ESSAYS IN EMPIRICAL DEVELOPMENT ECONOMICS

Tanmoy Majilla

546567-L-bw-Majilla
Processed on: 6-8-2020

This research was partially funded by the Research Innovation Facility (RIF) of International Institute of Social Studies.

MIX
Paper from responsible sources
FSC® C128610

ISBN 978-90-6490-120-1

Essays in Empirical Development Economics

Essays over empirische ontwikkelingseconomie

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
Rector Magnificus

Prof.dr. R.C.M.E. Engels

and in accordance with the decision of the Doctorate Board

The public defence shall be held on
18 September 2020 at 9.00 hrs

by

Tanmoy Majilla
born in Burdwan, India

International Institute of Social Studies

Erasmus University Rotterdam

Doctoral Committee

Doctoral dissertation supervisor
Prof. A.S. Bedi

Other members
Prof. A. Mukhopadhyay, Indian Statistical Institute, Delhi
Prof. N. Datta Gupta, Aarhus University
Prof. S. Klonner, University of Heidelberg

Co-supervisor
Dr M. Rieger

Contents

List of Tables

List of Figures

List of Appendices

Acknowledgements

First and foremost, I would like to thank my supervisors Arjun Bedi and Matthias Rieger for their continuous support throughout my Ph.D. journey. I can never thank them enough for their mentorship and understanding in this intellectual journey. Matthias, this thesis would not have been completed without your support, trust, and energy. I did my M.A. thesis with you. You are more of a friend than a supervisor. You are also my most frequent collaborator, and I doubt I can ever be so comfortable in collaborating with anybody else. I never hesitated to send you under-developed ideas, initial drafts, and, sometimes I feel I disturb you too much. But you have always been ready to support me and have kept pushing. I cannot tell you how much I have benefited from you. Thank you, for your patience, motivation, and enthusiasm.

It is a great privilege and honor to study with you Arjun. I offer my sincere appreciation for the learning opportunities with you. Arjun nobody could have understood my family obligations like you. I will always remain thankful to you. You are always willing to support. When I had to pass up few Ph.D. offers (you know why!) and had very little options left, you readily agreed to supervise me. Thank you for your willingness to work with me. I have the tendency to work on multiple projects concurrently. You taught me the benefit of remaining focused. Your guidance is the major reason why I have finished my Ph.D. on time. I am also grateful for your continuous support in this tough job market. It would not be enough even if I convey my endless gratitude to you.

I thank my thesis assessment committee, Nabanita Datta Gupta, Stefan Klonner and Abhiroop Mukhopadhyay for their support. I benefitted greatly from your excellent comments. I also thank Chris Elbers and Irene van Staveren for agreeing to be part of my extended thesis committee.

During the Ph.D. years, I had my child - Aharshi. Aharshi you are the reason why I never felt exhausted in this inevitably challenging task. You have completely changed my life. You bring in light in rough times. I can never be thankful enough for this. This thesis would never have seen daylight without your support, trust, respect, and love - Dipsa. Anything I write about you would be inadequate. Despite many pressures in our mundane life, supporting me has always been your priority. Any attempt to be thankful for your encouragement, care, and support when the times got rough would not be enough.

I would like to thank my parents for introducing me to this intellectually stimulating world. Mamoni you always wanted me to do something creative, and I can assure you this is just the beginning. I thank you for your love, understanding, and sacrifices that you have made to educate me. I would like to thank you Bapi. You undertook many of my family responsibilities in my absence with utmost care. I also thank my extended family members, specifically Mamima, for your support.

Last but not the least, I would like to thank ISS for providing such an intellectually stimulating, and just space. I would like to acknowledge the research innovation fund of the ISS for supporting my research. I have just named few of the people who were part of this journey. There are many others who have contributed to my intellectual journey. Thank you all for your support and trust.

Abstract

It is well documented that the economic status of men and women is not equal in most societies. However, South Asia stands out as an extreme case. The region is well-known for its strong son preference, it has a large number of 'missing women' and low female labor force participation. Motivated by the stark gender differences both within the household and at the point of labor market entry, this thesis examines issues that are pertinent to son preference and labor market entry of young women.

Son preference manifests itself in different ways, including unequal parental allocation of monetary and non-monetary resources within a household. However, direct empirical evidence on the unequal allocation of monetary resources on children within a household is limited.

Chapter 2 of this thesis use recently available gender-disaggregated household survey data on shadow educational expenditures (or private supplementary tuition) to examine parental allocation of educational resources in India (Chapter 2). The analysis shows a birth order disadvantage for later-born children. Girls face disadvantages in every birth order compared to their male peers. The pattern may be causally attributed to parental preferences for elder sons. Furthermore, the gender disparity in intra-household allocation of educational resources accounts for a substantial proportion of the gender gap in cognitive test scores.

Chapter 3 deals with similar issues in the context of Pakistan. The chapter demonstrates that the gender gap in mathematics in Pakistan may plausibly be explained by parental preferences for elder sons combined with family size. The gender gap in math does not exist at the age of five, but monotonically increases with the age of children. Similar to the situation in India, this chapter shows that birth order disadvantages for later-born children, and gender gaps at every birth order. The chapter further demonstrates that the gender gap is less likely to be observed in small families.

Chapters 2 and 3 document gender differences in educational expenditure and gender gaps in cognitive achievement. These differences are likely to translate into gender gaps in labor market achievements. Such disadvantages are exacerbated if there is demand-side discrimination in the labor market. In fact, a large number of studies have documented labor

market disadvantages for women. One of the key disadvantages experienced by women is associated with motherhood, and the incidence of motherhood penalty has been well-documented in high income countries, but less so in the context of developing countries. While motherhood is likely to exert a penalty in low-income countries, the magnitude of the penalty may vary depending on the (perceived) community gender norms. Based on an experimental approach, Chapter 4 of this thesis examines the extent of the motherhood penalty in urban India, and the link between community gender norms (patrilineal versus matrilineal communities) and the motherhood penalty. Chapter 5 extends the experiment to examine the extent to which access to childcare support mitigates the motherhood penalty.

The analysis shows a large motherhood penalty in India which is particularly pronounced for women belonging to patrilineal communities. In contrast, mothers from matrilineal communities face no such penalty. The extension in chapter 5 shows that signaling the availability of childcare at home leads to a partial reduction in the motherhood penalty in a patrilineal community. A common phenomenon in India and perhaps other developing countries is the widespread possibility of acquiring gray degrees or potentially bought degrees to combat disadvantages at labor market entry. The last chapter of this thesis examines the impact of gray degrees, or potentially bought academic credentials from legitimate universities, on callback rates to job applications using a resume experiment in India. The evidence show that applicants with gray degrees fare better – have higher callback rates, as compared to applicants with no degrees, but do worse as compared to applicants with authentic degrees. The evidence also shows that gray degrees have a larger positive impact on women applicants compared to their male counterparts.

Samenvatting

Het is een bekend gegeven dat mannen en vrouwen in de meeste samenlevingen geen gelijke economische status hebben. In Zuid-Azië is het statusverschil echter extreem groot. De regio staat bekend om een sterke voorkeur voor zonen, er is een groot aantal 'vermiste vrouwen', en de arbeidsparticipatie van vrouwen is er laag. De grote sekseverschillen, zowel binnen het huishouden als op het moment van toetreding tot de arbeidsmarkt, vormden de aanleiding voor dit onderzoek naar aspecten die te maken hebben met de voorkeur voor zonen en de toetreding tot de arbeidsmarkt van jonge vrouwen.

De voorkeur voor zonen komt op verschillende manieren tot uitdrukking, zoals door een ongelijke toewijzing van financiële en niet-financiële middelen door ouders binnen een huishouden. Er zijn echter slechts weinig directe empirische gegevens die wijzen op de ongelijke verdeling van financiële middelen over kinderen binnen een huishouden.

In hoofdstuk 2 van dit proefschrift wordt beschreven hoe ouders in India onderwijsmiddelen toewijzen. Dit is onderzocht op basis van naar sekse uitgesplitste onderzoeksgegevens over schaduwuitgaven aan onderwijs (of aanvullend privéonderwijs) die sinds kort beschikbaar zijn. Hieruit blijkt dat later geboren kinderen achtergesteld worden. Meisjes worden ongeacht de geboortevolgorde achtergesteld ten opzichte van jongens. Het patroon kan veroorzaakt worden door de voorkeur van ouders voor eerder geboren zonen. Bovendien is een substantieel deel van het sekseverschil in cognitieve testscores te wijten aan de sekseongelijkheid bij de toewijzing van onderwijsmiddelen binnen het huishouden.

Hoofdstuk 3 gaat over vergelijkbare kwesties in de context van Pakistan. Dit hoofdstuk laat zien dat de sekseverschillen op het gebied van wiskunde in Pakistan plausibel kunnen worden verklaard door de voorkeur van ouders voor oudere zonen in combinatie met de grootte van het gezin. Op de leeftijd van vijf jaar zijn er nog geen sekseverschillen op het gebied van wiskunde, maar deze worden steeds groter naarmate kinderen ouder worden. Uit dit hoofdstuk blijkt dat de geboortevolgorde nadelig is voor later geboren kinderen en dat er bij elke geboortevolgorde sprake is van sekseverschillen, net als in India. Verder blijkt dat de kans dat sekseverschillen worden waargenomen kleiner is in kleine gezinnen.

Hoofdstuk 2 en 3 beschrijven sekseverschillen in onderwijsuitgaven en cognitieve prestaties. De kans is groot dat deze verschillen zullen leiden tot sekseverschillen in prestaties op de arbeidsmarkt. Dergelijke verschillen worden nog verergerd als er sprake is van discriminatie aan de vraagzijde van de arbeidsmarkt. Uit een groot aantal studies is inderdaad gebleken dat vrouwen een achterstand op de arbeidsmarkt hebben. Een van de belangrijkste belemmeringen die vrouwen ondervinden, houdt verband met het moederschap. De benadeling op de arbeidsmarkt vanwege het moederschap is goed gedocumenteerd in landen met een hoog inkomen, maar in ontwikkelingslanden is dat minder het geval. Hoewel het moederschap in lage-inkomenslanden waarschijnlijk nadelig is, kunnen de (gepercipieerde) normen voor mannen en vrouwen binnen de gemeenschap van invloed zijn op hoe nadelig dit is. Hoofdstuk 4 van dit proefschrift beschrijft experimenteel onderzoek in stedelijk India naar hoe nadelig het moederschap is, en naar het verband tussen de normen voor mannen en vrouwen binnen de gemeenschap (patrilineaire versus matrilineaire gemeenschappen) en de nadeligheid van het moederschap. In dit experiment is ook onderzocht in hoeverre toegang tot kinderopvang het moederschapsnadeel vermindert. Dit deel van het onderzoek wordt beschreven in hoofdstuk 5.

Uit het onderzoek blijkt dat moederschap in India een groot nadeel is op de arbeidsmarkt, vooral voor vrouwen die tot patrilineaire gemeenschappen behoren. Moeders uit matrilineaire gemeenschappen ondervinden dit nadeel daarentegen niet. Uit het deelonderzoek in hoofdstuk 5 blijkt dat wijzen op de beschikbaarheid van kinderopvang thuis leidt tot een gedeeltelijke vermindering van het moederschapsnadeel in een patrilineaire gemeenschap.
Een veel voorkomend verschijnsel in India en misschien ook in andere ontwikkelingslanden is dat mensen dubieuze diploma's behalen of diploma's kopen om te compenseren voor nadelen bij de toetreding tot de arbeidsmarkt. Het voorlaatste hoofdstuk van dit proefschrift beschrijft een cv-experiment in India. Dit gaat over het effect van dubieuze of mogelijk gekochte diploma's van legitieme universiteiten op terugbelpercentages bij sollicitaties. Uit dit experiment blijkt dat sollicitanten met dubieuze diploma's het beter doen (hogere terugbelpercentages hebben) dan sollicitanten zonder diploma's, maar het slechter doen dan sollicitanten met echte diploma's. Ook blijkt uit de resultaten dat vrouwelijke sollicitanten meer profijt hebben van dubieuze diploma's dan mannelijke sollicitanten.[1]

[1] Dit hoofdstuk is gepubliceerd als Majilla, T., & Rieger, M. (2020). Gray University Degrees: Experimental Evidence from India. *Education Finance and Policy*, 15(2), 292-309.

Chapter 1
Introduction

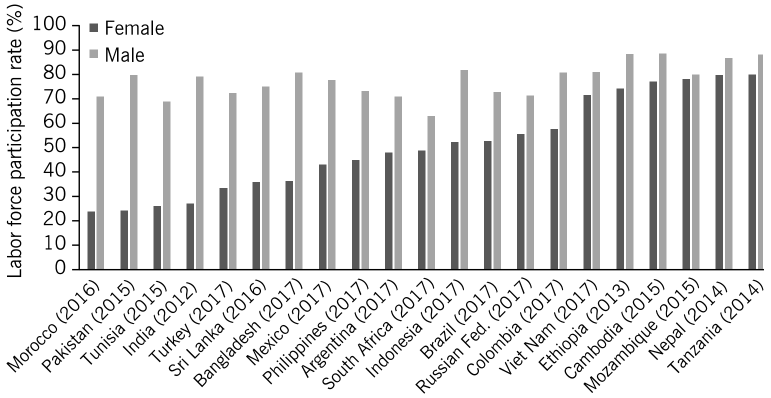
This dissertation consists of six essays in the field of empirical development economics. These essays are united by a clear intellectual theme - each chapter attempts to uncover different facets of disadvantages for women in developing countries, particularly in South Asia. The various chapters explore economically important and socially relevant disadvantages in different periods in women's lives. Methodologically, the first three chapters rely on observational data and attempt to exploit natural experiments while the last three chapters are based on field experiments. Chapter 2 focuses on intra-household resource allocation decisions, chapter 3 on gender gap in math test scores, chapters 4 and 5 on the motherhood penalty, and chapter 6 on the labor market impact of gray or potentially bought academic degrees. Although linked, the thesis consists of two parts. The first part comprising chapters 2 and 3 deal with pre-labor market gender disadvantages while chapters 4, 5 and 6 deal with gender disadvantages at the point of labor market entry.

A number of empirical observations regarding gender-based discriminations have triggered this dissertation. It has been well documented that parental investments play a crucial role in children's skill accumulation (Carneiro et al., 2013). In developing countries, boys outperform girls in educational and labor market achievements. Differential within-household parental investments are typically cited to explain this. In a number of developing countries parents prefer sons over daughters. For instance, mothers are less likely to breastfeed daughters (Jayachandran and Kuziemko, 2011), and, parents may make less health (Bhalotra et al., 2010) and educational investments (Zimmerman, 2012) in daughters. Given the importance of parental investments in skill accumulation, educational attainments and future labor market achievements of their children, it is important to understand what drives such within-household allocation of parental investments. Understanding these forces is also important when designing policies to mitigate gender gaps in academic and labor market achievements.

Likewise, in many developing countries a substantial number of women do not participate in the labor marker, and when they participate, they earn less compared to men. As depicted in

Figure 1.1, the gender gap in labor market participation is sharper in India as compared to many South-East Asian and African countries notwithstanding many favorable conditions.[2]

Figure 1.1: Labor Force Participation in Few Developing Countries



Source: Verick (2018)

Labor market penalties are often exacerbated with motherhood in many parts of the world (Kleven et al., 2018). While a number of studies have reported a motherhood penalty in the global north, the literature on such penalties in the global south is still in its infancy. The motherhood penalty in the global south may be even more important where becoming a mother is likely to further exacerbate gender-based discrimination.

Against this backdrop, the first three chapters document disadvantages experienced by girls in intra-household parental resource allocation and how such intra-household parental behavior may translate into gender disparity in educational achievements. The next two chapters document the motherhood penalty in India. In particular, respectively, they explore the interaction between the motherhood penalty and community gender norms (matrilineal vs. patrilineal), and, the impact of one potential mitigating factor - childcare support at home, on the motherhood penalty. The last chapter explores the gendered impact of gray degrees (or potentially bought degrees) on labor market entry.

The various chapters are linked in terms of the overall scope and agenda of this dissertation. I begin with (Chapters 2 and 3) gender discrimination during childhood, which adversely

[2] In India, 22% of women participate in the labor market, compared to more than 30% in Bangladesh, 50% in Indonesia, Philippines and Thailand, and more than 70% in Vietnam (Verick, 2018).

affects human capital accumulation. Even if there is no gender discrimination in the labor market, a gender gap in labor market achievements is inevitable as it may simply echo differences in pre-market skills. The conditions may worsen with the presence of demand-side discrimination in the labor market, as seems to be the case. However, as shown in chapters 5 and 6, women may find some strategies to ease, at least partially, the labor market penalties.

In the following paragraphs, I provide a brief summary of each of the substantive chapters.

In **chapter 2**, I explore intra-household allocation of parental educational resources in India. Studying intra-household allocation of resources is challenging as *direct* parental monetary expenditure on children is difficult to isolate from aggregate household expenditure. Due to such limitations, previous studies tend to examine parental allocation of resources *indirectly* from household expenditures or through other child outcomes. This article studies intra-household allocation of direct parental monetary expenditure on private supplementary tutoring or shadow education. I show a birth order disadvantage for later-born children in shadow education expenditure and find evidence of disadvantages for girls in every birth order. I attribute these patterns to the well-documented preference for elder sons, which is typical of the context, and I subsequently test several features which stem out of this preference. Shadow education appears to be effective in enhancing children's educational achievements with a larger effect for girls. These results indicate inefficient intra-household allocation of educational resources. The analysis also shows that a substantial proportion of intra-household disparity in shadow education expenditure translates into gender gaps in test scores.

In **chapter 3**, I document substantial gender gaps in mathematics test score using a large nationally representative dataset from Pakistan. The gender gap in mathematics has been documented in many contexts, yet little convincing evidence exists to explain it. I find that boys and girls have similar levels of achievement at the age of five, after which a monotonically increasing gap emerges. I also report a negative birth-order effect with boys outperforming girls at every birth order, but with a weaker gap in later-borns. I show that strong elder son preference, which skews parental resource allocation, is one of the underlying mechanisms of these gendered patterns. Elder son preference induces girls with elder brothers to do worse compared to those without. The gender gap is relatively more pronounced in larger families. In sum, elder son preference coupled with the adverse effects of family size explain the gender gap.

In **chapter 4**, co-authors Arjun Bedi, Matthias Rieger and I use a field experiment to study the effect of perceived gender norms on the motherhood penalty in the Indian labor market. We randomly reported motherhood on fictitious CVs sent to service sector jobs. We generated variation in gender norms by signaling community origins of applicants. Employers are less likely to callback mothers relative to women or men without children. Mothers from North-East India experience a smaller motherhood penalty and those of matrilineal origin face no penalty, unlike those of patrilineal origin. We discuss our findings in relation to the influence of ethnicity, the Indian context and theories of discrimination.

While there is growing evidence that mothers are discriminated in terms of labor market entry, experimental evidence from the Global South or of underlying mechanisms is still scant. Building on chapter 4, in **chapter 5**, based on a CV experiment conducted in one large city in India, co-authors Arjun Bedi, Matthias Rieger and I examine whether access to childcare support may offset the motherhood penalty associated with labor market entry. We randomly varied motherhood, as well as a childcare support signals in online applications sent to service sector jobs. Indicating motherhood on a CV led to a 57% or 20 percentage point reduction in callback rates as compared to non-mothers. Signaling childcare support offset the motherhood penalty by 20% or 4 percentage points.

In **chapter 6**, co-author Matthias Rieger and I study the impact of gray degrees, or potentially bought academic credentials from legitimate universities, on callback rates to job applications using a resume experiment in India.[3] The experiment varied the type of degree (no, gray and authentic) in online applications to entry level jobs that require no university qualification. We find that gray degrees increase callback rates by 42% or 8%-points relative to having no degree. However, we also document that gray degrees fare on average worse than authentic degrees. These empirical patterns are consistent with a model where employers have beliefs about the authenticity of degrees and are discounting gray degree universities probabilistically. With respect to gender, the callback rate for women with gray degrees was higher as compared to their male peers. We discuss our findings with respect to the Indian context.

[3] This chapter has been published in *Education Finance and Policy*. 2020. 15(2), 292-309.

Chapter 2
Shadow Education, Intra-Household Resource Allocation and Educational Achievements in India

2.1 Introduction

Childhood circumstances lay the foundation for future achievements (Almond et al., 2018). Since, childhood is predominantly the domain of the family, understanding the relative status of children in the family and particularly how a child's status within a family affects intra-household allocation of resources is crucial (Dunbar et al., 2013). Any within-household inequality between children based on gender or other attributes, may have adverse lifetime effects.

In a number of developing countries, parents tend to favor boys. Empirical evidence shows that son preference alters parental allocation of resources across children. As a consequence, substantial gender differences in consequent economic and other (e.g., anthropometric) outcomes are inevitable. One strand of the literature investigates parental allocation of *non-monetary* resources such as visits to clinics, breastfeeding and time spent on childcare (Jayachandran and Pande, 2017; Barcellos et al, 2014; Jayachandran and Kuziemko, 2011). Another approach uses household expenditures on children-specific goods to examine gender differences in resource allocation within household (Deaton, 1997, 1989). Empirically, these two strands of the literature often report conflicting results. The first strand of the literature frequently reports evidence of gender specific differences in parental investments, in contrast, studies using the expenditure approach do not tend to find much gender disparity. This may well be because it has been challenging to isolate individual expenditures form aggregate household data. Exceptions are Zimmerman (2011), Aslam and Kingdon (2008) and Kingdon (2005). These studies use parental educational expenditures on children to examine intra-household resource allocation, but these studies report conflicting results. For instance, Zimmerman (2011) finds gender discrimination in educational expenditures in India and so do Aslam and Kingdon (2008) in Pakistan, while Kingdon (2005) fails to identify any such disparity in India. In addition to the paucity of individual level data, estimating within-household allocation of educational resources is in fact more challenging as education is free in most developing countries. Though these studies have made significant progress, a primary concern is the measurability of direct parental monetary investments on individual children.

This chapter tries to address these by making two primary contributions. The first is to explore gender differences in intra-household resource allocation by making use of (escalating) educational expenditures on children as the primary variable of interest. In particular, the chapter shows intra-household gender disparities in parental expenditures on private supplementary education or 'shadow education'. Then, I quantify the consequence of such intra-household inequality on educational achievements. To be specific, using a decomposition analysis, I quantify the relative importance of gender disparity in shadow educational expenditures on gender gaps in cognitive test scores.

This study draws on the three distinctive strengths of shadow education. First, a major advantage of looking at expenditure on shadow education is that it is a direct measure of parental allocation of monetary resources. Thus, a novel element of this study is the availability of data which captures direct parental monetary expenses on individual children. Second, in some contexts, including India, expenditure on shadow education accounts for a substantial proportion of total household income and may be more central to student learning than formal schooling (Bray, 2014). As will be discussed in detail later in the text, the data shows that shadow education expenditures accounts for around 5% of household annual income. Third, compared to other educational expenditures, shadow education captures parental choices more effectively. For instance, parents may choose between different schools in urban areas but school fess are regulated in private schools and free in public schools.[4] Thus, parents may not have full freedom to spend money on formal schooling according to their preferences.

My primary analysis is based on the Indian Human Development Survey II data collected in 2011-12. The study proceeds by showing that parents spend more on boys as compared to girls in every birth order, although the disparity weakens for later-born children. To be specific, firstborn girl children on average receive around 0.079 SD (or INR 211 yearly) less shadow education expenditures compared to firstborn boys, the disparity falls to 0.071 SD

[4] Rural parents may not even have such options. In most villages, there is only one public school available, and private schooling is an urban phenomenon.

(INR 189) for the second borns and drops further down the birth order. Additionally, I demonstrate a birth order disadvantage for later-born children. For instance, moving from firstborn boys to second born boys results in 0.036 standard deviation (INR 96) drop in shadow education expenditures. The birth order disadvantage increases to0.065 standard deviations between firstborn boys and third born boys. The birth disadvantage appears to be flatter for girls. I also observe similar patterns in weekly duration of shadow education attendance.[5] These patterns clearly indicate elder son advantage in intra-household resource allocation. The evidence further demonstrates heterogeneity in effects. The gender gap is sharper in families belonging to forward castes and those with children in private schools. In other words, families with better socioeconomic backgrounds exhibit greater gender discrimination.

To ensure that differential fertility selection and other confounding characteristics do not drive the results, I also include mother, household and neighborhood fixed effects in the baseline specification. The empirical patterns appear to be robust. Afterwards, I further replicate and obtain fairly similar results in a sample of mothers who are likely to have completed fertility.

Next I replicate the main results with a more recent, larger but less extensive Annual Status of Education Report (hereafter ASER), 2016 data. ASER lacks fertility data and the survey recruits only children between five to sixteen years of age. Thus, I do not observe precise birth orders. In spite of the above caveats, the birth order – gender patterns are remarkably consistent with the baseline estimates.

In interpreting these results, I examine whether the preference for sons, particularly for elder ones, may explain such a gendered pattern. I perform several tests: *First*, following Jayachandran and Pande (2017), I find that both boys and girls without elder brothers have an advantage over their peers with elder brothers, which points to strong elder son preference. Moreover, boys without elder brothers have an advantage over girls without elder brothers, which indicates within-household hierarchy in resource allocation. *Second*, I further demonstrate that families who prefer an extra child to be a *son* invest 0.052 SD less on their existing children compared to families with no such preference. In other words, these families

[5] In our data, most children attend shadow education conditional on at least one child in the family attends shadow education, indicating that within-household selection is not likely to be an issue here.

reserve more resources for their yet unborn *son* compared to what they would have kept for daughters. *Third*, boys and elder children are provided with more expensive tutors compared to girls and later-born children. Put differently, boys and elder children face higher unit prices for shadow education, and these higher prices may reflect better quality tutors. Besides, although not particularly a consequence of elder son preference, but this rather stems out of ideal family size effects, I find that parental resource allocation on shadow education decreases once the family achieves its ideal fertility size. In sum, intra-family allocation decisions appear to be the determining factor of such patterns.

I then quantify the explanatory power of such unequal intra-household distribution in children's educational achievements. Specifically, I examine to what extent the disparity in shadow education expenditures may account for the widely documented gender gap in cognitive test scores (Fryer and Levitt, 2010). The evidence comes from a decomposition analysis which identifies the contribution of various observable factors on the gender gap in test scores. First, I find a sharp gender gap in math, reading, and writing test scores. It appears that there is a robust positive correlation between shadow education expenditures and mathematics, writing, and reading test scores. Decomposing the data, I find that around 14% of the explained variations (composition effect) in gender gap in math test score may be attributed to the disparity in shadow education expenditures. The contributions appear to be smaller but still significant in writing and reading tests, amounting to 11% and 8%, respectively. The inequality contributions of the disparity in shadow education expenditures in test scores are highest among the firstborns, and then a decreasing pattern emerges down the birth orders. For instance, shadow education accounts for about one-third of the explained variation in the gender gap in math among the firstborns. The corresponding contributions among the firstborns are 27.84% and 16.49% in writing and reading scores. Moreover, shadow education expenditures explain substantial proportion of the unexplained gap (due to structure effect). Together, it appears that intra-household disparity in educational resources translates into gender gap in test scores, and consequently puts girls in disadvantage in educational attainments.

In interpreting these results, it is important to keep in mind the descriptive nature of the analysis. While the chapter shows evidence of elder son preference in shadow education expenditures, and some of its possible consequences, it is silent on the causal link between shadow education expenditures and cognitive test scores.[6] The chapter is also silent on the causes of elder son preference.[7] [8]

This chapter adds to several strands of the literature. *First*, and most directly, I contribute to the literature on son preference in low and middle income countries (Jayachandran and Pande, 2017; Bhalotra et al., 2010; Tarozzi and Mahajan, 2007; Kingdon, 2005; Deaton, 1997). This study is closer to the expenditure based approach that tends to estimate within-household gender disparity from aggregate household expenditures data (Deaton, 1997). I deviate methodologically by exploring resource allocation from individual data, and thus I am not compelled to indirectly estimate gender discrimination through an Engel curve approach. A few studies though examine son preference from individual data on educational expenditures (Zimmerman, 2012; Kingdon, 2005),[9] capturing educational expenditure data nonetheless is found to be challenging as formal education is free in most countries. In fact, exploiting the phenomenon of shadow education is the major advance which enable us to circumvent primary limitations in the literature. I further add to the existing evidence by showing parental preference for both sons and elder children, and thus elder son is found to be situated at the top of the hierarchy in intra-household resource allocation.

Second, I also contribute to a small economics literature on private supplementary tutoring. The literature is still in its infancy, although a few papers have examined shadow education in varied contexts (Dang and Rogers, 2015; Jayachandran, 2014; Lee, 2008). For instance, Jayachandran (2014) explores how shadow education incentivizes teachers in Nepal to teach less during school hours. As a consequence, she finds that students suffer in exams. Whereas the primary objective of Lee (2008), and Dang and Rogers (2012) is to test the quantity-quality tradeoff. Similar to them, this chapter shows that shadow education may be used as a

[6] However, there is abundant evidence on the importance of parental investments on children's economic success (see Almond et al., 2018).

[7] There is a sizeable literature on the potential sources of son preference in developing countries (see Carranza, 2014; Alesina et al., 2013).

[8] It is also important to know who in the family decides on investments in children. There is a large literature documenting that children enjoy more resources if mothers have the decision making power (Attanasio and Lechene, 2012; Browning et al., 2010). In the Indian context, advantaged mothers at least are better informed of child functioning (See Blunch and Datta Gupta, 2020a, 2020b).

[9] Few studies examine intra-household gender disparity in school choice in India. See Sahoo (2017) and Maitra et al. (2016).

measure of direct parental monetary investments, but I do so to study intra-household resource allocation.

Third, my results also speak to the literature on gender gaps in test scores, in particular mathematics test scores (Contini et al, 2017; Bharadwaj et al, 2016; Fryer and Levitt, 2010). The empirical evidence is remarkably consistent. Gender gaps in mathematics increase with years of schooling. Socio-economic background often fails to explain such gender gaps. I provide suggestive evidence that intra-household resource allocation may play a significant role. Specifically, I find that a substantial proportion of gender gaps in test scores may be attributed to disparity in shadow education expenditures. In fact, shadow education is more important when it comes to math as compared to other tests. While a significant portion of gender variation in math test scores cannot be explained, these results highlight one possible channel through which gender gaps in test scores may arise.

Fourth, this study is also related to the literature on parental investment and skill development. A number of studies identify and estimate production functions of skills (Attanasio et al., 2018; Falk et al., 2019; Cunha et al., 2010). One of the primary obstacles in the literature is that parental investments are inherently unobservable. My primary contribution to the literature is that I use direct parental educational investment that was previously unnoticed in the literature.

This chapter unfolds as follows: Section 2.2 provides background information, a brief introduction to shadow education, and a description of the data. Section 2.3 describes the empirical strategy and presents baseline estimates. Section 2.4 replicates baseline estimates using an alternative data set. Section 2.5 tests for several features in the data that may be attributed to elder son preference. Section 2.6 explores the impact of shadow education on test scores. Section 2.7 concludes.

2.2 Background and Data

Relevant Institutional Details

Shadow education is an integral part of a student's educational curriculum in many countries.[10] One strand of the literature argues that shadow education is going to be the norm in many parts of the world, rather than the exception (Bray, 2010). Nonetheless, it is difficult to conceptualize the exact nature of shadow education and it takes diverse shape in different contexts (Mori and Baker, 2010; Bray, 2010, 2009, 1999). Bray (1999) for instance identifies three dimensions of shadow education: supplementation, privateness and academic subjects. Shadow education is 'supplementary' to formal education and does not replace formal education. Tuitions are given on academic subjects only. These supplements are not provided by public providers but typically by a private entity, or quite often by public teachers in private settings. Nevertheless, these characteristics take different shapes in different places and are often adapted to the local setting. In the Indian context, arguably, it should be seen as private supplementary tutoring, especially out of school private tuition, and data used in this chapter capture this private supplementary tutoring.[11] Parental expenditure on shadow education constitutes a substantial proportion of household income. In fact, a typical middle class household often spends more money on shadow education than formal education (Roy, 2010).

Data and Summary Statistics

The data used in this chapter comes from the Indian Human Development Survey (henceforth IHDS) II (2011-12) and includes a sample of more than 52,000 school/college-going

[10] In Hong Kong, for example, around 72% of 12th grade students take shadow education (Zhan et al, 2014). The figures are significant and growing in new contexts where shadow education was relatively unknown, such as Armenia, Azerbaijan, Georgia (Kobakhidze, 2017; 2014), Egypt (Hartman, 2008), many states in the Mediterranean (Bray et al. 2013), Canada (Aurini and Davies, 2004) and increasingly in many European countries. In many contexts, the growth is substantial. For instance, shadow education in major Canadian cities has grown between 200-500% in the past 30 years (Aurini and Davies, 2004).

[11] From a parental view, investment in shadow education can be a compensating investment or a reinforcing one. Shadow education investment is also norm-driven. In other words, parents respond to the current trend of shadow education expenditures in their community and peers. In that sense shadow education may be interpreted as a positional good (Bray and Lykins, 2012). As a consequence, shadow education expenditures may not be driven purely by individual ability, rather by existing cultural and gender norms.

students.[12] My main variable of interest is shadow education expenditures. In addition, I also analyze the likelihood of shadow education attendance and weekly duration of shadow education.

To be included in the sample, children should be at school or in college. I restrict the sample to children on whom there is information (without missing information) about expenditure on private supplementary tutoring.[13]

[12] IHDS is a panel data with a first round conducted in 2004-05. In my analysis, I use the second round of data.

[13] To be precise, I drop 1641 observations which amounts to 3% of the total observations.

Table 2.1: Summary Statistics

Variables	Full Sample	Boys Sample	Girls Sample
Yearly Expenditures on Shadow Education (INR)	692.02	774.98	598.13
	(2660.93)	(2923.16)	(2325.25)
Weekly Duration of Shadow Education (Hour)	2.40	2.59	2.19
	(2.40)	(5.41)	(5.08)
Proportion of Students Taking Shadow Education	0.23	0.24	0.21
	(0.42)	(0.43)	(0.41)
Yearly Expenditures on Shadow Education for Students with Non-Zero Expenditures	3061.11	3190.82	2888.88
	(4906.09)	(5241.76)	(4415.78)
Age	12.06	12.17	11.94
	(4.88)	(4.95)	(4.80)
Math Test Score	standardized	0.071	-0.045
		(0.99)	(1.00)
Writing Test Score	standardized	0.029	0.001
		(0.99)	(1.00)
Reading Test Score	standardized	0.046	-0.007
		(0.98)	(1.00)
Standard[1]	6.51	6.52	6.49
	(4.31)	(4.32)	(4.31)
Private School	0.32	0.35	0.29
	(0.47)	(0.48)	(0.45)
Teacher Attendance	0.16	0.16	0.16
	(0.37)	(0.36)	(0.37)
Ever Repeated an Exam	0.08	0.09	0.07
	(0.27)	(0.28)	(0.26)
Mother Age	36.62	36.76	36.46
	(6.78)	(6.81)	(6.73)
Mother Education[2]	5.11	5.05	5.17
	(4.88)	(4.88)	(4.89)
Proportion of Families Desire Extra Son	0.20	0.18	0.22
	(0.40)	(0.39)	(0.42)
Desired Fertility	2.55	2.53	2.58
	(0.99)	(0.97)	(1.01)
Household Head's Education (No. of Years)	8.51	8.49	8.53
	(4.96)	(4.97)	(4.95)
Yearly Household Income (INR)	139714	141774	137382
	(248990)	(265770)	(228497)
Hindu	0.80	0.80	0.79
	(0.80)	(0.40)	(0.40)
Forward Caste	0.28	0.28	0.28
	(0.45)	(0.45)	(0.45)
Urban	0.35	0.35	0.35
	(0.48)	(0.48)	(0.48)

Notes: This table reports means of the dependent variables and covariates. Standard errors are reported in parentheses. Around 47% of the full sample constitutes girls. Average sample sizes (sample sizes vary across outcomes): Full 52360, Boys 27800, and Girls 24560.

(1) Grade of study: A bachelor's degree is coded as 15, above bachelor's as 16, and 1-12 grades are coded as 1-12 respectively.

(2) Education: A bachelor's degree is coded as 15, above bachelor's as 16, and 1-12 grades are coded as 1-12 respectively.

Table 2.1 reports summary statistics. On average, parents spend INR 692.02 (~ USD 12.82[14]) yearly on every child in shadow education, and the average duration of shadow education is 2.40 hours per week. Conditional on shadow education attendance, the expenditure amounts to INR 3061.11 per child, which is slightly lower than the yearly expenditures on formal schooling, INR 3695. In the Indian context, this expenditure is substantial. The annual shadow education expenditures constitute around 5% of annual household income in these families. In the data, nearly 23% of students take shadow education.[15]
The average family[16] size in the sample is 2.94, and around 47% of the sample are girls. The average age of the children in the sample is about 12.06 years.[17]

In the more recent ASER, 2016 survey data that I use in section 2.4, conditional on shadow education attendance, annually, parents spend INR 3205 per child on shadow education.[18]

Columns 2 and 3 in Table 2.1 show stark gender differences in shadow education expenditures and weekly duration of shadow education. In both cases, boys have an advantage over girls. For instance, on average, parents spend almost 30% more on boys as compared to girls. In absolute terms, annual household expenditures on shadow education averages INR 598 for girls and INR 775 for boys. In addition, girls are 3%-points less likely to be engaged in shadow education. I also observe gender disparity in the type of schooling. To be precise, 35% boys are reported to be in private schooling as compared to 29% for girls. Girls tend to come from slightly better educated families, and families with higher ideal family size. In other words, families on average with lower ideal family size are more likely to have more boys than girls. Another noteworthy difference is whether a child has ever repeated in an exam. Girls are less likely to repeat an exam than boys. To be specific, around

[14] I use an exchange rate of INR 54/ USD for 2011.

[15] The phenomenon is rapidly growing. The data collected in 2011-12. I may expect the pattern to be reversed in more recent data.

[16] Throughout, family size refers to the number of children per mother. Another potential candidate, perhaps a better fit in a context of extensive joint families, is the number of children present in the household. However, I follow the literature and measure family size at the mother level. In addition, I do not account for expected number of unborn children. As I discuss later, this may substantially limit the ability to control for family size.

[17] The corresponding sample size by birth order and gender is reported in Table A1 in the Appendix A.

[18] With an average of 2.11 children between five to sixteen in the survey data, parents on average spend INR 6763 on shadow education. However, parents are likely to spend more on children above sixteen, the expenditures reported here are a lower bound of parental expenditures on shadow education.

7% of girls have ever repeated an exam as compared to 9% for boys. Most other background variables are fairly similar in magnitude.

Overall, the evidence in Table 2.1 indicates that on average parents spend less on girls as compared to boys and girls are less likely to be enrolled in shadow education as compared to boys. This suggests that girls are in disadvantageous position in intra-household resource allocation, and parents clearly favor boys over girls in allocating monetary resources. In the next section, I provide formal evidence of whether these descriptive patterns hold after controlling for confounding characteristics by investigating different facets of son advantage.

2.3 Son Preference and Parental Resource Allocation

A. Empirical Framework

My objective is to explore intra-household allocation of shadow education expenditures by gender. Specifically, I examine such gender differences across birth orders. As I will be using another data to check the robustness of the empirical patterns, I standardize shadow education expenditures to have zero mean and standard deviation one. I consider the following specification:

(2.1) $$Y_i = \alpha + \sum_{j=2}^{4} \beta_j BO_{ij} + \sum_{j=1}^{4} \delta_j BO_{ij} \times Girl_i + \gamma X_i + \varepsilon_i,$$

where, Y_i is the shadow education expenditure of individual i, and BO_{ij} is a dummy for birth order j. Here β captures coefficients on the birth order variables, δ captures coefficients on the birth order and girl interaction variables, and, ε is an error term. X is a vector of background variables that includes a set of child, mother, household and school-specific covariates. X includes child age and its quadratic and standard of current study. I also control for two school level variables – private school dummy, and another dummy for whether teachers attend school regularly. I include maternal literacy, age and a quadratic in maternal age. Other covariates are household specific, e.g., caste, religion, income, education of the head of household and dummy for urban residence. Standard errors are adjusted for clustering at the primary sampling unit (PSU) level.

In specification (2.1), β and δ are the primary parameters of interest. The identification assumption is that the birth order effects reflect prenatal and postnatal environments, but genetic makeup remains constant. In fact, treating birth orders as a natural experiment is a common identifying assumption (Black et al., 2018). The crux of my identification strategy is thus to compare children at different birth orders who are otherwise genetically similar. However, later-born children are more likely to be observed in larger families. As a first response, I augment (2.1) to control for family size. Formally, I estimate a baseline model

(2.2) $$Y_i = \alpha + \sum_{j=2}^{4} \beta_j BO_{ij} + \sum_{j=1}^{4} \delta_j BO_{ij} \times Girl_i + \gamma X_i + \mu S_i + \varepsilon_i$$

where *S* denotes family size. While one may control for family size, it is quite challenging to observe precise family size in the data as most families may not have completed their fertility yet. In addition, larger families are more likely to be poor, less educated and rural.[19] In other words, fertility and family characteristics are highly correlated. I have four strategies to enhance (2.2) to ensure that such differential fertility selection and confounding characteristics do not drive the results. These are controlling for observable maternal characteristics, controlling for observable household characteristics, controlling for observable neighborhood characteristics, and replicating estimates in a sample of families who are likely to have completed fertility. While the first three strategies directly control for confounding characteristics, the fourth specifically addresses the potential omitted variable bias due to unobserved family size.

To elaborate, my primary approach to control for factors that are fixed within a family is to include mother fixed effects, and thus difference out time-invariant family characteristics including family size. This approach will further remove any residual association between birth order and family factors. Essentially, this approach generates comparisons between

[19] The average family size in rural and urban samples is 3.04 and 2.75, respectively. Likewise, average annual household income in the rural sample is INR 116,707, much lower than INR 182,592, that an average urban household earns annually.

siblings within the same family by examining the extent to which differences in shadow education expenditures between two siblings is due to birth order and gender.

Inclusion of mother fixed effects addresses confounds due to factors that are fixed within families. It cannot address confounds due to household characteristics.[20] Fertility behavior is highly correlated with household factors, and one may also expect a close association between birth order and household characteristics. To address these, I include household fixed effects. By doing so, I control for within household fixed factors, and also remove any association between birth order and household characteristics. Perforce, I compare children within the same household to estimate birth order and gender effects.

Specification (2.2) may suffer from confounding neighborhood factors. Shadow education can be extremely norm-driven, and the culture may vary substantially across neighborhoods. For instance, shadow education is highly prevalent in some areas, but may not be of the same magnitude in other areas. Such area-specific factors may drive the results. I do not have detailed variables to control for neighborhood characteristics. I control for neighborhood specific fixed factors by including PSU (primary sampling unit) fixed effects. In rural areas, PSU is a village, and in urban areas it denotes a neighborhood. This approach eliminates any such area-specific confounding characteristics.

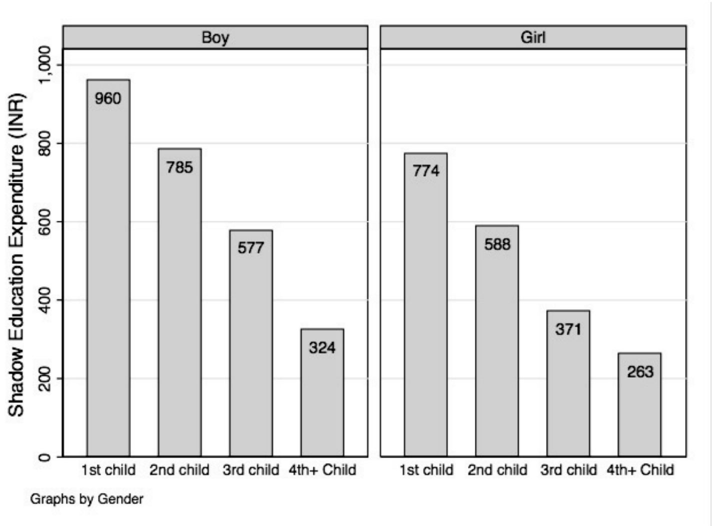
To implement the fourth strategy, I select a sub-sample of mothers who are likely to have completed fertility. Families who are yet to complete fertility may keep (or save) resources for future births, which may affect the estimates. To that end, I make an attempt to select a sample of families with completed fertility in two ways: *First,* I use a survey question to select a completed fertility sample. IHDS II asked women whether they wanted any more children. My subsample consists of those women who answered negatively. Additionally, I include those women who expressed that either they or their husbands were sterilized, and those who were not fertile anymore. *Second,* I use biological age of mothers to proxy for completed fertility. Specifically, women who crossed forty years of age are likely to have completed fertility. Eventually, the completed fertility sample consists of the union of these two subsamples.

[20] In the case of nuclear families, a household is equivalent to a family. In the case of joint families there is a distinction between the nuclear family and the household/joint family.

In the data a large number of children do not attend shadow education. If there is within household selection on shadow education attendance, the estimates could be biased. Nonetheless, only a small number of children in the data do not attend shadow education when at least one child in the family attends shadow education, indicating that shadow education attendance is a household level choice rather than a choice based on individual children.[21] Likewise another concern is selection on schooling type (i.e., public vs private), in particular when parents have to pay tuition fees in private schools compared to almost free public schools.[22] In other words, in such a situation shadow education spending could be conditioned on within household choice of school type.[23] Similar to shadow education attendance, few children go to private schools when any other sibling is in public school.[24]

B. Results

Figure 2.1: Shadow Education Expenditures



[21] In the data, there are 2046 (or 4% of the sample) such children in multi-child families.

[22] In the state of West Bengal at least, public schools collect a very small fee.

[23] For instance, parents may send some children to private schools and compensate others who are in public schools by purchasing shadow education, or vice versa.

[24] To be precise, only 415 (or 0.01% of the sample) children in multiple children families are in public school when at least one child in the family is in a private school.

Notes: This figure shows unconditional yearly shadow education expenditures by birth order and gender.

I start by graphically documenting parental allocation of resources on shadow education by birth order and gender. Figure 2.1 plots mean yearly shadow education expenditures by birth order segregated by gender. Irrespective of gender, I observe a birth order disadvantage for later-born children in shadow education expenditures. For instance, firstborn boys attract INR 175 or 22.29% more household shadow education expenditure compared to second born boys. Upon further examining these patterns by gender, I observe a sharp gender difference in every birth order. For instance, parents outspend on boys compared to girls by 24.03% (INR 186), 33.50% (INR 197) and 55.53% (INR 206) in first three birth orders.

Table 2.2: Main Results

Dep. Var. Shadow Education Expenditures (SD)	(1)	(2)	(3)	(4)	(5)	(6)
Girl	-0.066***					
	(0.011)					
Girl × 1st Child		-0.079***	-0.093***	-0.093***	-0.079***	-0.083***
		(0.018)	(0.030)	(0.029)	(0.017)	(0.020)
Girl × 2nd Child		-0.071***	-0.063**	-0.067**	-0.068***	-0.073***
		(0. 018)	(0.028)	(0.027)	(0.018)	(0.019)
Girl × 3rd Child		-0.055***	-0.060	-0.062	-0.051***	-0.057***
		(0.018)	(0.037)	(0.037)	(0.019)	(0.018)
Girl × 4th+ Child		-0.022	-0.027	-0.028	-0.024	-0.023
		(0.015)	(0.026)	(0.026)	(0.018)	(0.015)
2nd Child		-0.036*	-0.017	-0.015	-0.001	-0.040
		(0.020)	(0.030)	(0.029)	(0.020)	(0.021)
3rd Child		-0.065**	-0.006	-0.002	-0.021	-0.070**
		(0.025)	(0.041)	(0.040)	(0.026)	(0.025)
4th+ Child		-0.040	0.024	0.027	0.017	-0.045
		(0.025)	(0.053)	(0.051)	(0.026)	(0.026)
Forward Caste	0.140***	0.140***	-	-	0.023	0.141***
	(0.02)	(0.020)			(0.021)	(0.021)
Hindu	0.046***	0.046***	-	-	0.040**	0.045**
	(0.02)	(0.016)			(0.018)	(0.017)
Urban	0.150***	0.150***	-	-	-	0.152***
	(0.02)	(0.023)				(0.024)
Private School	0.019	0.018	0.107*	0.102*	0.101***	0.020
	(0.02)	(0.018)	(0.059)	(0.057)	(0.024)	(0.018)
Standard	0.027***	0.028***	0.041***	0.041***	0.041***	0.027***
	(0.00)	(0.004)	(0.008)	(0.008)	(0.004)	(0.004)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes
Mother Fixed Effect	No	No	Yes	No	No	No
Household Fixed Effect	No	No	No	Yes	No	No
PSU Fixed Effect	No	No	No	No	Yes	No
Completed Fertility	No	No	No	No	No	Yes
Observations	45148	45148	41660	41660	45148	43184

Notes: Every column reports a separate linear regression. Standard errors reported in parentheses are robust to within primary sampling unit (PSU) clustering. The dependent variable, shadow education expenditures is standardized. 2nd and 3rd child are indicators for children whose birth orders are 2 and 3, and, 4th+ indicates children with birth order 4 or later. In columns 3, 4 and 5, caste, religion and urban/rural effects are absorbed by mother, household and PSU fixed effects respectively. Other controls always include child's age and its square, mother age and its square, maternal education, household income, household head's education, family size and a dummy for teacher attendance at school. Backward caste refers to those belong to the scheduled caste category. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Next, I report estimates based on specification (2.2) in Table 2.2. In column 1, I show that parents spend 0.066 SD (INR 175.43)[25] less on girls compared to boys, and the effect is statistically significant. In column 2, I explore gender disparity across birth orders. Two patterns in shadow education expenditures are immediate from the estimates in column 2: *first*, I observe sharp gender differences and this disparity is consistent across birth orders, and *second,* the evidence further shows that later-born children attract less shadow education expenditures. Column 2 of Table 2.2 shows that coefficients of birth order and girl interaction terms are negative and significant. In other words, parents spend less on girls than boys in every birth order, indicating son advantage in intra-household allocation of shadow education expenditures. For instance, I observe that yearly shadow education expenditure is on average 0.079 SD (INR 211.17) less for firstborn girls than firstborn boys. The effect is significant. In contrast to evidence in Figure 1, once I control for confounding factors, the inequality tends to fall with birth orders. For instance, moving from firstborns to second-borns, the magnitude of gender difference estimate falls from 0.079 SD to 0.071 SD, then further drops to 0.055 SD and 0.022 SD for children at third, and fourth or later birth orders respectively.

I also estimate the role of birth orders in shadow education allocation. Column 2 of Table 2.2 depicts male birth order disadvantages (i.e., coefficients of *2^{nd} child*, *3^{rd} child* and *4^{th+} child* are negative). I observe a decreasing shadow education expenditure by birth orders. For instance, parents on average spend yearly 0.036 SD less on second born boys compared to firstborn boys. This unequal distribution is further inflated to 0.065 SD between firstborn third born boys, but eases to 0.040 SD between first and fourth or later born boys. The drop in shadow education expenditures appear to be non-linear. For instance, the gap between second born boys and on third born boys is 0.029 SD (*3^{rd} Child - 2^{nd} Child*), but fourth or later-born boys gain by 0.025 SD as compared to third borns.

Furthermore, the birth order disadvantage appears to be flatter for girls. For instance, parents on average spend 0.028 SD less yearly on second born girls than the first one (i.e., *2^{nd} Child + Girl × 2^{nd} Child - Girl × 1^{st} Child*). Moving from firstborns to third borns, girls lose, on average, 0.041 SD in shadow education expenditures.

[25] The empirical patterns are invariant to the dimension on which the outcome is measured. The corresponding estimates for shadow education expenditures in monetary value is reported in Table A2 in the Appendix A.

A parallel analysis with weekly duration of shadow education, reported in Table A3 in the Appendix A, show qualitatively similar patterns. Taken together, these observed patterns presented above clearly indicate elder son advantage in intra-household parental resource allocation.

In column 3 of Table 2.2, I control for factors fixed within family by adding mother fixed effects. The estimates for birth order – girl interactions increase from -0.079 SD to -0.093 SD for the firstborns. However, the estimates tend to decline for later-born children. The coefficients remain statistically significant, although effects are estimated less precisely. Therefore, by using sibling comparison for identification, the girl disadvantage tends to be smaller compared to the baseline estimates. Likewise, the magnitude of the birth order disadvantage decreases, but at the cost of significant loss in precision. Taken together, the estimates are in general smaller in magnitude, yet, the overall empirical patterns remain qualitatively similar.

The household fixed effects estimates reported in next column 4 are markedly close to the mother fixed effect estimates.[26] The empirical patterns remain similar after controlling for PSU fixed effects (column 5). In fact, the PSU fixed effect estimates are closer to the baseline estimates.[27]

The last set of estimates pertains to effects based on the sample of mothers with completed fertility. Since 96% of the children are included in the sample with completed fertility, the effects reported in column 6 are similar to those reported in column 2..[28] Together, the evidence suggests that heterogeneous family selection is not likely to drive the results.

Turning now to other regressors, column 2 shows that children from better socioeconomic backgrounds spend more on shadow education. Specifically, forward caste children spend 0.14 SD more compared to backward caste children. I find similar advantage for urban

[26] This is perhaps not surprising as household fixed effects are the same as mother fixed effects in nuclear families. The sample consists of 27,821 unique families, and 24,156 unique households.

[27] I also estimated a specification controlling for mother, household and PSU fixed effects simultaneously, and observed fairly similar results. Estimates are available on request.

[28] I also replicated using a sample without biological age restriction and got qualitatively similar results.

children; parents on average spend 0.15 SD more on children than their rural counterparts. Furthermore, Hindus spend more on shadow education compared to non-Hindus.

C. Heterogeneous Effects

Next I ask whether the empirical patterns discussed above differ depending on family circumstances. For instance, son preferences may differ between higher and lower caste families, and may also depend on other group identities. For example, the rural-urban divide is well-documented in the current context. It would be interesting to quantify such birth order disadvantages and girl disadvantage separately for these groups. Table 2.3 presents baseline estimates by allowing for heterogeneous effects by different groups. The patterns are qualitatively similar across groups but vary in magnitudes.

Table 2.3: Heterogeneity in Effects

Dep. Var. Shadow Education Expenditures	Rural (1)	Urban (2)	Forward Caste (3)	Backward Caste (4)	Private School (5)	Public School (6)
Girl × 1st Child	-0.080***	-0.080**	-0.145***	-0.055***	-0.076***	-0.083***
	(0.018)	(0.039)	(0.042)	(0.019)	(0.028)	(0.024)
Girl × 2nd Child	-0.069***	-0.076**	-0.136***	-0.048**	-0.123***	-0.050***
	(0.021)	(0.034)	(0.045)	(0.019)	(0.041)	(0.023)
Girl × 3rd Child	-0.050**	-0.073*	-0.111*	-0.042***	-0.158***	-0.013
	(0.020)	(0.039)	(0.064)	(0.013)	(0.049)	(0.015)
Girl × 4th+ Child	-0.019	-0.037	-0.010	-0.028*	-0.088***	0.004
	(0.015)	(0.035)	(0.039)	(0.015)	(0.033)	(0.015)
2nd Child	0.001	-0.089**	-0.064	-0.022*	0.027	-0.067**
	(0.023)	(0.037)	(0.053)	(0.018)	(0.046)	(0.026)
3rd Child	-0.032	-0.096**	-0.048	-0.060***	0.065	-0.125***
	(0.029)	(0.047)	(0.084)	(0.017)	(0.059)	(0.023)
4th+ Child	-0.022	-0.023	-0.080	-0.022	0.065	-0.083***
	(0.029)	(0.051)	(0.072)	(0.023)	(0.053)	(0.026)
Observations	29432	15716	12761	32387	14392	30756

Notes: Every column reports a separate linear regression. Standard errors reported in parenthesis are robust to within primary sampling unit (PSU) clustering. Controls always include child's age and its square, current standard of study, mother age and its square, maternal education, household income, household head's education, family size, dummy for teacher attendance at school, and Hindu. ***p < 0.01; **p < 0.05; *p < 0.10.

Columns 1 and 2 report baseline estimates for rural and urban children separately.[29] I observe that girl disadvantage is greater for urban children. Also of interest are separate estimates for forward and backward caste children. Likewise, I find that girl disadvantage is larger for forward caste children compared to their backward caste peers (columns 3 and 4). For instance, forward caste parents spend 0.145 SD and 0.136 SD less on firstborn and secondborn girls compared to firstborn and second born boys. By comparison, such disparity amounts to 0.055 SD and 0.048 SD for firstborns and second borns respectively in backward caste families. These estimates are highly significant. The gender gap appears to be steeper for private school children (columns 5 and 6 of Table 2.3).

[29] The sample split could be endogenous.

One striking pattern revealed in the analysis is that families with better backgrounds exhibit greater gender discrimination.[30] These families also have higher incidence of shadow education. Shadow education expenditures, in fact, can be quite different from investments in normal goods. Shadow education is often interpreted as a positional good in the literature. For instance, in analyzing shadow education expenditures, Bray and Lykins (2012) note that expenditure depends on "whether the amount is adequate relative to the amounts held by peers and competitors" (P. 68). Thus, it could be that these families invest more on their children, but do so more selectively. Another possible explanation could be that these families may have higher future (job market) expectations on their children, specifically on sons. Considering the highly competitive (urban) job market and low female labor force participation in India, these families invest more on their children's future, and do so selectively, seems to be a plausible explanation.[31] However, I do not have such expectation data to test this.[32]

2.4 Evidence from Alternative Data

Next, I try to replicate baseline estimates in an alternative more recent data. I use the ASER data collected in 2016. My analysis with the alternative data can shed light on whether the observed patterns are specific to the IHDS data I am using. Annual Status of Education Report (ASER) is a survey on educational achievements of children between 5 and 16 years of age, conducted every year by a nongovernmental organization, Pratham. ASER enrolls all the children in the household irrespective of their schooling status. I provide some summary statistics in Table A4 in Appendix A. Before going to the estimates, I note few caveats of the ASER data. First, the ASER 2016 data is more recent and larger, but less extensive. For instance, ASER does not collect data on religion, caste, household income and many other important background variables. Second, ASER recruits all children between the age of five

[30] This pattern is not unusual in the South Asian context. A large literature on gender discrimination finds similar pattern (Chamarbagwala and Ranger, 2010; Srinivasan and Bedi, 2008; Agnihotri, 2003). Richer communities and families are more likely to have skewed sex ratios, practice dowry, and have lower female labor force participation as compared to poorer families.

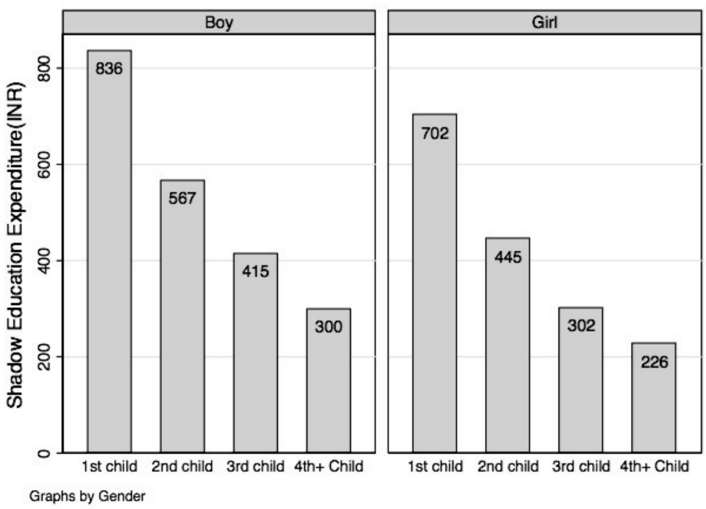
[31] Molina (2020) for instance show that labor market opportunities drive gender differences in human capital decisions in Mexico.

[32] Another possibility could be that parental intra-household resource allocation, at least some part of it, is driven by the migration prospects of the children (see Shrestha and Palaniswamy (2017) for an example of how parental resource allocation responds to increased migration prospects of men in Nepal).

to sixteen in a household. Thus, the exact birth order and family size cannot be observed as ASER recruits only children between 5 and 16 years of age.

With these caveats in mind, I construct birth orders within children in the age range 5 to 16 . Consequently, I can compare the empirical patterns but magnitudes are not directly comparable to those based on the IHDS data.

Figure 2.2: Shadow Education Expenditures in ASER data



Notes: This figure shows mean yearly shadow education expenditures in ASER data by birth order and segregated by gender.

I first graphically demonstrate that the empirical patterns appear to be markedly similar to the baseline analysis. Figure 2.2 presents the intra-household allocation of shadow education expenditures across birth orders and gender. I find both birth order disadvantages for later-born children, and greater expenditure on boys in every birth order. Next, I replicate the baseline specification with ASER, 2016 data. Column 1 of Table 2.4 shows that the coefficients of birth order – girl interactions are remarkably close to the baseline estimates. To be specific, the gender gap appears to be 0.071 SD and 0.057 SD among first and second borns. I also find a clear birth order disadvantage for later-born children. In columns 2 and 3, I difference out fixed factors by including mother and household fixed effects respectively. Together, the empirical patterns appear to be robust to the use of an alternative dataset.

Table 2.4: Baseline Estimates with ASER Data

Dep. Var. Shadow Education Expenditures	(1)	(2)	(3)
Girl × 1^{st} Child	-0.071^{***}	-0.077^{***}	-0.079^{***}
	(0.005)	(0.010)	(0.009)
Girl × 2^{nd} Child	-0.057^{***}	-0.039^{***}	-0.039^{***}
	(0.005)	(0.008)	(0.007)
Girl × 3^{rd} Child	-0.048^{***}	-0.047^{***}	-0.046^{***}
	(0.007)	(0.012)	(0.010)
Girl × 4^{th+} Child	-0.027^{***}	-0.033^{*}	-0.033^{*}
	(0.010)	(0.018)	(0.015)
2^{nd} Child	-0.033^{***}	-0.026^{**}	-0.026^{**}
	(0.005)	(0.011)	(0.009)
3^{rd} Child	-0.013^{*}	0.019	0.020
	(0.008)	(0.018)	(0.015)
4^{th+} Child	-0.011	0.043	0.044^{*}
	(0.012)	(0.026)	(0.022)
Mother Fixed Effect	No	Yes	No
Household Fixed Effect	No	No	Yes
Observations	346231	247420	247420

Notes: Every column reports a separate linear regression. Standard errors reported in parentheses are robust to within village clustering. 2^{nd} and 3^{rd} child are indicators for children whose birth orders are 2 and 3, and, 4^{th+} indicates children with birth order 4 or later. All estimates include controls for child age and its square, current standard of study, mother age and its square, parental education, family size, and dummies for private school, electricity and toilet at home and presence of graduate in the household. In columns 2 and 3, the effects of some of the variables is absorbed by mother and household fixed effects respectively. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

2.5 Elder Son Preference and Desired Family Size

In the previous two sections, I observed a birth order disadvantage in shadow education expenditures and a sharp gender disparity in every birth order. In addition, I also observe that inequality in shadow education expenditures is more pronounced in early birth orders, pointing to an elder son advantage. To interpret these results, I examine whether the observed empirical patterns in shadow education expenditures may be attributed to the well-documented elder son preference typical in the current context. To that end, I test several predictions that stem out of elder son preference.

Prediction 1: *If parental resource allocation is due to elder son preference, both boys and girls without an elder brother would receive more shadow education expenditures than their counterparts with an elder brother. Moreover, there would be less shadow education expenditure on girls with an elder brother than boys with an elder brother.*

Son preference predicts that girls without an elder brother should be in an advantageous position as sons extract more resources in the sibling rivalry game. On the other hand, sex biased fertility behavior further indicates that girls without an elder brother may be in disadvantageous position for two reasons: *first*, if these girls have a younger brother, then the younger brother being the elder son in the family could draw more resources, and, *second*, if the family lacks a son, they may try again for a son and in turn may keep a chunk of their resources for the unborn son. The net effect depends on the relative strength of one mechanism over the other.

The second part of prediction 1 implies: girls with elder brother would have different outcomes than boys with elder brother due to son preference. Girls with elder brother face two adverse positions in the intra-household hierarchy – being 'girl' and having an elder brother, however, for boys, such gender disadvantage does not exist although they may suffer from a birth order disadvantage.

To test the above prediction, I estimate the following specification:

(2.3) $$Y_i = \alpha + \sum_{j=2}^{4} \beta_j BO_{ij} + \alpha_1 EB_i + \alpha_2 EB_i \times Girl_i + \alpha_3 Girl_i + \gamma X_i + \mu S_i + \varepsilon_i$$

where, EB_i is a dummy which takes on a value of 1 if child i does not have an elder brother and zero otherwise.

Table 2.5: Son-Biased Behavior

Dep. Var. Shadow Education Expenditures	(1)	(2)	(3)
Girl	-0.048***		-0.068***
	(0.015)		(0.012)
No Elder Brother	0.033**		
	(0.016)		
No Elder Brother × Girl	-0.027		
	(0.023)		
Son Desire		-0.052***	
		(0.015)	
∆Ideal			0.007
			(0.013)
1 (∆Ideal ≥ 0)			0.009
			(0.018)
1 (∆Ideal ≥ 0) × ∆Ideal			-0.010
			(0.017)
p-values			
No Elder Brother + No Elder Brother × Girl = 0	0.77		
Girl + No Elder Brother × Girl = 0	0.00		
Observations	45148	45148	43411

Notes: Every column reports a separate linear regression. Standard errors reported in parentheses are robust to within primary sampling unit (PSU) clustering. 2nd and 3rd child are indicators for children whose birth orders are 2 and 3, and, 4th+ indicates children with birth order 4 or later. Controls always include child's age and its square, current standard of study, mother age and its square, maternal education, household income, household head's education, family size and dummies for private school, teacher attendance at school, Hindu, forward caste and urban residence. No Elder Brother is a dummy for those children who do not have elder brother. Son Desire is a dummy if the family desires more sons. ∆Ideal represents distance from ideal fertility size, i.e., (birth order – desired fertility size), and, 1 (∆Ideal ≥ 0) is a dummy if the family has already reached the desired family size. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Column 1 of Table 2.5 presents estimates of specification 2.3. The coefficient of no elder brother dummy (α_1) is the main coefficient of interest, capturing differential treatment by families in the presence of elder sons. The coefficient of no elder brother dummy is positive. To be specific, boys without an elder brother benefit from an average 0.033 SD more in yearly expenditure. The joint effect *No Elder Brother + Girl × No Elder Brother* (i.e., α_1 + α_2 in specification 2.3) captures the net effect of 0.006 SD on girls lacking an elder brother. It is positive but statistically insignificant. Girls without elder brothers are at least qualitatively in a better position compared to their peers with an elder brother.[33] Thus the effect of no-elder-brother is concentrated on boys. For girls, two mechanisms may have counterbalanced each other. As I will show in the next prediction, those families that want the next child to be a *son* keep away resources from their existing children. The data further exhibits that girls with an elder brother receive fewer resources compared to boys with an elder brother: the net effect *Girl + Girl × No Elder Brother* is significantly negative and amounts to 0.075 SD. Elder son preference elucidates that there is a within household hierarchy with elder sons are positioned at the highest level followed with other sons and then girls.

Prediction 2: *Children in families that still desire a son spend less on shadow education compared to their counterparts in families with no such gender preference*.

Families may reveal that they desire a son for for two potential reasons.[34] One possibility is that they are yet to reach their desired sex ratio. Another possibility could be that they have a preference for sons. If the choice is due to son preference, then son preference predicts that would keep more resources for their unborn son compared to what they would have kept for daughters. In turn, children in these families may have less shadow education expenditures compared to families with no such gender preference. On the other hand, if a family prefers a son to fulfill its desired sex ratio, they may keep resources for the extra child irrespective of the gender of the child. Consequently, I may not observe different (gendered) expenditure pattern in these families.

[33] This pattern is in contrast to Jayachandran and Pande (2017) who found a net negative effect for girls without elder brothers on health outcomes.

[34] The survey instrument asked women to state their sex preference for a future child. In our data, around 28% of the women wanted their next child to be a son, compared to just 6% who wanted the next child to be a daughter.

Column 2 of Table 2.5 reports estimate of the baseline specification with an additional dummy for son desire, but without a birth order - girl interaction. The coefficient is negative and statistically significant. Families that prefer their next child to be a son spend 5% of a standard deviation less on each existing child. In other words, children living in households where parents desire a son as opposed to those families where parents prefer daughters or are gender neutral, are likely to experience lower annual expenditure on shadow education.

Prediction 3: *Sons and older children are sent to more expensive tutors*.

In other words, on average parents not only spend less on girls as compared to boys, but girls are sent to less expensive tutors. In contrast, sons have more parental resources in general, but additionally they are sent to more expensive tutors, and consequently I can assume to better quality tutors. Put differently, boys and girls face different unit prices, and these unit prices may reflect quality of such private tutors.

The shadow education market has a wide range of service providers.[35] Some providers are more expensive than others. Parents not only choose how much to spend on shadow education, but also choose the quality of the providers as measured by hourly tuition rates. Even if both boys and girls take similar duration of shadow education, parents can provide sons with more expensive tuition.[36]

I estimate private tuition expenditures for 100 hours[37] in the pooled sample, and additionally in rural and urban areas separately. Figure A1 in Appendix A plots the cost of 100 hours of tuition expenditures segregated by gender and birth order in pooled, rural and urban sub-samples respectively. I observe a sharp gender disparity in unit prices. Table 6 reports the estimates. Column 1 shows that parents spend INR 24.40 less on girls for 100 hours of tuition

[35] In June 2018, I conducted ten key informant interviews with private tutors and several interviews with students and parents in a district town in West Bengal. I found a vibrant shadow education market with multiple providers. I found providers who gave tuition in multiple subjects, providers who specialized in a single subject, providers who took only primary schools students to those who only teach to college students. The perceived quality of the tutors differs widely as do their fees.

[36] These private tutors can also provide intensive in-home tuition at higher fees. In general, in most cases I find that these intensive trainings are given in student's home, however, tutors can also give tuition in their own places but with limited and exclusive students. In general, this is often the case that parents provide their 'good' children with these intensive tutoring.

[37] For ease of interpretation, I scaled-up the unit of analysis to 100 hours.

as compared to boys. Disparities are more prominent in the urban sample, perhaps due to the availability of a wider range of service providers in urban areas. Column 3 in Table 2.6 reports that in the urban sample the tuition rates per 100 hours in the case of girls are INR 33.68 lower than those for boys, though the magnitude drops after controlling for background variables (as presented in column 4). I further observe negative birth order effects, indicating that later born children are sent to less expensive tutors.

Table 2.6: Shadow Education Expenditures in 100 Hours (in INR)

Expenditures in 100 hours (INR)	Pooled Sample		Urban Sample		Rural Sample	
	(1)	(2)	(3)	(4)	(5)	(6)
Girl	-24.40**	-10.52	-33.68**	-24.97*	-20.07*	-6.69
	(9.90)	(11.92)	(15.85)	(17.63)	(12.16)	(15.43)
Birth Order	-29.10***	-4.56	-38.31***	-13.67	-20.98***	-3.98
	(3.96)	(5.69)	(8.89)	(10.53)	(4.33)	(6.72)
Controls	No	Yes	No	Yes	No	Yes
Observations	52360	45558	18284	15932	34076	29626

Notes: Every column reports a separate linear regression. Standard errors reported in parenthesis are robust to within primary sampling unit (PSU) clustering. The depended variable is in monetary term, Indian Rupee. All estimates include controls for child's age and its square, current standard of study, mother age and its square, maternal education, household income, household head's education, family size and dummies for private school, teacher attendance at school, Hindu, forward caste and urban residence. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Desired fertility size plays a crucial role in strengthening son preference. For instance, prediction 2 presents a case for an interaction between the desired sex ratio and desired family size so I expect:

Prediction 4: *Parental expenditure on shadow education declines after the family achieves the ideal family size*.

Parents have a desired family size, and any forward-looking parent would keep resources for their unborn children till they reach that ideal. After the family has reached its target, they may need to cut the resources, simply because they may have exhausted all their planned resources.[38] To examine the behavior at the margin of ideal family size, I built on Chakravarty (2015) and Jayachandran and Kuziemko (2011) to estimate the following specification:

(2.4) $$Y_i = \alpha + \beta_1 Girl_i + \beta_2 \Delta Ideal_i + \beta_3 1(\Delta Ideal_i \geq 0) + \beta_4 1(\Delta Ideal_i \geq 0) \times \Delta Ideal_i + \gamma X_i + \mu S_i + \varepsilon_i,$$

where, Y_i is the within-household shadow education expenditure on child i, $\Delta Ideal_i$ is defined as (birth order$_i$ – ideal fertility size), measuring distance from the ideal family size, and $1(\Delta Ideal_i \geq 0)$ is a dummy which takes a value 1 if $\Delta Ideal$ is non-negative and zero otherwise. I am interested in the vector of β coefficients: β_2 captures the effect on shadow education expenditures as the family approaches ideal fertility rate, β_3 captures the effect once a family reaches the ideal fertility rate, and, β_4 identifies the effect of distance between birth order and ideal fertility once the family has exceeded the ideal family size.

Column 5 of Table 2.5 presents the estimates. I find parents spend additionally 0.007 SD per child that brings the family closer to the ideal fertility size. Once the family reaches the ideal fertility size, there is a gain of 0.009 SD yearly expenditure but, the expenditure declines by 0.010 SD per child after the family exceeds the ideal fertility size. The estimates are statistically insignificant.

Taken together, it appears that the observed elder son advantage may plausibly be attributed to elder son preference.[39] Next I explore the consequences of elder son preference on children's educational achievements.

[38] Jayachandran and Kuziemko (2011) observed similar maternal behaviour in breastfeeding practices in India.

[39] The gendered pattern may also potentially be due to parental safety concerns. Specifically, parents may be reluctant to send their daughters and younger children to private tutors late in the evening. While this could be a plausible explanation, expenditure patterns such as predictions in this section cannot be explained by safety concerns.

2.6 Shadow Education and Educational Achievements

The previous sections have shown substantial gender disparity in parental allocation of shadow education expenditures. Having uncovered this, a question that follows is – what is the extent to which unequal distribution in shadow education resources affects children's educational attainments? IHDS II randomly tested around 10,000 children on mathematics, reading and writing abilities. The math test consists of three modules.A child gets one point each if he/she can recognize numbers, subtract and divide. Thus, a child can score between zero and three. In contrast, the reading test consists of four modules, and the writing test has two. Because of this variation in testing patterns, I standardize test scores to have zero mean and a standard deviation of one.

Table 2.1 presents descriptive statistics. I observe sharp gender differences in test scores. For instance, girls on average score 12% of a standard deviation less in math compared to boys, although smaller in magnitude, I also observe gender gaps in reading and writing test scores. These patterns raise the natural question, that is, to what extent do intra-household allocation of educational resources (proxied by shadow education expenditures) influence differences in test scores. In Table A5 in the Appendix, I show that shadow education is positively related to test scores.

Shadow Education and Gender Gap in Test Scores

To what extent may differences in test scores be attributed to disparity in shadow education expenditures? To that end, I implement an Oaxaca-Blinder decomposition analysis. I perform a twofold decomposition analysis which separates the test score differentials into an explained part that is explained by group differences in the covariates (composition effect), and an unexplained part (structure effect). The counterfactual parameter is estimated from pooled regressions for both groups. The Oaxaca-Blinder decomposition results are shown in Table 1.7. Panel A reports the decompositions into explained and unexplained parts, and, panel B and C further reports the contributions of some of the covariates to the explained and unexplained variations, respectively. The differences in test scores by gender are sizeable. Girls score 12.5% of a standard deviation less in math, and 4.6% and 6.7% of a standard deviation less in writing and reading scores than their male counterparts. Column 1 of panel B further show that only 32.8% of the gender differences in math test scores can be explained by covariates. The explained part may look small, but this is consistent with the existing literature on gender differences in math test scores (Bharadwaj et al., 2016; Fryer and Levitt, 2010). By contrast, a substantial proportion of the variation in writing and reading test scores may be explained by covariates. For instance, a substantial 82.61% and 59.70% of variation in writing and reading test scores may be explained by covariates.

Table 2.7: Oaxaca- Blinder Decomposition of Test Scores

	Math Scores	Writing Scores	Reading Scores
	(1)	(2)	(3)
Panel A: Decomposition			
Gender Difference	0.125***	0.046***	0.067***
	(0.020)	(0.020)	(0.020)
Explained	0.041***	0.038***	0.040***
	(0.011)	(0.009)	(0.010)
Unexplained	0.083***	0.007	0.027
	(0.017)	(0.018)	(0.017)
Panel B: Percentage contributions to explained variation			
Shadow Education	13.53	11.15	7.98
Mother Education	15.45	13.34	8.72
Caste	2.52	3.87	2.18
Panel C: Percentage contributions to unexplained variation			
Shadow Education	14.31	131.68	27.37
Observations	10185	10132	10222
Other Controls	Yes	Yes	Yes

Notes: Oaxaca – Blinder decomposition. Controls include child's age and its square, current standard of study, mother age and its square, maternal education, household income, household head's education, family size and dummies for private school, teacher attendance at school, Hindu, forward caste and urban residence. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

How much of the explained variation in test scores may be attributed to disparities in shadow education expenditures? Panel B of Table 2.7 shows that shadow education can explain around 14% of the explained gender differences in math test results. The corresponding contributions in writing and reading test scores are 11% and 8% respectively. In fact, shadow education has most explanatory power in test scores relative to other parental and household level variables. For instance, caste explains only a miniscule part of the explained gender gap, whereas, only maternal education has a relatively higher weight as compared to shadow education expenditures. The findings suggest that the gender gap in math, compared to writing and reading scores, can more assertively be attributed to disparities in intra-household resource allocations compared to other subjects.[40]

Shadow education also accounts for a substantial proportion of unexplained gaps in test scores. In other words, shadow education has a higher return for boys as compared to girls. Specifically, around 14.31% of the unexplained variation in math test scores may be explained by gender gap in returns to shadow education expenditures (Panel C of Table 2.7).[41]

In Appendix B, I show that decomposing the gender gap using a Recentered Influence Function regression method (Firpo et al., 2009), and re-weighting function (Firpo et al., 2018) to estimate the counterfactual scenarios yield fairly similar results.

Table 2.8: Oaxaca- Blinder Decomposition by Birth Order

	Math Scores	Writing Scores	Reading Scores
	(1)	(2)	(3)
Percentage contribution of shadow education to explained variations in gender gap			
1st Child	30.75	27.84	16.49
2nd Child	10.71	7.36	8.56
3rd Child	11.29	3.37	2.25

Notes: Oaxaca – Blinder decomposition by birth order. Controls include child's age and its square, current standard of study, mother age and its square, maternal education, household income, household head's education, family size and dummies for private school, teacher attendance at school, Hindu, forward caste and urban residence. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Table 2.8 reports decomposition results by birth orders. The contribution of shadow education is highest among the firstborns. Column 1 suggests that around one-third of the

[40] This should be interpreted in parallel to the literature on gender gaps in math, which suggests that gender differences in math test scores are most persistent compared to other tests (Bharadwaj et al., 2016; Fryer and Levitt, 2010). Thus, unequal shadow education allocation may be an important causal mechanism behind the well-documented gender gaps in math. In chapter 4, I show that elder son preference coupled with family size effects can plausibly explain gender gaps in math test scores in Pakistan.

[41] This is perhaps due to the fact that boys are sent to better quality tutors as we observe in Table 2.6.

explained variation in the gender gap in math among the firstborns may be attributed to disparities in the shadow education expenditures. As one moves down the birth orders, the relative importance of shadow education declines, but nonetheless remain substantial. Family backgrounds matter relatively more among the later-borns. One can observe similar birth order patterns in writing and reading test scores.

In sum, disparities in shadow education expenditures may play an important role in driving gender gaps in test scores. However, in understanding the decomposition results, it is important to keep in mind the descriptive approach of the analysis. This is essentially a partial equilibrium exercise that abstracts from several general equilibrium considerations.[42] Most importantly, although I show that shadow education and test scores are positively related, I cannot rule out the simultaneous determination of shadow education expenditures and test scores. Hence, the analysis should not be imbued with a causal interpretation, and should be understood as an accounting exercise.

2.7 Conclusion

This chapter examined the unequal distribution of parental educational investments across children within households. One of the major limitations of the literature has been that few studies actually measure direct parental monetary expenditures on individual children. I overcome this restriction by examining direct parental expenditure on shadow education. The analysis reveals four key findings. First, there is a birth order disadvantage for later-born children in shadow education expenditures. Second, girls are disadvantaged within families. Shadow education expenditures on them are relatively lower at every birth order. Third, intra-household resource allocation favours the oldest son, followed by other sons and then girls. Fourth, based on a decomposition analysis, I found that disparities in shadow education expenditures account for a substantial proportion of the gender gaps in test scores. The contribution is highest among the firstborns, and declines for lower birth orders.

These results motivate policies promoting gender equality at the micro level. Parents need to be incentivized to treat their children equally. In modern economies, the quality of human

[42] Although I show that shadow education and test scores are positively related, I cannot rule out the simultaneous determination of shadow education expenditures and test scores. See Fortin et al. (2011) for a review of decomposition methods in Economics.

capital is the backbone of economic growth and development. Such disparities already have and will continue to have a severe impact on India's long run growth and development.

Chapter 3: Son Preference, Family Size and the Gender Gap in Mathematics in Pakistan

3.1 Introduction

While extensive empirical evidence suggests that boys outperform girls in standardized math tests, the underlying reasons remain unclear. The literature fails to identify convincing individual-level determinants, though cross-country studies tend to put forward social and cultural factors (Fryer and Levitt, 2010). A few studies also point to underlying biological differences (Nollenberger et al., 2016). In spite of a long standing debate, the literature is yet to reach a consensus. In this chapter I show that elder son preference together with family size can plausibly explain the gender gap in math in Pakistan. The evidence assembled in this chapter supports the idea that within-country social and cultural conditionings are the primary factors underlying the gender gap.

Math test scores have received much attention for their ability to predict labor and educational attainments in important ways. For instance, math test scores and math based curricula are a good predictor of future income and likelihood of a college degree in the rewarding science, technology, engineering and math (STEM) fields (Card and Pyne, 2017). Although, the gender gap in educational achievements has narrowed, and has been reversed in countries like the United States (Goldin et al., 2006), it persists in lower and middle income countries. Gender gap in math tests may be more important in developing countries, where girls suffer from limited resources in addition to adverse stereotyping. The current study context, Pakistan, suffers from a sharp gender division in educational and labor market achievements.[43] Given the strong correlation of math-based test scores and future labor market and educational achievements, Pakistan provides an interesting context to study.

A growing literature has documented a widening gender gap in math test scores in many contexts. The empirical findings are remarkably consistent in that the gender gap monotonically increases with years of schooling (Contini et al., 2017; Bharadwaj et al., 2016; Fryer and Levitt, 2010). In fact, the literature suggests that the gender gap does not exist

[43] Using the Pakistan Integrated Household Survey 2002 data, Aslam (2009) calculated that around 26% of women participate in the labor force as compared to 88% for men

before children enter schools. Several explanations are put forward e.g., social, biological and cultural factors, with no conclusive determinants. For instance, Fryer and Levitt (2010) find that observable explanatory variables are unable to explain the gap, instead gender-biased environments appear to be relevant in a cross-country comparison. Nollenberger et al. (2016) in a related cross-country analysis further emphasize the role of cultural beliefs about the role of women in society as a potential explanation. Few studies attempt to uncover micro-level determinants. Carlana (2019) finds that gender gap in math performance is associated with teachers' gender stereotype in Italy – the stronger the gender stereotypes the sharper the gender gap. Conversely, Bharadwaj et al. (2016) conclude that only self-assessed ability appears to explain the gender gap in Chilean data. Their study further indicates that differential parental investments may play a role. Another study by Dossi et al. (2019) shows that socialization processes at home and gender-biased environments can explain a substantial part of the observed gender gap in math tests in the US. Similar gender-biased environments, specifically parental gender preferences, which affect educational and health outcomes, has been widely documented in many developing countries.[44] Though evidence in chapter 2 indicates that within-household parental preferences may contribute to the gender gap in math, establishing the causal link remains an open question.

Based on data from 200,000 children, my primary findings are in line with the existing evidence. I find a sharp gender gap in math test scores. The gap does not exist at the age of five, but monotonically increases with age. For instance, the gender gap amounts to around 0.009 to 0.024 standard deviations for children aged below ten, but increases to 0.088-0.136 SD for older children. Moreover, consistent with earlier findings, socio-economic and school level variables do little to explain the gender gap. Unlike previous literature, however, I also observe a similar pattern in reading test scores.

In what follows, I investigate how differential parental preferences and resource allocation explain the gender gap in math test scores. To that end, I explore if son preference, particularly a preference for elder sons, is one of the drivers of the gender gap. In many Asian countries parents tend to prefer boys (Jayachandran and Pande, 2017). In chapter 2, I have shown an elder son preference in parental within-household allocation of educational

[44] See Jayachandran and Pande (2017); Bhalotra et al. (2010); Tarozzi and Mahajan (2007).

resources in India.[45] Using a similar approach, I show several features in test scores in the data that may be traced back to elder son preference. *First*, the data exhibits a birth order gradient in math test scores. In other words, test scores monotonically decrease down the birth order, indicating birth-order advantages for elder children. *Second,* girls score less than boys in every birth order. Moreover, the gender gap tends to weaken down the birth order, implying an advantage for elder sons. These conclusions are robust to the inclusion of parent, household and school fixed effects. *Third*, I observe a qualitatively similar pattern in test score distribution. For example, older firstborn girls are 3%-points less likely to achieve the maximum test score compared to firstborn boys.[46] *Fourth*, girls without older brothers appear to score better compared to their peers with older brothers. These features are in fact consistent to the observed patterns in child height in India (Chapter, 2; Jayachandran and Pande, 2017).

Furthermore, test scores appear to be negatively related to family size, suggesting a quantity-quality tradeoff (QQ). The gender gap is less likely to be observed in single-child families. The gap is more pronounced in multi-child families. Moreover, gender gaps tend to exist only for children at early birth orders in larger families, indicating a combined effect of elder son preference and family size.

The evidence also shows differential parental investment in *direct* monetary expenditure on private tuition. I find that parents tend to prefer boys in monthly private tuition expenditure, in addition the gender gap in private tuition expenditure appears to be higher for older children. This significantly strengthens the argument of differential parental investment as the underlying driver of gender gap.

I discuss and rule out several potential alternative mechanisms, most notably inherent biological differences by gender or birth order and biased testing environments that can disproportionately affect children.

---

[45] Chapter 2 further show that a substantial proportion of disparity in private tuition expenditures may translate into a gender gap in test scores.

[46] In addition, among top scorers, 63.44% are boys.

This chapter relates to a number of strands in the literature. *First*, this study is primarily related to the literature that explains gender gaps in math test scores (Carlana, 2019; Dossi et al, 2019; Contini et al, 2017; Bharadwaj et al, 2016; Nollenberger et al., 2016; Fryer and Levitt, 2010). Most previous studies have not provided clear and causal within-country evidence. Standard individual-level variables often fail to fully explain the gender gap. Dossi et al. (2019) is closest to my study. The authors show that girls raised in a male-biased family in the US score on average 3%-points less on math tests compared to their peers in other families. I attempt to associate a similar socio-cultural underpinning typical to the study context with the gender gap in math. Thus, this study fills a gap by offering a plausible mechanism at the micro-level.

Second, this work adds and complements a growing literature that documents son preference and its consequences, primarily in Asian countries. Son preference governs diverse parental behaviors including breastfeeding practices (Jayachandran and Kuziemko, 2011), allocation of educational resources (see chapter 2), and unequal child outcomes such as anthropometric status (Jayachandran and Pande, 2017; Bhalotra et al., 2010; Tarozzi and Mahajan, 2007). This study identifies and explores an additional consequence of son preference.

Third, this chapter feeds into the literature on the QQ tradeoff (Angrist and Levy, 2010; Black et al., 2005; Becker and Lewis, 1973). Although, I do not test for a QQ tradeoff *per se* but rather examine the indirect effect of the QQ tradeoff in explaining gender gaps in math test scores. I find a substantial role for family size in explaining test scores. The fact that the gender gap diminishes in small families suggests that parental resource allocation tends to discriminate against girls relatively more in larger families.

The chapter is organized as follows. Section 4.2 introduces the data and presents summary statistics. Section 4.3 models gender gaps in math test scores. Section 4.4 attempts to test for elder son preference. Section 4.5 explores the role of family size. Section 4.6 documents an unequal distribution of parental education expenditures. Section 4.7 discusses potential alternative explanations, and section 4.8 concludes the chapter.

3.2 Data and Descriptive Statistics

The empirical analysis in this chapter is based on the Annual Status of Education Report (ASER), 2016 data. The survey was conducted by Pratham, a nongovernmental organization. ASER is primarily a survey of educational achievements of children between 5 and 16 years of age. ASER is unique in the sense that the survey is conducted in children's home and enrolls all the eligible children in home irrespective of their current school enrollment status. I have data on children currently enrolled in schools, dropped out, and including those who never enrolled in schools. Furthermore, ASER collets a whole spectrum of individual, parental, household and school level data. I have data on more than 200,000 children.

The ASER math test consists of four questions. If a child is a beginner or could not answer any question she gets a score of 1. The four questions deal with number recognition between 0-9, 10-99, subtraction and division. The scores are coded as 1 if a child answers correctly, otherwise coded as 0. A child can have a maximum score of 5 and a minimum 1. In addition, ASER also conducts a reading test. In case of the reading test, ASER asks 4 questions[47] which are coded like the math score. That is, a child can have a maximum reading score of 5 and a minimum of 1.

[47] These are whether the child can read letters, words, sentences and story.

Table 3.1: Summary Statistics

Variables	Pooled Sample	Boys	Girls
	(1)	(2)	(3)
Math Test Scores	3.00	3.14	2.80
	(1.55)	(1.52)	(1.57)
Reading Test Scores	2.95	3.09	2.75
	(1.59)	(1.56)	(1.61)
Age	8.78	8.92	8.59
	(3.83)	(3.82)	(3.83)
Standard	3.02	3.24	2.73
	(2.67)	(2.73)	(2.55)
Proportion of School Enrolment	0.70	0.75	0.63
	(0.46)	(0.43)	(0.48)
Family Size	3.67	3.63	3.74
	(1.46)	(1.45)	(1.46)
Mother Age	36.26	36.31	36.19
	(7.30)	(7.33)	(7.26)
Mother Education	2.09	2.00	2.20
	(3.73)	(3.67)	(3.80)
Father Age	40.88	40.92	40.83
	(7.85)	(7.88)	(7.81)
Father Education	4.02	3.94	4.13
	(4.80)	(4.78)	(4.83)
Pacca House	0.22	0.22	0.23
	(0.42)	(0.42)	(0.42)
Own Home	0.90	0.90	0.90
	(0.30)	(0.30)	(0.30)
Electricity at Home	0.84	0.84	0.85
	(0.36)	(0.37)	(0.36)
Own Computer	0.15	0.15	0.16
	(0.36)	(0.36)	(0.36)
Receiver of Social Safety Scheme	0.16	0.16	0.16
	(0.40)	(0.37)	(0.37)
Private School	0.25	0.24	0.26
	(0.43)	(0.43)	(0.44)
Mixed Gender School	0.14	0.13	0.16
	(0.35)	(0.34)	(0.37)
English Medium School	0.25	0.24	0.28
	(0.44)	(0.43)	(0.45)
Toilet in School	0.66	0.64	0.69
	(0.47)	(0.48)	(0.46)
Library in School	0.17	0.16	0.18
	(0.37)	(0.36)	(0.38)
Electricity in School	0.64	0.62	0.67
	(0.37)	(0.48)	(0.47)
Number of Students in School	193.40	190.28	198.28
	(192.56)	(189.88)	(196.61)
Number of Teachers in School	7.34	7.19	7.58
	(7.04)	(6.95)	(7.18)

Notes: This table reports means of the dependent variables and covariates. Standard errors are reported in parentheses. Around 43% of the pooled sample consists of girls. Average sample sizes: Pooled 248,760, Boys 141,601, and Girls 107,159.

Table 3.1 presents summary statistics. Column 1 shows that the average raw math test score is 3.00. Column 2 and 3 report a sharp gender division in test scores. For instance, on average in math, girls score 0.35 points less than boys, which amounts to 0.216 SD.[48] Unlike other study contexts, gender gap is also evident in reading test score. The gender gap in reading test scores amounts to 0.34 points or 0.212 SD.

The remainder of Table 3.1 presents summary statistics for background variables. Around 43% of the children are girls, reflecting skewed sex ratios in the Pakistani population. 70 % of the children are currently enrolled in schools. The ASER data is particularly useful as nearly 30% of children have either dropped-out or never enrolled in schools. Nearly 25% of children go to private schools and a similar proportion is enrolled in English-medium schools. Average family size is 3.67. Fathers tend to be older and more educated than mothers. I observe gender gaps in most background variables. Girl children are more likely to come from families with better educated parents with better household conditions, although the differences are small. Consistent with this, girls have advantages in most school level variables. For instance, girls are more likely to be enrolled in private, English medium schools, and, in girl sample, schools have more toilets, libraries and electricity, and importantly, on average more teachers. Interestingly, girls are 3%-points more likely to be enrolled in mixed-gender schools, although in general, enrolment in mixed-gender school is low with only 14% of the children enrolled in mixed-gender schools. Prima facie, it is puzzling that in spite of better family and school conditions, girls have lower math and reading scores. In the next section, I examine this pattern in more detail.

3.3 The Gender Gap in Math Test Scores

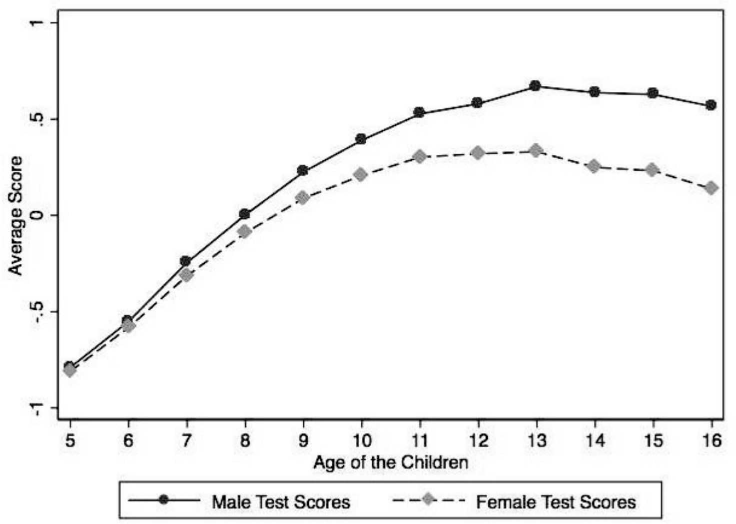
First, I graphically present the dynamics of the male-female gap in mathematics test scores as a function of child age. Following the previous literature, I standardize test scores to have zero mean and standard deviation one. In Panel A of Figure 3.1, I plot mean test scores by age in years and gender, and, panel B plots the corresponding gender gap. Two facts are immediate; *first* the math test score monotonically increases till the age of 13 and then falls marginally, and, *second*, the gender gap in math test scores also monotonically increases till the age of 14, falls marginally at the age of 15 and increases again at the age of 16. The

[48] Compared to other study contexts, the gender gap is on the higher side here.

gender gap rises from a small gap at the age of 5 to more than 0.4 SD at the age of 16,[49] which translates into an annual 0.037 SD increase in the gap. The pattern is consistent across schools: I do not find any significant differences in gender gaps by school type.[50] I observe a similar pattern in reading scores in Figure F1 in the Appendix F. Both boys and girls start at fairly similar level at the age of 5, but similar to the math test, a monotonic pattern emerges with age. The ground lost by girls amounts to 0.44 SD at the age of 16, or an annual fall back of 0.038 SD.[51]

Figure 3.1: Math Test Scores (SD) by Gender

Panel A

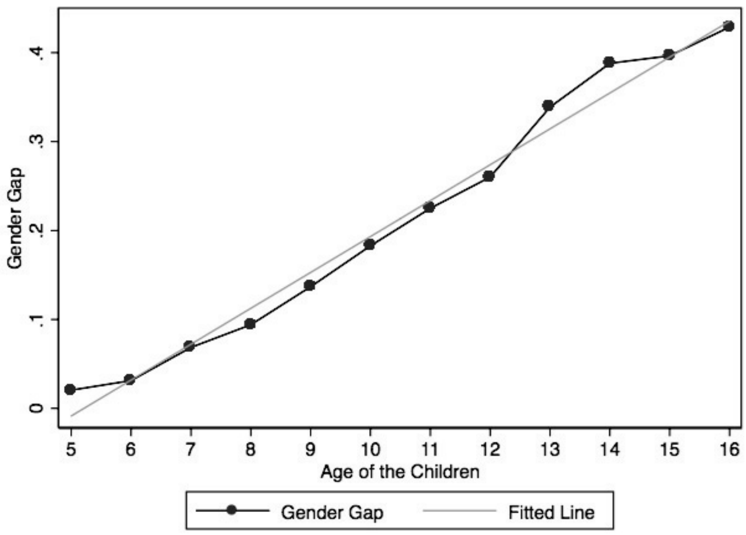


[49] Raw math scores at age 5 are 1.63 for boys and 1.60 for girls. Raw scores increase at the age of 16 to 3.82 for boys and 3.13 for girls.

[50] Plots by school are available on request.

[51] Raw reading scores at age 5 are 1.55 for boys and 1.52 for girls. The reading scores increase at the age of 16 to 3.81 for boys and 3.11 for girls.

Panel B

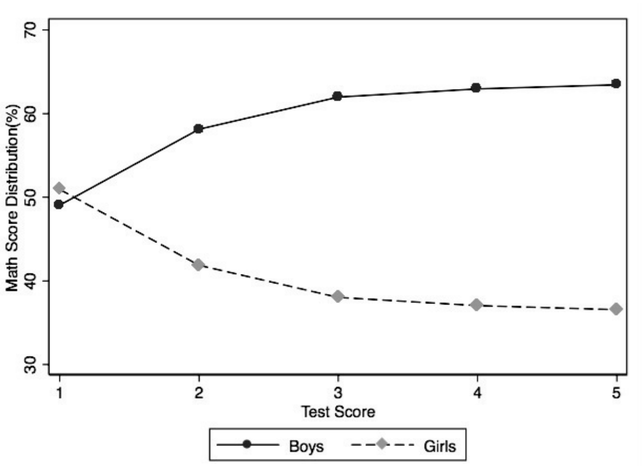


Notes: This figure depicts mean gender gap (boy-girl) in normalized math test scores by age.

Figure 3.2 shows the distribution of math achievement by gender. Girls are 26.88%-points less likely to earn the highest score of 5 as compared to boys. The gap monotonically increases with test scores.[52]

[52] The exact distribution in percentage is shown in Table C1 in the Appendix C.

Figure 3.2: Math Score Distribution



Notes: This figure depicts math score distribution (in percentage) by gender. Test scores are not standardized.

Next, I present a series of estimates of the gender gap in math scores in Table 4.2. I estimate the following specification:

(3.1) $$Y_i = \delta + \beta\, Girl_i + \gamma X_i + \mu_i$$

where Y_i is the standardized ASER math test score of child i, Girl is a dummy for girl children, X is a vector of individual, parental, household and school level covariates, and, μ is the error term. β is the primary parameter of interest. X includes child age and its quadratic, standard of current study and a dummy for school enrolment status. I also control for a few school level variables – student and teacher capacity, dummies for private, co-educational and English medium schooling, and, dummies for the availability of toilet, library and electivity in the school. I include maternal and paternal literacy, age, a quadratic in age and sibling size. Other covariates are household specific, e.g., dummies for home ownership, pacca house, electricity, computer and recipient of social safety schemes. For the sake of brevity and ease of interpretation, the analysis is segregated by two age groups: 5 to less than 10 years (younger group), and 10 to 16 years (older group). If a student has uninterrupted study she is most likely to be in fourth grade or below in the younger group, and, tenth grade or below in the older group, respectively. [53]

[53] In the data, 97% of the children in the younger group are in grade 4 or below.

Table 3.2: Gender Gap in Math Test Scores

Dep. Var. Math Test Scores	Younger Age Group (5 to less than 10 Years)				Older Age Group (10 to 16 years)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Girl	-0.015***	-0.024***	-0.013**	-0.009*	-0.117***	-0.136***	-0.091***	-0.088***
	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)	(0.006)	(0.005)
Standard	0.319***	0.282***	0.309***	0.274***	0.134***	0.101***	0.139***	0.110***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)	(0.002)	(0.002)	(0.002)
Mother Age	0.012***	-	0.016***	0.008*	0.017***	-	0.017***	0.002
	(0.003)		(0.004)	(0.004)	(0.004)		(0.005)	(0.005)
Mother Schooling	0.059***	-	0.066***	0.013	0.058***	-	0.060***	-0.000
	(0.007)		(0.009)	(0.009)	(0.010)		(0.011)	(0.011)
Father Schooling	0.039***	-	0.028***	0.013	0.030***	-	0.015	0.026**
	(0.006)		(0.008)	(0.008)	(0.008)		(0.009)	(0.009)
Pacca House	0.030***	-	-0.007	0.00	-0.002	-	-0.014*	0.015
	(0.005)		(0.006)	(0.007)	(0.006)		(0.007)	(0.008)
Private School			0.108***	-			0.129***	-
			(0.008)				(0.009)	
Mixed- Gender School			-0.028***	-			0.021**	-
			(0.007)				(0.008)	
English School			0.077***	-			0.029***	-
			(0.007)				(0.008)	
Family Size	-0.012***	-	-0.010***	-0.007***	-0.024***	-	-0.021***	-0.014***
	(0.001)		(0.002)	(0.002)	(0.002)		(0.002)	(0.002)
Child Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parent Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Currently Enrolled	No	No	Yes	Yes	No	No	Yes	Yes
Parent Fixed Effect	No	Yes	No	No	No	Yes	No	No
School Fixed Effect	No	No	No	Yes	No	No	No	Yes
Observations	98660	95261	69920	69920	97035	94581	68592	68592

Notes: Every column reports a separate linear regression. Standard errors reported in parentheses are robust to within village unit clustering. Controls include child, parent and household level covariates, and columns 4,5,9 and 10 additionally include school level variables. ***p < 0.01; **p < 0.05; *p < 0.10.

Columns 1 – 4 of Table 3.2 present estimates for the younger age group, and columns 5 – 8 for the older age group. A comparison of columns 1 and 5 confirms that gender gaps in the ASER math test scores increase with age. To be specific, younger girls on average score 0.015 SD lower as compared to boys while the corresponding gender gap is 0.117 SD in the older group.

Next, columns 2 and 6 report estimates with household and parent fixed effects. By doing so, I difference out unobserved parental and household confounding factors. The estimates remain fairly stable. The gender gap increases to 0.024 SD for younger children, and for the older children, the corresponding gap is 0.136 SD. The test scores may also differ depending on the schooling and school environments may disproportionality affect different school-going children. Columns 3 and 7 present estimates for a sub-sample of currently enrolled students after controlling for school level characteristics. The estimates are similar to the full sample, and are robust to the inclusion of school fixed effects in columns 4 and 8. I also estimate age-specific regressions for each age between 5 and 16 and find qualitatively similar results.[54]

Overall, the estimates suggest that the gender gap remains substantial, despite the inclusion of individual, parental, household and school variables. This is remarkably consistent with previous evidence (Bharadwaj et al., 2016; Fryer and Levitt, 2010).

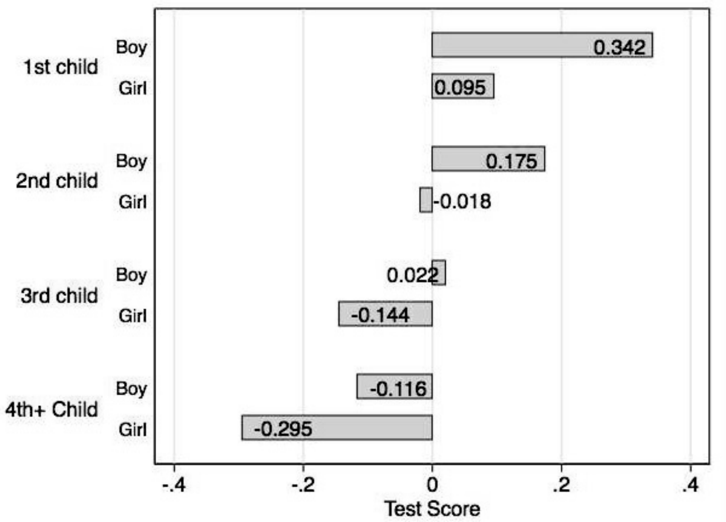
3.4 Birth Order, Son Preference and the Gender Gap in Math Test Scores

In this section, I explore if the well-documented elder son preference typically observed in the region explains the gender gap. Empirical evidence suggests that parents tend to favor elder sons in many South-Asian countries. Son preference has been documented in breastfeeding practices (Jayachandran and Kuziemko, 2011), intra-household allocation of educational resources (see chapter 2), and has been shown to cause gender gaps in child height (Jayachandran and Pande, 2017). Accordingly, parental resource allocation may be viewed as the product of elder child preference and preference for sons. Thus, if parents allocate resources according to elder son preference, the data should exhibit a birth order

[54] Results are available on request

gradient in test scores, and moreover, girls should be disadvantaged at every birth order. Additionally, I may expect the gender gaps to be stronger in the case of early birth orders.

Figure 3.3: Math Test Scores by Birth Order and Gender



Notes: This figure shows mean (normalized) math test scores by birth order and gender.

In Figure 3.3, I show a birth order gradient in math test scores. Test scores monotonically decrease as birth order increases, and boys outperform girls at every birth order. In addition, the gender gap tends to decline as the birth order increases. Next I investigate the pattern in more detail. To that end, I estimate the baseline specification:

$$Y_i = \delta + \sum_{j=2}^{4} \alpha_j BO_{ij} + \sum_{j}^{4} \beta_j BO_{ij} \times Girl_i + \gamma X_i + \varepsilon_i \qquad (3.2)$$

Where, Y_i is the individual ASER math test scores, and BO_{ij} is a dummy which takes the value 1 for birth order j. Here α captures coefficients of birth orders and β captures coefficients of birth orders and girl interactions. X_i is a vector of background variables, and, ε_i is an error term.

The identification of (3.2) relies on the fact that birth order is a natural experiment, the genetic makeup remains constant in all biological children in the family. However, children at later birth orders are more likely to be observed in larger families. To alleviate this concern to some extent, I control for family size in all estimates. Yet, it cannot control for unobserved associations between birth order and other confounding characteristics. In particular, fertility is highly correlated with family and household factors. I address such endogeneity concerns by including family and household fixed effects. By doing so, I difference out factors fixed within families and households, and also remove residual associations between birth orders and family or household factors.[55]

[55] A household may contain multiple nuclear families.

Table 3.3: Birth Order and Gender Gap

Dep. Var. Math Test Scores	Younger Age Group				Older Age Group			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2nd Child	-0.027***	-0.039**	-0.033***	-0.028***	-0.032***	-0.074***	-0.036***	-0.044***
	(0.007)	(0.015)	(0.009)	(0.008)	(0.008)	(0.015)	(0.009)	(0.007)
3rd Child	-0.069***	-0.098***	-0.078***	-0.068***	-0.046***	-0.102***	-0.044***	-0.045***
	(0.010)	(0.023)	(0.011)	(0.010)	(0.011)	(0.024)	(0.013)	(0.011)
4th+ Child	-0.096***	-0.107***	-0.119***	-0.082***	-0.020	-0.110***	-0.014	-0.011
	(0.013)	(0.031)	(0.015)	(0.012)	(0.016)	(0.036)	(0.018)	(0.016)
Girl × 1st Child	-0.024**	-0.044**	-0.027**	-0.021*	-0.128***	-0.178***	-0.105***	-0.109***
	(0.009)	(0.019)	(0.010)	(0.009)	(0.010)	(0.018)	(0.012)	(0.011)
Girl × 2nd Child	-0.015	-0.038***	-0.015	-0.016	-0.109***	-0.117***	-0.075***	-0.063***
	(0.008)	(0.014)	(0.009)	(0.008)	(0.011)	(0.016)	(0.013)	(0.011)
Girl × 3rd Child	-0.004	-0.010	0.004	0.010	-0.090***	-0.106***	-0.066***	-0.069***
	(0.008)	(0.015)	(0.010)	(0.009)	(0.013)	(0.020)	(0.016)	(0.014)
Girl × 4th+ Child	-0.014	-0.012	-0.010	-0.007	-0.132***	-0.138***	-0.113***	-0.104***
	(0.008)	(0.014)	(0.010)	(0.008)	(0.017)	(0.027)	(0.021)	(0.019)
Parent FE	No	Yes	No	No	No	Yes	No	No
School FE	No	No	No	Yes	No	No	No	Yes
Currently Enrolled	No	No	Yes	Yes	No	No	Yes	Yes
Observations	98660	95261	69920	69920	97035	94581	68592	68592

Notes: Every column reports a separate linear regression. Standard errors reported in parentheses are robust to within village unit clustering. All estimations include controls for child, parent and household level variables, and columns 3,4,7 and 8 additionally include school level variables. 2nd and 3rd child are indicators for children whose birth orders are 2 and 3, and, 4th+ indicates children with birth order 4 or later. In columns 2 and 6 main effects are absorbed by parent fixed effects. In columns 4 and 8 main effects are absorbed by school fixed effect. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Table 3.3 presents estimates for equation (3.2). Column 1 shows a male birth-order gradient in math test scores. The coefficients associated with birth orders are negative and monotonically increase down the birth order. Second and third born boys in the younger age group score on average 0.027 SD and 0.069 SD less than firstborn boys. In the older age group, second born boys score 0.032 SD less than firstborn boys, and moving from firstborn to third and fourth+ born boys results in 0.046 SD and 0.020 SD drops in estimates for test scores. I find a fairly similar birth order gradient for girls. For instance, in column 1, I observe that second born girls from the younger age group score on average 0.018 SD lower compared to firstborn girls (calculated as *2nd Child + Girl × 2nd Child – Girl × 1st Child*). The effects are 0.049 SD and 0.086 SD lower for third and fourth+ born girl children compared to firstborn girl children. Effects tend to be lower for older girls.

The above evidence indicates a birth order gradient both for boys and girls, pointing to an advantage for elder children. Next I investigate gender gaps across birth orders. For the younger age cohort, the coefficients on the gender and birth order interactions reported in Table 3.3 are negative, and the estimates for gender gap tend to decline with birth order. The gender gap is -0.024 SD among the younger firstborn children. The estimate tends to weaken for the second-borns, and, disappears for third-borns and fourth+ borns. However, the gender gap remains persistent and sizeable across birth orders among older children. The weakening of the gender gaps as one moves down the birth order emphasizes the advantageous position of elder sons.

In columns 2 and 6 of Table 3.3, the main effects are absorbed by both parents and household fixed effects in the sample of younger and older children respectively. The estimated effects appear to be on the higher side as compared to their baseline estimates, although the patterns remain similar. For instance, firstborn younger girls score 0.044 SD less than their firstborn peers, while for the older children the gap is 0.178 SD. Moving from first to second borns the gender gap amounts to 0.038 SD and 0.117 SD for younger and older children. I also observe a sharper birth order gradient.

In columns 3 and 7, I replicate the baseline estimates in a subsample of currently enrolled students. Afterward, in column 4 and 8 I add school fixed effects in the baseline specification (2), and thus difference out confounding factors fixed within schools. The estimates appear to be robust to the inclusion of school fixed effects.

To supplement the evidence, Table C2 in the Appendix C shows that gender gaps in test score distribution are concentrated among older children. I find that firstborn girls from older age groups are 3%-points less likely to score 5 as compared to firstborn boys, the corresponding gap is 2%-points for second-borns. However, for younger children the gaps are not sizeable.

The above evidence clearly indicates an elder son advantage. To further investigate the validity of elder son preference as a plausible explanation for the gender gap, I hypothesize that the gender gap ceases to exist or at least weakens if girls lack an elder brother. There is theoretical ambiguity on whether girl children without elder brothers have an advantage. Girls without elder brother should have an advantage. However if these girls have younger brothers, then by virtue of being the eldest son, the younger brother may attract more resources. In addition, if such families lack a son, they will try again for a son, and in turn may save up resources for the unborn son.[56] The net effect depends on the strength of one mechanism over another. To uncover the net effect, I estimate the following specification:

(3.3) $$Y_i = \alpha + \beta_1 Girl_i + \beta_2 No\_Elder\_Brother_i + \beta_3 Girl_i \times No\_Elder\_Brother_i + \gamma X_i + \varepsilon_i,$$

where, $No\_Elder\_Brother_i$ is a dummy for whether the child lacks an elder brother. Estimates are presented in Table 3.4. The estimates suggest that boys without elder brothers have clear advantages in test scores (the coefficient of *No Elder Brother*). The net effect for girls without elder brother (*No Elder Brother + Girl × No Elder Brother*) tends to be positive. In other words, girls without elder brothers tend to score better as compared to girls with elder brothers. The effect is higher for older girls. The effect amounts to 0.016 and 0.020 SD respectively in pooled and currently enrolled sample for the older children. The effects are statistically significant for older children, but not for younger children. Comparatively, girls without elder brothers tend to score lower as compared to boys without elder brothers. The net effect, *Girl + Girl × No Elder Brother* which captures the difference, is negative and tends to be statistically significant only for the older age group.

[56] For instance, Jayachandran and Pande (2017) found a negative impact on girls of lacking an older brother on child height in India. By contrast, in chapter 2 I find a net positive effect on private tuition expenditure.

Table 3.4: Son Preference

Dep. Var. Math Test Scores	Younger Age Group		Older Age Group	
	(1)	(2)	(3)	(4)
Girl	-0.018**	-0.016*	-0.125***	-0.102***
	(0.006)	(0.008)	(0.010)	(0.012)
No Elder Brother	0.004	0.003	-0.001	-0.004
	(0.005)	(0.006)	(0.006)	(0.006)
Girl × No Elder Brother	0.006	0.007	0.017	0.024
	(0.008)	(0.010)	(0.010)	(0.013)
p -values				
No Elder Brother + Girl × No Elder Brother = 0	0.10	0.20	0.06	0.08
Girl + Girl × No Elder Brother = 0	0.04	0.21	0.00	0.00
Currently Enrolled	No	Yes	No	Yes
Observations	98660	69920	97035	68592

Notes: Every column reports a separate linear regression. Standard errors reported in parentheses are robust to within village unit clustering. All estimations include controls for child, parent and household level variables, and column 2 and 4 additionally include school level variables. ***$p < 0.01$; **$p < 0.05$; *$p < 0.10$.

In sum, elder son preference seems to partially explain the gender gap, and plausibly at least for the younger age group. In the older age group, the gender gap weakens down the birth order but remains negative and statistically strong for later-borns.

3.5 Family Size and Gender Gap

Section 3.3 provided some evidence that individual test scores are negatively related to family size, which may indicate a QQ tradeoff (Becker and Lewis, 1973). In this section, I examine the dynamics of the gender gap in math test scores by family size. Given the negative relationship between family size and test scores, an obvious natural progression is to estimate the birth-order, gender effects by family size, which also allow for heterogeneous effects by family size.

Table 3.5: Family Size and Gender Gap in Math

Dep. Var. Math Test Scores	Younger Age Group			Older Age Group		
Family Size	N=1	N=2	N=3	N=1	N=2	N=3
	(1)	(2)	(3)	(4)	(5)	(6)
Girl × 1st Child	-0.023	0.004	-0.053**	-0.032	-0.088***	-0.078***
	(0.025)	(0.016)	(0.018)	(0.035)	(0.021)	(0.018)
Girl × 2nd Child		0.022	-0.015		-0.058*	-0.041*
		(0.016)	(0.014)		(0.028)	(0.020)
Girl × 3rd Child			0.006			0.044
			(0.015)			(0.028)
2nd Child		-0.064***	-0.068***		-0.017	-0.043**
		(0.015)	(0.014)		(0.020)	(0.014)
3rd Child			-0.124***			-0.085***
			(0.018)			(0.024)
Observations	2436	11363	20125	1821	7927	17400

Notes: Every column reports a separate linear regression. Standard errors reported in parentheses are robust to within village unit clustering. All estimations include controls for child, parent, household and school level variables. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 3.5 presents a series of estimates conditional on family size. Columns 1 and 4 show that the gender gap ceases to exist statistically in single child families for both younger and older age group. In other words, in a single-child family where parents need not allocate resources based on their gender preferences, and the girl children are not compelled to be in the sibling rivalry game, boys and girls tend to perform equally.

By contrast, a gender gap emerges in two-child families in the older age group, and persists in three-child families. The effect sizes are lower compared to their baseline estimates though. The male-female gaps are 0.088 SD and 0.078 SD for older firstborns in two-child and three-child families respectively. For younger children only firstborn girls from three-child families tend to have a lower score than firstborn boys.

In a strong multi-child society like Pakistan, the observed single-child families are most likely to have additional children, which may drive our result. However, what does remain clear is that the gender gap tends to be small to non-existent in small families. In multi-child families, an additional child increases competition for resources, and may induces parents to allocate resources according to their preferences. I observe that the girl disadvantage is more prevalent in bigger families. Together with the birth order gradient, the evidence suggests that elder son preference coupled with QQ tradeoff plausibly explain the observed gender gap. However, the adverse effect of family size on gender gap is more prevalent for older children compared to their younger peers.

3.6 Unequal Distribution of *Direct* Parental Resources

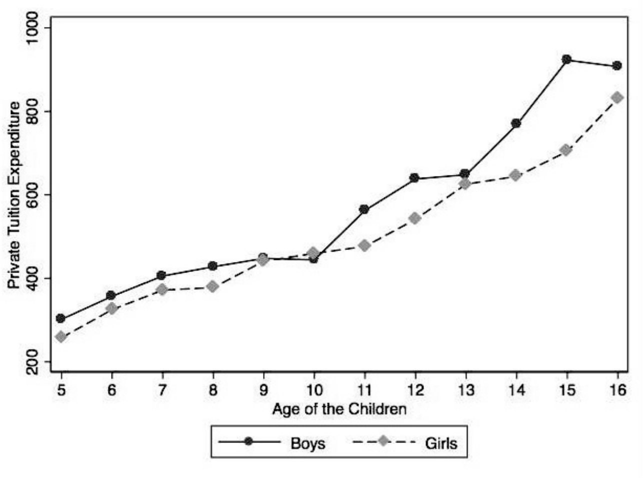
So far I have *indirectly* shown the potential effect of parental resource allocation manifested in child test scores. A number of studies, however, have looked into the gendered allocation of direct parental monetary resources. In chapter 2, I show a similar birth order-gender pattern in parental allocation of private tuition expenditure as observed in the case of math tests in this chapter.

In what follows, I investigate whether the particular patterns can be observed in direct parental educational expenditure on private supplementary tuition. My investigation on private tuition expenditures further allows me to examine the increasing gender gap by age. Although private tuition has become a norm in many Asian countries,[57] in the current context only 8% of children attend private tuition.[58]

[57] For instance, around 23% of Indian students were taking private tuition in 2012 (see Chapter 2). The phenomenon is rapidly expanding and reaching new continents where private tuition was relatively unknown. In Hong Kong, for example, around 72% of 12th grade students engage in private tuition (Zhan et al, 2014).

[58] Private tuition expenditure also seems to be highly effective in this context. For instance, 62.34% of students at the 90th quantile of private tuition expenditure score 5, and 87.3% of them score 3 or more. Even 67.74% of students at the 10th quantile score 3 or more, although, only 28.04% of them earn the top score of 5.

Figure 3.4: Yearly Private Tuition Expenditures



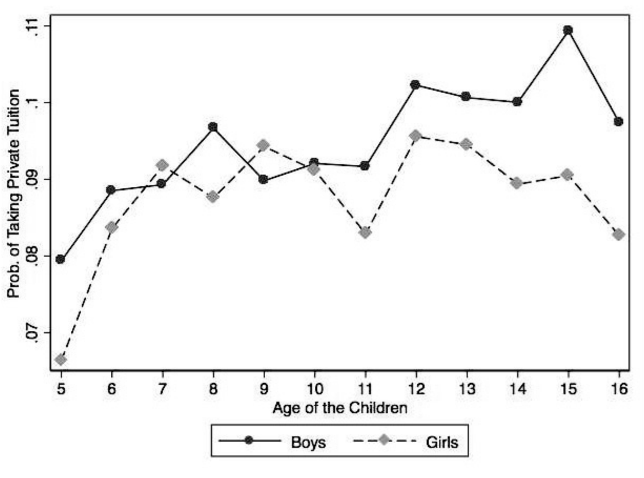
Notes: This figure depicts mean yearly private education expenditures in Pakistan Rupees.

First, I plot monthly private tuition expenditures by gender varying with age of the children in Figure 3.4. Parents tend to prefer boys over girls, and the gap tends to be higher for older children with occasional dips.[59] For instance, the gender gap does not exist at the age of ten and thirteen. In Table C3 in Appendix C, I show a fairly similar birth order-gender pattern that I observe for math test scores in private tuition expenditure. Next, I plot private tuition attendance by age in Figure 3.5. I find a clear gender gap in private tuition attendance, and the gap is clearly evident for older children, specifically for children aged 10 years or older. Together, differential parental resource allocation by gender appears to worsen with age of the children strengthen the argument of differences in intra-household resource allocation on private educational expenditures as one of the drivers of the gender gap in math, and may well be a potential explanation for the increasing gender gap.[60]

[59] This is in fact surprising considering the fact that returns to an additional year of schooling in Pakistan is substantially higher for women compared to men (Aslam, 2009). To be precise, Aslam (2009) finds that labor market returns of an additional year of schooling ranges between 7% - 11%, and 13% - 18% for men, and women respectively. Together, if educational attainments are positively related to private tuition expenditures, parental within-household resource allocation could be irrational.

[60] Private tuition is a choice variable and considering that very few parents invest in private tuition indicates that these families may be qualitatively different from the rest of the population.

Figure 3.5: Probability of Taking Private Tuition



Notes: This figure depicts the probability of taking private tuition by gender.

3.7 Alternative Explanations

So far I have argued that elder son preference together with the adverse effect of family size may plausibly explain the gender gap. There are, however, some alternative mechanisms that could explain the gender gap. In this section, I discuss a set of competing explanations for the gender gap and show that these explanations may be ruled out.

A number of lab and field experiments highlight that women shy away from competition (Niederle and Vesterlund, 2011), and that such underlying preferences could explain the gender gap in math test scores. In other words, the gender gap may be the outcome of inherent gaps between boys and girls. However, in our data there is no gender gap in single child families, indicating that inherent gender differences in competitiveness is unlikely to be a plausible determinant. In fact, another strand of the literature argues that underlying differences in competitive preferences are not inherent, rather an outcome of socialization processes. For instance, women from matrilineal communities are found to be as competitive as men from patrilineal communities, pointing at the role of gender norms (Gneezy et al., 2009). In our data, disparity in parental resource allocation seems more plausible and appears to be much more consistent than any explanations attributed to the differences in inherent differences in economic preferences.

Likewise, another competing explanation is that ability itself may be unequally distributed across birth orders, which may drive the observed results. In fact, several studies have documented that firstborns are healthier, have higher IQs, higher earnings and higher educational achievements (Black et al., 2018; Kristensen and Bjerkedal, 2007; Kantarevic and Mechoulan, 2006). Associated with the personality theory of Alfred Adler, the hypothesis is that elder children are more responsible and have a taste for power, in comparison, later borns are more open and sociable, and the middle children are more competitive and ambitious (Black et al., 2018). However, these firstborn premiums are itself a product of sibling rivalry. For instance, the first few years of life are very important for overall development. Being the only child during early years, firstborns get the full attention of their parents. So, the birth order gradient that can be potentially attributed to the Adlerian personality hypothesis, itself may be attributed to sibling rivalry. Another strand of the literature examines the link between birth order pattern and inherent biological differences in ability. Although, our data do not allow us to formally test for the existence of any inherent gap in ability by birth order that can be attributed to biology, existing evidence suggest that biological factors in fact tend to favor later borns (Black et al., 2018).[61] For instance, exploiting the fact that some families adopt children in Sweden, the paper by Black et al. (2018) found that the negative birth order effect in non-cognitive ability is entirely driven by social birth order, and, biological birth order favors later-born children. Consistent with my hypothesis, the authors suggest uneven parental treatment as the underlying mechanism behind differential birth order development in non-cognitive abilities. In support, they found that parents are less likely to discuss schoolwork with later-born children. Nevertheless, even if ability is unequally distributed across birth orders, our argument of intra-household gendered resource allocation holds.

One criticism in the literature pertains to sample selection (Muñoz, 2018). To be precise, previous work tends to examine the gender gap in children enrolled in schools, however in

[61] Recall, I show that the empirical pattern is robust to the inclusion of parent fixed effects. In other words, the within-family comparison is driven by birth order effects as all the siblings are not systematically genetically different.

many contexts, boys in particular are more likely to drop out of schools early.[62] The selection problem is not likely to drive our results. The ASER data in principle do not suffer from selection issues as the ASER data is based on all the children in a family irrespective of whether a child is enrolled in school or not. Moreover, ASER conducts tests at the home of the children, so the testing environment is unlikely to disadvantage dropped-out children.

3.8 Discussion and Conclusion

The gender gap in mathematics has been observed across many contexts, and the gap appears to monotonically increase with age. However, there is very little evidence on micro-level causal mechanisms underlying these gender patterns. Using a large representative survey from Pakistan, I argue that elder son preference coupled with family size can be a plausible and important mechanism to explain the gender gap.

I uncovered several facts. First I documented substantial gender gaps in math test scores, which rose with age. Consistent with previous evidence, individual and socio-economic covariates failed to explain the gap. I observed a birth order gradient in math test scores, and girls underperformed in every birth order compared to boys. These patterns are concordant with the well-documented elder son preference particularly in South Asia. Girls without elder brothers tend to score better than girls with elder brothers, suggesting elder son preference. Furthermore, I observed that the gender gap does not emerge in single-child families but in multi-child families, indicating the negative effect of family size. The evidence showed that parents allocate educational resources differentially according to birth order and gender. Furthermore, parents spent less on private supplementary tuition for daughters, and the gender gap in private tuition expenditure was higher among older children. Taken together, the results point at the micro-level social and cultural conditioning behind such gender gaps.

While the interactions between son preference and family size do seem to account for gender gaps in math, there are other potential channels which may be driving the observed patterns. Most pertinently, my arguments are unlikely to fully explain the increasing gender gap by child age. This opens up avenue for further research in this area. Here I lay out two potential mechanisms that may be tested.

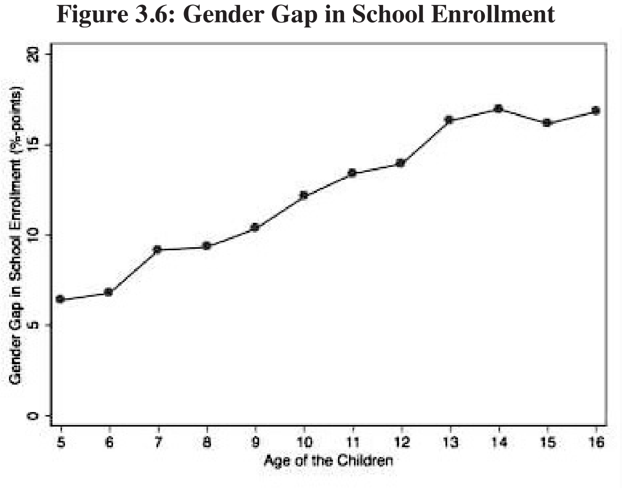
[62] Muñoz (2018) estimates that sample selection explains between 50% - 60% of the gender gap.

One potential mechanism may be the phenomenon of early girl marriage in Pakistan. Although the mean female marital age has increased over time, in 2012-13 almost 48% of currently married women between 15-24 were married before the age of 18 (Nasrullah et al., 2014b). The prevalence was 50.1% in 2006-07 with around 24% of women entering marriage before the age of 16 (Nasrullah et al., 2014a). In many Asian countries, marriage market matching places greater emphasis on physical attributes of girls rather than their cognitive and potential labor market opportunities (Bovet et al., 2018). If the marriage market puts little to no or even negative weight on educational achievements and attainment, girls close to marital age may tend to reduce the effort they put into studying. By contrast, marriage market puts substantial weight on male labor market potential (Bhatti and Jeffery, 2012). Although, I cannot test this with the data at hand, one potential indirect outcome is that girls are more likely to drop-out of school once they reach near-marital age. We may expect a gradual increase in girls-boys gap in school drop-out. In Figure 4.6, I show an increasing boys-girls gap in school enrollment. There is a 17%-points gender gap in school enrolment at the age of 16 as compared to a smaller 7% gap at the age of 5. I also show a similar gradually increasing girls-boys gap in school drop-out in Figure 3.7.[63] Unlike in many other contexts where boys drop-out of schools early, our data reveal a reverse pattern.[64]

[63] The gender gap in drop-out is small compared to gaps in school enrollments. Thus, most of the gender gap in current enrolment could be driven by never enrolled students.

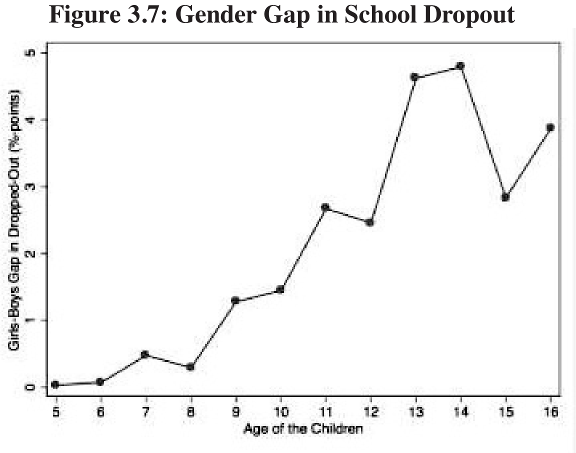
[64] Both boys and girls start at a similar level at the age of five, but there is a gap of 4%-points at the age of 16. However, girls and boys drop-out of school for different reasons. Boys primarily drop-out to engage in economic activities, girls in South-Asia leave school primarily because of early marriage. The argument thus can be plagued by the extent of economic opportunities for boys, and thus it would also be interesting to isolate natural trends.

Figure 3.6: Gender Gap in School Enrollment



Notes: This figure shows boys-girls gap in school enrollment.

Figure 3.7: Gender Gap in School Dropout



Notes: This figure shows *girls-boys* gap in school dropout.

Another promising avenue is to explore the role of self-assessed ability. Bharadwaj et al. (2016) for example found that self-assessed ability is the primary determinant of the observed gender gap in Chile. In our data the gender gap tends to be higher for older children – suggesting acquired skills may be important. In other words, in a dynamic environment, once the gender gap starts to emerge boys may put more effort in skill acquisition. Thus, any initial gender gap can be self-sustaining and in fact deteriorate if it affects self-esteem. Discriminatory societal norms may be fully internalized and acted upon once girls observe the gender gap, which perhaps leads to a vicious cycle.

Chapter 4

Gender Norms and the Motherhood Penalty: Experimental Evidence from India

(with Arjun Bedi and Matthias Rieger)

4.1 Introduction

Around the world, a substantial proportion of women do not participate in labor markets. If women do work, they tend to earn less than men, and face entry barriers in certain jobs or challenges in terms of climbing the career ladder. While gender gaps have narrowed, they remain large in some regions of the world and are often attributed to motherhood (Weichselbaumer and Winter-Ebmer, 2005; Goldin, 1994, 2014; Goldin et al., 2017; Klasen and Pieters, 2015; Verick, 2014; Das and Zumbyte, 2017).

Another strand of the literature has examined if differences in underlying preferences may explain gender gaps. For instance, if appetite for competition varies between men and women (Croson and Gneezy, 2009; Niederle and Vesterlund, 2011; Charness and Gneezy, 2012; Geraldes, 2018), or between women with and without children, this could in turn influence selection into certain types of jobs (Cassar et al., 2016). However, this raises a fundamental question as to what fashions these gender differences. A growing experimental literature has turned to the role of culture and society (Gneezy et al., 2009; Hoffman et al., 2011; Andersen et al., 2013; Cadsby et al., 2013; Jayachandran, 2015; Klonner et al., 2019). Some of these studies speak to the heated nature versus nurture debate using cross-cultural experiments. For instance, Gneezy et al., (2009) compare competitive preferences of men and women living in a patrilineal (the Maasai in Tanzania) and a matrilineal (the Khasi in India) community. In matrilineal cultures such as the Khasi, maternal grandmothers head households, and eventually transmit (ancestral) wealth and power to their youngest daughters. After marriage, Khasi women do not move to their husbands' families while Khasi men frequently join their wives' households.[65]

[65] The youngest daughter of a Khasi family inherits ancestral property, is the head of the family and after marriage her husband joins her natal family. In the case of older daughters, they may form separate households with their husbands.

Husbands tend to have limited say over resources and it is not unusual for men to take on stereotypically "female" tasks such as childcare.[66] Intriguingly, women are as competitive in experiments as men if they live in such a matrilineal society (see Gneezy et al., 2009). These results have been corroborated and extended by Andersen et al. (2013) and Klonner et al. (2019). Based on the same experiments as used by Gneezy et al. (2009), Klonner et al. (2019) report a monotonic relationship between patriarchy and gender differences in competitiveness in India's Northeast.[67] They conclude that "patriarchal norms suppress women's economic potential by making them compete too little" (Klonner et al. 2019, p.3). However, it is unclear if such culturally induced gender differences are recognized in the labor market and have a bearing on labor market outcomes.

In order to carve out the potential links between gender, culture and actual labor market outcomes, we focus on a key event in many women's lives, motherhood. As mentioned above, labor markets tend to penalize mothers in terms of wages and job opportunities (Budig and England, 2001; Anderson et al., 2002; Gangl and Ziefle, 2009; Benard and Correll, 2010; Budig and Hodges, 2010; Budig et al., 2012; Goldin et al., 2017). Notably, Correll et al. (2007, p.1298) hypothesize that mothers are often discriminated against as compared to non-mothers, as employers may consider them "less competent and less committed to their jobs." The authors find that, in the United States, (exogenously) reporting motherhood on CVs halved callback rates to actual job applications.[68] Perceptions of working mothers tend to reflect "patriarchal" stereotypes. Benard and Correll (2010, p.1) write that "highly successful mothers" are seen as "less warm, less likable, and more interpersonally hostile." Put differently, patriarchal norms shape the image of the "ideal" mother. Culture determines if mothers should or even may participate in labor markets (Budig et al., 2012).

[66] See for instance a report in the Guardian, January 2011: "Where women of India rule the roost and men demand gender equality", available at: https://www.theguardian.com/world/2011/jan/18/india-khasi-women-politics-bouissou [Accessed 28 August 2018]. Roy (2018, p.283) writes, *"Unlike the other patriarchal societies, the father has little authority in a Khasi family"*.

[67] Klonner et al. (2019) report that men from the patriarchal/patrilineal Karbi tribe compete 50 percent more often than women, while amongst the Dimasa tribe which is duo- or bilineal, men compete about 18 percent more often than women, while in the case of the Khasi tribe, which is matrilineal, men are 15 percent *less* likely to compete as compared to women.

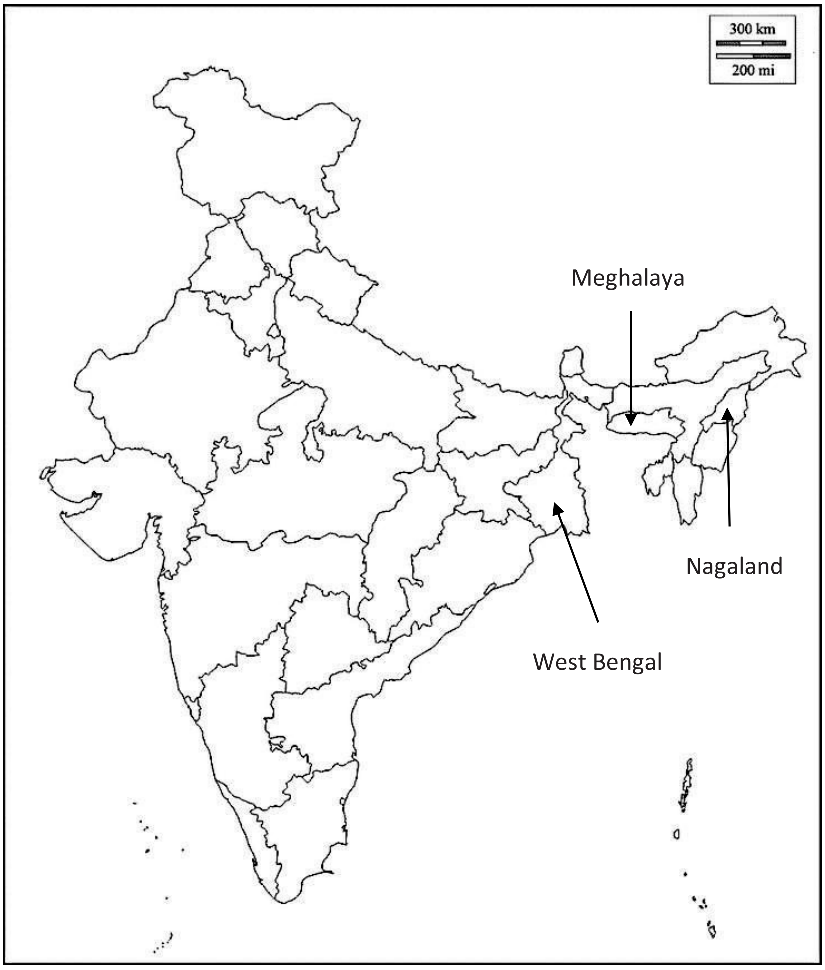
[68] In fact, controlling for qualifications, women without children did better than men without children. A more recent CV experiment in Sweden found no differences in callback rates across gender and/or parenthood (Bygren et al., 2017).

Bringing together the literature on the motherhood penalty and the effect of culture on gender competitiveness, we hypothesize that mothers from more empowered communities, and even more so from matrilineal societies are less likely to face a motherhood penalty. With respect to the former, employers are likely to value their competiveness, cultural background and supportive household arrangements, for instance, when it comes to childcare and are likely to view them as "*more* competent and *more* committed to their jobs."[69] In subsequent sections, based on key informant interviews, we comment on whether employers indeed have such perceptions (even if they do not know the explicit cultural/organizational underpinnings of these communities).

This chapter examines the labor market success, as measured by interview callback rates of mothers and non-mothers as a function of two community origins: (i) We hypothesize that women from Northeast India are less vulnerable to a motherhood penalty as it is well-known that they are more empowered compared to women from the rest of India (Ladusingh and Singh, 2006; Jayachandran and Pande, 2017). For instance, Ladusingh and Singh (2006, p.67) state: *"The social status of women in the Northeast India is high relative to that of women in many parts of the country where purdah and caste based rules restrict their activities."* Additional information to support the idea of greater empowerment of women from the Northeast is provided in Table 4.1. (ii) We differentiate between women from matrilineal (Khasi) and patrilineal (Naga, Bengali) societies located in Northeast and East India (see Figure 4.1) and hypothesize that women from matrilineal communities are less likely to experience a motherhood penalty.

[69] Reversing the earlier quote from Correll et al. (2007, p.1298).

Figure 4.1: Map of India –Location of communities



Similar to Correll et al. (2007), the chapter is based on a CV experiment and building on Gneezy et al. (2009), proposes a cross-cultural identification strategy. We quantify if employers (regardless of their own societal origin) differentiate between applications sent by mothers and non-mothers *within* origin societies. Our fictitious applicants are mothers or non-mothers of Khasi (Northeast India, matrilineal), Naga (Northeast, patrilineal) or Bengali (East India, patrilineal) origin. This allow us to examine the effect of empowerment by comparing interview callback rates for mothers from Northeast India with mothers from East India, as well as the implicit effect of culture by comparing Khasi to Bengali and Naga mothers.

To execute the experiment, which was conducted in two rounds, we searched for entry-level jobs in call centers or business process outsourcing (BPO) and in the financial sector, across three Indian cities.[70] In the first round we sent three female CVs, with no prior work experience, to each job posting. In a second round, to examine the potentially moderating effect of experience, CVs did indicate experience. Furthermore, in both rounds, to net out overall effects of gender and community we also sent male CVs from each of the three communities. In total, we sent 1,276 CVs (957 female, 319 male) to 319 job openings. To complement the experiments and to enhance our understanding of the findings we conducted 12 key informant interviews with current or former (human resource) managers (six), with head hunters (five) and one academic.

Our chapter makes three contributions to the literature. *First*, we provide causal evidence on the societal origin and labor market success of women with and without children. We build on previous gender experiments across cultures (Gneezy et al., 2009; Hoffman et al., 2011; Andersen et al., 2013; Cassar et al., 2016). Rather than focusing on preferences, we examine the downstream effects of culture on labor market success in the context of one of the most important dimensions of gender and labor markets, namely motherhood.[71]

Second, we add to the broader literature on female labor market participation in developing countries. Many factors influence whether women work or not in developing countries, including changes in income per capita, the structure of the economy, fertility trends, education levels, and social policies (Verick, 2014; Gaddis and Klasen, 2014; Bloom et al., 2009; Mammen and Paxon, 2000; Goldin, 1994). In some countries, most notably India, female labor force participation is lagging behind favorable economic and demographic dynamics (Klasen and Pieters, 2015). To the best of our knowledge, there is no experimental evidence on the labor market consequences of motherhood and gender norms in a developing

[70] We focused on entry-level jobs and these sectors for two main reasons. First, these sectors offer a steady and relatively large volume of job advertisements. Related literature also underlines the importance of the chosen sectors. For instance, Jensen (2012, p.754) notes that the BPO field *"…has grown rapidly in India over the past decade, creating a significant number of new, high-paying job opportunities, particularly for women."* Second, focusing on entry level jobs allowed us to examine whether work experience translated into an advantage for mothers.

[71] In this chapter we focus on the motherhood penalty. It is also possible that firms' treatment of fathers and non-fathers varies across origin communities.

country setting. Such evidence is likely to be useful in motivating and designing childcare and maternity policies.

Third, in addition to identifying the labor market effects of gender, culture and motherhood, our rich setting allows us to examine the effect of ethnicity. At the outset, it is important to emphasize the potentially differential effect of two aspects – ethnicity and gender. Both the Khasi and Naga are from Northeastern India and while women from this part of the country are considered to be more empowered and have a better status as compared to women from the rest of the country (Ladusingh and Singh, 2006) at the same time, both men and women from the Northeast are likely to be discriminated in the labor market as compared to "mainland" Indians. While there have been changes in recent years, there remains a perception that people from the Northeast are not part of mainstream India.[72] Consistent with this notion, there are regular reports of discrimination in the housing market against Northeasterners in cities such as Delhi (McDuie-Ra, 2013; Irfan, 2011). This is also supported by one of our key informants (female, 40 years), who articulated, that women from the Northeast who migrate to Delhi or Bangalore face discrimination in the housing market as they are thought to be morally questionable and don't have the same moral values as mainstream Indians so there is xenophobia [interviewed on May 26, 2018]. However, there is no credible evidence on the extent of such ethnic-based discrimination in the labor market. Thus, our chapter feeds into a growing experimental literature on labor market discrimination in emerging and developing countries (Banerjee et al., 2009; Siddique, 2011; Galarza and Yamada, 2014, 2017; Beam et al., 2017).

To preview our results, we find that mothers are substantially less likely to receive callbacks (14%-points). This effect varies considerably across communities. Relatively more empowered Northeastern women average a smaller motherhood penalty (5.68%-points; p-value=0.10). Mothers from *patrilineal* East India and Northeast India are affected strongly (-29.48%-points; p-value=0.00) and mildly (-9.12%-points; p-value=0.08), respectively.

[72] There are various reasons for this including geographical distance, differences in physical features, language and religion among others. The seven Northeastern states are connected to the rest of India through a narrow corridor. A number of them have a long history of insurgency and have demanded independence from the Indian state. In term of their physical features, Northeasterners have a greater affinity to East Asians and it is not uncommon that they are racially labelled as such. Furthermore, according to data from Census 2011, Christianity is the dominant religion in a number of the Northeastern states (74.6% in Meghalaya and 88% in Nagaland). While some of these aspects, may enhance xenophobic attitudes to people from the Northeast these features - missionary education, English language skills - are also held responsible for the greater freedom and mobility of women from the Northeast as compared to women from "mainland" India.

Mothers from *matrilineal* Northeast India face no discernible penalty (-2.27%-points; p-value=0.67). Interestingly, we do not find gender differences in callback rates for male and female applicants *without* children. Consistent with findings from the US (Correll et al., 2007), gender differences materialize only due to motherhood. In the second round of the experiment where we added experience to all CVs, qualitatively similar patterns emerge, although, the magnitude of the adverse motherhood effect is smaller. With regard to ethnicity we find that women from the Northeast receive substantially fewer callbacks as compared to Bengalis. This gap arises mainly due to differences in callback rates in the financial sector.

The chapter is organized as follows: Section 4.2 provides contextual information which draws on the existing literature as well as on key informant interviews and outlines the empirical strategy. Section 4.3 presents the main results and related robustness checks. Section 4.4 discusses the findings in relation to theories of discrimination, as well as with respect to recent Indian policy changes including the 2016 amendments to the maternity benefits act.

4.2 Context and Empirical Strategy

We implemented a field experiment to test for the effect of reporting motherhood on callback rates to job applications in three Indian cities and two industry sectors. Our aim was to examine motherhood effects conditional on community origin and ethnicity. To underpin the empirical approach, this section provides details on the context, the choice of communities, the selection of jobs, the design of applicant profiles as well as treatments, experimental procedures and the econometric model.

Context and Selection of Communities

We first picked matrilineal and patrilineal societies from Northeast and East India, respectively. For the matrilineal treatment we chose the Khasi community. The Khasi community which is based in and around the city of Shillong and in the Khasi hills in the Northeastern Indian state of Meghalya was chosen for two reasons. *First*, it is well-documented that Khasi women enjoy greater privileges as compared to women from patrilineal communities. For instance, according to Nongbri (2006, p. 168): "*Throughout the ages, the Khasis have lived in a casteless and classless society where every kind of labour is* *respected. Men and women work and talk together freely. Everyone knows that he or she is equal with others in the society*."[73] More recently, experimental evidence has shown that women from the Khasi community are as competitive as men from other patrilineal societies (Gneezy et al., 2009; Andersen et al., 2013; Klonner et al., 2019). We thus expect that gender-related treatments such as motherhood are likely to have a lower effect on callback rates. *Second*, amongst the handful of matrilineal societies in India, the Khasi community is one of the largest and perhaps most well-known across India (the first row of Table 4.1 provides female population sizes).[74]

For the patrilineal treatment, as well as to identify the effect of ethnicity, we selected the Bengali community from the Eastern Indian state of West Bengal. Although there have been changes, Bengali society may be characterized as patriarchal, patrilineal and patrilocal. However, simply comparing Bengali women and Khasi women is not straightforward as there may be discrimination against people from Northeastern India which may confound or drive the heterogeneous impacts of motherhood across Bengali and Khasi CVs. We address this issue in two ways. *First*, we selected an additional patriarchal, patrilineal, and patrilocal community, the Naga, who are also from Northeast India, and are physically similar to the Khasi, both groups are predominantly Christian and English is the official language and medium of instruction in educational institutions in both Meghalaya and Nagaland. To emphasize, although Naga society is characterized as patriarchal, patrilineal and patrilocal, women tend to enjoy a higher social position and are more independent than women from patriarchal societies located in "mainland" India (Naga Women's Union et al., 2018) but enjoy less freedom and experience greater restrictions as compared to Khasi women (Ellena and Nongkynrih, 2017).[75] Thus, we expect that the motherhood penalty for Naga women to be lower than that for Bengali women but higher than that for Khasi women. *Second*, we sent out male CVs from all three communities to document overall callback rates. Figure 1 shows the location of the three communities on a map of India.

---

[73] Nongbri (2006, p.168) goes on to write, "*The Khasi woman is no mere chattel of the family of men. No feminist movement is required to free her from bondage. She is the glorified person, free to act*."

[74] The Khasis in particular are well-known for their matrilineal traditions and there are numerous media reports in both the national and the international press about their cultural practices. See for instance the reference in footnote 2. There are other matrilineal groups in Meghalaya such as the Garo and the Jaintia but they are not as large in number as the Khasis (Roy, 2018).

[75] The Naga community has a strong warrior heritage and traditionally boys were more highly valued than girls. According to Ellena and Nongkynrih (2017), even up to the late '60s, some Naga groups practiced headhunting.

With this set-up at hand, we can decompose callback rates by gender, motherhood and community origin. The implicit expectation is of course that community characteristics and signals carry over when people migrate domestically for work. Indeed, the reasons to hire women from the Northeast may well be due to their community traits and signals. For instance, amongst other reasons, one of our key informants (female, 50 years) argued that for service-sector positions there was a bias towards hiring women from the Northeast, because women from the Northeast are not fazed by challenges, they are able to deal with long working hours, they have a calmer temperament, they are non-confrontational" [Interview conducted on January 10, 2018].

Table 4.1: Female status in applicant communities

Region		*East India*	*Northeast India*	
	India	West Bengal (Bengali, Patrilineal)	Nagaland (Naga, Patrilineal)	Meghalaya (Khasi, Matrilineal)
Nr. of women (*in millions*, 2011 Census)	587.58	44.47	0.95	1.48
Women ever worked	0.42	0.30	0.24	0.81
Willing to work	0.61	0.64	0.75	0.95
Husband decided number of children	0.92	0.92	0.36	0.75
Husband beats if wife leaves without permission	0.51	0.47	0.12	0.15

Note: Data are from the women questionnaires in the Indian Human Development Survey II 2011-12. Average sample sizes (sample sizes vary slightly across outcomes): India 39291, West Bengal 2385, Nagaland 48, Meghalaya 67.

Evidence from survey data and key informant interviews supports the idea of differences in the status of women across the three communities. Table 4.1 shows the non-negligible population sizes of the three communities and contains information on some pertinent statistics. Only about 30% of Bengali women (state of West Bengal) have some work experience, which is well below the national average of 42% and much below the Khasi average of 81%. More importantly, in terms of attitudes, both Naga and Khasi women are more willing to work (75% and 95%, respectively) than Bengali women (64%).[76] Women from the Northeast have more say over the number of children. Finally, about half of the Bengali women report a community norm of violence against women if they leave the house without the permission of their husband compared to less than 15% for Naga and Khasi women. Thus, in terms of a number of proxies for women's empowerment such as mobility,

[76] The exact question in the survey is, "If you found a suitable job, would you be willing to work?"

say in decision making and potential attachment to the labor market, there is a clear difference between women from the Northeast and women from Bengal. Key informant interviews also highlighted differences between Bengali women and women from the Northeast. According to a former (male, 70 years) human resource manager who had worked in Bengal and in the Northeast, “these people [from the Northeast] are the products of missionary education, they have a stronger work ethic, they will find a way to work, they have stronger family support; whereas, Bengali women will expect sympathy; gender equality is higher in the Northeast …. their [Northeast women] attitude is better” [interview conducted on January 6, 2018]. Another key informant, a former (male, 48 years) recruiter for an international BPO firm mentioned that unlike women from other regions of the country, women from the Northeast were flexible and happy to work, day or night [interview conducted on January 8, 2018].

Selection of Job Market and Postings

We focused on job markets in three of India’s most cosmopolitan cities, that is, Delhi, Mumbai and Chennai. All three cities have residents from the three communities. There is survey evidence which suggests that 48% of all Northeast people residing in Indian cities live in Delhi (NESCH, 2011 as quoted by McDuie-Ra, 2013, p.1629 and Irfan, 2011).

We used the most popular Indian job website to search and apply for openings in these three locations. Women in urban India are most likely to work in the service sector and the job website features a steady volume of service-sector positions. We focused on low- to medium skilled jobs in two broad sectors: (i) Business Process Outsourcing (BPO) and call center jobs, (ii) Banking/Finance/Insurance. Both sectors feature a steady and large volume of job ads required for the experiment. We selected jobs that were open to both experienced and inexperienced applicants.[77] In terms of sector-specific characteristics, as compared to jobs in the financial sector, BPO and call center jobs require greater flexibility in working hours, customer contact is not face-to-face and spoken English language skills/accents are more highly valued.

[77] We have saved screenshots of all the positions to which resumes were sent. These are available on request.

Design of Applicant Profiles

Based on input from a human resource consultancy firm, we designed several fictitious resumes. Our aim was to build comparable CVs across applicants and most importantly clearly signal community origins. All CVs provided a current address in the respective job market city, and also a permanent address in the home community. For the latter, we picked thee cities from each community – Siliguri, Shillong and Kohima for Bengali, Khasi and Naga applicants, respectively.[78] We also used names which are typical for each of the three communities. All fictitious participants had the same education level, graduated from comparable colleges and acquired their high school education in English medium schools in their native places. In India, there is a strict hierarchy of academic disciplines with the hard sciences situated at the top. We assigned three comparable, relatively less prestigious academic subjects to our applicants – Political Science, Sociology and History. All our applicants were legally married and aged 25 to 28. Given that we sent three female CVs to each job posting, profile details and CV format could not be "identical." This should not be a concern, given that we estimate the within community impact of motherhood. And while some characteristics such as age may be related to labor market outcomes, these differences are limited. An alternative would have been to vary some CV characteristics (e.g. age, type of degree) across jobs and within applicants, however this would have further increased the already large number of CVs (27) and the complexity of data collection. In the first round of data collection, the applicants had no prior job experience while in the second round of data collection we sent out the same CVs with about two years of relevant job experience.

Community Signaling

We signaled community origins in five ways. First, we used typical Khasi, Naga or Bengali names and provided a permanent address, and details on schooling and college which indicated their respective home states. Second, current addresses on all the CVs indicated C/O (care of). In the case of Khasi CVs, it was the applicant herself while the Bengali and Naga CVs featured the names of husbands. Third, the permanent addresses of the applicants mention the names of parents, which is not uncommon in India. In the case of Khasi CVs, we used D/O (daughter of) and used a female (mother's) name and in the case of Naga and Bengali CVs, we used C/O and used a male (father's) name. Fourth, Khasi applicants had the same surname as their mother, while the patrilineal applicants had the same surnames as their

[78] The main Bengali city is Kolkata. However we picked Siliguri to better match the size and status of the other cities.

husbands. Finally, in the case of Khasi CVs, we also mentioned that Khasi was the native language.

Motherhood Treatment and Procedure

We reported motherhood (1 child between 2-2.5 years of age) allowing for *within* job posting variation (at least one mother and non-mother per job posting). We focused on mothers with just *one* child given the potentially large changes associated with first motherhood and more practically not to encumber the experiment with additional complexity. Varying the number of children is likely to trigger interesting but equivocal considerations among employers.[79] Thus, there were six possible combinations to assign the motherhood treatment to the three female CVs. Before searching for jobs and sending out the CVs, we randomly determined the sequence in which these six combinations were applied throughout the ensuing experiment. To each job posting, we also randomly sent out *one* of three additional male CVs (without reporting fatherhood). This allows us to examine overall differences in callback rates amongst the different communities.

In total, we used twenty seven CVs – nine (three mother, three non-mother and three male) in each of the three cities. Each CV was assigned a unique email id and phone number to record callbacks. All CVs are available on request from the authors for bona fide researchers and purposes.[80]

Data collection, that is, sending out CVs and recording outcomes, took place between July and September of 2017. The experiment was conducted in two rounds and our overall sample consists of 1276 applications (male, female) sent to 319 job openings. In the first round, our target sample size was at least 200 female applications per community.[81] In a smaller, second round, we sent out CVs *with* job experience for a total of 90 applications per community. Table 4.2 summarizes realized sample sizes by communities, and mother and non-mother treatments. In total, we sent out 957 female applications across 258 firms. In addition, we

[79] When evaluating CVs of mothers, an employer may among other things place a lower probability on another pregnancy for a mother with two or more children compared to a mother with just one child. Future work could investigate the potential effects of reporting more than one child.

[80] We do not present the CVs in the appendix as real addresses and schools were used.

[81] We carried out power calculations for a test of two proportions using the "pwr" package in R. Setting Cohen's h to 0.4 (small to medium effect), power to 80% and significance level to 5%, the proposed sample size was 200 for each community. In practice, we slightly exceeded this target.

sent out 229 (1st round) and 90 (2nd round) male applications well-balanced across sectors, communities and cities.

Table 4.2: Sample sizes for female sample

	No prior job experience (1st round experiment)		Experienced (2nd round experiment)		
	Non-mother	Mother	Non-mother	Mother	Total
Bengali	115	114	44	46	319
Khasi	113	116	46	44	319
Naga	116	113	44	46	319
Total applications (job openings)	**344**	**343**	**134**	**136**	**957 (258)**
… broken down by place and sector:					
Chennai	117	111	44	46	318
Delhi	112	116	45	45	318
Mumbai	115	116	45	45	321
Call center, Business Process Outsourcing (BPO)	182	178	63	72	495
Finance, banking, insurance	162	165	71	64	462

Econometric Model

We first visualize results based on unconditional means and then estimate corresponding econometric models of the form:

$$y_{ij} = \alpha + \beta_1 Mother_{ij} + \beta_2 Naga_{ij} + \beta_3 Khasi_{ij} + \beta_4 Delhi_{ij} + \beta_5 Mumbai_{ij} + \beta_6 BPO_{ij} + \delta_2 Naga_{ij} \times Mother_{ij} + \delta_2 Khasi_{ij} \times Mother_{ij} + \epsilon_{ij},$$

where, y_{ij} is a binary callback variable for applicant *i* and job opening *j*. $Mother_{ij}$, $Naga_{ij}$, $Khasi_{ij}$ are binary variables indicating whether an applicant is a mother or not, whether an applicant is Khasi or Naga, respectively. The excluded categories are non-mothers and Bengalis. $Delhi_{ij}$, $Mumbai_{ij}$, and BPO_{ij} are indicators for Delhi, Mumbai and the BPO/Call Center sector. Excluded categories are Chennai and Finance/Banking. We include interactions between Naga (Khasi) and motherhood ($\ldots \times Mother_{ij}$). The usual error term is denoted ϵ_{ij}. Standard errors are adjusted for clustering at the job posting level.

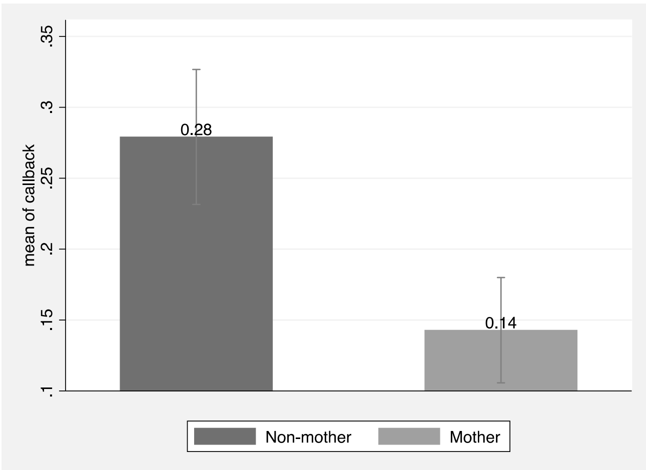
4.3 Results

We first present simple differences in mean callback rates across communities, then regression-based estimates and finally sector and city-specific estimates.

Baseline Results

Figure 4.2 shows callback rates for non-mothers and mothers without prior work experience. The average callback rate is 21%. However, mothers receive substantially fewer callbacks (14%) than non-mothers (28%). In other words, reporting motherhood on CVs halves callback rates. The motherhood treatment effect amounts to -14%-points and is precisely estimated (p-value=0.00).

Figure 4.2: The impact of motherhood on callback rates for women *without* prior job experience (Δ -13.62%-points, p-value=0.00, n=687)

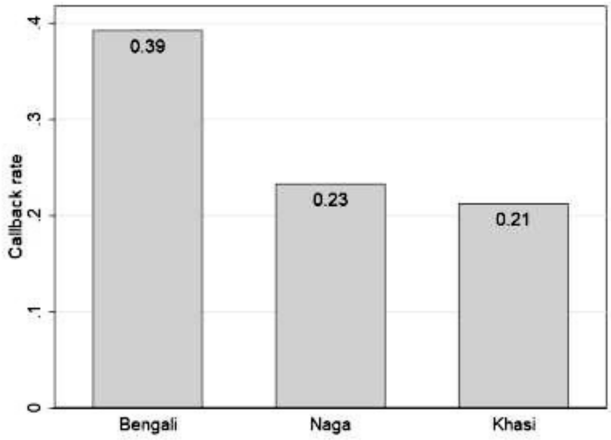


Note: P-value stems from linear regression-based t-tests adjusted for clustering at the job posting level (229 jobs).

Motherhood treatment effects vary considerably between Northeast and East India, as well as across patrilineal and matrilineal applicants. However, before discussing these effects, it is useful to consider baseline callback rates across communities, that is, callback rates for applicants without children and no prior work experience. Figure 4.3, Panel A reports results for women and Panel B for men. This exercise allows us to net out community-specific and gender effects from the motherhood penalty. There is a clear hierarchy. Bengali applicants receive about twice the number of callbacks as compared to Naga and Khasi applicants. This is consistent with the expectation, as discussed in the introduction, that individuals from the Northeast experience discrimination. At the same time there is no statistically significant difference in callback rates between Nagas and Khasis. Furthermore, callback rates for female and male applicants are similar (compare Panels A and B). These "baseline" patterns which show no gender differences but sharp community-based differences lend credibility to our strategy of comparing motherhood effects between Naga and Khasi women to identify the effect of patrilineal versus matrilineal culture on labor market outcomes.

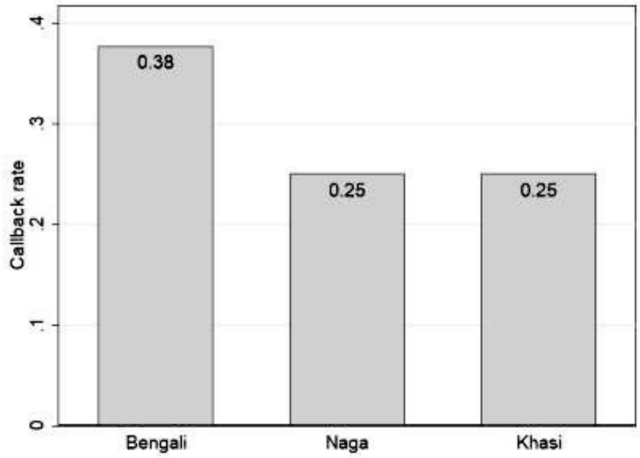
Figure 4.3: Baseline callback rates for Bengali (East India, patrilineal), Naga (Northeast India, patrilineal) and Khasi (Northeast India, matrilineal) women and men (*without* children/ *without* prior job experience)

Panel A: Women (non-mothers)



Δ p-values (N=344)
Bengali vs. Naga: 0.00
Bengali vs. Khasi: 0.00
Naga vs. Khasi: 0.68

Panel B: Men (non-fathers)



Δ p-values (N=229)
Bengali vs. Naga: 0.10
Bengali vs. Khasi: 0.09
Naga vs. Khasi: 1.00

Note**:** P-values stem from linear regression-based t-tests adjusted for clustering at the job posting level (229 jobs, Panel A) or heteroscedasticity (Panel B).

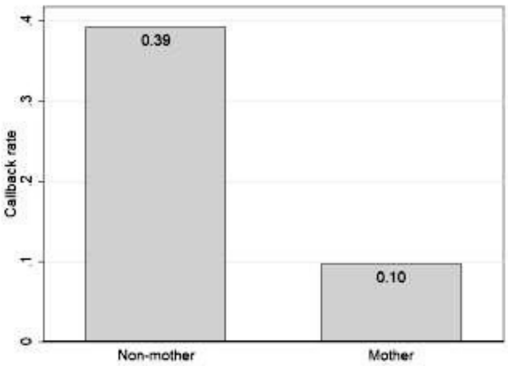
Figure 4.4 illustrates that the motherhood penalty decreases as empowerment of women increases (moving from Panel A, B to C). Panel A shows that amongst Bengali women, the treatment effect associated with motherhood is almost -30%-points (p-value=0.00). This is a very large effect with Bengali mothers experiencing a low 10% callback rate. In fact, this is the lowest callback rate in our experiment across communities and applicants (both male and female). Panel B shows qualitatively similar but smaller effects for Naga women. This smaller effect may be attributed to the general perception that women from Northeast India are more empowered compared to the rest of India (Ladusingh and Singh, 2006; Jayachandran and Pande, 2017). The motherhood penalty amounts to 9%-points (p-value=0.08). This is still a sizeable reduction of about 40%. Panel C shows no motherhood penalty for Khasi women. Note that we obtain qualitatively similar conclusions on the overall motherhood penalty as well as the three community-specific ones if p-values are based on randomization inference (see Alwyn, 2019).[82] Additionally, combining the estimates for

[82] We calculated p-values with randomization inference using the command *ritest* by Heß (2017) in STATA. The randomization inference p-value associated with the overall motherhood penalty taking into account firm strata is 0.00 (in line with the standard p-value in Figure 2). We further ran regressions in split samples corresponding to the three panels in Figure 4. These randomization inference p-values are close to ones reported in the main text and figure. They are 0.00, 0.097 and 0.745 among the Bengali, Naga and Khasi samples, respectively. These patterns alleviate statistical inference concerns related to small samples in an experimental setting (see Alwyn, 2019).

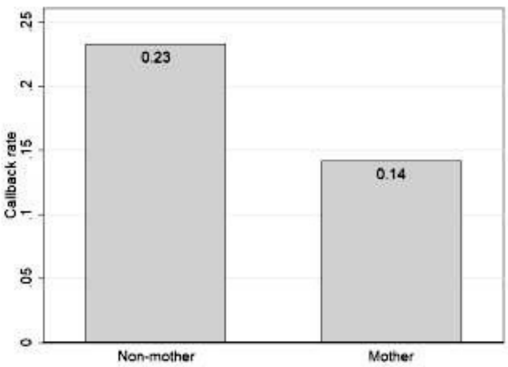
women from the Northeast yields a motherhood penalty of 5.68%-points (p-value=0.10). In sum, motherhood penalties are lower for women from the Northeast and only affect women from patrilineal communities.

Figure 4.4: The impact of motherhood on callback rates for Bengali (East India, patrilineal), Naga (Northeast India, patrilineal) and Khasi (Northeast India, matrilineal) women (*without* prior job experience)

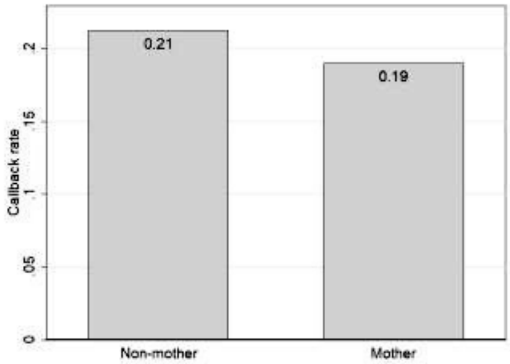
Panel A: Bengali – East India, Patrilineal (Δ -29.48%-points, p-value=0.00, n=229)



Panel B: Naga – Northeast India, Patrilineal (Δ -9.12%-points, p-value=0.08, n=229)



Panel C: Khasi – Northeast India, Matrilineal (Δ -2.27%-points, p-value=0.67, n=229)



Note: P-values stem from linear regression-based t-tests adjusted for heteroscedasticity.

Regression Results

Table 4.3 reports results from a linear probability model. Column 1 reports estimates adjusting only for community effects (Naga and Khasi). The coefficient associated with the motherhood treatment is -14%-points. In column 2, we add dummies for cities and sector of employment. The motherhood effect is unaffected by these additional covariates, which is unsurprising given experimental balancing. In column 3, we further include interaction terms between motherhood and communities. The excluded category is a Bengali mother. Consistent with Figure 4.4, we document large negative motherhood effects for Bengali women (30 %-points) and to a lesser extent for Nagas (9%-points), but there are no negative motherhood effects for Khasi women. The interaction term between Khasi applicants and motherhood (27%-points) statistically offsets the negative main effect of motherhood.

Table 4.3: Linear probability model

Dep. var. Callback	(1)	(2)	(3)	(4)	(5)
Mother	-0.14***	-0.14***	-0.30***	-0.08*	-0.18*
	(0.03)	(0.03)	(0.05)	(0.04)	(0.09)
Group (Bengali is excl.)					
Naga	-0.06**	-0.06**	-0.16***	-0.07	-0.12
	(0.02)	(0.02)	(0.05)	(0.04)	(0.09)
Khasi	-0.04	-0.04	-0.18***	-0.11**	-0.21**
	(0.03)	(0.03)	(0.05)	(0.04)	(0.08)
Mother x Naga			0.21***		0.11
			(0.08)		(0.14)
Mother x Khasi			0.27***		0.19
			(0.08)		(0.13)
City (Chennai is excl.)					
Delhi		0.06	0.05	0.07	0.06
		(0.05)	(0.05)	(0.08)	(0.08)
Mumbai		0.05	0.06	0.09	0.09
		(0.05)	(0.05)	(0.09)	(0.09)
Sector (Finance is excl.)					
Call center/BPO jobs		0.03	0.02	0.11	0.11
		(0.04)	(0.04)	(0.07)	(0.07)
Constant	0.31***	0.26***	0.34***	0.23***	0.28***
	(0.04)	(0.05)	(0.06)	(0.08)	(0.10)
P-values:					
Mother = - Mother x Naga			0.09		0.40
Mother = - Mother x Khasi			0.61		0.94
N	687	687	687	270	270
Prior job experience		No		Yes	

Note**:** Linear probability model. Standard errors in brackets below point estimates are clustered at the job posting level (229 jobs in columns 1-3; 90 jobs in columns 4 and 5). Significance levels are denoted *p<0.1, **p<0.05, ***p<0.01.

So far, we have analyzed motherhood effects for women with no prior job experience. In the second round we sent out a smaller set of applications with the same CVs but with two years of job experience.[83] As shown in column 4, it does seem that experience weakens the motherhood penalty. However, we still find a motherhood penalty of 8%-points which is statistically significant at the 10% level and while the effect is smaller in magnitude (by 6%-points) as compared to CVs with no experience, the difference is not statistically significant

[83] In the case of male CVs reporting prior experience, callback rates amount to 43% for Bengali, 21% for Naga and 19% for Khasi applicants. In other words, male community patterns are qualitatively similar to those stemming from the first experiment without job experience.

(p-value=0.31). In column 5, we find very similar qualitative patterns across communities but the estimates are not precise. Across the two samples, tests for equality of the main effect as well as the community interactions fail to reject the nulls at conventional levels.[84] Similar to the results based on the sample without experience, Khasi applicants do not experience a motherhood penalty. The main effect of motherhood (-18%-points) is completely offset by the corresponding interaction term (19%-points). The motherhood effect for Nagas amounts to -7%-points, but it is imprecisely estimated.

Heterogeneity by Sector[85]

Table 4.4 provides pooled estimates based on both rounds of data collection (column 1 and 2), which display the same patterns as discussed above, and sector-specific estimates (columns 3 to 6). Column 3 shows that there is a motherhood penalty in the call center/BPO sector (17%-points) but there are no negative effects associated with belonging to the Northeast. The motherhood-community interaction specification (column 4) confirms the sizeable motherhood penalty while at the same there is a significant and positive interaction term for women from the Khasi community (column 4). In the finance sector the motherhood penalty is lower (7%-points) but women from the Northeast experience an additional penalty of -9 to -11%-points. The motherhood penalty varies across communities with Bengali mothers experiencing a substantial penalty, while there are no negative effects for mothers from the Naga community and perhaps even a small premium for Khasi mothers.

[84] Tests for equality of coefficients in Table 4.3 across inexperienced (column 3) vs. experienced (column 5) samples: Mother, p-value=0.30; Mother x Naga, p-value=0.54; Mother x Khasi, p-value=0.60.

[85] For the sake of completeness, we do explore sector and city specific heterogeneity. However, sample sizes are small as our ex-ante power calculations were not based on providing sector and city specific-estimates.

Table 4.4: Pooled sample (experienced and inexperienced) – Heterogeneity by sector

Dep. var. Callback	(1)	(2)	(3)	(4)	(5)	(6)
Mother	-0.12***	-0.26***	-0.17***	-0.26***	-0.07**	-0.27***
	(0.02)	(0.05)	(0.03)	(0.06)	(0.03)	(0.07)
Naga	-0.06***	-0.15***	-0.01	-0.05	-0.11***	-0.24***
	(0.02)	(0.04)	(0.03)	(0.06)	(0.03)	(0.06)
Khasi	-0.06***	-0.19***	-0.04	-0.11*	-0.09**	-0.26***
	(0.02)	(0.04)	(0.03)	(0.06)	(0.04)	(0.07)
Mother x Naga		0.18***		0.10		0.26**
		(0.07)		(0.09)		(0.10)
Mother x Khasi		0.25***		0.16*		0.34***
		(0.07)		(0.09)		(0.10)
Call center/BPO jobs	0.05	0.05				
	(0.04)	(0.04)				
Inexperienced	-0.02	-0.02	-0.06	-0.06	0.02	0.02
	(0.04)	(0.04)	(0.06)	(0.06)	(0.06)	(0.05)
Constant	0.26***	0.34***	0.36***	0.40***	0.22***	0.34***
	(0.05)	(0.06)	(0.07)	(0.08)	(0.06)	(0.08)
P-values:						
Mother = - Mother x Naga		0.07		0.02		0.81
Mother = - Mother x Khasi		0.70		0.14		0.26
Sample	Full		Call Center/BPO		Finance	
N	957		495		462	

Note**:** Linear probability model. City dummies not shown. Finance/Banking and Bengali are excluded categories in columns 1 and 2. Standard errors in brackets below point estimates are clustered at the job posting level (319 jobs in column 1 and 2; 165 jobs in columns 3 and 4; 154 jobs in columns 5 and 6). Significance levels are denoted *p<0.1, **p<0.05, ***p<0.01.

Heterogeneity by City

In Table 4.5, we split samples by cities. The coefficient associated with motherhood is negative across all locations and specifications but varies across communities.

Table 4.5: Pooled sample (experienced and inexperienced) – Heterogeneity by city

Dep. Var. Callback	(1)	(2)	(3)	(4)	(5)	(6)
Mother	-0.20***	-0.40***	-0.12***	-0.20**	-0.05	-0.19**
	(0.05)	(0.07)	(0.04)	(0.09)	(0.03)	(0.08)
Naga	-0.05	-0.15*	-0.07*	-0.11	-0.06*	-0.18**
	(0.04)	(0.08)	(0.04)	(0.08)	(0.03)	(0.07)
Khasi	-0.00	-0.18**	-0.06	-0.15*	-0.12***	-0.20***
	(0.04)	(0.08)	(0.04)	(0.08)	(0.04)	(0.07)
Mother x Naga		0.23**		0.07		0.23**
		(0.11)		(0.14)		(0.11)
Mother x Khasi		0.37***		0.17		0.18*
		(0.12)		(0.12)		(0.11)
Call center/BPO	0.01	0.02	0.07	0.07	0.08	0.06
	(0.07)	(0.06)	(0.07)	(0.07)	(0.06)	(0.06)
Inexperienced	-0.01	-0.02	-0.04	-0.04	-0.01	-0.01
	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)
Constant	0.36***	0.44***	0.33***	0.38***	0.22***	0.30***
	(0.08)	(0.08)	(0.08)	(0.10)	(0.07)	(0.09)
P-values:						
Mother = - Mother x Naga		0.04		0.13		0.57
Mother = - Mother x Khasi		0.75		0.71		0.96
City Sample	Delhi		Mumbai		Chennai	
N	318		321		318	

Note**:** Linear probability model. Finance/Banking is an excluded categories. Standard errors in brackets below point estimates are clustered at the job posting level (106 jobs in columns 1 and 2; 107 jobs in columns 3 and 4; 106 jobs in columns 5 and 6). Significance levels are denoted *p<0.1, **p<0.05, ***p<0.01.

In Delhi, column 1 reports that mothers experience a penalty of 20%-points. Column 2 shows that the motherhood effect is particularly pronounced for Bengali women (-40%-points), is smaller for Naga applicants (-17%-points) while Khasi women do not experience a motherhood penalty. In Mumbai, the overall effect of motherhood amounts to -12%-points (column 3). The community-motherhood interaction terms indicate that at least qualitatively Khasi and Naga mothers suffer less. However, these interaction terms are imprecisely estimated. In Chennai, the overall motherhood penalty amounts to an insignificant 5%-points (column 5). This small effect masks heterogeneity by community origins. Bengali mothers face a penalty of 19%-points, while we find offsetting effects for both Naga and Khasi women.[86]

[86] Appendix Tables D1-D3 provide a detailed breakdowns of callback rates by gender, city and sectors in the inexperienced, experienced and pooled samples, respectively.

4.4 Discussion and Conclusion

In contrast to, but building on the existing literature (e.g., Gneezy et al., 2009, Andersen et al. 2013, Klonner et al., 2019), which has focused on gender, culture and competitive preferences, this chapter examined the downstream effect of culture on labor market success in the context of an important dimension of gender and labor markets, namely motherhood. Perforce, our study focused on two sectors in three mega-cities and represents only a tiny share of the labor market in each of these cities. This constrains the external but not the internal validity of our findings.

Our findings are consistent with the gender and culture literature which has shown a negative monotonic relationship between patriarchy and women's competitiveness. We found strong evidence of a motherhood penalty in callback rates to job applications in India. This penalty was starkly mediated by origin (Northeast and East India) and culture (patrilineal versus matrilineal).

We do not have information on hiring preferences of managers/firms where CVs were sent, but results are consistent with the views of a number of key informants (head hunters and human resources managers) who argued that ***(i)*** it is important for mothers to have child-care support "unless there is family support we clearly prefer to hire non-mothers" [male, 48 years, interviewed on January 8, 2018]; "we ask mothers who have young children who will take care of the child and if there is a mother-in-law or a full-time maid that increases chances of hiring" [female, 30 years, interviewed on May 20, 2018] and ***(ii)*** that mothers from the Northeast are more flexible and have a stronger work commitment, in part, due to greater child-care support - "they have stronger family support" [male, 70 years, interviewed on January 6, 2018]; "they are more efficient - if they have children they know how to manage it" [female, 50 years, interviewed on January 10, 2018]. The negligible motherhood penalty for Khasi women is consistent with the idea of stronger family support for their labor market roles as compared to women from patrilineal societies.

While our argument here and the argument in the existing literature is that a community's organizational system (patrilineal, matrilineal or the degree of patriarchy), manifests itself in differences in competitiveness, family support, independence, flexibility and efficiency, and results in lower motherhood penalties for certain groups, it is not essential for employers to know these origins. As long as potential employers have interacted with members of different communities and have experienced differences in their ability to handle job market needs it is likely to translate into the pattern of motherhood penalties that have been obtained in the chapter.

There was also suggestive evidence that job experience may moderate the motherhood penalty. We did not find noteworthy gender differences in callback rates when we differentiated by community origins. In other words, gender norms are most relevant on the labor market when it comes to motherhood. We also documented differences in callback rates across ethnic groups (Khasi and Naga versus Bengali) and found that these differences were concentrated in the Finance/Banking sector.

While it is not our intention to formally test whether differences in callback rates may be attributed to taste-based (Becker, 1971) or statistical discrimination (Phelps, 1972), our estimates link to both forms of discrimination. With regard to motherhood, if fully prejudiced, employers would discriminate against mothers regardless of their cultural background (matrilineal versus patrilineal) or ethnic origins. In contrast, statistical discrimination would predict that employers may discriminate against mothers, but may use observable signals of community or ethnic origin to proxy for unobservable traits such as competitiveness. The erosion of the motherhood penalty for women coming from a matrilineal background and the substantially lower motherhood penalty for women from the Northeast, who are generally considered more empowered as compared to women from other parts of India, points to statistical discrimination based on visible traits.

With regard to ethnic-based differences in callback rates, if differences are mainly driven by animus towards people from the Northeast then callback rates should not vary substantially across job sectors. However, we find that in the BPO/Call Centre sector where there is limited face-to-face client interaction and traits such as English speaking skills and flexibility (late night shifts) are relatively more important, women from the Northeast face no discrimination, while in a sector (finance/insurance) where employee-client interactions are more likely, employers tend to favor Bengalis.[87] If we assume that clients prefer to interact with Bengalis then employers may favor Bengalis even if they themselves are not prejudiced. Overall, the difference in ethnicity-specific callback rates across sectors also tends to support statistical discrimination.

Although we cannot pin-point the drivers of city-level differences in motherhood penalties, it is perhaps not surprising that Delhi, which is probably the most patriarchal and woman-unfriendly of the three cities in our study, stands out. For instance, according to the 2011 Census Delhi's child sex ratio (number of girls per 1000 boys, 0 – 6 years of age) is 873, which is well below ratios in Mumbai (914), Chennai (951) and India as a whole (902).[88] Female-male literacy gaps are larger in Delhi (80.96% and 90.98%) compared to Mumbai (86.45% and 91.48%) and Chennai (86.55% and 93.86%). These statistics reflect gender attitudes, and underlie the larger motherhood penalty in Delhi.

Our experimental evidence complements previous analyses of the "puzzling" Indian labor market. Klasen and Pieters (2015) report that the labor market participation rate of women in urban India has been stuck at around 18% over the period 1987 to 2011 despite drops in fertility and increases in female education. The authors argue that men's education and household income have risen starkly so that women, despite having higher levels of education may choose to stay at home. Thus far, one important dimension that has received far less attention is motherhood and related gender norms. Our findings echo a recent paper by Das and Zumbyte (2017, p.5) pointing to a strong role of motherhood norms and the lack of modern childcare in India: *"... women who are not perceived as fulfilling the role in the traditional sense are censured, either overtly or covertly, both within the home and outside."* Analyzing several rounds of employment surveys and controlling for a host of observables, Das and Zumbyte find that the odds ratio of employment among non-mothers compared to mothers (with at least one child under the age of 6) was 1.4 in 2011. In our experiment, the baseline odds ratio (Figure 4.2) is 2.3. Our outcome variable does not directly translate into

[87] One of our key informants, a former call center/BPO recruiter (male, 48 years) mentioned that for call center positions his firm preferred to hire women from the North-East as compared to Bengali women as it was easier to train women from the North-East to modify their English accents [interviewed on January 8, 2018]. Another key informant, an experienced recruiter (female, 50 years) argued that she does not discriminate but "filters" while recruiting. She went on to elucidate that for jobs that require sales and marketing skills she prefers women from Delhi and Mumbai while for jobs that require numerical skills she prefers women from the South [interviewed on January 10, 2018].

[88] Census data are available online at https://www.census2011.co.in/ [Accessed 28 August 2018]

actual employment and our experimental setting focused on specific sectors and cities, but the sizeable motherhood effect size that we find squares with this relatively large observational estimate.

As countries such as India develop, qualified women will be increasingly drawn into the expanding sectors investigated in this study such as BPO (Jensen, 2012). Even if fertility levels are on the decline, just one child may substantially penalize women on the labor market. This chapter finds that a supportive culture may mitigate this penalty. In addition, well-thought out and supportive labor market policies may also help, although it is unlikely that recent changes in the maternity benefits act in India are likely to do so.[89] Our empirical work took place soon after the law came into force and it is unlikely to have had a bearing on our results. However, more generally, while the act may have been well-intentioned, the entire burden of providing paid leave, organizing and setting up a crèche as well as hiring additional workers is placed on the employer. Given such a construction it is likely that a consequence of the act is a reduction in the probability of hiring women of child-bearing age as well as downward pressure on women's wages as firms may try to recover costs.[90,91]

[89] The Nordic countries are good examples (see Datta Gupta et al., 2008).

[90] The new act came into force in April 2017 and it entitles women to maternity benefits in the form of fully paid absence from work for a duration of 26 weeks, of which 8 weeks may be taken before birth of her child and the rest later. The act applies to all establishments that employ 10 or more employees and if an establishment has more than 50 employees, it is mandatory to have a crèche on the work premises. To be eligible a woman must have worked for at 80 days in the same establishment during the past 12 months. The main change in the revised act was the extension of the duration of benefits from 12 to 26 weeks.

[91] A number of articles in the popular press commented on the potentially adverse effects of the passage of the law for young women. For instance, see https://www.livemint.com/Opinion/XXInpbtQzgRWwe28GBr9aM/Opinion-Indias-wrong-approach-to-paid-maternity-leave.html (accessed on December 1, 2019) and https://www.financialexpress.com/opinion/indias-maternity-benefits-law-will-do-more-bad-than-good-here-is-why/1225954/ (accessed on December 1, 2019).

Chapter 5
Does Signaling Childcare Support on Job Applications Reduce the Motherhood Penalty?

(with Arjun Bedi and Matthias Rieger)

5.1 Introduction

While gender differences in labor market indicators are tapering off in many parts of the world (Goldin, 2014), the presence of women and even more so of mothers in labor markets remains low in many developing countries.[92] One much discussed case is India (Klasen and Pieters, 2015; Das and Zumbyte, 2017; Verick, 2018), where less than 30% of women participate in the labor force notwithstanding improvements in socio-economic and demographic conditions in recent years. The country's overall female labor force participation rate has even dropped from 22 percent in 1987 to 17 percent in 2011 (Klasen and Pieters, 2015), with the labor force participation of women typically falling around first motherhood (see chapter 4).[93]

There are several supply and demand side reasons for this decline. One potential reason why a non-prejudiced employer may be less likely to hire mothers is the belief that due to childcare and other family obligations, mothers are inflexible and are less dedicated to their jobs. This is likely to place mothers, especially in the case of jobs with traditional "9 to 5" work arrangements, at a severe disadvantage. Several studies suggest that mothers experience grave difficulties in balancing job and family demands and require supportive work arrangements (Goldin, 2014; Anderson et al, 2003).[94] Theoretically, such behavior may be related to statistical discrimination against mothers. If statistical discrimination, motivated

[92] There are wide variations in female labor force participation across the developing world. For example, in North Africa less than 30% of women aged 25 years or older participate in the labor market. In South Asia, the female labor force participation ranges from less than 30% in Pakistan to 80% in Nepal. See Verick (2018) for more details.

[93] According to data from the Indian National Sample Survey (NSS), the gap between the labor force participation of mothers and married non-mothers in urban area was 7.5 percentage points in 2011 (Das and Zumbyte, 2017).

[94] Based on NSS 2011 data, in the Indian context, a majority (70%) of married women (aged 25-55) reported that they stay out of the labor market due to domestic obligations.

primarily by the perceived (by employers) inflexibility of mothers and their lack of commitment is a primary factor holding back mothers from participating in the labor market, one way of improving their labor market prospects would be to signal their flexibility to employers. If mothers can signal flexibility, for example, by indicating childcare support at home, then the motherhood penalty should be smaller. Using an experimental approach, where one may credibly control for observable confounding characteristics and abstract from self-selection of applicants into certain jobs, the aim of this article is to identify the potentially mitigating impact of childcare support on the motherhood penalty.

While access to childcare may be expected to mitigate the motherhood penalty, it is possible that if employers are prejudiced against mothers/women (Becker, 1971) or do not believe that access to childcare is strong enough to undo traditional gender norms and patriarchal expectations, then, childcare availability may not have a substantial effect. In a similar vein, Benard and Correll (2010) argue that if mothers show a very strong commitment to paid work, and display traditionally masculine qualities, they may experience 'normative discrimination', that is, they may be "viewed as less warm and more interpersonally hostile (e.g., more selfish, cold, and devious) than other types of workers" (p. 617), which implies that mothers who violate traditional norms and display stronger labor market commitment may continue to experience discrimination as compared to other employees. Their argument is that "evidence of workplace competence and commitment will not eliminate discrimination but merely alter its mechanism" (p. 622), that is, from statistical or status-based discrimination to "normative discrimination".

In terms of empirical evidence on the motherhood penalty, a growing literature, mostly from sociology but increasingly in economics, finds that mothers experience disadvantages in terms of wages, promotions and job opportunities in the Global North (Anderson et al., 2002; Benard and Correll, 2010; Budig et al., 2012; England et al., 2016; Goldin et al., 2017; Gallen, 2018; Kleven et al., 2019a).[95] For instance, in Germany each child reduces a woman's wages by 16% - 18% (Gangl and Ziefle, 2009). In Denmark, having children

[95] Datta Gupta and Smith (2002) show a significant loss of human capital accumulation during the childbirth periods.

creates a gender gap in earnings by as much as 20% in the long run (Kleven et al., 2019b).[96] A related experimental literature has explored the motherhood penalty using fictitious CVs sent to actual job advertisements. In a CV experiment, Correll et al. (2007) found that by randomly adding at least one child to otherwise similar CVs, callback rates fell by 50%. Few such experiments have been implemented in the Global South. In chapter 4, we examined the labor market implications of motherhood and the effect of cultural norms (mothers from patrilineal versus matrilineal backgrounds) in India (Delhi, Mumbai, Chennai). Based on callback rates to fictitious job applications in the service sector, we found that the average callback rate was 14% for mothers as opposed to 28% for non-mothers. We also found that the motherhood penalty was restricted to mothers from patrilineal backgrounds while mothers from matrilineal backgrounds (the Khasis from Northeast India), where "every kind of labour is respected" (Nongbri, 2006), did not experience a penalty.[97]

In this chapter, we experimentally test for the mitigating impact of childcare support on the motherhood penalty. We employ a CV audit approach with random variation in CV characteristics and evaluate the impact of these characteristics on the likelihood of a callback for an interview. Specifically, we sent three applications to each job opening – one non-mother applicant and two mother applicants. In one of the mother CVs, we explicitly reported the availability of strong childcare support at home.[98] Building on our previous study, we searched for entry-level jobs in two relatively new and flourishing sectors in the Indian economy – that is, call centers/business process outsourcing (BPO) and financial sector firms in Delhi. We selected Delhi because it offers the most patriarchal context with the strongest motherhood penalty amongst the three cities in chapter 5, while at the same time it has a vibrant job-market in the two selected sectors.[99] All our fictitious applicants had a Bachelor's

[96] In a cross-country analysis Kleven et al. (2019a) find heterogeneity by culture. For example, they find that English speaking countries display a long-run child penalty of 31-44%, while in Germanic countries that penalty reaches 51-61%.

[97] In her work on the matrilineal Khasi community, Nongbri (2006, p. 168) writes, "Throughout the ages, the Khasis have lived in a casteless and classless society where every kind of labour is respected. Men and women work and talk together freely. Everyone knows that he or she is equal with others in the society"

[98] For instance, the skills section of the CV stated - Flexible and ready to adapt to corporate needs or Flexible and ready to adapt to corporate needs (strong childcare support at home).

[99] The other two cities were Mumbai and Chennai. In Mumbai and Chennai, mothers experienced penalties of 12 and 5 percentage points, respectively compared to a motherhood penalty of 20 percentage points in Delhi.

degree and had two years of relevant work experience. In total, we sent 450 CVs to 150 job openings.

This study is related to two strands of the literature. *First*, it is related to the literature on the labor market impact of motherhood (Angelov et al., 2016; Kleven et al., 2019a; Kuziemko et al., 2018). By investigating the effect of childcare support on mothers' labor market success we address a potential channel through which the motherhood penalty arises and may be weakened. This study is the first to experimentally quantify the effect of childcare support on callback rates in the Global South.[100] *Second*, our research links to the wider literature on women's participation in labor markets in emerging countries, in particular when it comes to demand side aspects. While there is a sizeable literature which examines the impact of education and fertility on female labor supply (Guo et al., 2018; Heath, 2017; Agüero and Marks, 2010; Bloom et al., 2009; Angrist and Evans, 1998) the influence of demand side factors on women's labor market outcomes have been relatively less investigated.

To preview our results, *first*, we find a large motherhood penalty, and *second*, we find that the childcare support signal reduces the motherhood penalty. On average, mothers' callback rates were 20 percentage points lower as compared to non-mothers. This translates to a 57% decrease in the callback rate. The childcare support signal significantly dampens the motherhood penalty by 4 percentage points or by 20%. The childcare support signal appears to be at least qualitatively more effective in the case of financial firms as compared to the BPO/Call center sector. This difference may be attributed to the greater credibility of the child care signal in the case of financial firms which operate during traditional working hours as compared to call center firms. Taken together, the evidence presented here suggests that childcare support has a modest, albeit a discernible moderating impact on the motherhood penalty.

This chapter is organized as follows: Section 5.2 outlines the research design. Section 5.3 presents results, and Section 5.4 discusses the findings and concludes the chapter.

[100] We were unable to find papers that have used an experimental approach to examine the effects of signalling child care support in countries from the Global North. There are related papers, for instance, Aranda and Glick (2014) that have (experimentally) studied the effect of signalling strong labor market commitment - devotion to work as opposed to family, in the case of mothers in the Spanish labor market.

5.2 Experimental Design

This section outlines the experimental design. Our primary aim was to examine the effects of motherhood and childcare support on callback rates. We randomly reported both treatments in fictitious CVs allowing for *within* applicant as well as *within* job postings variation in an Indian city and two industry sectors. We then recorded if each application received a response from employers. Data collection took place between February and April of 2019.

Hypotheses

We test two hypotheses: First, mothers are less likely to receive callbacks as compared to non-mothers (H1). Second, a childcare support signal reduces the mother penalty (H2).

Overview of the Experiment

We collected several anonymous CVs of real job applicants from a human resource consultancy firm. Based on these CVs we composed 9 fictitious CVs. To reduce the possibility that a potential employer may know the schools, colleges or firms where people in the fictitious CVs had studied or worked, all the applicants were from the East Indian city of Kolkata applying to jobs in the North Indian city of Delhi. All applicants had the same number of years of education and received degrees from qualitatively similar colleges and high schools in similar academic streams – Political Science, Sociology and History. All CVs had a residential address in Delhi and each applicant had a distinctive phone number and email contact to record callbacks. All the applicants signaled that they were legally married and reported an age between 25 and 27. Two-thirds (300) of the applications indicated the presence of a young child (mother) while the remainder did not indicate the presence of young children (non-mothers).

We looked for and applied to jobs through the most popular Indian job website (*naukri.com*). We restricted ourselves to low- to medium skilled service sector jobs. Building on our previous work, we chose jobs in two broad sectors which feature dynamic job markets: (i) Business Process Outsourcing (BPO)/ Call Center, and (ii) Banking/Finance jobs.[101]

[101] The service sector contributes around 61.5% of the India's total GDP, and IT/ Fintech are amongst the fastest growing sectors providing over $ 155 billion in gross value addition. (see https://www.investindia.gov.in/team-india-blogs/service-sector-india-paradigm-shift) [Accessed December 22, 2019].

BPO/Call Center jobs demand greater flexibility compared to the finance sector; night shifts are often required in BPO/ Call Center jobs but not in the finance sector.

Childcare Support Treatment

We signaled the availability of childcare support at home by adding a section on skills and competencies in all CVs. In half (150) of the mother CVs, we mentioned 'flexible **(strong childcare support at home)**' as one of the skills and competencies. 'Flexibility' was differently framed in the three CV formats. For instance, in one CV we mentioned 'willingness to work in different shifts' to signal such flexibility. In the other two CVs we mentioned 'flexible and ready to adapt to corporate needs' and 'flexible in working hours'. We allowed within (CV) profile variation of applicant identities. To do so, we first composed CVs of three applicants with unique identities. Subsequently, each of these three CVs had three versions/profiles: non-mother, mother and mother with childcare support. By doing so, we had nine CVs in total. In total, there were six potential combinations of these three applicant profiles. We randomly assigned one combination to each job opening/firm.

In sum, every job opening received three applications from three different applicants with different CV formats/profiles that were randomly assigned. By doing so, our design enables us to randomly vary profiles and to rule out the possibility that our results are due to specific CV formats or applicant profiles (e.g. name, address).

In total, we sent 450 applications to 150 job openings/firms.[102] These 450 applications were equally divided between two sectors. Table 5.1 reports sample sizes by treatments and sectors.

[102] We targeted about 150 jobs over the fieldwork period. Given our previous work, we are comfortably powered for a small to medium effect.

Table 5.1: Sample Size

	Non-mother	Mother	Mother+Childcare
Total Job Applications	150	150	150
By Sector			
BPO/ Call Center	75	75	75
Banking/ Finance	75	75	75

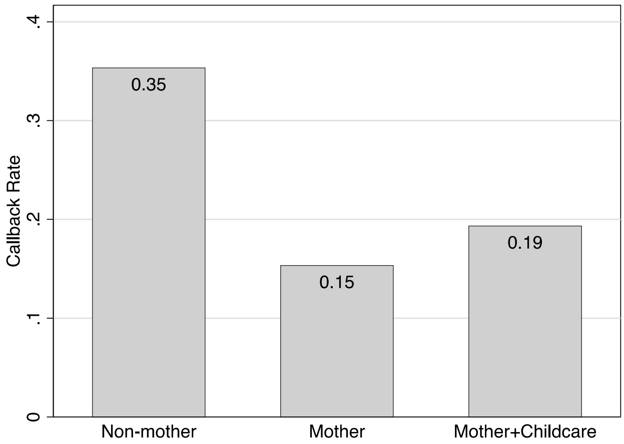
5.3 Results

We begin by presenting our results graphically, and thereafter provide impacts of reporting motherhood and signaling childcare based on estimating linear probability regression models specified as,

$$y_{ipj} = \alpha + \beta_1 Non\_mother_{ipj} + \beta_2 Childcare_{ipj} + \delta_p + \theta_j + \epsilon_{ipj}, \qquad (5.1)$$

where, y_{ipj} is a binary callback variable for applicant i, profile p, sent to job opening j. and Non_mother_{ipj} and $Childcare_{ipj}$ are binary variables indicating whether an applicant is a mother or not and whether an applicant has access to child care or not, respectively. Some of the specifications include profile (δ_p) and job opening (θ_j) fixed effects for robustness checks, while ϵ_{ipj} is an error term.

Figure 5.1: Callback Rates



Δ p-value (N=450)

Non-mother vs. Mother: 0.00

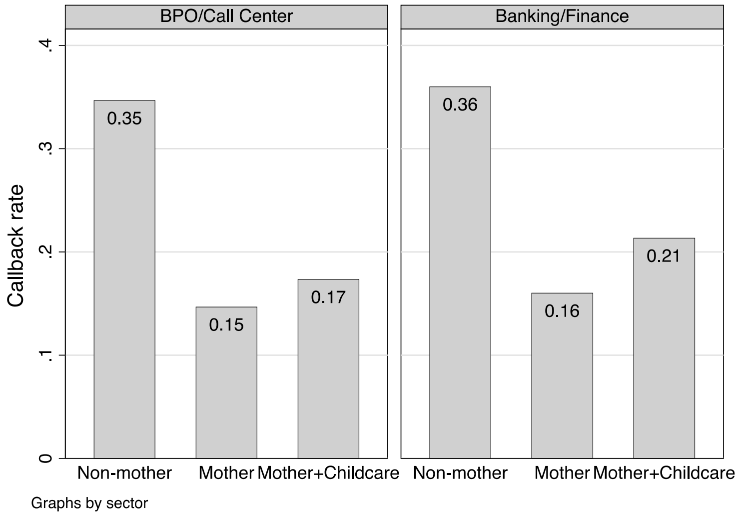
Mother vs. Mother + Childcare: 0.01

Non-mother vs. Mother + Childcare: 0.00

Note: P-values stem from linear regression-based t-tests adjusted for clustering at the job opening level.

Overall, 23.3% of applications received a callback. Figure 5.1 breaks down callbacks by treatments. We document a sizeable motherhood penalty. The non-mother callback rate was 35% as compared to a mother callback rate of 15%. This amounts to a decrease of more than a half or 20 percentage points (p-value = 0.00). Consistent with the second hypothesis, mothers signaling childcare support experience a lower penalty (at 16 percentage points). The treatment reduced the baseline motherhood penalty by 20% or 4 percentage points (p-value = 0.01).

Figure 5.2: Callback rates by sector



BPO/ Call Center
Δ p-value (N=225)
Non-mother vs. Mother: 0.00
Mother vs. Mother + Childcare: 0.16
Non-mother vs. Mother + Childcare: 0.00

Banking/ Finance
Δ p-value (N=225)
Non-mother vs. Mother: 0.00
Mother vs. Mother + Childcare: 0.05
Non-mother vs. Mother + Childcare: 0.00

Note: P-values stem from linear regression-based t-tests adjusted for clustering at the job opening level.

With regard to callback rates by sector (see Figure 5.2), non-mother applicants recorded callback rates of 35% and 36% in BPO/ Call Center and Banking/ Finance (compare Panel A and B), respectively. The motherhood penalty was also similar for both sectors – with mothers experiencing a 20 percentage point reduction in both sectors. While there are no differences in baseline callback rates or in the motherhood penalty across sectors, it seems that the childcare support signal is qualitatively more effective in the Banking/ Finance sector. Childcare support eases the motherhood penalty by 5 percentage points in the Banking/ Finance sector, compared to a reduction of 2 percentage points in the BPO/ Call Center sector.

Next, we provide results from linear probability models (1) where the dependent variable indicates whether an individual received a callback (=1) or not (=0). Table 5.2 summarizes our estimates. Consistent with Figure 6.1, we document a sharp motherhood penalty in Column 1.

Table 5.2: Linear Probability Model

Dep. var. Callback	(1)	(2)	(3)	(4)
Mother is excl.				
Non-mother	0.20***	0.20***	0.20***	0.20***
	(0.03)	(0.03)	(0.03)	(0.04)
Mother+Childcare	0.04**	0.04**	0.04***	0.04**
	(0.02)	(0.02)	(0.02)	(0.02)
Sector (Finance is excl.)				
BPO/ Call Center		-0.02	-0.02	-
		(0.06)	(0.06)	
Constant	0.15***	0.16***	0.16***	0.15***
	(0.03)	(0.04)	(0.04)	(0.02)
Profile Fixed Effects	No	No	Yes	No
Job Opening Fixed Effects	No	No	No	Yes
N	450	450	450	450

Note: Every column reports a separate linear regression. Standard errors are reported in parentheses and are clustered at the job opening level. *p<0.1, **p<0.05, ***p<0.01

We find a statistically significant motherhood penalty of 20 percentage points. In addition, motherhood with childcare support yields a callback rate that is 4 percentage points (p-value = 0.01) higher than that of mothers without such support, and accordingly 16 percentage points less than that of non-mothers. This implies that the childcare signal reduces the motherhood penalty by about a fifth. Considering the very low-cost signal, this may be regarded as a substantial effect, although moderate in size compared to the overall motherhood penalty. In Column 2, we add an indicator for the sector. Our estimates are robust to this inclusion. This is perhaps not surprising because of the random assignment of treatment arms. Our findings are equally robust to the inclusion of profile and job opening/firm fixed effects (Columns 3 and 4). This implies that our results are not driven by variations in the CV profiles sent to different job openings or variations in the specific needs of the job opening but by variations in CVs sent to the same job opening.

In Table 5.3, we split the sample by sectors. Column 1 shows a motherhood penalty of 20 percentage points in the BPO/Call Center sector. The motherhood penalty for mothers with childcare support is 3 percentage points lower than for those without childcare support, although the effect is not statistically significant. We find a similar motherhood penalty in the Finance sector. In this sector, childcare support reduces the motherhood penalty by 5 percentage-points, and the effect is statistically significant.[103]

Table 5.3: Heterogeneity by Sector

Dep. var. Callback	(1)	(2)
Mother is excl.		
Non-mother	0.20^{***}	0.20^{***}
	(0.05)	(0.05)
Mother+Childcare	0.03	0.05^{**}
	(0.02)	(0.03)
Constant	0.15^{***}	0.16^{***}
	(0.04)	(0.04)
Sector	BPO/ Call Center	Banking/ Finance
N	225	225

Note: Every column reports a separate linear regression. Standard errors are reported in parentheses and are clustered at the job opening level. *p<0.1, **p<0.05, ***p<0.01.

5.4 Discussion and Conclusion

We investigated how employers value childcare support for job applicants who are mothers as compared to mothers without explicit mention of childcare support and compared to non-mothers who are otherwise similar. Using a CV audit study, we examined the impact of reporting childcare support at home on the probability of receiving a phone call in response to a job application in an Indian city and two industry sectors. Our study has two main findings. First, similar to our previous study, we document a strong motherhood penalty, and second, while childcare support reduces the motherhood penalty, it does so only partially. Whilst the evidence clearly indicates that the lack of childcare support is one of the channels through which the motherhood penalty manifests itself, in the context of Delhi, it has a modest effect.

[103] Differences in the mother plus childcare coefficient are not statistically different across sectors (p-value = 0.41).

Although, the aim of the chapter is not to test theories of labor market discrimination *per se*, we comment on the link between our results and potential sources of discrimination. First, consistent with Benard and Correll's (2010) conception of 'normative discrimination,' we find that displaying access to child care as a signal of a mother's labor market attachment does not eliminate the motherhood penalty, although there is a reduction. Second, while access to reliable child care may be expected to reduce the motherhood penalty, it remains an imperfect substitute. It seems to be the case that employers are only partially convinced about the flexibility and commitment rendered by access to child care and the presence of both, statistical discrimination (Phelps, 1972) emanating from deep-rooted gender norms - other than childcare responsibilities, and taste-based discrimination (Becker, 1971) against mothers, may not be ruled out. This chapter probed the effect of access to child care in mitigating the motherhood penalty. It is possible that a stronger signal, which focuses not only on access to child care but which shows that a mother is willing to sacrifice family for work will be more successful at crowding out statistical discrimination (see Aranda and Glick, 2014). For instance, in their paper on the Spanish labor market, where business students were asked to pose as human resource managers and provide job recommendations, Aranda and Glick (2014) found that self-reported personal priorities statements which signaled a mother's strong commitment and devotion to work as opposed to family completely crowded out the motherhood penalty.

While our study is able to show the mitigating effects of access to childcare at the point of labor market entry, it provides insights for just one (large) city and two (important) sectors. Furthermore, it was not possible to indicate the type of childcare support in a CV or to have a discussion on the childcare provider at hand which would have been possible in an interview setting. It is possible that employers hiring decisions are sensitive to the type of child support on hand – that is, whether it is provided by a professional childcare worker/crèche with fixed timings as opposed to care provided by a grandparent or another family member. To investigate such differences, one possibility maybe to employ lab-based CV audit studies and inform subjects about the nature of childcare providers. Despite its shortcomings and potential avenues for further research, this chapter shows that a low-cost signal - one-line in a CV - is able to play a role in influencing the decisions of employers in favor of mothers.

Chapter 6

Gray University Degrees: Experimental Evidence from India

(with Matthias Rieger)[104]

6.1 Introduction

It is an open secret that academic degrees can be bought in India. Local media have widely reported on this phenomenon with headlines ranging from "Degrees on sale: Jaipur study centers offer bachelor degree to PhD for money" (Kumar et al. 2011) to "Fake degree scam: No sweat, you can get a university degree in 10 days" (Ullas and Prasher 2013), as well as "PhDs, Bachelor's degrees on sale in Punjab" (Chowdhary 2011).

How are degrees bought? To answer this question, we collected qualitative data through a local market review, as well as through interviews with agents and potential buyers in the state of West Bengal. Agents and intermediaries handle (parts of) the process. They advertise their services in local newspapers, the Internet, flyers and railway coaches, often using ambiguous language due the illegality of such services (see one such example in Figure 6.1). Most of the advertised degrees originate from privately funded universities in other states and often by means of distance education.[105] Students can obtain degrees within as little as two months without even sitting in exams. However, the market and the offered service packages are very diverse. For instance, one interviewed agent summarized the service as follows:

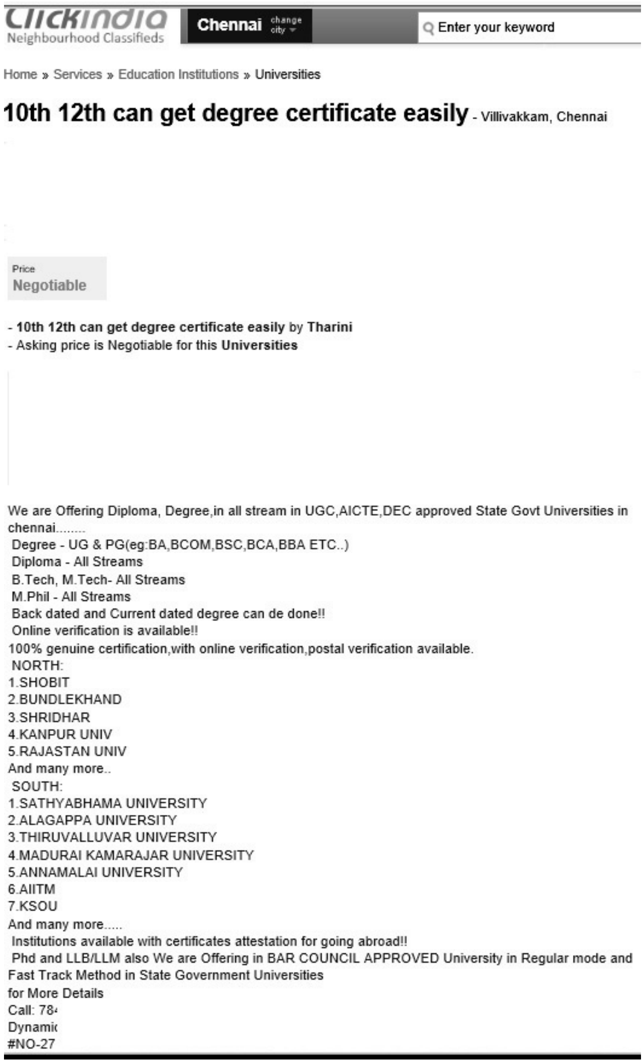
> *"We manage her [the student's] signature much before the exam on the answer sheet […]. During exam days she can send anybody to sit in the exam. The only*

[104] This chapter has been published as Majilla, T., & Rieger, M. (2020). Gray University Degrees: Experimental Evidence from India. *Education Finance and Policy*, 15(2), 292-309.

[105] Most Indian states feature at least one state-funded open and distance learning university. These are 'open' in the sense that admissions are not selective. The largest and federally funded example of such an institution is the Indira Gandhi National Open University (IGNOU) with more than 4 million students. Note that the Indian distance education system is not an 'online' system. Rather students take part-time degrees and are not required to attend lectures. They receive self-study materials in hard copy. They may attend optional lectures, in some cases complete take-home assignments and finally take exams (for more details, see Gupta and Gupta 2012; Gaba and Li 2015).

requirement is that the person taking the exam needs to be female if the original student is female. She may wish to write something or not. We manage a certificate." [translated from Bengali by the authors, interview dated 27 July 2016]

Figure 6.1: Example of an online advertisement to buy degrees



Note**:** This online advertisement is presented as an example and was not used in the experiment. Contact details were blanked. The original is available at: http://www.clickindia.com/detail.php?id=136428806 [Accessed 14 July 2017]. Here 10^{th} is referred to 10^{th} standard, i.e., the board exam after 10 year of studies whereas 12^{th} is referred to 12^{th} standard, i.e., the next board exam after 12 years of study. Student typically sit in the 10^{th} and 12^{th} standard exams at the age of 16 and 18 respectively. The categorization of universities is geographical, clustered in North and South India.

In this chapter, we focus on the most straightforward rationale for buying a degree, namely boosting job market success as proxied by callback rates to applications. Employers cannot perfectly differentiate between potentially bought and authentic or more credible credentials. However, they are likely to discount degrees from institutions that offer both authentic and bought degrees, as they cannot tell if a student has gained human capital. This chapter tries to quantify the extent to which this is the case. To the best of our knowledge, there has been no quantitative/experimental work on gray degrees in India or other developing countries. There is also no reliable data on the functioning, size and extent of the market other than anecdotal evidence. More broadly, we contribute to the literature on the value of different type of college degrees on the labor market: Notably, Darolia et al. (2015) and Deming et al. (2016) performed large correspondence studies to examine the value of for-profit college degrees in the US. Darolia et al. (2015) found no effect of such degrees relative to community college degrees or even to having no college degree at all. In contrast, Deming et al. (2016) document a negative effect for some online degrees compared to nonselective public institutions.

We focus on questionable institutions where it is possible to buy degrees. These universities also issue valid and legal degrees. In what follows, we therefore refer to these as *gray* degrees. Informed by a simple conceptual model, we examine if employers discount such questionable degrees in applications to low skilled, entry-level jobs advertised on online platforms.[106] Preparatory qualitative work informed the design of the experiment, which took place in the state of West Bengal. We first identified several universities from which gray degrees could easily be bought and we picked three of these. These universities tend to be distant from the local job market (West Bengal) and are clearly questionable; for instance, a simple Google search would reveal that at least one of them was involved in a degree scam. We also searched for comparably low ranked local (control) universities that issue strictly authentic degrees (details on the choice of institutions are given below). We then performed a resume study similar in procedure to those found in the broader labor economics literature (Bertrand and Mullainathan 2004; Carlsson and Rooth 2007; Correll et al. 2007; Pager et al. 2009; Kaas and Manger 2012). We picked job advertisements in sectors that required neither specific skill training nor work experience or academic degrees. This is the relevant segment of the job market for our analysis, allowing us to test the impact of gray degrees on callback rates compared to *both* having no degree and authentic degrees (from universities that clearly

[106] In our local context, firms and applicants tend to use the term CV, which in the US context corresponds to a resume. We use the term resume throughout.

do not issue gray degrees). In other words, we can examine how gray degrees boost employment chances compared to no degrees, and if they are discounted by employers.[107] In practice, we sent three resumes to 132 identified job openings, varying the type of degree. To assess the influence of gender, we sent female and male resumes to mixed sector jobs, and only female (male) resumes to female (male) dominated sector jobs. We then recorded callback rates for interviews.

To preview our results: Resumes featuring a gray degree average 8%-points more callbacks than those with no degree. This amounts to a 42% increase in the number of callbacks. This difference is significant only for and driven by female applicants. These gendered patterns may be specific to the sampled jobs (and gendered industry differences) as well as the Indian context. That said, at least qualitatively we see similar patterns for men and women with an advantage of gray degrees compared to no degrees. Authentic degrees always fare better than no degrees (in particular among female applicants) and on average than gray ones. Note that these main findings are robust to controlling for sector, firm and resume/profile fixed effects. In a heterogeneity analysis, we also classified gray degrees into low, medium and high ranked degrees based on the perceived ease of acquiring these degrees according to agents and buyers in qualitative interviews. The lowest rank university has been involved in scams as discussed on the Internet. We see a corresponding empirical ranking of gray degrees. Reporting a low ranked degree is statistically the same as reporting no degree at all. Conversely, there are no statistical differences in callback rates between authentic and high ranked gray degree universities. Our results imply that employers have consistently formed beliefs. They appropriately discount questionable degrees as a function of degree rank.

We rationalize these results in three ways: First, a degree from a questionable university is better than no degree at all. This finding is compatible with a simple model (outlined in Section 6.2) where employers cannot verify the authenticity of a degree from a questionable university but at least assign some corresponding probability. Second, authentic degrees are clearly preferred by employers compared to no or gray degrees, as indicated by the highest rate of callbacks. Taken together these findings indicate that, at least in part, the veracity of

[107] In India it is perfectly common that job seekers with university credentials apply to jobs that only require high school degrees due to the competiveness of the job market, while high skilled jobs are normally occupied by the applicants from elite and established universities. The Indian media has covered this phenomenon: "PhD holders apply for SSC's clerical posts in West Bengal" (*Times of India*, 20 September 2016). Available at: http://timesofindia.indiatimes.com/city/kolkata/PhD-holders-apply-for-SSCs-clerical-posts-in-West-Bengal/articleshow/54432085.cms [Accessed 13 July 2017]

degrees is judged well by employers. Third, we can understand our findings in the light of the literature on fake degrees (Brown 2006; Grolleau et al. 2008; Attewell and Domina 2011). Attewell and Domina (2011) argue that "those who are blocked from attaining degrees through normal means are those most likely to employ false credentials" (pp.59).[108] This is backed by what one potential buyer of a gray degree stated during our qualitative interviews:

> "*I need to take care of my baby, cook, collect water and take care of my parents-in-law... I simply do not have time to study. The duty of a married woman is to take care of her household. If I start studying, who will take care of the household?" [translated from Bengali, interview dated 18 July 2016]*

The remainder of this chapter is structured as follows: Section 6.2 gives a simple conceptual model, details some qualitative insights and describes the experiment. Section 6.3 presents the results. Section 6.4 discusses the findings, while section 6.5 concludes with some policy implications.

6.2 Empirical Strategy

In this section, we first briefly outline a conceptual model to motivate the experiment and to provide some predictions. Then we outline the experimental design as informed by qualitative information.

Conceptual Model

Applicants report degrees or no degrees on their resumes.[109] Universities denoted as i vary in their quality Q_i. A fraction ($1\text{-}x_i$) of reported degrees is "bought." This fraction captures that it is possible to both buy and legitimately earn a degree from a gray university. Employers evaluate resumes and then decide whether to interview the candidate at unit cost c. They hold beliefs about the fraction or prevalence of earned degrees x_i for each institution. The unit quality of the degree Q_i positively correlates with the quality of the candidate. The lower the fraction of bought degrees ($1\text{-}x_i$) and the higher Q_i, the higher the likelihood of a callback. In

[108] Degrees also convey "a certain prestige or social status" which may motivate buyers (see Groulleau et al. 2008, pp.680).

[109] We thank an anonymous referee for suggesting this model.

sum, callback probability or expected productivity minus interview costs can be written as $y_i = x_i Q_i - c_i$.

In this simple model, differences in callback rates stem from employers' perceived differences in x_i. Say, the employer decides between two candidates from universities j and k, so she compares $x_j Q_j - c_j$ and $x_k Q_k - c_k$. If we assume equal institutional quality $Q_j = Q_k$ and interview costs $c_j = c_k$; further, if university k is associated with a lower prevalence of bought degrees so that $(1\text{-}x_j) > (1\text{-}x_k)$, then university k will be preferred over j. As long as the fraction of bought degrees is less than one, then some degree will be better than none. To sum up, the results from our experiment below are interpretable in terms of beliefs. Employers assign a probability that a degree from a gray institution is "bought" rather than earned. As a result, authentic degrees should receive most callback rates, followed by gray and then no degrees.

Overview of the Experiment

Our aim was to test the impact of those universities from where it is possible to buy gray degrees on callback rates to job applications relative to authentic degree providing colleges and having no degree at all. Based on insights from qualitative data collection efforts, we ran a resume experiment from July to September 2016 in West Bengal. We sent 396 applications to 132 job postings. We applied to each of the job postings using three generic male or female applicant profiles. For each job posting, the three profiles were arbitrarily distributed to the following three groups: (1) no university degree; (2) gray university degree; (3) authentic university degree. Our minimum target sample was at least 188 applications.[110] We budgeted fieldwork time to achieve this target and in practice we exceeded it.

Background

The design of the experiment required input from qualitative interviews, since there is little information on the functioning of the gray degree market. The qualitative data were collected in two steps by the first author of the study in the state of West Bengal: First, we reviewed the local gray degree market and looked for flyers to identify universities from which degrees could be bought. Second, we did qualitative interviews using snowball sampling with sellers and agents in the gray degree market, as well as with prospective buyers (for details on the

[110] We carried out power calculations for a test of two proportions using the "pwr" package in R. Setting Cohen's h for a binary outcome to 0.5 (medium effect), power to 80% and significance level to 5%, the required sample size was 188.

ethnographic methodology and results, see Majilla 2016). More specifically, we gathered advertisements for suspect degrees and approached agents pretending to be prospective buyers. We also interviewed actual and potential buyers and accompanied them to the agents. In total, five interviews were done with agents and 23 with potential and actual buyers.

How does this gray degree market function in practice in our study area? A university may have a distance education department, which is legally entitled to provide degrees under the name of the university. A university in a given state (private universities are always state universities, i.e., they are established and governed by laws of the state government) can have study centers across the state. Federally funded universities can also have study centers in other states. However, we did not find any federally funded university providing gray degrees. Study centers have links with agents/offices located in other states, most often located in state capitals. These offices have no legal authority to provide services to students. Recall, only proper study centers within the state have such legal authority. Study centers and these offices located in other state capitals cooperate with agents at the district level, and these in turn may work with agents at the block level.[111] Official study centers are meant to provide learning support and also administer exams in the presence of invigilators from the central University. Yet, in reality, we found that buyers of gray degrees never actually seem to visit these real and legal study centers.

We found that in most cases students contact local agents, who forward applications and requests to their contacts at a higher level within the state. Such agents are not common for authentic colleges and students in India. Of course, legitimate and official educational consultants do exist, and their primary customers are students looking for engineering degrees in other states and students looking for foreign degrees. These consultants are merely advising students, but they rarely have any influence on the admission process. Consultants for arts or social science degrees are less common. The Indian education system is very much prestige driven and students try to enter the most prestigious colleges they can within their reach and budget.

We found that gray degrees are "earned" through aforementioned informal agents. They not only guide students, they manage the entire "study" process and have substantial influence on

[111] In an administrative sense, a state consists of several districts, which consist of several blocks.

the university administration. They literally sell their ability to arrange gray degrees. In India, most undergraduate arts and social science degrees take three years and students sit in exams, either in semester mode or on a yearly basis after their coursework. Gray degrees in principle require coursework but if they are channeled through the distance education mode, then actual attendance in coursework is optional. We also found some agents who arrange gray degrees through full-time mode. We found most students do not actually sit in exams. In fact, most degrees are acquired through this distance education system. If students do need to sit in exams (as in the case of bachelor of law degrees), they just sign the answer sheet and somebody fills in responses for them; or sometimes answers are provided to them and they copy them onto the exam sheet.

In terms of academic streams, we found arts and social science gray degrees to be most prevalent as these are easier to manipulate. These degrees are cheaper compared to science subjects. We found most students were buying gray degrees either (and equally so) at the bachelor or master level. We found that students can get undergraduate gray degrees within as little as two months. However, it also important to note that many different business models and service packages exist in this market and one can naturally only get a small glimpse of these during qualitative research.

During the qualitative work, we identified several universities from which it was possible to (indirectly) buy gray degrees. Out of these we selected three universities that were popular among local agents. Our aim is not to "name and shame" these particular three universities, so we do not disclose their identities here. Details are of course available on request from the authors for replication and research purposes. These three universities are in three different cities and in three different regions. It is important to underline that all three universities also offer authentic degrees, earned through actual academic work. However, it is possible to easily access and buy these questionable or gray degrees with the help of agents. In these cases, students do not write exams and agents take care of the entire process. More specifically, all gray degrees used in our experiment can be accessed through distance learning. The chosen degrees cost around INR 18,000-20,000 (~USD 275-300). Note that costs are roughly similar across the three ranks. We would also like to note that the three universities tend to be away from the applicant's residence and the job market. For instance, the most questionable gray degree comes from a university that is roughly 1500 km away from the home address of our job applicants.[112] This is important since the qualitative data indicate that agents tend to collaborate with distant universities. Degrees from such universities should be discounted by employers. A local applicant could obtain an authentic degree from a nearby university.

We also ranked these three questionable universities. We did this based on the perceptions of agents. The lowest ranked gray degree comes from a university that easily cooperates with agents and has been involved in scams as documented on websites. Naturally, there are no hard quality measures or measures on the prevalence of bought degrees available, so rank-specific results have to be interpreted with caution. We need to assume that the difference in callback rates only stem from the relative prevalence of and beliefs about illegitimate degrees (as suggested by our conceptual framework).

Selection of Authentic (Control) Degrees

Recall that we compare resumes featuring gray degrees to both resumes with authentic university degrees and only high school degrees. Our aim is to compare institutions that are of similar quality, so that differences in callback rates can be attributed to the existence of a market for bought degrees.

To this end, we prepared a list of colleges in the state of West Bengal offering authentic degrees, which we thought would be comparable to the identified universities offering gray degrees. We had a four step strategy to select those used in the experiment. First, we showed our list of authentic institutions and those known for gray degrees to a human resource consultancy firm and asked a representative to rank and match institutes in terms of job market prospects. Second, we repeated the same exercise with two gray degree agents operating in the local market. Third, we asked several buyers to examine our list. However, their knowledge of gray degree universities (apart from their own) turned out to be limited and we did not use their input. In the end, we had a list of five comparable authentic and five gray degree institutions and we picked three of each for the experiment.

[112] We do not have data on the number of students travelling out of their state. Our anecdotal and qualitative evidence suggests that only top students move out of states for arts or social science degrees, mainly to study at elite colleges in big cities. One exception appears to be engineering, where many students move out of their state.

Further Education Characteristics

We did not mention whether degrees were obtained through distance education or not (which is often the case for gray degrees) on resumes. Mentioning distance learning just for gray degrees may have confounded our results. However, one limitation is that we cannot test for the impact of reporting distance education on callback rates, which may differ between authentic and gray degrees, and which may provide an additional signal to the employer. Further, we had to make a choice on the academic discipline itself. Based on insights from the qualitative interviews, we opted for two comparable and relatively less prestigious academic disciplines in the Indian context, namely history and political science. Finally, all resumes featured comparable secondary and high school degrees located near the applicants' residential address.

Selection of Job Postings and Location

We used three popular online portals in India to search for jobs.[113] These portals allow applicants to create a profile, upload a resume and apply to job postings. We made sure not to apply to the same job twice across portals.

We sent applications to entry level and relatively low-skilled jobs. We only selected job postings that were open to both inexperienced and experienced applicants. As these jobs did not require a university degree,[114] we can test if gray degrees provide a competitive advantage over having no degree at all. More specifically, we picked job postings ranging from sales, in particular insurance or personal loan sales to administrative support, clerical support, call centers as well as medical representatives. We did not design the experiment to be able to detect impacts by specific sectors. Rather, we simply selected and classified jobs into female (N=99), male (N=99) and mixed sector jobs (N=198). We gendered the resumes of applicants according to this threefold classification. This allows us to examine patterns by the gender of the applicant and the gender of the sector. For instance, administrative jobs are typically held by females, while medical sales representatives are mainly male. Conversely, call centers are known to have a mixed work force. One limitation of our design and the

[113] The large majority of jobs were chosen and applied to through one of these three platforms. So we cannot break down results by job portal.

[114] Unfortunately, we could not find hard data on the percentage of employees without university credentials in the types of jobs included in our experiment.

small sample is that we cannot clearly disentangle gender differences (if any) from job and industry differences.

We exclusively applied to jobs in the state of West Bengal (in east India), where also the qualitative fieldwork took place. We mainly applied to jobs based in Kolkata, and in some cases to jobs in district towns within the state of West Bengal. Related, we excluded job postings that would have required applicants to relocate. Our experiment closely mirrors what we found in the qualitative interviews with agents and buyers. Applicants in the local job market can either obtain an authentic local degree or buy degrees outside of their home state (in this case, West Bengal). The three gray universities were in three different regions (East-central, West and South India) and located in mid to small-sized cities. The limitation of our design is that we cannot separate out the effect of geography from the overall effects of gray degrees.

Design of Applicant Profiles

We designed the resumes with three considerations in mind: First, we wanted to create realistic applicants for the chosen jobs and sectors. We obtained typical and real resumes from a human resource consultancy. These resumes had Pan-India coverage, had been used in actual applications and were aimed at jobs where no work experience was necessary. Second, we needed to pick names for applicants that suited the Indian context. Discrimination by caste, gender and ethnicities is well researched and documented in the literature (Carlsson and Rooth 2007; Pager et al. 2009). For instance, there is evidence that employers discriminate using the names of job applicants (Banerjee and Knight 1985; Deshpande 2011). To minimize these kinds of sources of statistical noise and competing drivers of callback rates, we selected upper caste Hindu names. Furthermore, in India there is also prejudice and discrimination based on the state of origin. So we used only Bengali names. Finally, in mixed-sex sectors we randomly used male and female names. In the case of female (male) dominated sectors, we only applied with female (male) resumes.

Experimental Procedure and Data Collection

In total we used eighteen resumes: nine male and nine female resumes with three different names and each with three different types of education levels: gray, authentic and no university degree (high-school only). For each job, we sent three resumes. Across jobs we varied the following features in a balanced way: (i) high school name and residential address and (ii) name and gender of applicants (for mixed gender sector jobs).

We recorded callback rates through e-mails and phone calls. Every resume had a unique e-mail and mobile phone number. We did not apply to jobs requiring applicants to directly call or visit the company. Further we used authentic residential addresses for all resumes; however for practical reasons and due to the fictitiousness of applicants we could not record responses by post.[115] An applicant's home address may come with labor market discrimination. Bertrand and Mullainathan (2004) found that applicants from dominantly black neighborhoods in the US receive relatively less callbacks than those from white neighborhoods. To minimize such effects, we used residential addresses located in semi-urban areas in West Bengal that are mainly inhabited by Hindus. Sample sizes by degree and gender are summarized in Table 6.1.

Table 6.1: Sample sizes by university degree and gender

University degree	No	Gray	Authentic	Total
Male applicant	66	66	66	198
Female applicant	66	66	66	198
Total applications	132	132	132	396

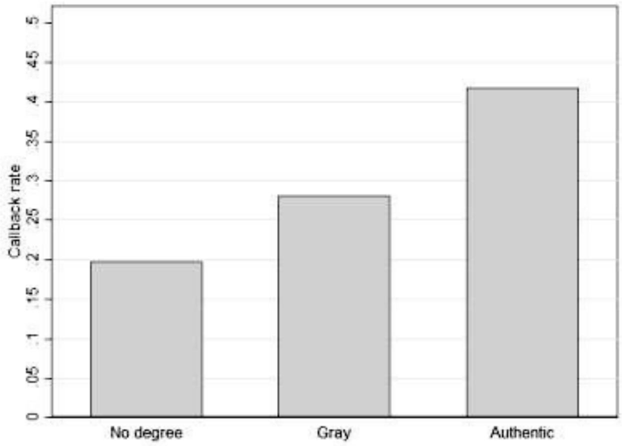
6.3 Results

Gray degrees significantly increase callback rates compared to having no degree. The impact is moderate in size and stronger among female sector jobs and applicants. In addition, authentic degrees have a much larger, relative impact on callback rates. However, we can document heterogeneous impacts of gray degrees according to their ranking (as a function of the subjective ease of obtaining them). In particular, the positive impact of the high ranked gray degree is not statistically different from the one associated with authentic degrees, while the effect associated with the lowest rank one is insignificant. In what follows, we first present the main findings graphically, and then investigate the robustness and heterogeneity of the effects in a regression model.

[115] In any case, it is unusual for employers to use regular postal services for jobs advertised on online platforms.

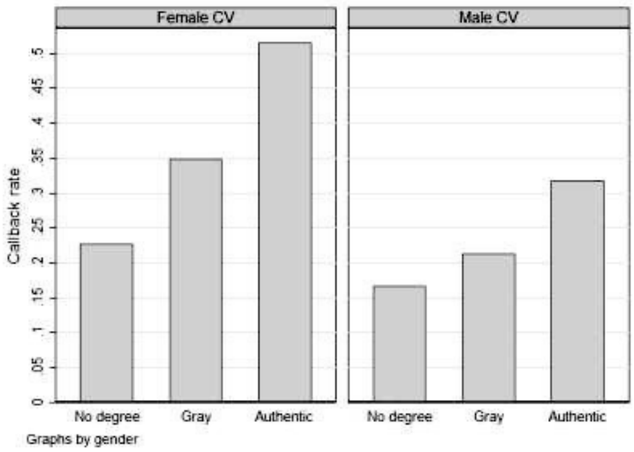
Figure 6.2: Callback rates by type of university degree

Panel A: Overall callback rates



No vs. Gray degree: p-value=0.07
Gray vs. Authentic degree: p-value=0.00
No vs. Authentic degree: p-value=0.00

Panel B: Callback rates by gender



No vs. Gray degree: p-value=0.09; p-value=0.45
Gray vs. Authentic degree: p-value=0.01; p-value=0.07
No vs. Authentic degree: p-value=0.00; p-value=0.01

Note**:** N=396 (see Table 1 for a breakdown by degree and gender). P-values stem from pairwise regression-based t-tests adjusted for clustering at the job posting level. *In Panel B*, the first p-value refers to differences in the female and the second p-value to differences in the male sample.

Overall, 30% of our applications received a callback. Figure 6.2 breaks down callback rates by the type of degree. Panel A show a clear hierarchy: resumes with no degree received callbacks in 19.70% of all applications, while those with a gray and authentic degree averaged 28.03% and 41.67%, respectively. The impact of a gray degree on callback rates compared to no degree amounts to 8.33%-points (p-value=0.07, t-statistic=1.82, N=264).[116] This amounts to a 42% increase. In comparison, the impact of authentic degrees over no degrees is 21.97%-points (p-value=0.00, t-statistic=4.90, N=264). Finally, the difference between gray and authentic degrees amounts to -13.64%-points (p-value=0.00, t-statistic=-3.29, N=264).

We can also differentiate callback rates by degree and gender of the applicant (see Panel B, Figure 6.2). One limitation is that gender differences may in part be driven by sectoral differences. Further, our experiment was not designed to test for gender differences efficiently. That said, a similar ranking of degrees emerges: resumes with authentic degrees receive more callbacks than those with gray or no degrees. However, differences are stronger among female applicants. Overall, female receive more callbacks than males (36.36% vs. 23.23%). And the difference in callback rates between female resumes with gray and no degrees amounts to 12.12%-points (p-value=0.09, t-statistic=1.72, N=132). The corresponding difference for male resumes is small in magnitude (4.55%-points) and insignificant (p-value=0.45, t-statistic=0.77, N=132).

[116] Throughout we present linear regression-based differences in means tests adjusted for clustering at the job posting level.

Table 6.2: Regression estimates of the effects of types of university degrees on callback rates

Dep. var. Callback	(1)	(2)	(3)	(4)
Degree type:				
(no degree excl. category)				
Gray	0.08*	0.08*	0.08*	
	(0.05)	(0.05)	(0.05)	
Gray (low rank)				0.04
				(0.08)
Gray (medium rank)				0.06
				(0.08)
Gray (high rank)				0.15*
				(0.08)
Authentic	0.22***	0.22***	0.22***	0.22***
	(0.04)	(0.04)	(0.04)	(0.05)
Male applicant		-0.10		-0.09
		(0.10)		(0.10)
Male sector		-0.06		-0.06
		(0.08)		(0.08)
Female sector		-0.04		-0.04
		(0.09)		(0.09)
Constant	0.20***	0.28***	0.20***	0.28***
	(0.03)	(0.07)	(0.03)	(0.07)
Effect equality (p-value)				
Gray=Authentic	0.00	0.00	0.00	0.01;0.04;0.41
Profile dummies		x		x
Job posting fixed effects			x	
N	396	396	396	396

Note: Linear probability models. Standard errors in brackets under point estimates are clustered at the job posting level. In column 4, we report tests of the equality between effects associated with the three gray degrees (low, medium, high) and the authentic degree. Mixed sector is the excluded category in columns 2 and 4. Significance levels are denoted *p<0.1, **p<0.05, ***p<0.01.

Table 6.2 summarizes estimates stemming from a linear probability model[117] where the dependent variable takes on one in the case of a callback and zero otherwise. Standard errors are clustered at the job posting level. Column 1 presents a stripped-down model with no covariates. “No degree” is the excluded category. The impact associated with gray degrees relative to no degree amounts to 8%-points. The corresponding impact of authentic degrees is 22%-points. Column 2 gauges the sensitivity of these marginal effects to the inclusion of a

[117] Probit or logit models yield similar results and are available on request.

gender dummy, as well as sector and applicant profile/resume[118] dummies. Point estimates associated with the main variables of interest are extremely stable, confirming that randomization and a balanced sample were achieved in practice. The point estimates associated with gender and the type of sector are insignificant. Column 3 shows that our main findings are also robust to the inclusion of firm/job opening fixed effects. Point estimates and standard errors are stable. Findings are robust within and across firms.

Until now, we have estimated average effects of gray degrees on callback rates. However, in the experiment, we employed three different (subjective) ranks of gray degrees. Column 4 shows the results of this heterogeneity analysis. The effects associated with these degrees are all positive. The effect of low and medium ranked degrees is however insignificant and small. Only the effect of the high ranked gray degree is large and significant (at the 10% level). What is more, we find that authentic degrees have lost some of their competitive edge over gray degrees. We can no longer statistically differentiate between the effects of authentic vs. high-ranked gray degrees (see p-values at the bottom of the table).[119] In sum, callback rates are monotonically increasing with gray degree rank. This is suggestive evidence of proper market discounting by firms with respect to the ease with which gray degrees can be bought.

[118] These dummies account for effects of the 6 specific resumes (beyond the effects of gender and degree type).

[119] Also, the difference between authentic and medium ranked degrees in terms of callback rates is insignificant. However the *economic* difference is sizeable. Therefore we might be lacking power to test this difference.

Table 6.3: Gender heterogeneity

Dep. var. Callback	(1)	(2)	(3)	(4)	(5)	(6)
Applicants		Female			Male	
Degree type:						
(no degree excl. category)						
Gray	0.12*	0.12	0.12	0.05	0.06	0.03
	(0.07)	(0.11)	(0.10)	(0.06)	(0.08)	(0.09)
Authentic	0.29***	0.27**	0.30***	0.15***	0.18**	0.12
	(0.07)	(0.10)	(0.10)	(0.05)	(0.08)	(0.07)
Constant	0.23***	0.21***	0.24***	0.17***	0.12**	0.21***
	(0.05)	(0.07)	(0.08)	(0.05)	(0.06)	(0.07)
Effect equality (p-value)						
Gray=Authentic	0.01	0.14	0.01	0.07	0.16	0.26
Sector(s)	Female,Mixed	Female	Mixed	Male,Mixed	Male	Mixed
N	198	99	99	198	99	99

Note**:** Linear probability models. Standard errors in brackets under point estimates are clustered at the job posting level. Significance levels are denoted *p<0.1, **p<0.05, ***p<0.01.

Table 6.3 shows separate regression results for female and male applicants. In columns 1 and 4, we show full sample results for female and male applicants, respectively. In the remaining columns, we provide gender-specific results by industry sectors. The effects of gray degrees within gender are similar across the various sub-samples. However, we lose precision in the smaller sectoral samples.[120] We find that the effect of gray degrees on female callback chances is statistically significant only (at the 10% level) in column 1, yet always economically important and larger among female applicants (12%-points in columns 1-3 compared to 3-6%-points in columns 4-6). Finally, note that authentic degrees fare (at least qualitatively so) better than gray degrees and no degrees. It is also interesting to note that the effect of authentic degrees is large among female applicants.

6.4 Discussion

Applicants with degrees from authentic colleges fare significantly better compared to both having no degree and degrees from questionable universities in our simple audit study. However, questionable degrees from gray universities can partially compensate the lack of

[120] We did not design the experiment to be able to document precise gender differences. We cannot efficiently disentangle gender and gendered-industry differences.

authentic credentials; and this finding is concentrated among female applicants and female-sector jobs. These gender differences may be explained in that callback rates are higher for women in the first place and that women in India tend to have lower education levels than men and also tend to apply to lower skilled jobs. A related explanation for the gender differences in callback rates may stem from the fact that employers understand the difficulty of Indian women to accumulate human capital. Consequently, they reward any degree (gray or authentic) obtained by women relatively more.

We rationalize our results in the sense that employers hold beliefs about the probability that a degree from a gray institution is bought rather than earned. It is clear that employers discount degrees from questionable universities. Table 6.4 breaks down callback rates at the firm level: 13% of firms called back applicants with degrees from authentic and gray universities and ignored applicants with no degree (see column 1). At the same time, 14% of firms only called back applicants with degrees from authentic universities. Conversely, 5% of firms called back applicants with authentic and no degrees, ignoring those with questionable degrees. These patterns are stronger for female resumes (compare columns 2 and 3). Our findings are thus consistent with a simple conceptual model (see section 2) and suggest that most employers lump gray degrees within a broad degree category. Simply put, employers prefer a degree (gray/questionable or authentic) over no degree at all. They cannot perfectly verify the authenticity of a degree from a questionable university but assign a positive probability (however small that may be) that the degree is nevertheless authentic.

Table 6.4: Callback rates at the job posting/firm level

	(1)	(2)	(3)
		Sample	
Firm-level response	Total	Male	Female
No applicant	0.48	0.59	0.38
Only no degree	0.05	0.03	0.06
Only gray degree	0.05	0.06	0.05
Only authentic	0.14	0.11	0.17
Only no degree and gray	0.00	0.00	0.00
Only no degree and authentic	0.05	0.06	0.05
Only gray and authentic	0.13	0.08	0.18
All applicants and degrees	0.10	0.08	0.12
Number of jobs	132	66	66

Note**:** This table shows the distribution of callback rates *within* job postings.

Our qualitative insights complement our experimental findings. One agent summarized the economics of gray bachelor of law degrees (LLB):

> *"Let's talk about an LLB. You cannot do this on a part time basis. But honestly, tell me, is it worth spending three years doing an LLB unless you go to a good university? We are providing all LLBs at [Rupees] 60,000…If you pay, then you'll find that your answer script is ready and you just need to sign." [translated from Bengali, interview dated 12 July, 2016]*

Do the benefits of gray degrees really exceed the costs? While our experiment suggests that gray degrees can increase callback rates by 8%-points, this does not inform about costs and benefits. That said, the gray degrees in our analysis cost about INR 20,000 (~USD 300), which can be around a fifth of a yearly salary[121] in an entry-level job. But costs vary depending on the degree level, subjects and the extent to which the degree is managed by agents. For example, a LLB degree costs around INR 60,000 (~USD 900) compared to INR 18,000 for a BA in History (~USD 270). The costs of such degrees are actually quite steep compared to the positive impacts on callback rates. Further research should examine the motivations of buyers beyond callback rates, including longer-term economic advantages and social motives.

We would like to point out some immediate limitations of our study and sample: first, we only studied the impact on callback rates. We cannot document if these callbacks translate into getting an offer and keeping a job. Related, we do not have data on wages for holders of different degrees. Second, we focused on a small set of universities from which gray degrees can be obtained and on just one Indian state. Third, we focused on jobs advertisements that do not require higher education degrees. We do expect a higher scrutiny of resumes for higher ranked jobs that require academic credentials.

Our experimental design also has several limitations: Most importantly, we cannot parcel out whether employers make inference about an applicant's productivity based on the choice of schools – both in terms of size, type, mode of instruction (online, distance etc.) and location.

[121] According to our qualitative interviews, a yearly salary in the low skilled sector could be around 8000 to 10000 INR (USD 1500 to 1880). Also note that the unskilled minimum wage in West Bengal is 5962 INR/Month (Government of West Bengal, 2017, available at: https://www.wblc.gov.in/sites/default/files/synopsys/01-07-2017/agriculture.pdf [Accessed 27 August 2017]).

For instance, gray degree institutions in our experiment tend to be relatively far away, and employer perceptions of productivity may vary depending on the distance between home and institution. Furthermore, we find relatively strong effect of gray degrees among women. However, we do not distinguish between mothers and non-mothers, which is an important dimension in the Indian labor market (Das and Zumbyte, 2017). Having a gray degree (potentially involving distance education) may tell the employer that the (female) applicant has children at home and may not have had time to obtain an authentic degree (in line with the above quote from the qualitative interviews). This in turn could imply that the applicant has lower commitment to the job, and will only be available part time (Correll et al 2007). Several further limitations of our design are worth mentioning: we cannot provide more disaggregated results by study subject (for instance, history compared to political sciences), as well as job board, firm or sectoral characteristics (see Deming et al. 2016 for a very thorough design along these lines in the case of post-secondary qualifications the U.S.).

We would like to highlight avenues for future empirical and theoretical work: First, it would be interesting to quantify the screening costs for firms in the (expanding) presence of gray degrees. Second, it would be insightful to study to which extent employers make inference about human capital when reviewing applications (with different degrees). One possibility would be to randomly report additional skills (unrelated to university education) on the resume. Third, gray degrees may theoretically have predictive value of job productivity. For instance, Weaver (2018) documents that bribes to obtain a public sector job in developing countries may actually increase welfare, as the willingness to pay a bribe is positively predictive of subsequent productivity on the job.

6.5 Conclusion

In conclusion, mentioning a gray degree from a questionable university on one's resume improved callback rates relative to reporting no degree in our resume experiment. However, employers favored degrees from institutions that offer only authentic degrees. From our results, several policy implications emerge. First, it seems clear that authorities need to regulate and control the Indian higher education sector more thoroughly. We found that services relating to gray degrees are advertised quite openly; thus, either the legal framework is too lax or means to curb such practices are not adequately deployed. Some limited policy measures have been introduced such as the abolition of the inefficient 'Distance Education

Council' and the establishment of the 'Distance Education Bureau' as the new regulatory authority. However, the effectiveness of this new bureau has yet to be demonstrated. Further, authentic colleges need to keep distinguishing themselves clearly from gray institutions. They could, for instance, report gray degrees and agents operating in their areas to local and national authorities. Lessons may come from how other countries such as the U.S. regulate diploma mills or online colleges. In the U.S., states govern and regulate the education sector and some of them have issued "negative lists of unapproved, unaccredited, or illegal providers" (USNEI 2007, pp.1). Likewise, Indian authorities could list gray degree-issuing institutions in publicly available databases. This may help employers discount questionable degrees even more and may lower the incentives of students to buy degrees.

546567-L-bw-Majilla
Processed on: 6-8-2020

Chapter 7
Summary and Remarks

This dissertation used experimental and non-experimental approaches to analyze and deal with two parallel but connected issues in policy relevant empirical development economics – intra-household parental resource allocation, and labor market discrimination. A unifying theme was that each essay attempted to uncover disadvantages experienced by women in different periods of their lives. The thesis offers five sets of results. First, it shows intra-household gender disparity in parental allocation of shadow education expenditures. It shows a clear birth order gradient and disadvantages for girls at every birth order. Second, differential intra-household resource allocation together with the negative effect of family size may plausibly explain gender gaps in math test scores. These patterns clearly indicate elder son preference. Third, it shows a labor market penalty for mothers, but the existence and the magnitude of the penalty depends on community gender norms. Fourth, the penalty eases if mothers can signal availability of childcare at home. Finally, the thesis finds a labor market advantage for those with gray undergraduate degrees as compared to having no degree.

Although, these chapters document disadvantage for women, the fifth and sixth chapters in particular show some promising results. The fifth chapter shows that childcare support at home may reduce motherhood penalty, and thus strengthens the case for policies that work towards providing childcare support. The sixth chapter also stands out as it explores a potential, although illegal, coping strategy that women may adopt to compensate their disadvantage. In particular, the chapter shows that gray degrees are more effective in helping women applicants as compared to their male peers. While the chapter shows that gray degrees are more beneficial for women, the result needs to be understood keeping in mind the context. In other words, attempting to acquire a gray degree itself is an outcome of multiple disadvantages, and unlikely to be a policy solution.

The evidence presented here suggests a clear need to promote gender equality from early childhood. While there is no dearth of interventions in the country at the central or the state level which privilege girls and women, their effectiveness and the manner in which various policies fit together, whether they are suitable and whether they may be bundled to enhance their effectiveness has rarely been systematically and carefully examined.

For instance, in January 2015, at the federal level, the Indian government launched its '*Beti Bachao, Beti Padhao*' (Save the daughter, educate the daughter) campaign. This is a behavioral change communication campaign which aims to "ensure girls are born, loved and nurtured without discrimination, educated and raised to become empowered citizens of this country with equal rights".[122] The scheme has three main objectives. First, it aims to prevent gender biased sex selective eliminations. Second, aims to ensure the survival and protection of the girl child, and, third, enhancing education and participation of the girl child. While the policy hits the right note, the reach and resources mobilized for the campaign are limited.[123] There are a number of programs across Indian states that incentivize parents by providing cash transfers or in kind support for "the girl child". However, as shown in the second chapter, son preference is relatively greater in higher socioeconomic status families. Thus, there is no reason to suspect that poverty (or liquidity constraints to be more specific) leads to son preference and accordingly the usefulness of such cash transfer policies may be limited.[124] Furthermore, parental concern over future labor market returns could explain differential resource allocation. Thus, there is a need for policies targeting female labor force participation. Unfortunately, to the best of my knowledge there is no coherent policy which systematically attempts to increase female labor force participation in India, although a number of the new digitally-oriented service opportunities emerging in India are perhaps more women friendly. Increasing female labor force participation also needs to deal with demand-side discrimination against women and especially mothers. A campaign such as '*Beti Bachao, Beti Padhao*', but for adult females may help.

---

[122] For instance, see:
https://web.archive.org/web/20141105190442/http://wcd.nic.in/tender/Beti_bachao_beti_padhao_campaign_24072014.pdf accessed on 17 April 2020.

[123] Considering the magnitude of discrimination against girls in India, the scheme started with miniscule funding of 100 crore INR (~ 18.51 million USD). However, it has gradually increased over time. The federal budget of 2020-21 allocated 28,600 crore INR (3.44 billion USD) for women specific programs which includes 'Beti Bachao, Beti Padhao'. While this is a huge increase over time, but it is still quite small on a per capita basis. See https://www.financialexpress.com/budget/budget-2020-beti-bachao-beti-padhao-a-big-success-scheme-yielded-tremendous-results-says-fm/1848214/ accessed on 23 May 2020.

[124] For a specific case, see the analysis of the Girl Child Protection Scheme in Tamil Nadu. In their analysis, Srinivasan and Bedi (2009) argue that the scheme assumes that only poor families are anti-daughters.

A promising result that this dissertation uncovers is that women from matrilineal communities faces no motherhood penalty. Motherhood penalty is most discernible for mothers from patrilineal communities. Thus women empowerment, proxied by matrilineal community origin but many employers read it as a sign of independence in our research, may ease the motherhood penalty. Once women have more resources through labor force participation or by means of other policy benefits, such independence may follow.[125]

Childcare support has been found to be an important mitigating factor. Unfortunately India does not have a childcare support program as such. The country's flagship schemes e.g., Anganwadi services and Pradhan Mantri Matru Vandana Yojana are *per se* not child care support schemes. These schemes primarily provide support in cash or kind to help with children's nutrition. The Government of India has a National Crèche Scheme. However, the scheme is yet to reach scale and its reach has declined over time.[126]

Together, the results presented in this thesis motivate an integrated and well thought-out set of evidence-based policies that may be used to address the multidimensional disadvantage for women. The plethora of existing policies indicates that policy makers recognize the problems, but as is quite common implementation and effectiveness lags behind. A final issue is that most of the existing policies (cash transfers, nutritional support, free access to public school) are primarily intended for families in weaker socioeconomic circumstances. While it is ethically justified to target disadvantaged families, issues discussed in this thesis are not confined to any particular social strata and call for policies and actions that may spark deeper social change.

[125] The Nordic countries are remarkably successful in increasing the labor force participation of mothers. This is primarily due to the family friendly policies (Datta Gupta et al., 2008).

[126] As of June 2019, only 7,930 Crèches were functional across the country with about 300,000 beneficiaries. Over time the number of creches has been declining. https://timesofindia.indiatimes.com/india/why-the-number-of-creches-has-dropped-sharply-since-2017/articleshow/67765125.cms, accessed on 28 May 2020.

Bibliography

Agnihotri, S. B. (2003). Survival of the girl child: Tunnelling out of the Chakravyuha. *Economic and Political Weekly*, 4351-4360.

Agüero, J. M., & Marks, M. S. (2011). Motherhood and female labor supply in the developing world evidence from infertility shocks. *Journal of Human Resources*, *46*(4), 800-826.

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, *128*(2), 469-530.

Almond, D., Currie, J., & Duque, V. (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature*, *56*(4), 1360-1446.

Alwyn, Y (2019) Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results, *The Quarterly Journal of Economics*, 134(2): 557–598.

Anderson, D. J., Binder, M., & Krause, K. (2002). The motherhood wage penalty: Which mothers pay it and why? *American Economic Review*, *92*(2), 354-358.

Anderson, D. J., Binder, M., & Krause, K. (2003). The motherhood wage penalty revisited: Experience, heterogeneity, work effort, and work-schedule flexibility. *ILR Review*, *56*(2), 273-294.

Andersen, S, Ertac, S, Gneezy, U, List, J A, and Maximiano, S (2013) Gender, Competitiveness, and Socialization at a Young Age: Evidence from a Matrilineal and a Patriarchal Society, *Review of Economics and Statistics*, 95(4): 1438-1443.

Angelov, N., Johansson, P., & Lindahl, E. (2016). Parenthood and the gender gap in pay. *Journal of Labor Economics*, *34*(3), 545-579.

Angrist, J. D., & Evans, W. N. (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review*, *88*(3), 450-477.

Aranda, B. & Glick, P. (2014). Signaling devotion to work over family undermines the motherhood penalty, *Group Processes & Intergroup Relations, 17 (1)*, 91-99.

Angrist, J., Lavy, V., & Schlosser, A. (2010). Multiple Experiments for the Causal Link Between the Quantity and Quality of Children. *Journal of Labor Economics*, *28*(4), 773-824.

Aslam, M. (2009). Education gender gaps in Pakistan: Is the labor market to blame? *Economic Development and Cultural Change*, *57*(4), 747-784.

Aslam, M., & Kingdon, G. G. (2008). Gender and household education expenditures in Pakistan. *Applied Economics*, *40*(20), 2573-2591.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., & Rubio-Codina, M. (2018). Estimating the production function for human capital: Results from a randomized control trial in Colombia (No. w20965). *National Bureau of Economic Research*.

Attanasio, O. P., & Lechene, V. (2014). Efficient responses to targeted cash transfers. *Journal of political Economy*, *122*(1), 178-222.

Attewell, P., & Domina, T. (2011). Educational imposters and fake degrees. *Research in Social Stratification and Mobility*, *29*(1), 57-69.

Aurini, J., & Davies, S. (2004). The transformation of private tutoring: Education in a franchise form. *Canadian journal of sociology/Cahiers canadiens de sociologie*, 419-438.

Banerjee, A, Bertrand, M, Datta, S, and Mullainathan, S (2009) Labor Market Discrimination in Delhi: Evidence from a Field Experiment, *Journal of Comparative Economics*, 37(1): 14-27.

Banerjee, B., & Knight, J. B. (1985). Caste discrimination in the Indian urban labour market. *Journal of development Economics*, *17*(3), 277-307.

Barcellos, Silvia Helena, Leandro S. Carvalho, and Adriana Lleras-Muney. "Child gender and parental investments in India: Are boys and girls treated differently?" *American Economic Journal: Applied Economics* 6, no. 1 (2014): 157-89.

Beam, E, Hyman, J, Theoharides, C B (2017) The Relative Returns to Education, Experience, and Attractiveness for Young Workers, *IZA Discussion Paper*, No. 10537.

Becker, G S (1971) *The Economics of Discrimination*, 2nd ed, Chicago: University of Chicago Press.

Becker, G. S. (1960). An Economic Analysis of Fertility. In *Demographic and Economic Change in Developed Countries,* ed. Gary Becker. Princeton, N.J.: Princeton University Press.

Becker, G. S., & Lewis, H. G. (1973). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy*, *81*(2, Part 2), S279-S288.

Benard, S. & Correll, S. J. (2010). Normative discrimination and the motherhood penalty, *Gender & Society*, *24(5)*, 616-646.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American economic review*, *94*(4), 991-1013.

Bhalotra, S., Valente, C., & Van Soest, A. (2010). The puzzle of Muslim advantage in child survival in India. *Journal of Health Economics*, *29*(2), 191-204.

Bharadwaj, P., De Giorgi, G., Hansen, D., & Neilson, C. A. (2016). The gender gap in Mathematics: Evidence from Chile. *Economic Development and Cultural Change*, *65*(1), 141-166.

Bhatti, F., & Jeffery, R. (2012). Girls' schooling and transition to marriage and motherhood: exploring the pathways to young women's reproductive agency in Pakistan. *Comparative Education*, *48*(2), 149-166.

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

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2010). Small family, smart family? Family size and the IQ scores of young men. *Journal of Human Resources*, *45*(1), 33-58.

Blunch, N. H., & Datta Gupta, N. (2020a). Mothers' health knowledge gap for children with diarrhea: A decomposition analysis across caste and religion in India. *World Development*, *126*, 104718.

Blunch, N. H., & Datta Gupta, N. (2020b). Mothers' health knowledge for children with diarrhea: who you are or who you know? *Review of Economics of the Household*, 1-34.

Bloom, D. E., Canning, D., Fink, G., & Finlay, J. E. (2009). Fertility, female labor force participation, and the demographic dividend. *Journal of Economic Growth*, *14*(2), 79-101.

Bongaarts, J. (1994). Population policy options in the developing world. *Science*, *263*(5148), 771-776.

Bovet, J., Raiber, E., Ren, W., Wang, C., & Seabright, P. (2018). Parent–offspring conflict over mate choice: An experimental study in China. *British Journal of Psychology*, *109*(4), 674-693.

Bray, M. (2014). The impact of shadow education on student academic achievement: Why the research is inconclusive and what can be done about it. *Asia Pacific Education Review*, *15*(3), 381-389.

Bray, M., & Lykins, C. (2012). *Shadow Education: Private Supplementary Tutoring and Its Implications for Policy Makers in Asia* (No. 9). Philippines: Asian Development Bank.

Bray, M., Mazawi, A. E., & Sultana, R. G. (Eds.). (2013). *Private Tutoring Across the Mediterranean*. Rotterdam: Sense Publishers.

Brown, G. M. (2006). Degrees of doubt: Legitimate, real and fake qualifications in a global market. *Journal of Higher Education Policy and Management*, *28*(1), 71-79.

Browning, M., Chiappori, P. A., & Lechene, V. (2010). Distributional effects in household models: Separate spheres and income pooling. *The Economic Journal*, *120*(545), 786-799.

Budig, M J, and England, P (2001) The Wage Penalty for Motherhood, *American Sociological Review*, 66(2): 204-225.

Budig, M J, and Hodges, M J (2010) Differences in Disadvantage Variation in the Motherhood Penalty Across White Women's Earnings Distribution, *American Sociological Review*, 75(5): 705-728.

Budig, M. J., Misra, J., & Boeckmann, I. (2012). The Motherhood Penalty in cross-national perspective: The Importance of work–family policies and cultural attitudes, *Social Politics: International Studies in Gender, State & Society*, *19(2)*, 163-193.

Bygren, M, Erlandsson, A, and Gähler, M (2017) Do Employers Prefer Fathers? Evidence from a Field Experiment Testing the Gender by Parenthood Interaction Effect on Callbacks to Job Applications, *European Sociological Review*, 33(3): 337-348.

Cadsby, B C, Servátka, M, and Song, F (2013) How Competitive Are Female Professionals? A Tale of Identity Conflict, *Journal of Economic Behavior & Organization*, 92: 284-303.

Carneiro, P., Løken, K. V., & Salvanes, K. G. (2015). A flying start? Maternity leave benefits and long-run outcomes of children. *Journal of Political Economy*, *123*(2), 365-412.

Card, D., & Payne, A. A. (2017). High school choices and the gender gap in STEM (No. w23769). *National Bureau of Economic Research*.

Carlana, M. (2019) Implicit Stereotypes: Evidence from Teachers' Gender Bias. *The Quarterly Journal of Economics*. Advance Access.

Carlsson, M., & Rooth, D. O. (2007). Evidence of ethnic discrimination in the Swedish labor market using experimental data. *Labour Economics*, *14*(4), 716-729.

Carneiro, P., Lokshin, M., & Umapathi, N. (2017). Average and marginal returns to upper secondary schooling in Indonesia. *Journal of Applied Econometrics*, *32*(1), 16-36.

Carneiro, P., Meghir, C., & Parey, M. (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association*, *11*(suppl_1), 123-160.

Cassar, A, Wordofa, F, and Zhang, Y J (2016) Competing for the Benefit of Offspring Eliminates the Gender Gap in Competitiveness, *Proceedings of the National Academy of Sciences*, 113(19): 5201-5205.

Casterline, J. B., & Sinding, S. W. (2000). Unmet need for family planning in developing countries and implications for population policy. *Population and development review*, *26*(4), 691-723.

Carranza, E. (2014). Soil endowments, female labor force participation, and the demographic deficit of women in India. *American Economic Journal: Applied Economics*, *6*(4), 197-225.

Cleland, J., Bernstein, S., Ezeh, A., Faundes, A., Glasier, A., & Innis, J. (2006). Family planning: the unfinished agenda. *The Lancet*, *368*(9549), 1810-1827.

Chakravarty, A. (2015). Gender-Biased Breastfeeding in Egypt: Examining the Fertility Preference Hypotheses of Jayachandran and Kuziemko (2011). *Journal of Applied Econometrics*, *30*(5), 848-855.

Chamarbagwala, R., & Ranger, M. (2010). A multinomial model of fertility choice and offspring sex ratios in India. *The Journal of Development Studies*, *46*(3), 417-438.

Charness, G, & Gneezy, U, (2012) Strong Evidence for Gender Differences in Risk Taking, *Journal of Economic Behavior & Organization*, 83(1): 50-58.

Chowdhury, V. (2011). PhDs, Bachelor's degrees on sale in Punjab. *NDTV,* 06 August. Available at: http://www.ndtv.com/india-news/phds-bachelors-degrees-on-sale-in-punjab-463936 [Accessed 16 October 2016]

Contini, D., Di Tommaso, M. L., & Mendolia, S. (2017). The gender gap in mathematics achievement: Evidence from Italian data. *Economics of Education Review*, *58*, 32-42.

Correll, S.J., Benard, S. & Paik, I., (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology*, *112*(5), 1297-1338.

Croson R, & Gneezy, U (2009) Gender Differences in Preferences, *Journal of Economic Literature*, 47(2): 1-27.

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, *78*(3), 883-931.

Dahl, G. B., Løken, K. V., Mogstad, M., & Salvanes, K. V. (2016). What is the case for paid maternity leave?. *Review of Economics and Statistics*, *98*(4), 655-670.

Dahl, G. B., & Moretti, E. (2008). The demand for sons. *The Review of Economic Studies*, *75*(4), 1085-1120.

Dang, H. A. H., & Rogers, F. H. (2015). The decision to invest in child quality over quantity: Household size and household investment in education in Vietnam. *The World Bank Economic Review*, *30*(1), 104-142.

Darolia, R., Koedel, C., Martorell, P., Wilson, K., & Perez-Arce, F. (2015). Do employers prefer workers who attend for-profit colleges? Evidence from a field experiment. *Journal of Policy Analysis and Management*, *34*(4), 881-903.

Das, M. B. & Zumbyte, I. (2017). The motherhood penalty and female employment in urban India, *Policy Research Working Paper*, No. 8004: World Bank, Washington, DC.

Datta Gupta, N., & Smith, N. (2002). Children and career interruptions: the family gap in Denmark. *Economica*, *69*(276), 609-629.

Datta Gupta, N., Smith, N., & Verner, M. (2008). The impact of Nordic countries' family friendly policies on employment, wages, and children. *Review of Economics of the Household*, *6*(1), 65-89.

Deaton, A. (1997). *The analysis of household surveys: A Microeconometric approach to development policy*. Baltimore: Johns Hopkins University Press for the World Bank.

Deaton, A. (1989). Looking for boy-girl discrimination in household expenditure data. *The World Bank Economic Review*, *3*(1), 1-15.

Deming, D. J., Yuchtman, N., Abulafi, A., Goldin, C., & Katz, L. F. (2016). The value of postsecondary credentials in the labor market: An experimental study. *American Economic Review*, *106*(3), 778-806.

Deshpande, A. (2011). The Grammar of Caste Economic Discrimination in Contemporary India. Delhi: Oxford University Press.

Dossi, G., Figlio, D. N., Giuliano, P., & Sapienza, P. (2019). Born in the Family: Preferences for Boys and the Gender Gap in Math. *National Bureau of Economic Research*, (No. w25535).

Dunbar, G. R., Lewbel, A., & Pendakur, K. (2013). Children's resources in collective households: identification, estimation, and an application to child poverty in Malawi. *American Economic Review*, *103*(1), 438-71.

Ellena, R, and Nongkynrih, K (2017) Changing gender roles and relations in food provisioning among matrilineal Khasi and patrilineal Chakhesang Indigenous rural People of North-East India, *Maternal and Child Nutrition,* 13(s3):1-14.

England, P., Bearak, J., Budig, M. J., & Hodges, M. J. (2016). Do highly paid, highly skilled women experience the largest motherhood penalty? *American Sociological Review*, *81(6)*, 1161-1189.

Ebenstein, A. (2010). The “missing girls” of China and the unintended consequences of the one child policy. *Journal of Human Resources*, *45*(1), 87-115.

Falk, A., Kosse, F., Pinger, P., Schildberg-Hörisch, H., & Deckers, T. (2019). Socio-economic status and inequalities in children's IQ and economic preferences. Unpublished.

Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, *77*(3), 953-973.

Firpo, S. P., Fortin, N. M., & Lemieux, T. (2018). Decomposing wage distributions using recentered influence function regressions. *Econometrics*, *6*(2), 28.

Fortin, N., T. Lemieux and S. Firpo (2011) 'Decomposition Methods in Economics', 'Decomposition Methods in Economics', *Handbook of Labor Economics,* Vol. 4. pp. 1-102. Elsevier.

Fryer Jr, R. G., & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, *2*(2), 210-40.

Gaba, A., & Li, W. (2015). Growth and development of distance education in India and China: a study on policy perspectives. *Open Praxis*, *7*(4), 311-323.

Gaddis, I, and Klasen, S (2014) Economic Development, Structural Change, and Women’s Labor Force Participation: A Re-Examination of the Feminization U Hypothesis, *Journal of Population Economics*, 27(3): 639-81.

Galarza, F B, and Yamada, G (2014) Labor Market Discrimination in Lima, Peru: Evidence from a Field Experiment, *World Development*, 58: 83-94.

Galarza, F B, and Yamada, G (2017) Triple Penalty in Employment Access: The Role of Beauty, Race, and Sex, *Journal of Applied Economics*, 20(1): 29-47.

Gallen, Y. (2018). Motherhood and the gender productivity Gap. Unpublished.

Gangl, M., & Ziefle, A. (2009). Motherhood, labor force behavior, and women’s careers: An empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States. *Demography*, *46*(2), 341-369.

Geraldes, D. (2018). Women dislike competing against men. Available online at grape.org. pl/wpcontent/uploads/2016/08/Geraldes paper. pdf.

Gneezy, U., Leonard, K. L., & List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, *77*(5), 1637-1664.

Goldin, C (1994) The U-shaped Female Labor Force Function in Economic Development and Economic History. *NBER Working Paper Series*, No. 4707.

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

Goldin, C. (2014). A grand gender convergence: Its last chapter. *American Economic Review*, *104*(4), 1091-1119.

Goldin, C., Kerr, S. P., Olivetti, C., & Barth, E. (2017). The expanding gender earnings gap: Evidence from the LEHD-2000 Census, *American Economic* Review, 107(5), 110-14.

Grolleau, G., Lakhal, T., & Mzoughi, N. (2008). An introduction to the Economics of Fake Degrees. *Journal of Economic Issues*, *42*(3), 673-693.

Guilmoto, C. Z. (2009). The sex ratio transition in Asia. *Population and Development Review*, *35*(3), 519-549.

Gupta, D., & Gupta, N. (2012). Higher education in India: structure, statistics and challenges. *Journal of education and Practice*, *3*(2).

Hartmann, S. (2008). *The informal market of education in Egypt: private tutoring and its implications*. Johannes Gutenberg-Universität Mainz.

Heath, R. (2017). Fertility at work: Children and women's labor market outcomes in urban Ghana. *Journal of Development Economics*, *126*, 190-214.

Heß, S (2017) Randomization inference with Stata: A guide and software, *Stata Journal*, 17(3): 630-651.

Heckman, J., Pinto, R., & Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, *103*(6), 2052-86.

Heckman, J. J., & Vytlacil, E. (2005). Structural equations, treatment effects, and econometric policy evaluation 1. *Econometrica*, *73*(3), 669-738.

Hoffman, M, Gneezy, U, and List, J A (2011) Nurture Affects Gender Differences in Spatial Abilities, *Proceedings of the National Academy of Sciences*, 108(36): 14786-14788.

Irfan, H, (2011) North-East Residents in Delhi Facing Bias: Report. India Today, April 18 2011. Available at: http://indiatoday.intoday.in/story/report-says-north-east-residents-in-delhi-face-humiliation/1/135561.html [Accessed on November 30, 2017]

Jayachandran, S. (2014). Incentives to teach badly: After-school tutoring in developing countries. *Journal of Development Economics*, *108*, 190-205.

Jayachandran, S, (2015) The Roots of Gender Inequality in Developing Countries, *Annual Review of Economics*, 7: 63–88.

Jayachandran, S., & Kuziemko, I. (2011). Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India. *The Quarterly journal of economics*, *126*(3), 1485-1538.

Jayachandran, S., & Pande, R. (2017). Why are Indian children so short? The role of birth order and son preference. *American Economic Review*, *107*(9), 2600-2629.

Jensen, R (2012) Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India, *Quarterly Journal of Economics*, 127(2): 753-792.

Kaas, L., & Manger, C. (2012). Ethnic discrimination in Germany's labour market: A field experiment. *German economic review*, *13*(1), 1-20.

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

Kingdon, G. G. (2005). Where has all the bias gone? Detecting gender bias in the intrahousehold allocation of educational expenditure. *Economic Development and Cultural Change*, *53*(2), 409-451.

Klasen, S, & Pieters, J. (2015) What Explains the Stagnation of Female Labor Force Participation in Urban India? *World Bank Economic Review*, 29(3), 449-478.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., & Zweimüller, J. (2019a). Child penalties across countries: Evidence and explanations. *AEA Papers and Proceedings*, 109, 122-126.

Kleven, H., Landais, C., & Søgaard, J. E. (2019b). Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, *11*(4), 181-209.

Klonner, K, Pal S, Schwieren, C (2019) Social Norms, Women's Status and Gender Differences in Competitiveness: Evidence from Field Experiments in India's Northeast. https://www.isid.ac.in/~epu/acegd2019/papers/SumantraPal.pdf. Accessed on December 15, 2019.

Kobakhidze, M. N. (2014). Corruption risks of private tutoring: case of Georgia. *Asia Pacific Journal of Education*, *34*(4), 455-475.

Kobakhidze, M. N. (2017). (Un) regulating Shadow Education in West Asia: Armenia, Azerbaijan, Georgia and Turkey. In Kennedy, KJ & Lee, JCK (Eds.). *Routledge International Handbook of Schools and Schooling in Asia*. London: Routledge.

Kristensen, P., & Bjerkedal, T. (2007). Explaining the relation between birth order and intelligence. *Science*, *316*(5832), 1717-1717.

Kumar, S., Khan, A., & Khan, A. (2011). Degrees on sale: Jaipur study centres offer bachelor degree to PhD for money., *indiatoday.in,* 5 October. Available at: http://indiatoday.intoday.in/story/degrees-on-sale-in-jaipur-phd-in-3-days/1/153852.html [Accessed 16 October 2016]

Kuziemko, I., Pan, J., Shen, J., & Washington, E. (2018). The mommy effect: Do women anticipate the employment effects of motherhood? (No. w24740). *National Bureau of Economic Research*.

Ladusingh, L, Singh, C H (2006) Place, Community Education, Gender and Child Mortality in North-east India, *Population, Space and Place*, 12(1): 65-76.

Lee, J. (2008). Sibling size and investment in children's education: An Asian instrument. *Journal of Population Economics*, *21*(4), 855-875.

Maitra, P., Pal, S., & Sharma, A. (2016). Absence of altruism? female disadvantage in private school enrollment in india. *World Development*, *85*, 105-125.

Majilla, T. (2016). Gray Academic Degree Market in India: A Mixed Method Study. *Institute of Social Studies* - Social Policy for Development (SPD). Available at: http://hdl.handle.net/2105/37158 [Accessed 1 March 2017]

Mammen, K, Paxson, C (2000) Women's Work and Economic Development, *Journal of Economic Perspectives*, 14(4): 141-164.

McDuie-Ra, D (2013) Beyond the 'Exclusionary City': Northeast Migrants in Neo-liberal Delhi, *Urban Studies*, 50(8), 1625-1640.

Molina, T. (2020). Pollution, ability, and gender-specific investment responses to shocks. *Journal of the European Economic Association,* Forthcoming.

Mori, I., & Baker, D. (2010). The origin of universal shadow education: What the supplemental education phenomenon tells us about the postmodern institution of education. *Asia Pacific Education Review*, *11*(1), 36-48.

Muñoz, J. S. (2018). The economics behind the math gender gap: Colombian evidence on the role of sample selection. *Journal of Development Economics*, *135*, 368-391.

Mwaikambo, L., Speizer, I. S., Schurmann, A., Morgan, G., & Fikree, F. (2011). What works in family planning interventions: a systematic review. *Studies in family planning*, *42*(2), 67-82.

Naga Women's Union, International Work Group for Indigenous Affairs, Henry Martyn Institute (2018) *The Place of Women in Naga Society*, Christian Literature Centre, Guwahati, Assam.

Nasrullah, M., Muazzam, S., Bhutta, Z. A., & Raj, A. (2014a). Girl child marriage and its effect on fertility in Pakistan: findings from Pakistan Demographic and Health Survey, 2006–2007. *Maternal and child health journal*, *18*(3), 534-543.

Nasrullah, M., Zakar, R., & Zakar, M. Z. (2014b). Child marriage and its associations with controlling behaviors and spousal violence against adolescent and young women in Pakistan. *Journal of Adolescent Health*, *55*(6), 804-809.

NESCH (2011) North east migration and challenges in national capital cities. NESCH, Delhi. Available at: http://www.nehelpline.in/wp-content/uploads/2015/06/NE-Migration-Challenges-Research-Report.pdf [Accessed on February 15, 2018]

Niederle, M., & Vesterlund, L. (2011). Gender and competition. *Annual Review of Economics*, *3*(1), 601-630.

Nollenberger, N., Rodríguez-Planas, N., & Sevilla, A. (2016). The math gender gap: The role of culture. *American Economic Review*, *106*(5), 257-61.

Nongbri, M. W. L. (2006). Basic foundations of Khasi culture, continuity and change; North-Eastern Hill University, http://hdl.handle.net/10603/62028.

Pager, D., Bonikowski, B., & Western, B. (2009). Discrimination in a low-wage labor market: A field experiment. *American sociological review*, *74*(5), 777-799.

Phelps, E S (1972) The Statistical Theory of Racism and Sexism, *American Economic Review*, 62(4): 659-661.

Roy, A (2018) Discord in Matrilinearity: Insight into the Khasi Society in Meghalaya. *Society and Culture in South Asia*, 4(2): 278-297.

Sahoo, S. (2017). Intra-household gender disparity in school choice: Evidence from private schooling in India. *The Journal of Development Studies*, *53*(10), 1714-1730.

Siddique, Z (2011) Evidence on Caste Based Discrimination, *Labour Economics*, 18(Supplement 1): S146-S159.

Shrestha, S. A., & Palaniswamy, N. (2017). Sibling rivalry and gender gap: Intrahousehold substitution of male and female educational investments from male migration prospects. *Journal of Population Economics*, *30*(4), 1355-1380.

Srinivasan, S., & Bedi, A. S. (2008). Daughter elimination in Tamil Nadu, India: A tale of two ratios. *The Journal of Development Studies*, *44*(7), 961-990.

Srinivasan, S., and Bedi, A.S. (2009). Girl child protection scheme in Tamil Nadu: An Appraisal. *Economic and Political Weekly*, XLIV (48), 10-12.

Tarozzi, A., & Mahajan, A. (2007). Child nutrition in India in the nineties. *Economic Development and Cultural Change*, *55*(3), 441-486.

Tsui, A. O. (2001). Population policies, family planning programs, and fertility: The record. *Population and Development Review*, *27*, 184-204.

Ullas, S. S., & Prasher, G. (2013). Fake degree scam: No sweat, you can get a university degree in 10 days. *The Times of India,* 24 September. Available at: http://timesofindia.indiatimes.com/city/bengaluru/Fake-degree-scam-No-sweat-you-can-get-a-university-degree-in-10-days/articleshow/22956406.cms [Accessed 1 March 2017]

USNEI. (2007). Accreditation and Quality Assurance: Diploma Mills and Fraud. *International Affairs Office, U.S. Department of Education*. Available at: https://www2.ed.gov/about/offices/list/ous/international/usnei/us/fraud.doc [Accessed 1 June 2018]

Weaver, J. (2018). Jobs for sale: Bribery and misallocation in hiring. *manuscript, University of Southern California*. Available at: https://www.dropbox.com/s/6bktr62uondsenl/JeffWeaverJMPfinal.pdf?dl=0 [Accessed 9 April 2018]

Verick, S. (2018). Female labor force participation in developing countries. IZA World of Labor, 2014, 87v2.

Yeung, W. J. J., Desai, S., & Jones, G. W. (2018). Families in southeast and South Asia. *Annual Review of Sociology*, *44*, 469-495.

Yum, M., Kim, S., & Tertilt, M. (unpublished). Status Externalities and Low Birth Rates in Korea.

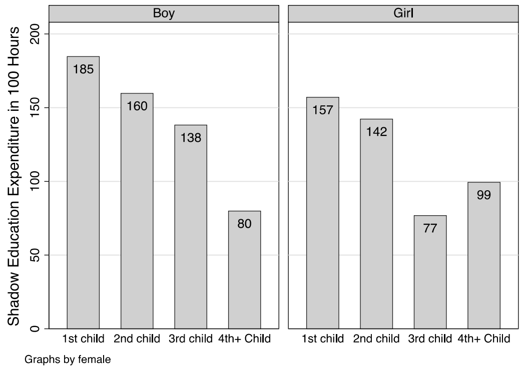
Zhan, S., Bray, M., Wang, D., Lykins, C., & Kwo, O. (2013). The effectiveness of private tutoring: Students' perceptions in comparison with mainstream schooling in Hong Kong. *Asia Pacific Education Review*, *14*(4), 495-509.

Zimmermann, L. (2012). Reconsidering gender bias in intrahousehold allocation in India. *Journal of Development Studies*, *48*(1), 151-163.

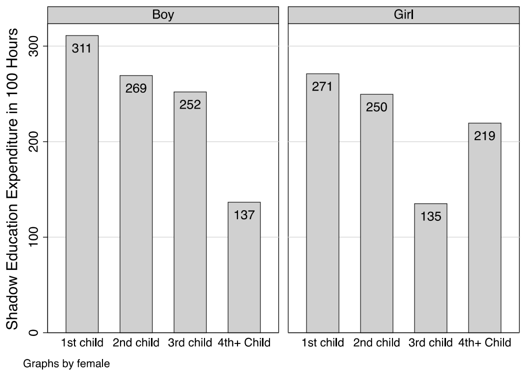
Appendix A

Figure A1: Shadow Education Expenditures in 100 Hours

Panel A: Pooled Sample



Panel B: Urban Sample



Panel C: Rural Sample

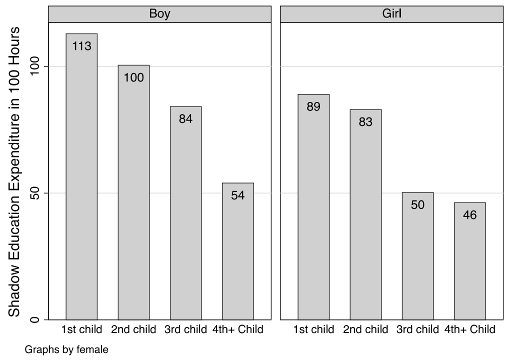


Table A1: Sample Size by Birth Order and Gender

Birth Order	Full Sample	Boys	Girls
1^{st}	21299	11124	10175
2^{nd}	17299	9300	7999
3^{rd}	8630	4625	4005
4^{th+}	5132	2751	2381

Notes: Sample size by gender in full sample - Boys: 27800 and Girls: 24560.

Table A2: Estimates for Shadow Education Expenditures (INR)

Dep. Var. Shadow Education Expenditures (INR)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girl	-175.43^{***}						
	(29.55)						
Girl × 1st Child		-235.76^{***}	-211.17^{***}	-247.39^{***}	-246.60^{***}	-209.70^{***}	-219.60^{***}
		(46.92)	(49.71)	(75.42)	(71.86)	(46.52)	(54.71)
Girl × 2nd Child		-202.31^{***}	-189.43^{***}	-168.31^{**}	-177.54^{***}	-180.27^{***}	-194.47^{***}
		(45.08)	(45.68)	(62.38)	(56.89)	(48.21)	(47.01)
Girl × 3rd Child		-226.08^{***}	-147.57^{***}	-158.50	-164.77	-134.80^{***}	-151.35^{***}
		(47.40)	(34.90))	(117.47)	(115.06)	(39.22)	(35.81)
Girl × 4th+ Child		-66.61^{***}	-59.58^{*}	-72.83	-74.10	-63.24	-61.84^{*}
		(33.64)	(29.99)	(59.05)	(57.67)	(43.75)	(30.70)
2nd Child		-173.53^{***}	-96.42	-45.45	-38.70	-17.36	-107.75
		(47.04)	(75.16)	(82.67)	(76.37)	(80.45)	(78.13)
3rd Child		-410.14^{***}	-172.89^{**}	-14.94	-5.36	-54.89	-186.31^{**}
		(59.16)	(70.56)	(135.36)	(125.61)	(82.17)	(74.08)
4th+ Child		-678.38^{***}	-106.50	63.22	71.96	46.10	-119.81
		(51.61)	(71.66)	(180.08)	(159.30)	(101.04)	(73.76)
Other Controls	Yes	No	Yes	Yes	Yes	Yes	Yes
Mother Fixed Effect	No	No	No	Yes	No	No	No
Household Fixed Effect	No	No	No	No	Yes	No	No
PSU Fixed Effect	No	No	No	No	No	Yes	
Completed Fertility	No	No	No	No	No	No	Yes
Observations	45148	52360	45148	45148	45148	45148	43184

Notes: Every column reports a separate linear regression. Standard errors reported in parenthesis are robust to within primary sampling unit (PSU) clustering. 2nd and 3rd child are indicators for children whose birth orders are 2 and 3, and, 4th+ indicates children with birth order 4 or later. In columns 4, 5 and 6 main effects are absorbed by mother, household and PSU fixed effects respectively. Controls always include child's age and its square, current standard of study, mother age and its square, maternal education, household income, household head's education, family size and dummies for private school, teacher attendance at school, Hindu, forward caste and urban residence. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Table A3: Weekly Duration and Probability of Shadow Education

Dependent Variable	Weekly Shadow Education
2^{nd} Child	-0.07
	(0.14)
3^{rd} Child	-0.39^{***}
	(0.12)
4^{th+} Child	-0.24
	(0.18)
Girl × 1^{st} Child	-0.44^{***}
	(0.10)
Girl × 2^{nd} Child	-0.49^{***}
	(0.13)
Girl × 3^{rd} Child	-0.46^{***}
	(0.11)
Girl × 4^{th+} Child	-0.12
	(0.18)
Other Controls	Yes
Observations	37634

Notes: Standard errors reported in parenthesis are robust to within primary sampling unit (PSU) clustering. 2^{nd} and 3^{rd} child are indicators for children whose birth orders are 2 and 3, and, 4^{th+} indicates children with birth order 4 or later. Controls always include child's age and its square, current standard of study, mother age and its square, maternal education, household income, household head's education, family size and dummies for private school, teacher attendance at school, Hindu, forward caste and urban residence. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Table A4: Summary Statistics – ASER data

Variables	Full Sample	Boys Sample	Girls Sample
Yearly Shadow Education Expenditures	643.15	701.07	582.82
	(2065.94)	(2132.93)	(1991.98)
Proportion of Students Taking Shadow Education	0.20	0.22	0.18
	(0.40)	(0.41)	(0.39)
Yearly Expenditures in Shadow Education for Students with Non-Zero Expenditures	3205.49	3247.46	3154.42
	(3613.71)	(3578.30)	(3655.75)
Child Age	10.31	10.21	10.41
	(3.28)	(3.27)	(3.30)
Mother Age	34.27	34.24	34.30
	(7.78)	(7.83)	(7.72)
Mother Education[1]	8.06	8.08	8.05
	(3.09)	(3.09)	(3.09)
Father Education[1]	8.86	8.87	8.85
	(3.21)	(3.21)	(3.21)
Standard	5.23	5.12	5.34
	(2.97)	(2.93)	(2.99)
Private School	0.31	2.53	2.58
	(0.46)	(0.97)	(1.01)
Electricity at House	0.82	0.82	0.82
	(0.38)	(0.39)	(0.38)
Toilet at House	0.54	0.54	0.54
	(0.50)	(0.50)	(0.50)
Household has a Graduate	0.23	0.23	0.23
	(0.42)	(0.42)	(0.42)

Notes: This Table reports means of the dependent variables and covariates. Standard errors reported in parenthesis. Around 49% of the full sample constitute Girls.
(1) Education: Bachelor degrees is coded as 15, Master as 17, and 1-12 grades are coded as 1-12 respectively.

Table A5: Association Between Shadow Education and Test Scores

	Math	Writing	Reading
	(1)	(2)	(3)
Shadow Education	0.183***	0.133***	0.095***
	(0.038)	(0.026)	(0.030)
Observations	10046	9995	10083

Notes: Every column reports a separate linear regression. All estimates include controls for child's age and its square, current standard of study, mother age and its square, maternal education, household income, household head's education, and dummies for private school, teacher attendance at school, ever repetition in exams, Hindu, forward caste and urban residence. To account for the unobserved ability shocks, I assume that parents (partly) predict this through children's educational achievements. In the specification, I use a dummy which takes the value 1 if the children have 'ever repeated' in any exam. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Appendix B: Recentered Influence Function (RIF) Decomposition

In this section, I decompose the test scores using Recentered Influence Function regression method suggested by Firpo et al. (2009).

The joint distribution between test scores t, the covariates X that include shadow education expenditures, and a categorical variable $G \in (0,1)$ where 1 represents a girl child is $f_{T,X,G}(t_i, x_i, G_i)$. The cumulative distribution of T conditional on G is

(B.1) $F_T^G(t) = \int F_{T|X}^G(T|x)dF_X^G(x)$

The difference between boys ($G = 0$) and girls ($G = 1$) for a given distribution statistics ν is given by:

(B.2) $\Delta\nu = \nu_0 - \nu_1 = \nu\left(\int F_{T|X}^0(T|X)dF_X^0(X)\right) - \nu\left(\int F_{T|X}^1(T|X)dF_X^1(X)\right)$

To estimate (B.2), I use the RIF[127] regressions developed by Firpo et al. (2018). Specifically,

(B.3) $\nu_i = \int E(RIF(T; \nu(F_t^i)|X)dF_X(X) = X^{i'}\hat{\beta}_i \quad for\ i \in (0,1)$

Following Oaxaca-Blinder decomposition in the main text, I decompose the test scores at mean.[128] The test scores distribution under the counterfactual scenario is given by:

(B.4) $F_T^c = \int F_{T|X}^1(T|X)dF_X^0(X) \cong \int F_{T|X}^1(T|X)dF_X^1(X)\Phi(X)$

where $\Phi(X)$ is a re-weighting function that can be identified as:

---

[127] The Recentered Influence Function of a distribution statistics ν is defined as:

$$\nu(F_T) + \frac{\partial\nu(F_T \to H_T)}{\partial\epsilon}$$

where the last past is the influence function (IF). IF is a directional derivative that shows how the distribution statistics would response to a small change in the distribution.

[128] The RIF mean of any value of T is simply itself. One can do RIF decomposition analysis at any distributional statistics. See Firpo et al. (2018) for more detail.

(B.5) $\Phi(X) = \frac{dF_X^0(X)}{dF_X^1(X)} = \frac{dF_{X|T}(T=0|X)}{dF_{X|T}(T=1|X)} = \frac{p}{1-p}\frac{1-P(T=1|X)}{P(T=1|X)}$

where p is the number of girls and, $P(T = 1|X)$ is the conditional probability of someone with characteristics X is a girl. I obtain the reweighting function using a probit model. The decomposition is then defined as:

(B.6) $\Delta\nu = X^{0'}(\hat{\beta}_0 - \hat{\beta}_c) + (X^0 - X^c)'\hat{\beta}_c + (X^c - X^1)'\hat{\beta}_1 + X^{c'}(\hat{\beta}_c - \hat{\beta}_1)$

where the first two components are the aggregate unexplained part (structure effects) and last two terms correspond to the explained part (composition effects). First and third parts correspond to pure structure and composition effects respectively. The second term is the reweighting error and the last term is the specification error.

The decomposition results are shown in Table B1. Reassuringly, disparities in shadow education can be accounted for substantial portions of gender gaps in test scores. To be specific, around 20%, 18%, and 13% of the explained gendered variations (composition effects) in the math, writing, and reading test scores can be attributed to the disparities in shadow education expenditures. Shadow education also contribute in unexplained variations (coefficient effects). Boys experience higher returns in math and writing test scores, but not in reading test scores where girls have clear advantage. The specification errors are insignificant, indicating that more flexible models are not needed. Taken together, the compositional effects of shadow education expenditures on test scores tend to be robust irrespective of the decomposition methods.

Table B1: RIF Decomposition of Test Scores

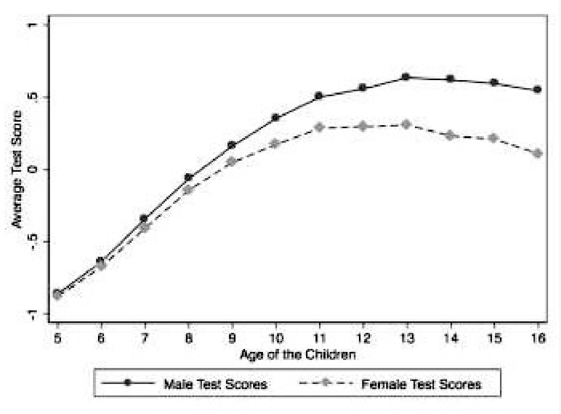
	Math		Writing		Reading	
	Composition	Coefficient	Composition	Coefficient	Composition	Coefficient
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Decomposition						
Aggregate	0.045***	0.080***	0.041***	0.004	0.044***	0.024
Decomposition	(0.00)	(0.02)	(0.00)	(0.02)	(0.00)	(0.02)
Pure Composition	0.046***		0.042***		0.045***	
	(0.00)		(0.00)		(0.00)	
Pure Coefficient		0.083***		0.007		0.026
		(0.02)		(0.02)		(0.02)
Specification Error	-0.001		-0.001		-0.001	
	(0.00)		(0.00)		(0.00)	
Reweighting Error		-0.003		-0.002		-0.002
		(0.01)		(0.01)		(0.02)
Panel B: Percentage contributions in pure composition effects						
Shadow Education	19.87	5.21	17.83	59.29	12.53	-3.42

Notes: Recentered Influence Functions decomposition. Composition and Coefficient refer to the decomposition due to composition (explained) and coefficient (unexplained) effects. Controls always include child's age and its square, current standard of study, mother age and its square, maternal education, household income, household head's education, and dummies for private school, teacher attendance at school, Hindu, forward caste and urban residence. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Appendix C

Figure C1: Reading Test Scores by Gender

Panel A



Panel B

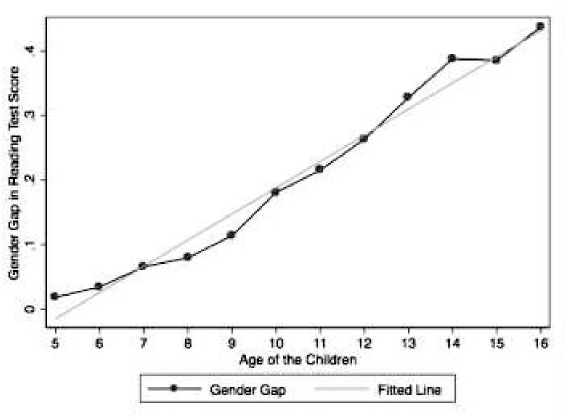


Table C1: Math Score Distribution

Math Score	**One**	**Two**	**Three**	**Four**	**Five**
	(1)	(2)	(3)	(4)	(5)
Boy (%)	49.06	58.13	61.98	62.95	63.44
Girl (%)	50.94	41.87	38.02	37.05	36.56

Table C2: Son Preference in Math Score Distribution

Dep. Var. Prob of Math Score	Younger Age Group		Older Age Group	
Math Score	**Five**	**Four**	**Five**	**Four**
	(1)	(2)	(3)	(4)
Girl × 1st Child	-0.003	0.003	-0.031***	-0.027***
	(0.003)	(0.004)	(0.005)	(0.004)
Girl × 2nd Child	0.003	-0.004	-0.019***	-0.029***
	(0.002)	(0.003)	(0.005)	(0.004)
Girl × 3rd Child	0.003	0.003	-0.015*	-0.031***
	(0.002)	(0.003)	(0.007)	(0.006)
Girl × 4^{th+} Child	0.003*	-0.001	-0.030***	-0.026***
	(0.002)	(0.002)	(0.008)	(0.007)
2nd Child	-0.010**	0.003	-0.009	-0.008*
	(0.003)	(0.004)	(0.004)	(0.004)
3rd Child	-0.011**	-0.007	-0.013*	-0.007
	(0.003)	(0.004)	(0.006)	(0.005)
4^{th+} Child	-0.012**	-0.003	-0.002	-0.018**
	(0.004)	(0.005)	(0.008)	(0.007)
Observations	143087	143087	104021	104021

Notes: Every column reports a separate linear regression. Standard errors reported in parenthesis are robust to within village unit clustering. All estimations include controls for child, parent, household and school level variables. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$.

Table C3: Private Tuition Expenditure

Dep Var. Private Tuition Expenditure		
	(1)	(2)
Girl × 1st Child	-0.011	0.020
	(0.01)	(0.02)
Girl × 2nd Child	-0.003	-0.027*
	(0.01)	(0.02)
Girl × 3rd Child	0.003	-0.012
	(0.01)	(0.02)
Girl × 4th+ Child	-0.011**	-0.016
	(0.00)	(0.03)
2nd Child	-0.034***	-0.008
	(0.01)	(0.01)
3rd Child	-0.056***	-0.006
	(0.01)	(0.02)
4th+ Child	-0.049***	0.025
	(0.01)	(0.02)
Age Group	Younger	Older
Observations	143087	104021

Notes: Every column reports a separate linear regression. The dependent variable is the standardized yearly private tuition expenditure. Standard errors reported in parenthesis are robust to within village unit clustering. All estimations include controls for child, parent, household and school level variables. $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Appendix D

Table D1: Breakdown of callback rates (in %) by city, gender and sector (inexperienced sample)

		Delhi		Mumbai		Chennai	
		BPO	Finance	BPO	Finance	BPO	Finance
Panel A: Female CV							
Bengali	Non-Mother	40.00	43.48	40.00	50.00	29.17	37.50
	Mother	0	7.69	15.00	8.00	12.5	15.00
Khasi	Non-Mother	33.33	23.53	30.00	18.75	19.05	4.76
	Mother	13.63	31.58	20.00	19.05	21.05	6.67
Naga	Non-Mother	35.00	28.57	38.82	20.00	17.65	5.56
	Mother	5.00	13.64	11.11	16.67	26.09	1.11
Panel B: Male CV							
Bengali		35.29	50.00	30.77	41.67	33.33	35.71
Khasi		50.00	14.29	37.50	7.69	25.00	11.11
Naga		36.36	10.00	45.45	8.33	33.33	15.38

546567-L-bw-Majilla
Processed on: 6-8-2020

Table D2: Breakdown of callback rates (in %) by city, gender and sector (experienced sample)

		Delhi		Mumbai		Chennai	
		BPO	Finance	BPO	Finance	BPO	Finance
Panel A: Female CV							
Bengali	Non-Mother	57.14	44.44	42.86	14.29	33.33	40.00
	Mother	12.50	0	50.00	25.00	0	20.00
Khasi	Non-Mother	25.00	14.29	22.22	16.67	33.33	0
	Mother	28.57	25.00	16.67	22.22	0	20.00
Naga	Non-Mother	50.00	0	50.00	20.00	40.00	1.11
	Mother	22.22	14.29	22.22	20.00	20.00	16.67
Panel B: Male CV							
Bengali		66.67	40.00	66.67	33.33	33.33	50.00
Khasi		25.00	0	37.50	0	25.00	0
Naga		25.00	25.00	25.00	25.00	20.00	11.11

546567-L-bw-Majilla
Processed on: 6-8-2020

Table D3: Breakdown of callback rates (in %) by city, gender and sector (experienced and inexperienced sample)

		Delhi		Mumbai		Chennai	
		BPO	Finance	BPO	Finance	BPO	Finance
Panel A: Female CV							
Bengali	Non-Mother	44.44	43.75	40.74	36.84	30.33	39.10
	Mother	3.57	5.26	25.00	12.12	9.09	16.67
Khasi	Non-Mother	30.76	20.83	27.59	18.18	22.22	3.22
	Mother	17.24	29.63	19.23	20.00	14.29	10.00
Naga	Non-Mother	38.46	18.18	35.71	20.00	22.73	7.41
	Mother	10.34	13.79	14.81	17.65	24.24	12.50
Panel B: Male CV							
Bengali		40.00	47.06	37.50	38.10	33.33	38.89
Khasi		43.75	10.00	37.50	6.67	25.00	9.09
Naga		31.58	14.29	40.00	12.50	33.33	13.64

TANMOY MAJILLA

https://sites.google.com/view/majilla
majilla@iss.nl

International Institute of Social Studies
Erasmus University Rotterdam
Kortenaerkade 12, 2518 AX Den Haag
The Netherlands
Cell: (+91) 9748917906

EDUCATION

Ph.D. Economics, Erasmus University Rotterdam, Netherlands, 2020 (submitted)

M.A. Development Studies, International Institute of Social Studies of Erasmus University Rotterdam, Netherlands, December 2016

M.A. Economics, Delhi School of Economics, University of Delhi, 2012

B.A. (Hons.) Economics, Rabindra Bharati University, Kolkata, 2009

FIELDS OF INTEREST

Primary: Development Economics, Labor Economics
Secondary: Behavioral Economics, Economics of Education, Family Economics

PAPERS

"Shadow Education, Intra-Household Resource Allocation and Educational Achievements in India"

"Gender Norms and Motherhood Penalty: Experimental Evidence from India" (with Arjun Bedi and Matthias Rieger), *IZA Discussion Paper No. 11360*.

"Shadow Education and the Quantity-Quality Tradeoff in India"

"Who Benefits from Private Schools in India?"

"Son Preference, Family Size and the Gender Gap in Mathematics in Pakistan"

"Does Signaling Childcare Support on Job Applications Reduce the Motherhood Penalty?" (with Arjun Bedi and Matthias Rieger)

(Pre- PhD) PUBLICATION

"Gray University Degrees: Experimental Evidence from India" with Matthias Rieger. 2020, *Education Finance and Policy,* 15(2), 292-309. https://doi.org/10.1162/edfp_a_00268
Based on M.A. Thesis

WORK IN PROGRESS

"Workplace Arrangements and Female Labor Supply" (with Arjun Bedi and Matthias Rieger)

PRESENTATIONS

2020 Glasgow, CTRPFP Kolkata, IZA – Institute of Labor Economics, The World Bank, Nottingham
2019 UNU WIDER, Erasmus, IIM Bangalore, Indian Statistical institute: Kolkata
2018 Bonn, Maastricht, Jakarta

FELLOWSHIPS AND AWARDS

Research Innovation Fund, ISS of EUR Rotterdam (€ 12720), 2018 – 19
Erasmus PhD Field Work Support Grant (€ 2500), 2019
ISS of EUR Rotterdam Research Fund ((€ 5000), 2017
Netherlands Fellowship Program, Full Scholarship for MA at ISS Hague, 2015-16

WORK EXPERIENCE

Executive Director, Anannya Sabha (a civil society organization), 2014-15
Assistant Manager, IDBI Bank Ltd. (Economist at Treasury), 2012-14

LANGUAGES

Bengali (Native), English (Fluent), Hindi (Fluent)

COMPUTER SKILLS: STATA, MATLAB, R

ABSTRACTS

Shadow Education, Intra-Household Resource Allocation and Educational Achievements in India

Studying intra-household allocation of resources is challenging as direct parental monetary expenditures on individual children is difficult to isolate from aggregate household expenditures. Due to such limitations, previous studies tend to examine parental allocation of resources indirectly from household expenditures or through other non-monetary investments. This article studies intra-household allocation of direct parental monetary expenditures on private supplementary tutoring or shadow education. I show a birth order disadvantage for later-born children in shadow education expenditures and find evidence of disadvantages for girls in every birth order. I attribute these patterns to the well-documented preference for elder sons, which is common in India, and I subsequently test several features which stem out of this preference. The analysis also shows that intra-household disparity in shadow education expenditures accounts for a substantial part of the gender gaps in cognitive test scores. The inequality contributions from the disparities in shadow education expenditures in test scores decrease as one moves down the birth order.

Gender Norms and Motherhood Penalty: Experimental Evidence from India (with Arjun Bedi and Matthias Rieger)

This paper uses a field experiment to study the effect of perceived gender norms on the motherhood penalty in the Indian labor market. We randomly reported motherhood on fictitious CVs sent to service sector jobs. We generated variation in gender norms by signaling community origins of applicants. Employers are less likely to callback mothers relative to women or men without children. Mothers from North-East India experience a smaller motherhood penalty and those of matrilineal origin face no penalty, unlike those of patrilineal origin. We discuss findings in relation to the influence of ethnicity, the Indian context and theories of discrimination.

Shadow Education and the Quantity-Quality Tradeoff in India

This article presents evidence of a quantity-quality tradeoff using plausibly exogenous variation in family size due to elder son preference. Most previous studies tend to examine the relationship between family size and child quality *indirectly* through the *consequences* of the allocation of parental resources, e.g., educational and labor market achievements. In this article, I examine the quantity-quality tradeoff in *direct* parental monetary allocation of resources in shadow education expenditures (or private tuition). Exploiting the elder son preference typically observed in the study context, I instrument fertility with gender of the first child, two-girl and same sex birth. Subsequently, IV estimates show a parity specific quantity-quality tradeoff. However, the effects are heterogeneous and vary in both magnitude and sign. Specifically, in a subpopulation with high unobserved desire for a multi-child family, quantity and quality are complements.

Who Benefits from Private Schools in India?

While the growth of private schooling in India is phenomenal, little is known on what types of students benefit most. In this paper, I investigate attendance decision and heterogeneity in returns to private schools on test scores by exploiting exogenous variation in private school availability. Drawing on a large sample of school going children, I estimate marginal treatment effects in an environment of essential heterogeneity. In addition to a robust positive returns to private schooling across study populations, the results indicate a reverse selection into gains with respect to both observed and unobserved characteristics. Private schools generate large gains for girls and disadvantaged children, but they are less likely to attend private schools. Likewise, exposure effects are inversely related to the unobserved desire to attend private schools. These findings suggest that private school attendance serves as an equalizer - leveling the playing field. Liquidity constraints restrict private school attendance of hard-to-reach children, but they are equally motivated, if not more.

Son Preference, Family Size and the Gender Gap in Mathematics in Pakistan

The gender gap in mathematics has been documented in many contexts, yet little convincing evidence exists to explain it. In this paper, I document a substantial gender gap in mathematics test scores using a large nationally representative dataset from Pakistan. I find that boys and girls have similar levels at the age of five, after which a monotonically increasing gap emerges. I also report a negative birth-order effect with boys outperforming girls in every birth order, but with a weaker gap in later-borns. I show that strong elder son preference, which skews parental resource allocation, is one of the underlying mechanisms of these gendered patterns. Elder son preference induces girls with elder brothers to do worse compared to those without. The gender gap is relatively more pronounced in larger families. In sum, elder son preference coupled with the adverse effects of family size plausibly explain the gender gap.

Gray University Degrees: Experimental Evidence from India (with Matthias Rieger), 2020, *Education Finance and Policy,* 15(2), 292-309.

Scams involving university degrees are flourishing in many emerging markets. This paper studies the impact of gray degrees, or potentially bought academic credentials from legitimate universities, on callback rates to job applications using a resume experiment in India. The experiment varied the type of degree (no, gray and authentic) in online applications to entry level jobs that require no university qualification. We find that gray degrees increase callback rates by 42% or 8%-points relative to having no degree. However, we also document that gray degrees fare on average worse than authentic degrees. These empirical patterns are consistent with a model where employers have beliefs about the authenticity of degrees and are discounting gray degree universities probabilistically. We discuss our findings with respect to the Indian context.

Does Signaling Childcare Support on Job Applications Reduce the Motherhood Penalty? (with Arjun Bedi and Matthias Rieger)

There is substantial evidence that due to perceived childcare obligations, mothers are disadvantaged on labor markets. To what extent can childcare support ameliorate such a disadvantage? To answer this question, we ran a CV experiment in a large Indian city and examined whether signaling access to childcare support may offset the motherhood penalty associated with labor market entry. We randomly varied motherhood, as well as a childcare support signal in online applications sent to service sector jobs in Delhi. Indicating motherhood on a CV led to a 57% or 20 percentage point reduction in callback rates for interviews as compared to non-mothers. A simple childcare support signal (in the form of one-line in a CV) offsets the motherhood penalty by 20% or 4 percentage points. We interpret the findings taking into account the Indian context and with respect to potential sources of discrimination.

