| **F-Stat (Cutoff2)** | 24.52 | | | |
| **Clusters** | 561 | | | |
| **Observations** | 149818 | | | |

Note: It presents the first stage estimates to check for weak identification. Flexible population controls refer to linear, squared and cubic terms of village population. The village control variables included in the specifications are number of households in the village, distance from nearest town, distance to district headquarter, distance from nearest railway station, source of drinking water in village, village connected to health sub center, primary health clinic or dispensary. Standard errors clustered at the district level are reported in brackets.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

**Table 3.4: Effect of PMGSY on infant mortality - FRDD**

| | Cutoff=500 | | | Cutoff=1000 | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Dependent Variable** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** | **Infant Mortality** |
| **Road** | -0.018 | -0.018 | -0.0175 | -0.024 | -0.029* | -0.024 |
| | (0.012) | (0.012) | (0.013) | (0.014) | (0.015) | (0.015) |
| **Village Controls** | NO | YES | YES | NO | YES | YES |
| **Mother-Child Controls** | NO | NO | YES | NO | NO | YES |
| **Flexible Population Controls** | YES | YES | YES | YES | YES | YES |
| **District FE** | YES | YES | YES | YES | YES | YES |
| **Number of clusters** | 561 | 561 | 561 | 561 | 561 | 561 |
| **Observations** | 153519 | 149818 | 138369 | 153519 | 149818 | 138369 |

Note: This is a Fuzzy RD estimate of the effect of road construction on infant mortality for cutoff=500 and cutoff=1000. Treatment is defined as child in a village being affected by PMGSY by 2008, time of DLHS survey. The mother-child control variables included in the specifications are sex of child, whether born in multiple birth, mother's age at birth, birth order, birth month, religion, caste, education of mother, and being poor. Flexible population controls refer to linear, squared and cubic terms of village population. The village control variables included in the specifications are number of households in the village, distance from nearest town, distance to district headquarter, distance from nearest railway station, source of drinking water in village, village connected to health sub center, primary health clinic or dispensary. Standard errors clustered at the district level are reported in brackets.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 3.5: Effect of PMGSY on neonatal mortality - FRDD

| | Cutoff=500 | | | Cutoff=1000 | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Dependent Variable** | **Neonatal Mortality** | **Neonatal Mortality** | **Neonatal Mortality** | **Neonatal Mortality** | **Neonatal Mortality** | **Neonatal Mortality** |
| **Road** | -0.005 | -0.006 | -0.007 | -0.013 | -0.017 | -0.013 |
| | (0.009) | (0.009) | (0.009) | (0.011) | (0.011) | (0.012) |
| **Village Controls** | NO | YES | YES | NO | YES | YES |
| **Mother-Child Controls** | NO | NO | YES | NO | NO | YES |
| **Flexible Population Controls** | YES | YES | YES | YES | YES | YES |
| **District FE** | YES | YES | YES | YES | YES | YES |
| **Number of clusters** | 561 | 561 | 561 | 561 | 561 | 561 |
| **Observations** | 153519 | 149818 | 138369 | 153519 | 149818 | 138369 |

Note: This is a Fuzzy RD estimate of the effect of road construction on neonatal mortality for cutoff=500 and cutoff=1000. Treatment is defined as child in a village being affected by PMGSY by 2008, time of DLHS survey. The mother-child control variables included in the specifications are sex of child, whether born in multiple birth, mother's age at birth, birth order, birth month, religion, caste, education of mother, and being poor. Flexible population controls refer to linear, squared and cubic terms of village population. The village control variables included in the specifications are number of households in the village, distance from nearest town, distance to district headquarter, distance from nearest railway station, source of drinking water in village, village connected to health sub center, primary health clinic or dispensary. Standard errors clustered at the district level are reported in brackets.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Table 3.6: Effect of PMGSY on sex ratio at birth

| | OLS | | | FRDD | |
|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
| **Dependent Variable** | **Male** | **Male** | **Male** | **Male** | **Male** |
| **Treatment** | -0.005 | 0.015 | 0.015 | -0.01 | -0.049* |
| | (0.005) | (0.023) | (0.025) | (0.035) | (0.029) |
| **Cutoff** | | | | 1000 | 500 |
| **Village and Mother Controls** | YES | NO | NO | YES | YES |
| **Child Controls** | YES | YES | YES | YES | YES |
| **Flexible Population Controls** | NO | YES | YES | YES | YES |
| **District FE** | YES | NO | NO | YES | YES |
| **Mother FE** | NO | YES | YES | NO | NO |
| **District-time trends** | NO | NO | YES | NO | NO |
| **Number of clusters** | 561 | 120765 | 120765 | 561 | 561 |
| **Observations** | 138368 | 142606 | 142606 | 138369 | 138369 |

Note: These provide the regression estimates of the effect of road construction on probability of being a male in both OLS and Fuzzy RD framework. For the OLS framework, a child is said to be treated if his/her mother was affected by PMGSY in the year of conception. The child control variables included in the specifications are sex of child, whether born in multiple birth, mother's age at birth, birth order, and birth month. Additional mother specific control variables are religion, caste, education of mother, and being poor. Flexible population controls refer to linear, squared and cubic terms of village population. The village control variables included in the specifications are number of households in the village, distance from nearest town, distance to district headquarter, distance from nearest railway station, source of drinking water in village, village connected to health sub center, primary health clinic or dispensary. Standard errors clustered at the district level are reported in brackets.

*** Significant at 1% level, ** significant at 5% level, * significant at 10% level.

## Conclusion

This research studied the impact of macroeconomic policies on child health outcomes and demography in the context of developing countries. I look at heterogeneous impacts across agents and how initial levels of endowments affect the accrual of benefits presented by a change in economic policy. I study the response of economic agents to these policies, and determination of optimal household decisions regarding health and labor market participation.

The findings of the first chapter reinstate how the same economic policy may have differential effects on different sections of the population. In studying the effect of African Growth and Opportunity Act (AGOA) in 30 different sub-Saharan African countries, I find heterogeneity in the effects of the policy on rural and poor mothers in lowering their probability of infant deaths after the policy is implemented. Similarly, at the macro level, I do not find significant reductions in infant deaths for oil exporting countries, supporting the widely cited hypothesis of "resource curse". On an average, I find a significant reduction in infant and neonatal deaths almost immediately after the policy is implemented in these nations. With many developing countries relying on trade openness to boost their economic growth, this study brings forth this discussion to economic development and opens up new channels of effective policy intervention to improve human capital outcomes of the society.

The second chapter emphasizes on the importance of well-rounded policy interventions. We find evidence of negative mortality selection in India indicating survival of weaker

kids due to neonatal interventions. Moreover, given the patrilineal structure, there is evidence of son preference and differential effects by birth order. If weaker kids survive due to successful neonatal interventions, without provision of appropriate care afterwards, it would worsen the case of malnutrition.

Lastly, the third chapter presents evidence of no short run impact of a policy change like rural road construction on the health outcomes of children in India. This chapter highlights the fact that only access to infrastructure may not be enough in improving outcomes. Complementary policies of provision of means of transportation and reducing financial constraints may help accrue full benefits of provision of infrastructure, especially for the rural and poor.

This research presents evidence of effectiveness of different kinds of policies on improving child health outcomes and shows that policies may have heterogeneous effects, policies need to be complemented with other policies for better effectiveness, as well as, there may be some lag between policy implementation and realization of benefits for the society.