Conditional Cash Transfers and Parental Investment in Daughters: Evidence from India

Sagar Wadhwa*

Abstract

Many states in India rely on conditional cash transfer schemes, to reduce discrimination against females. These programs aim to increase the perceived value of the girl child in general. In this paper, I study the impact of one such program on vaccination status and birth intervals following the eligible girls. Using District Level Household Survey (DLHS) data, I employ a difference-in-difference strategy, with boys as the control group. I find that the probability of vaccination for the eligible girls goes up by 11 percentage points relative to boys in rural areas. However, I do not find evidence of a differential change in the birth intervals following the eligible girls. My results suggest that while parents respond to the direct incentives for such schemes, they do not seem to respond by making investments which are not directly incentivized.

*Sagar_wadhwa@brown.edu, Economics Department and Population Studies and Training Center, Brown University, Providence, RI 02906. I am grateful to Andrew Foster, Bryce Steinberg and Daniel Bjorkegren for their help and advice throughout this project. I also thank Emily Oster, John Friedman, Kenneth Chay and Matt Turner for useful comments. I am also thankful to Girija Borker, David Glick and Vaishali Garga for useful comments and constant encouragement. I also benefited from comments at the Applied Microeconomics Lunch and Development Breakfast at Brown University.

1 Introduction

Gender discrimination exists in various forms in many parts of India. The sex ratio of 943 females per 1000 males[1] (Census 2011) shows the presence of excess female mortality. Gender differences also exist in the investments parents make in their children, including immunization (Pande and Yazbeck 2003), breastfeeding (Jayachandran and Kuziemko 2011), nutrition (Caldwell and Caldwell 1990, Arnold et al. 1998), birth intervals following the child (Maitra and Pal 2001)[2] etc. The reasons for this discrimination against females are both economic and non-economic. Among economic reasons, dowry (the payment made from the bride's family to the groom and his family at the time of the marriage), is one of the major reasons why girls are considered a liability in India (Arnold et al. 1998, Miller 1981, Harris 1993, Gupta et al. 2003). The patrilocal marriage system[3], also gives more incentives to parents to invest in their sons versus daughters. There are also non-economic reasons for son preference including the desire to continue the family name (Jayachandran 2015) and among Hindus who are a majority in India, the son performs the last rites in case of death of a parent (Chakraborty and Kim 2010). Equality between males and females is a "core development objective as it enhances productivity and improves development outcomes for the next generation" (World Development Report 2012) and it is also one of the Millennium Development Goals.

To make progress towards this objective of gender equality, many states have designed policies that incentivize parents to take better care of their daughters[4]. Many of these policies condition the incentives on various forms of investment in the girls such as immunization, school enrollment etc. to bridge the male-female gaps in these areas. One of the ostensible goals of these policies is also to increase the value of the girl child in general. One way one could see that happening is if there is an increase in other investments too that are not directly incentivized by such schemes. The idea is that if parents start valuing daughters more, they may start making investments in their daughters which they might have made only in their sons. It is important to know whether and to what extent, these schemes are successful in promoting the incentivized as well as non-incentivized investments in the girl child.

In this paper, I evaluate the impact of one such program, called *Ladli*[5], in the state of Haryana. Ladli scheme was launched in 2005. It provides financial assistance to all the *second* daughters in a family born on or after 20th August, 2005. The program requires the girls to be immunized and enrolled in school at appropriate age. An amount of Rupees 5,000[6]

[1] The estimated global average is 984.

[2] Verma et al. (1990) find that average interval was 1.74 years for women who had a prior girl child as compared to 2.55 years for women who had a prior boy child. The probability of a child being followed by a very short interval (less than 18 months) has also been found to be higher for girls as compared to boys (International Institute for Population Sciences (IIPS) 1995, 2000, 2007). Other papers also show that in settings with son preference, parents who just had a son are more likely to wait longer to have the next child (Rindfuss et al. 1982, Trussell et al. 1985, Arnold, Choe, and Roy 1998, Retherford and Roy 2003).

[3] A system in which the girl moves to her husband's house after marriage.

[4] Some examples are Kanya Jagriti Jyoti Scheme in Punjab (1996), The Girl Child Protection Scheme in Tamil Nadu (1992), Ladli in Delhi(2008).

[5] The word *Ladli* literally means beloved daughter

[6] This amount is roughly half of the monthly per capital income in Haryana.

(approximately \$85) per year is deposited, with an interest rate of around 8% in the name of the second daughter and her mother. The accumulated amount of Rupees 96,000 (approx. \$1600) is redeemable once the girl is 18 years old, provided she is unmarried. To assess how the conditional transfer affects the incentivized investment, I study the vaccination status of second daughters and to assess how the incentives affect the value of the child, I study the birth-spacing between the second daughter and the next child.

Shorter birth intervals (generally shorter than 2 years) have been found to have perverse effects on both the subsequent child as well as on the previous child (WHO report on birth-spacing 2005). Buckles and Munnich (2012) using US data find that shorter birth intervals decrease test scores of the older sibling by 0.65 SD. A short subsequent interval can have negative effects on the child in at least three ways (Muhuri and Menken 1997). Due to the negative relationship between breastfeeding and probability of conceiving, children with a shorter subsequent birth interval are likely to be breastfed less (Jayachandran and Kuziemko 2011). Secondly, if the mother becomes pregnant soon after the birth of a child, she would not be able to provide adequate care to the earlier child. Short subsequent intervals can "strain the quality and quantity of maternal care" devoted to the index child (Palloni and Tienda 1986). If two children are spaced closely together, there is also likely to be competition for limited family resources (Omran 1981). Girls may be even more affected by this competition effect if the next child is a boy as parents may be more likely to spend resources on the newborn boy. Therefore, I interpret a longer birth interval to be indicative of increased investment in previous child.

Even though Ladli requires only investment in terms of immunization, school enrollment and delay in marriage, there are reasons why it may also lead to investment in non-incentivized outcomes. This could happen if there is complementarity between vaccination and other investments in the health production function of the child. Vaccination improves the health of the girl child by making her immune to more diseases which could potentially raise the return to other investments in the child as well. This complementarity between investments can create an incentive for the parents to invest in other ways (for example, have longer subsequent birth intervals) in the child as well (Cunha and Heckman 2007, Aizer and Cunha 2012).

Another mechanism through which the policy may lead to higher non-incentivized investments in the child can be through the income effect. The transfer received from this policy could increase the demand for daughters' health[7]. The birth interval after the birth of the second daughter could also increase as a result of the mortality effect of shorter birth-spacing. Since the girl needs to be alive until 18 years of age to receive the benefit of the scheme, parents might breastfeed her more and wait longer for the next child to increase the likelihood of her survival.

The policy could affect birth-spacing through the relative preference for sons versus daugh-

[7]Since the amount is received only in future and if the parents are credit constrained, they may not be able to make monetary investments in the children. However, using birth interval as a measure of investment helps in that regard since waiting longer after the birth of a child is largely a non-monetary investment.

ters. If the primary reason for a stronger preference for a son is the economic liability of dowry associated with a daughter, the scheme, by providing part of the dowry expenditure, may be able to affect this preference and the birth-spacing after a second daughter may move closer to what the spacing would be after a son is born in the family.

However, it is possible that the policy may not be able to affect these investments. There may be a lack of complementarity between vaccination and other investments in the health of the girl. The amount of transfer received in the future may not affect investment decisions today. If the son preference is sufficiently strong, providing financial support for daughters may not be able to affect the relative preference for sons versus daughters. Hence, a priori, it is not clear if the policy would have an effect on how long parents wait for the birth of the next child.

For this analysis, I use data from three rounds of District Level Household Surveys (DLHS), DLHS-2 (2002-04), DLHS-3 (2007-08) and DLHS-4 (2012-13). The DLHS is one of the largest demographic surveys in India. I primarily use DLHS-4 as the post-Ladli data and DLHS-2 as the pre-Ladli data. I use whether or not the child has a vaccination card as my primary measure of vaccination. I calculate the birth intervals in months using reported birth month and birth year of children.

To identify the impact of the program, I estimate intent to treat effect using a difference-in-difference strategy. Specifically, I compare the post-program period gap between girls who have exactly one elder sister and no brother (i.e. the second daughters in the family but with no brother) and boys who have exactly one elder sister and no brother to the same gap in the pre-program period. The identifying assumption is that between the pre-program period and the post-program period, the only thing that changed for second daughters *relative* to the control group of boys is that the second daughters were eligible for the conditional transfer in the post-policy period. I limit my analysis to this sub-sample only to make the comparison between boys and girls who have the same number and sex-composition of siblings, as there is evidence that the number and sex-composition of siblings of a child affect the probability of vaccination of that child (Singh and Parsuraman 2014, Pande 2003). This empirical strategy allows me to take out the effect of any trends common to both males and females in vaccination and birth intervals. However, since this empirical strategy does not take out the effect of any national level policy or initiative targeted only towards girls, I compare my results to the results for the same sub-sample in the neighboring state, Punjab[8].

For vaccination, I estimate a linear probability model for the child having a vaccination card. For estimating the effect on birth intervals, I use a discrete hazard model. Since not all mothers have had the third birth by the end of the survey, estimating the birth interval duration using OLS can lead to bias (Foster et al 1986). Instead, I use the hazard analysis framework wherein I look at the probability of the birth of third child in a given month t, conditional on the third

[8]Since the results could also be due to some general change in attitude towards daughters, occurring at the same time in Haryana, I also do the analysis for the first daughters' sample in Haryana.

child not being born until month $t-1$.

I find evidence that in rural areas, the probability of having a vaccination card increases by around 11 percentage points for second daughters relative to the control group of boys, this accounts for nearly 100% of the gap in the pre-program period between the two groups. The same result holds when I use number of vaccines received by the child as another measure of vaccination. However, for birth intervals, even though there has been an increase in birth-spacing over time, I do not find evidence of a differential increase for the treatment group of girls relative to the control group of boys. These findings suggest that while people respond to the conditions for the conditional transfer, people do not necessarily respond in other favourable ways, that are not incentivized.

To explain the qualitative nature of my results, I develop a simple theoretical model in which the utility of the parent depends on the health of the child and the consumption good. I assume a CES health production function where vaccination and other investments in child's health can be either gross substitutes or gross complements. I model the conditional cash transfer as providing a subsidy to provide vaccination to the child. I show that the effect of this subsidy on other investments in child's health depends on the elasticity of substitution between vaccination and other investments, and how parents value the health of the child.

My paper makes the following contributions. There is a growing literature on the impact of conditional cash transfers (CCTs) in developing countries. There is evidence of increase in investments in children as a result of CCTs in many parts of South America (Gertler 2004, Lagarde et al 2007, Ranganathan et al 2012 among others). I contribute to this literature by estimating the impact of a unique conditional cash transfer program where the benefit accrues after a long lag and the type of conditionality is also different from other CCTs (i.e. the birth of a second daughter in the family and that daughter being unmarried until 18).

There has been some recent work on the impact of such long term conditional cash transfers on fertility. Mazumadar (2012) also looks at the impact of Ladli, on the probability of having a daughter in the family. Using DLHS-2 and DLHS-3 data and doing difference-in-difference with Punjab as the control group, she does not find a significant change in the probability of having a daughter as a result of the policy[9]. Even if there is no effect on the probability of having a daughter, it is important to study whether the conditonal transfer is successful in generating a positive change in behavior towards girls who are enrolled in the scheme.

A very closely related work is by Sinha and Yoong (2009) who study the effect of a similar program (*Apni Beti Apna Dhan* (ABAD)[10]) in Haryana on investment in human capital of daughters. They find positive effects of the scheme on sex ratio of living children, on post-natal health investments in the girl child and some mixed evidence on improvement in health status

[9]The paper finds an increase in the probability of having a daughter for border districts but it finds this result for the women who have 0 or 1 daughter and not for women who have exactly one daughter, and as the author herself admits, this result "does not clearly reflect the impact of the policy as the empirical evidence suggests that parents do not usually abort their first daughter".

[10]The scheme was launched in the state in October, 1994 and was in effect till 1998.

of the girl child.

Apart from looking at a different outcome of investment i.e. birth interval, my paper differs from Sinha and Yoong (2009) in two aspects: first, since they do not observe the BPL status of the family in the data, they construct proxy measures based on the assets available to estimate the eligibility status. They themselves note that the results rely on correctly specifying the eligibility status and they emphasize the need for further research to estimate the impact of such programs. For my paper, this problem does not arise as all the families of Haryana whose second daughter is born on or after 20th August, 2005 are eligible. Second, Ladli is also different from ABAD as it does not provide any cash at the birth of the girl child, all the money is given only when the child turns 18 while ABAD gave Rs. 500 within 15 days of the birth of the girl child. The income effect in the two cases can be quite different.

2 Background and Program

With a per capita income of around Rs. 120,000 (approximately $2,000)[11], Haryana is one of the richer states in India[12] but with only 879 females per 1000 males[13], the sex ratio in Haryana is one of the worst in the country. The literacy rate in Haryana is also 85% for males as compared to only 67% for females. Filmer, King and Pritchett (1998) show that the girls in Haryana are two times more likely to die as compared to boys between the ages of one and four. To address the problem of discrimination against girls, the state government has relied on financial incentives by launching *Apni Beti Apna Dhan* in 1994 and *Ladli* in 2005[14].

The objectives of the Ladli program were to "combat female feticide, increase number of girls in families, improve sex ratio, and raise status of girl child in the society". According to the scheme, all the residents of Haryana, whose second girl child is born on or after 20th August, 2005, are eligible for financial assistance of Rs. 5000 per annum for a period of 5 years (approximately $85) on the birth of second girl child. The money is conditional on parents ensuring proper immunization of the girl child and school enrollment at appropriate age[15]. The money is not given in terms of cash at that time, rather it is deposited in a Life Insurance Corporation (LIC) bond at an interest rate of roughly 8% in the name of second girl child and her mother which the girl can get in cash when she turns 18, provided she is unmarried. The amount at that time is likely to be approximately Rs. 96,000 ($1600). The amount given per year is roughly 50% of the monthly per capita income and the final amount is roughly 80% of the annual per capita income in the state. As of 2010, the total number of beneficiaries was 227, 295 (Sekher 2012). The rules of the scheme state that the applications for the program

[11] Statistical Abstract 2011-12, Haryana.

[12] India's per capita income was $1,570 in 2013 according to World Bank.

[13] Census 2011

[14] http://wcdhry.gov.in/new_schemes_f.htm

[15] The program states that both the daughters should be immunized. However, interactions with the people responsible for implementing the scheme at the district level have suggested that while parents were encouraged to immunize the first daughter as well, not immunizing the first daughter did not affect the enrollment in the program.

need to be made to the Anganwadi workers in rural areas and the Health Supervisor and health staff workers in the urban areas.

3 Data

For this paper, I primarily use data from 3 rounds of the District Level Household Surveys (DLHS), round 2 (2002-04), round 3 (2007-08) and round 4 (2012-13). DLHS is a large scale, repeated cross-section survey conducted in a representative sample of households throughout India. The DLHS, executed by Indian Institute of Population Sciences (IIPS), collects data on family health, utilization of public health services, demographic composition and socioeconomic characteristics including, religion, caste and asset ownership.

In Haryana, DLHS-2 interviewed 20,205 households between 2002-04, DLHS-3 interviewed 21,406 households between 2007-08 and DLHS-4 interviewed 33,772 households in 2013. The DLHS asks retrospectively about all births from reference date through the time of interview for ever-married women (ages 15-49). The reference period is generally 4-5 years before the survey. For example, for DLHS-3, survey asks about all births since 1st Jan, 2004. For DLHS-2, for half of the districts interviewed in 2002, reference priod is since Jan, 1999 and for the other half interviewed in 2004, the reference priod is since Jan, 2001[16]. The DLHS-4 round reports all birth since Jan, 2008.

Even though the scheme came out in 2005, my interactions with the officials implementing the scheme at district level suggest that the coverage of the scheme has been better since 2007. Therefore, in my primary specification, I use data on children in DLHS-4 (i.e. born since Jan 2008) as the post-program data and data on children in DLHS-2 (i.e. born before 2004) as the pre-program data. I also show results using 2005 as the cutoff period in the robustness section. For my analysis on vaccination, I use whether or not the child has a vaccination card as my primary measure. I use this measure as the parents would need to show the vaccination card to show that they have fulfilled the condition for the transfer. I also use the number of vaccines received by the child as another measure for robustness check. I use month and year of the birth of the child reported by the mother to construct birth intervals between the children. I also use National Family Health Survey (NFHS)[17] data (2005-06) for birth interval analysis.

3.1 Estimation sample

To compare the gap between girls who have exactly one elder sister and no brother (treatment group) and boys who have exactly one elder sister and no brother (control group) in the post-program period to the same gap in the pre-program period, I restrict my data to those mothers

[16]The DLHS-2 round was conducted in 2 phases where half of the districts throughout the country were surveyed in 2002 and other half were surveyed in 2004. For selection of districts for phase-1, random sampling was adopted within a state.

[17]The DLHS and the NFHS are similar in terms of the selection of respondents, the conduct of interviews, and the questionnaires used. Sample sizes, however, are much larger for the DLHS since it is also representative at the district level. I keep the treatment and control children born between 2000 and 2005 from the NFHS data.

who have exactly one daughter and no son before the reference period and have at least one child since the reference period. If the second child is a girl in the post-Ladli period, she is eligible for the scheme as she is the second daughter in the family. If the child is a girl in the pre-Ladli period, she is not eligible for the scheme and if the child is a boy, he is not eligible for the scheme in either of the rounds. I limit my analysis to this sub-sample only to make the comparison between boys and girls who have the same number and sex-composition of siblings, as there is evidence that the number and sex-composition of siblings of a child affect the probability of vaccination of that child (Singh and Parsuraman 2014, Pande 2003).

For birth intervals, I look at the probability of the birth of a child after the treatment group of girls and control group of boys. However, for the post-policy period (DLHS-4), the reference period was close to 70 months and for the pre-policy data (DLHS-2), it was close to 40 months. So, the probability of the mother having two births in the reference period (i.e. having a third child) in the post-policy period is higher than that in the pre-policy period. To take care of this issue and make the comparison more equal, I censor the post-policy birth intervals at 40 months i.e. if a mother has the third birth within 40 months of the birth of the second child, that interval remains unchanged but if a mother has the third birth in more than 40 months after the birth of the second child, I keep the interval at 40 months and mark it as censored. Finally, I also drop those mothers who report being sterilized and have not had a third child though I keep those who report being sterilized and have had a third child. With this, I am left with 2072 observations for the vaccination analysis and 1913 observations for the birth interval analysis.

3.2 Summary Statistics

In table 1, I show the summary statistics of the final estimation sample. The first panel shows the statistics for the dependent variables. As we can see, the probability of having a vaccination card is lower for females as compared to males in the pre-program period. However, in the post-program period, the probabilities are very close to each other for males and females. We can observe the same pattern in the number of vaccines received by the child[18]. The third row from panel A shows that the probability of being followed by a birth interval less than 24 months long is higher for a female than a male This pattern is true also for the post-program period although the probability declines for both males and females. The fourth row shows that the number of women who have the third child is higher if the second child is female as compared to if the second child is male.

The second panel shows the means and standard deviations for the various covariates for both treatment and control groups, before and after the program. The final column shows the difference-in-difference estimates for the covariates. These estimates are just the difference

[18] The variable number of vaccines includes vaccines for BCG (Bacillus Calmette–Guérin), DPT (Diphtheria, Pertussis and Tetanus), Polio, Measles, Hepatitis B and Vitamin-A, and takes values 0-6.

between the gap between treatment group and control group in the post-program period and the same gap in the pre-program period. Since my identifying assumption is that the only thing that is changing differently between the pre-period and the post-period for the treated group of girls relative to the control group of boys is that the girls are eligible for the scheme in the post-program period, in this column I test whether the covariates are changing differently for the treatment group relative to the control group. As we can see, only thing that seems to be changing differently is the probability of the household having a *pucca* house (a well-built house)[19]. The variables for the socio-economic status of the household including mother's education, father's education[20] and dummies for various assets show an increase over time, however, the increase is not different for the treatment group relative to the control group as the diff-in-diff coefficients are close to 0 and not significant.

4 Empirical strategy

The empirical strategy I use here is difference-in-difference. I compare the gap between girls who have exactly one elder sister and no brother and boys who also have exactly one elder sister and no brother in the post-program period to the gap between the same groups in the pre-program period. As mentioned earlier, I do this to compare boys and girls with the same number and sex composition of siblings. My estimating equation is as follows:

$$y_i = \alpha + \beta Post_i + \gamma Female_i + \delta Female_i * Post_i + \rho X_i + \epsilon_i$$

where y_i is the outcome variable for the child i, $Post_i$ is a dummy which is equal to 1 in the post-Ladli period and 0 otherwise, $Female_i$ is a dummy which is equal to 1 if the child is female and 0 otherwise, $Female_i * Post_i$ is the interaction of female dummy and the post-period dummy, X_i refers to the various socio-economic controls for the child i. The difference-in-difference coefficient is δ.

The identifying assumption here is the following:

$$E(\epsilon | Post, Female, Post * Female, X) = 0$$

In words, the identifying assumption is that the only thing that is changing differently between the pre-period and the post-period for the treated group of girls relative to the control group of boys is that the girls are eligible for the scheme in the post-program period. In the final column of table 1, I show that most of the observable variables are not changing differently for the treatment group relative to the control group. However, it could be the case that the unobservables are changing differently for the treatment group. While I cannot directly

[19] I control for the type of house as well as the other variables listed in table 1 in my analysis.

[20] Around 453 observations for mother's education and around 229 observations for father's education are missing in my data. While controlling for these variables, I replace the missing values by 0 and add a dummy for the missing values to see if the result is affected by replacing them by 0. The coefficient on the missing education dummy was never statistically different from 0 and my results are robust to whether or not I replace the missing values by 0. I replace them by 0 to increase precision of the estimates.

test for the differential change in unobservables, I try to provide some evidence against it by performing falsification checks in the results and robustness sections. The other assumption is that treatment and control groups should have parallel trends for the outcome variables before the policy was launched. I present the figures for the trends in the results section.

To interpret the results of the difference-in-difference as the impact of the policy, it should also be the case that there was no other scheme during this period specifically targeting the girls which could have led to the differential change in girls relative to boys. While *Janani Suraksha Yojana (JSY)*[21] was launched throughout India in 2005 and there is some evidence of an increase in vaccination due to JSY (Debnath, 2013), the scheme applied to both boys and girls and there is no evidence of a differential effect on girls due to JSY.

5 Hazard model for birth intervals

The question of interest here is whether parents wait longer for the birth of the next child after the birth of a girl when financial incentive is being given for that girl. One way to answer this question could be to use the duration between second daughter and the next child as a dependent variable and see whether and how that changes as a result of the policy. However, the problem with that approach is that not all mothers have had a third birth by the time survey ended. In other words, there is problem of censoring. The censored mothers may have a third birth after the survey ended or may not have a third birth at all and not including those would lead to selecting only shorter birth intervals and also, missing out on a lot of information.

To take care of this problem, I use hazard model analysis. Rather than using the duration between two events, hazard models look at the rate at which people move from one state to another and hence, also take the censored observations into account. Applying the hazard model terminology to the current context, failure would refer to the event where the third child is born and survival would refer to the event where the third child is not born. Following Foster et al (1986), I do discrete hazard model analysis where the hazard rate in any month is the proportion of women who have a third birth in that month out of those women who have two children (with the first child being a girl) and have not had a third birth by then.

For this analysis, I organize my data such that each woman contributes one observation for each month since the birth of the second child till she has a third birth or till the end of survey. Each observation consists of covariates and a binary dependent variable which takes value 1 if the third child is born in that month and 0 otherwise. For example, if a woman has her second birth in August, 2012 and she has her third birth in August, 2013, she would contribute 12 observations where for the first 11, the dummy variable for having third birth would be 0 and for the 12th observation, it would be 1. If the same woman did not have her third birth by the end of the survey i.e. August, 2013, she would still contribute 12 observations, with the

[21] The scheme incentivized women to deliver babies in medical facilities by giving a cash transfer.

difference that now the dummy variable for the third birth would be 0 in the 12th observation as well.

The estimating equation for this model is as follows:

$$logit(h) = \alpha + \beta * Female + \gamma * Post + \delta * Female * Post + \rho * X \tag{1}$$

where h is the hazard rate of having a third child, female is a dummy for second child being female, post is a dummy for the post-Ladli period and X refers to other covariates which could potentially affect the outcome variable. Here, α gives the log odds for having a third child when the second child is a male in the pre-Ladli period, β gives the change in log odds of having a third birth after the second child is a female as compared to a male in the pre-program period, γ gives the change in log odds of having a third child in the post-Ladli period which is common to both males and females and the parameter of interest here is δ which shows whether the odds of having a third child have changed differently over time after the treatment group as compared to the control group.

Note that since I am working here with probabilities of a third birth rather than the duration between second and third birth, a positive coefficient indicates a *decrease* in the length of birth interval and a negative coefficient indicates an *increase* in the length of birth interval. So, if the birth interval after the second girl child increased relative to the boy child, it will show up as a negative coefficient on the interaction term.

6 Vaccination Results

Since more than 60% of my sample is rural, I do my analysis for full sample and rural sample separately [22]. Focusing on rural areas separately also makes sense as there is evidence that not only are the vaccination rates lower in rural areas as compared to urban areas, but the gender gap in vaccination is also higher in rural areas (Pande and Yazbeck 2003). So, it is important to see whether the transfer is successful in incentivizing people to make better investment in their daughters in rural areas where the gender based discrimination is higher than urban areas.

Figure 1 shows the probability of having a vaccination card in rural areas by year of birth. The blue line shows the probability for males (i.e. second child who is male) and the red line shows the probability for second daughters.[23]The figure shows that even though the two lines are not completely parallel, there is a persistent gap in the probability of having a vaccination card before 2005. But the figure shows that the gap disappears after 2005 and the vaccination rates for girls catch up with boys.

Table 2 shows the vaccination results for the rural areas only. Here, I estimate a linear probability model. The first column shows the basic difference-in-difference specification. In

[22] Also, India is still largely a rural country, with close to 70% population living in rural areas (Census 2011).

[23] I use the two year averages to reduce the noise since I do not have enough number of observations every year. I show the one year figure in the appendix.

the first column, we can see that there is a gap in vaccination rates between males and females, indicated by the negative coefficient on female. It shows that in the pre-policy period, the probability of having a vaccination card is 11.8 percentage points lower if the second child is a female as compared to a male. The coefficient on post period shows that over time, there is an increase in the probability of having a vaccination card for both males and females. The difference in difference estimate is the coefficient on the interaction of female dummy and the post period dummy which shows that the vaccination rates have increased by almost 11.6 percentage points for second daughters *relative* to the control group of males which accounts for nearly 100% of the gap observed between males and females in the pre-program period. This coefficient is statistically different from 0 at the 5% significance level.

Now, it might be the case that the second daughters in the post period come from differentially wealthier households as compared to the second children who are males. The vaccination rates of the treatment group could be increasing due to this wealth effect and not as a result of the program. However, as shown in table 1, apart from the probability of *pucca house*, there is no evidence of a differential change in the assets for the treatment group relative to the control group. To test if these results are due to differential wealth effects, I control for type of house and other household assets and amenities in column 2. The difference in difference estimate falls only slightly between column 1 to column 2 and is significant at 5% level. It could also be the case that the effect is coming from some districts which are developing in vaccination and child care and is not a result of the scheme. To check this, I also control for district dummies in column 3 to use within district variation and as we can see, the difference in difference estimate does not change much by adding district fixed effects.

Table 3 shows the results for the total sample. The coefficient on female dummy shows that the probability of having a vaccination card is lower for the treatment group of girls as compared to the control group of males in the pre-program period by 7.98 percentage points. The coefficient on post period dummy shows the evidence of an increase in the probability of having a vaccination card for both males and females over time. The difference-in-difference estimate, even though it is positive, it is not statistically different from 0. The coefficient falls slightly as I add more controls but I still cannot reject that there is no differential effect for females.

In table 4, I use the number of vaccines[24] as another measure of vaccination. Here, the first 3 columns give the results for the entire sample and the last 3 columns give the results for the rural sample. The coefficient on female shows that for the entire sample, second daughters on average receive around 0.45 vaccines fewer than the case if the second child is boy and for the rural areas, this gap is higher. The post period coefficient suggests an increase in the number

[24]The number of observations for these regressions is less than the number of observations for the vaccination card regression as some observations are missing for whether or not the child received a particular vaccination, in the DLHS-4 data. For some children, the reported answer to the question of a particular vaccine was that the mother 'does not know'. I treat those cases as not receiving that particular vaccine.

of vaccines for both males and females over time. The difference in difference estimate is 0.45 for the entire sample without any controls and it falls to 0.36 as district dummies and the household assets controls are added. For the rural sample, the difference-in-difference estimate is even higher and indicates that the number of vaccines received by the treatment group increase by 0.47 relative to the control group. These results, like the results using vaccination card as a measure, also suggest that the differential increase in vaccination for the treatment group has been nearly 100% of the gap in the pre-program period in both rural sample as well as the entire sample.

Since the pre-Ladli period and post-Ladli period in my analysis are many years apart, it could be that there is a national level change in the mindset of parents towards daughters which might be driving the result in the vaccination rates for second daughters in Haryana as well. To test for this, I run the same specification for the neighboring state Punjab which also has similar sex ratio as Haryana. Even though Punjab also launched a program called *Balri Rakhshak Yojana* in year 2005-06 to improve the sex ratio in the state, but as of 2008-2009, the program was widely unattractive, with only 212 beneficiaries since it was launched (Government of Punjab, 2009)[25]. So, if the results in Haryana are due to national level trend or some national level policy targeting only girls, they should hold for Punjab as well.

The results for Punjab are given in table 5. The first column shows the results for entire Punjab. The difference in difference estimate is 0.02 and is not statistically different from 0. Since the results for Haryana are stronger in the rural areas, column 2 shows the results for rural areas in Punjab. The coefficient is 0.008 and is not statistically distinguishable from 0. Column 3 and column 4 show the results for number of vaccines for entire Punjab and the rural areas respectively and here too, we cannot reject the hypothesis that the difference-in-difference estimate is 0.

It could also be the case that there is a change in mindset towards daughters over the period in Haryana only and the vaccination rates are going up for all girls in Haryana. To check if this is the case, I take another sample in Haryana where the first child is a son. In this first child son sample, I compare the vaccination rates for the second male child to the vaccination rates for the case when the second child is female[26]. These results for rural areas are given in table 5. Column 5 shows the results for the vaccination card and the column 6 for the number of vaccines and we can see that in both the columns, we cannot reject the null that the difference-in-difference estimate is equal to 0.

I find stronger evidence of increase in vaccination for rural areas as compared to the total sample. Apart from the fact that the male-female gap in rural areas is higher, another reason could be that the information flows about the program are likely to be better in rural areas as compared to urban areas. This is because in the rural areas, the scheme is being implemented

[25] http://pbplanning.gov.in/pdf/New%20Schemes%20for%20Girl%20Child.pdf

[26] I take this sample so that the children I am comparing are the second children in the family as my baseline treatment and control sample are also second children.

by the Anganwadi workers[27] (AWW) who are the residents of the same village, maintain a record of the births in the village and can therefore be more effective in informing and getting people enrolled in the program, due to which the coverage could be better in rural areas[28].

7 Results on birth spacing

The hazard curves for the birth interval analysis are given in figure 2. The x-axis shows the number of months since the birth of second child. The y-axis shows the conditional probability of the third child being born in month t, given the child is not born till month $t-1$. The top panel shows the hazard curves for the case where the second child is male, with the left one showing the hazard curve in pre-program period and the one on the right showing the hazard curve for the post-program period. The bottom panel shows the curves for the case where second child is female, with again the one on the left showing the graph for pre-program period and the one on the right showing it for the post-program period.

The probability of having a third birth is 0 for the first 9 months after the birth of the second child, as it should be. We also see that the hazard rate for having a third child is much higher if the second child is female as compared to if the second child is male. The hazard curve for females is always higher than hazard curve for males, in both pre and post program periods. This difference in the hazard curves for males and females clearly shows that parents are indeed having a different birth interval after the birth of a daughter as compared to after the birth of a son.

There also seems to be a decline in the conditional probability of a third birth in the post-program period for both males and females which suggests that birth intervals have increased for both males and females over the years. But if the birth intervals after the treatment group have increased differentially after the policy as compared to the control group, we should see that the decrease in the bottom panel from pre-program period to post-program period should be more than the decrease in the top panel from pre-program to the post-program period. However, it is hard to tell from this figure whether this is the case. So, to test it further, I take the ratio of hazard rates for males to hazard rates for females in both pre and post-program periods and plot them in figure 3.

In figure 3, the graph on the left is just the ratio of the graphs on the left of figure 2 and the one on the right is the ratio of the two graphs on the right of figure 2. Most of the data points are below 1 because the hazard curve for females lies above the hazard curve for males. If the decrease in hazard curves for females was more than the decrease for males, the hazard ratio graph for the post-program period should be above the pre-period graph. But the two

[27] Anganwadi means courtyard shelter, these Anganwadis were started as the mother and child care centers by the Indian government in 1975 as part of the Integrated Child Development Scheme. Some of the functions of these workers include providing care for the newborn babies, providing immunization and supplementary nutrition to children.

[28] One way to test this could be to look at the heterogeneity in outcomes by the availability of Anganwadi workers. However, in DLHS-2 (2002-04) data, the proportion of people in rural areas having access to Anganwadi workers is more than 90% in each district of Haryana and there is not much variation in the availability of Anganwadi workers.

curves look quite similar, suggesting that the birth intervals have not increased differentially for females as compared to males over the period.

To explore further, I show the Kaplan-Meier survival curves for the probability of third birth in figure 4. The x-axis is again the number of months since the birth of the second child. The y-axis shows the probability that the third child is *not born* till month t. Recall that, the failure in my analysis refers to the event of third child being born while the survival refers to the event of third child not being born. In figure 4, the blue line shows the probability that the third child is not born till month t when the second child is male in the pre-program period. The red line shows the probability for males in the post-program period. The green line shows the probability that the third child is not born till month t when the second child is a female in the pre-progam period and the orange line shows the same probability for females in the post-program period.

The probability that the third child is not born till 9th month is 1 for all cases and then, it starts falling. In this figure also, we can see that the survival curves for males are always higher than the survival curves for females, showing that for any duration after the birth of the second child, the probability that the third child is not born is higher when the second child is male as compared to female. We can also see that the survival curves for the post-program period lie above those for the pre-program period, for both males and females, again suggesting that the birth intervals have increased for both males and females from pre-program period to the post-program period.

The figure also shows that not all mothers have had a third birth as we don't have data in figure below the median for the females and below the 75th percentile for the males, suggesting that the probability of a third child being born if the second child is female is around double of the probability of a third child being born if the second child is male. We can see in the figure that the Kaplan-Meier curves do suggest that there is some evidence of an increase in birth interval after the second daughter relative to son, particularly in the 20-30 months interval as the gap in pre and post survival curves for females seems larger than the gap for males. I test whether the gap is statistically different from 0 in the hazard regression and the results are presented in the table 6.

I present the basic hazard model results in table 6. Column 1 shows the basic difference in difference specification. The coefficient on female shows that the log-odds of having the third birth in any given month increase by 0.91 if the second child is female as compared to male in the pre-policy period. This means that the odds of having a third birth in any given month after a female are $exp(0.91) = 2.48$ times the odds of a third birth after a male. The coefficient on post indicates a decline in the probability of having a third birth in any given month, again suggesting a lengthening of birth intervals over time. The difference in difference estimate is the coefficient on Female*Post. If the birth intervals over the time period increased differentially for the treatment group relative to the control group, this coefficient should have been negative but in fact, it is slightly positive and is not statistically different from 0. Columns 2 and columns 3 add district fixed effects and controls for mother's education, father's education, mother's age, mother's age at marriage and household assets and amenities. The difference-in-difference estimate increases slightly but is again, not statistically different from 0. The final column shows the marginal effects for the specification in column 3.

Since the evidence of increase in vaccination is stronger in rural areas, I also estimate the birth interval specification separately for rural areas. The results are presented in table 7. For rural areas also, we can see that the odds of a third birth are higher in any given month if the second child is a girl as compared to a boy. From the coefficient on Post dummy, we also see some evidence of an increase in birth intervals over time but the coefficient is not statistically different from 0. The difference-in-difference coefficient is even higher than the one for the entire sample suggesting that the increase in the birth intervals after females over time has actually been less relative to males but we cannot reject that the coefficient is equal to 0.

It could be the case that some national level shock negatively affected the subsequent birth intervals only after females during this time period in which case even if there is a positive effect of the program on birth intervals after females, my difference in difference specification would not be able to detect that. To check if this is the case, I estimate the hazard model regression for the same sample in Punjab. If it is the case that some national shock negatively affected the birth intervals following the girl children, the difference in difference coefficient for Punjab should be positive and significantly different from 0. The results are presented in table 8. Columns 1 and 2 show the results for Punjab as a whole and only rural areas in Punjab respectively. These results are very similar to those in tables 6 and 7. The difference-in-difference estimate for Punjab is positive just like the Haryana sample but again, it is not statistically different from 0[29].

It could also be the case that some other unobservable shock in Haryana negatively affected the subsequent birth intervals after all females during this time period. In that case too, even if the policy had a positive effect on the birth intervals after second daughters, the effect would be canceled out by the negative shock to birth intervals after girls throughout the state. To check if this is the case, I estimate the difference-in-difference hazard model for the sample in Haryana in which the first child is a boy. The results for the full state and only rural areas are given in the columns 3 and 4 of the table. As we can see, the difference-in-difference estimate (0.11) is close to what we had in table 6 (0.08) and the estimate is not statistically different from 0.

Even though I cannot reject that there was no differential increase in the birth intervals following the eligible girls and difference-in-difference coefficient is never negative, the coefficient is not estimated very precisely and the standard error is large. However, even if I take the most

[29]The triple difference with Punjab as another control group gives a negative estimate (-0.29) but the standard error is still high (0.25).

negative estimate of the confidence interval on the difference-in-difference coefficient, I can rule out that the program reduced the pre-program period gap in birth intervals by more than 33%, while for vaccination, the program bridges the pre-program period gap by almost 100%.

8 Robustness

8.1 Vaccination robustness

To show that my results are not dependent on using DLHS-4 data as the post-program period data, I show the vaccination results using years after 2005 as the post-program data in table 9, column 1. The coefficient on the interaction of female dummy and the post period dummy shows that the probability of having a vaccination card increased by 10.8 percentage points for treatment group of girls relative to the control group of boys. This estimate if close to the difference-in-difference coefficient in table 2 (11.5 percentage points).

Figure 1 shows that the vaccination rates are trending up over time for both males and females and if the trend for females is different from the trend for males, it could be the case that the differential increase in vaccination rates for females is due to the differential trends only and is not due to the effect of the program. To test for this, I control for common linear trends and differential linear trends in columns 2 and 3 respectively. As column 2 shows, the linear year variable is highly significant, showing the increase in vaccination for both males and females with each year. However, there is not much change in the difference-in-difference estimate from column 1 to column 2.

In column 3, I also include the interaction of female dummy with the year variable to allow for differential trends for males and females. As we can see, the interaction of female dummy and year is very close to 0, suggesting that there is no evidence of differential trends for females relative to the control group of males. The difference-in-difference estimate increases slightly from column 2 to column 3, however, its standard error almost doubles as now the interaction of female dummy and the post-2005 dummy, and the interaction of female dummy with the year variable are highly correlated.

If the trends in vaccination rates are different for males and females only before the program was launched, then again, it could be the case that the differential increase in vaccination rates for females is because they were on a different trend and not as a result of the program. To test for it, I regress vaccination on year, interaction of female and year, and the female dummy only for the pre-2005 data. I then predict the residuals from this regression and use these predicted residuals as the dependent variable in my difference-in-difference specification. The results are given in column 4. Since the female dummy was already included in the regression for estimating residuals, the female dummy in column 4 does not show any gap between males and females in the pre-program period. The difference-in-difference coefficient is close to the one in column 1 and is significantly different from 0.

To summarize, I find no evidence of differential trends in vaccination either for all years or for only the pre-program years and the differential increase in vaccination for the treatment group of girls does not seem to be due to the difference in trends[30].

Another sample of boys and girls I can compare are those who have exactly one elder brother and one elder sister. These boys and girls also have the same number and sex composition of siblings and the girls in this sample are also second daughters in the family. However, I do not find any evidence of a differential change in vaccinations for girls in this sample (table shown in appendix). One potential reason for not finding any effect in this group could be that the sample size for this sample is less than half of size of my primary sample (608 as compared to 1283). Pande (2003) also finds that the boys and girls who have one elder brother *and* one elder sister have lower vaccination rates as compared to boys and girls who have one elder brother *or* one elder sister which suggests that these children might be treated differently than the children in my primary sample and the incentives don't seem to be leading to an increase in vaccination for these second daughters.

8.2 Birth Interval Robustness

The specification (1) assumes that the impact of explanatory variables on the log odds is constant for the entire duration. In this case, it implies that the odds ratio of having the next birth between females and males does not depend on the duration since the birth of the second child. However, the differential effect on females can differ by the duration since the birth of the second child. In particular, as we can see in the figure 4, there seems to be some evidence of differential effect on birth intervals after girls in the 20-30 months interval[31]. To test for the difference in the estimate of interaction of female dummy and post-program period dummy, I estimate the hazard model separately for different durations. Table 10 shows the results for the hazard models estimated for less than 20 months since the birth of the second child, 21-30 months since the birth of the second child and 31-40 months since the birth of the second child in columns 1,2 and 3 respectively. This specification assumes that the effect of the program on birth intervals is constant within each of these 10 months intervals and can differ across these intervals. As we can see from the table, the difference-in-difference estimate is negative for 21-30 months interval only, however, even for this interval, I cannot reject that the difference-in-difference estimate is equal to 0. I also cannot reject that the difference-in-difference estimates for the three intervals are equal to each other.

Since the survival curves for males and females, shown in figure 4 are not completely parallel, I try other specifications to account for this. Since the gap between males and females in the

[30] The effect in rural areas is also not distinguishable for households with different wealth. I test for wealth heterogeneity by creating wealth index based on principal component analysis of the assets dummies and I cannot reject that the effect is the same for bottom 50% and top 50%.

[31] Finding the evidence of an effect of the policy on birth intervals only in this duration would indicate that while there is no change in the birth intrvals for those who would wait less than 20 months for the birth of the third child after a girl is born but the people who would wait 20-30 months after the girl child, are now waiting longer.

pre-program period (i.e. the gap between blue and green lines) is increasing with the duration since the birth of the second child, I also include the linear interaction of the duration since the birth of second child with the female dummy, post-program period dummy and the interaction of female and post-program period. The results are given in table 11. The first column gives the results for the entire sample and the second column gives the results for the rural sample only. The coefficient on the interaction of female dummy and the durations since birth of the second child suggests some evidence of an increase in the male-female gap with the increase in duration. However, the interaction of difference-in-difference estimate with the durations variable is not statistically different from 0 and it also reduces the precision of the difference-in-difference estimate.

9 Model

In this section, I present a very simple model of investment in the health of children to explain the qualitative nature of my results. I assume that every household has one parent and one child. Parent's utility depends on child's health and consumption good.

Preferences: Let H denote the health of the child, C denote the consumption good. I also assume, for simplicity, that the utility function is additive separable in H and C, that is $\frac{\partial^2 U}{\partial H \partial C} = \frac{\partial^2 U}{\partial C \partial H} = 0$. This means that more consumption does not change the marginal utility that the parent gets from increase in health of the child and vice versa.

$$U(H, C)$$[32]

where $U_H > 0, U_C > 0, U_{HH} < 0$ and $U_{CC} < 0$.

Health production function: Let V denote the number of vaccines the child receives and I denote other investments in the health of the child.

$$H(V, I) = (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}} \tag{2}$$

I assume a Constant Elasticity of Substitution (CES) health production function where σ determines whether the inputs V and I are gross complements ($\sigma < 1$) or gross substitutes ($\sigma > 1$). I choose this function because of its flexible functional form which nests different types of production functions, Cobb-Douglas ($\sigma = 1$), perfect substitutes ($\sigma \to \infty$) and perfect complements ($\sigma = 0$)[33].

Each additional vaccine improve the health of the child by making her immune to an additional disease[34]. Other investments, for example, providing better nutrition, clean water and

[32] The utility function can also depend on the sex of the child, in which case, $U_b(H, C)$ would be the utility function if the child is a boy and it would be $U_g(H, C)$ if the child is a girl which can then allow for different investments in health of boy v/s health of girl. However, since Ladli scheme applies only to girls, I only focus on the utility function when the child is a girl.

[33] I choose CES also because of its previous use in the literature. Thompson (2014), Cunha et al. (2010) and Cunha and Heckman (2007) use CES production function to study how various genetic and environmental factors affect child development, including health status.

[34] It might appear that vaccination should affect only the probability that the child survives but conditional on survival, it should

having a longer birth interval after the child, also improve the health of the child. There can be complementarity or substitution between these inputs. For example, if the child is vaccinated, she won't get the disease, the body would be able to process nutrients better and the returns to better nutrition would be higher, generating the complementarity between vaccination and other inputs. They can also be substitutes. For example, cholera can be prevented either by giving the vaccine or by providing clean water. So, getting the child vaccinated against cholera may reduce the need for giving clean water to the child[35].

Budget Constraint: Let Y denote the household income, p_v denote the price of vaccination[36], p_i denote the price of other investments. I normalize the price of consumption good to be 1.

$$p_v * V + p_i * I + C = Y \tag{3}$$

The Ladli scheme provides a monetary transfer if the girl is vaccinated. I model this as a reduction in the price of vaccination, p_v. With the program, the budget constraint becomes:

$$(p_v - S)^{37} * V + p_i * I + C = Y \tag{4}$$

I am interested in how the investment in child's health changes as a result of this subsidy, i.e. $\frac{\partial I}{\partial S}$[38]. Setting up the Lagrangean and taking the first order conditions with respect to V, I, C, we have:

$$\frac{\partial U}{\partial H} * \frac{\sigma}{\sigma - 1} * (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{1}{\sigma-1}} * \frac{\sigma - 1}{\sigma} * V^{\frac{-1}{\sigma}} = -\lambda * p_v \tag{5}$$

$$\frac{\partial U}{\partial H} * \frac{\sigma}{\sigma - 1} * (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{1}{\sigma-1}} * \frac{\sigma - 1}{\sigma} * I^{\frac{-1}{\sigma}} = -\lambda * p_i \tag{6}$$

$$\frac{\partial U}{\partial C} = -\lambda \tag{7}$$

From (5) and (6), I get the following relationship between V and I:

$$V = \frac{p_i^\sigma}{p_v^\sigma} * I \tag{8}$$

not affect the health of the child. However, while in some cases, the vaccines might be for diseases which have a very high chance of the child dying but for some other disease, the chances of dying might be low but still the disease can affect the long term health of the child. One such example can be Pertussis, for which the chances of the child dying are comparatively low (1 in 100 of those who are infected),but the disease can lead to breathing problems, rapid cough and vomiting and these symptoms can last for a long time.

[35] The substitution and complementarity patterns can also be there if I model vaccination as only affecting the probability of survival of the child.

[36] Even though vaccination is provided at costless basis in government hospitals in India, there are other costs such as travel costs, waiting time and also information costs, where the parents may not know about the benefits of vaccinating the child, as also outlined in Pande and Yazbeck (2003). I assume that p_v includes all these costs.

[37] It is possible for amount of subsidy, S to exceed p_v but for simplicity I assume p_v to be greater than S. This assumption is not that unreasonable if we take p_v to be the sum of all costs required to fulfill the conditions of the scheme, which includes keeping the girl child alive till she is 18.

[38] I do not show the effect of S on vaccination, as for vaccination, both the substitution effect as well as the income effect lead to an increase in V as a result of the subsidy.

Using (6), (7), (8) and (3), I take the total differentiation with respect to p_v and get the following solution for $\frac{\partial I}{\partial p_v}$[39], the details of which can be found in the mathematical appendix.

$$\frac{-(\frac{U_H}{H}+U_{HH}\sigma)*\Big(V^{\frac{-1}{\sigma}}\frac{p_i^{\sigma}}{p_v^{\sigma+1}}I^{\frac{\sigma-1}{\sigma}}(V^{\frac{\sigma-1}{\sigma}}+I^{\frac{\sigma-1}{\sigma}})^{\frac{2}{\sigma-1}}\Big)-(\sigma-1)p_iU_{CC}I\frac{p_i^{\sigma}}{p_v^{\sigma}}}{-U_{HH}I^{\frac{-\sigma-1}{\sigma}}(V^{\frac{\sigma-1}{\sigma}}+I^{\frac{\sigma-1}{\sigma}})^{\frac{1+\sigma}{\sigma-1}}-p_i^2U_{CC}\Big(1+\frac{p_i^{\sigma-1}}{p_v^{\sigma-1}}\Big)} \tag{9}$$

Since $U_{HH}<0$, $U_{CC}<0$, $I>0$ and $V>0$, the denominator is always positive, so the sign of $\frac{\partial I}{\partial p_v}$ depends on the numerator only.

Result 1: If $\sigma<1$ and $\sigma<-\frac{U_H}{U_{HH}H}$, then $\frac{\partial I}{\partial p_v}<0$.

$$\underbrace{-\underbrace{(\frac{U_H}{H}+U_{HH}\sigma)}_{(+)}*\underbrace{\Big(V^{\frac{-1}{\sigma}}\frac{p_i^{\sigma}}{p_v^{\sigma+1}}I^{\frac{\sigma-1}{\sigma}}(V^{\frac{\sigma-1}{\sigma}}+I^{\frac{\sigma-1}{\sigma}})^{\frac{2}{\sigma-1}}\Big)}_{(+)}}_{(+)}-\underbrace{\underbrace{(\sigma-1)}_{(-)}\underbrace{U_{CC}}_{(-)}\underbrace{p_iI\frac{p_i^{\sigma}}{p_v^{\sigma}}}_{(+)}}_{(+)}$$

This shows that complementarity between V and I by itself is not sufficient for I to go up due to a reduction in p_v, the parent's valuation of child's health and the substitution pattern between H and C is also important. In particular, $-\frac{U_H}{U_{HH}H}$ is just equal to 1 if utility function is log-linear in H and C, in which case $\sigma<1$ is the only condition required to ensure that $\frac{\partial I}{\partial p_v}<0$. For further interpretation, if we consider the specific case in which the elasticity of substitution is constant between H and C, then the expression, $-\frac{U_H}{U_{HH}H}$ just gives that constant elasticity[40]. So, what this condition implies is that for I to go up as a result of a decrease in p_v, V and I should be complements and they should be more complementary than H and C. If there is complementarity between vaccination and other investments in the health production of the child, while health of the child and consumption good are substitutes in the utility function, this condition is naturally satisfied but if there is complementarity between H and C as well, then these conditions require that V and I should be more complementary than H and C. This makes sense as a reduction in p_v essentially implies a reduction in price of H and if the complementarity between H and C is higher than the complementarity between V and I, then the effect of a reduction in p_v can lead to an increase in C instead of an increase in I.

Result 2: If $\sigma>1$ and $\sigma>-\frac{U_H}{U_{HH}H}$, then $\frac{\partial I}{\partial p_v}>0$.

$$\underbrace{-\underbrace{(\frac{U_H}{H}+U_{HH}\sigma)}_{(-)}*\underbrace{\Big(V^{\frac{-1}{\sigma}}\frac{p_i^{\sigma}}{p_v^{\sigma+1}}I^{\frac{\sigma-1}{\sigma}}(V^{\frac{\sigma-1}{\sigma}}+I^{\frac{\sigma-1}{\sigma}})^{\frac{2}{\sigma-1}}\Big)}_{(+)}}_{(-)}-\underbrace{\underbrace{(\sigma-1)}_{(+)}\underbrace{U_{CC}}_{(-)}\underbrace{p_iI\frac{p_i^{\sigma}}{p_v^{\sigma}}}_{(+)}}_{(-)}$$

Here again, if utility function is log-linear in H and C, $\sigma>1$ is the only sufficient condition

[39] Note that $\frac{\partial I}{\partial S}=-\frac{\partial I}{\partial p_v}$, so a positive relationship between S and I would be a negative relationship between p_v and I and vice versa.

[40] If $U(H,C)=H^{\frac{\gamma-1}{\gamma}}+C^{\frac{\gamma-1}{\gamma}}$, then the constant elasticity of substitution between H and C is γ and the expression $-\frac{U_H}{U_{HH}H}=\gamma$.

for I to go down as a result of decrease in price of vaccination. Again, if we assume that the elasticity of substitution is constant between H and C, then the sufficient conditions for $\frac{\partial I}{\partial p_v} > 0$ are that V and I should be substitutes and they should be more substitutable than H and C. If vaccination and other investments are substitutes, while health of the child and consumption good are complements, then both the conditions hold naturally but if health of the child and consumption good are substitutes as well, then these conditions imply that vaccination and other investments should be more substitutable in the health production of the child as compared to H and C in the utility function of the parent.

Overall, my model predicts that there may not be the desired effect of the policy on other investments if the two conditions in result 1 do not hold, that is, if there is a lack of complementarity between vaccination and other investments or if the utility function of the parent is such that child's health and consumption good are more complementary than vaccination and other investments in the health production function.

10 Relation between vaccination and birth intervals

As shown in the model, the effect of the policy on birth intervals depends on the elasticity of substitution in the health production function between vaccination and birth intervals. While direct evidence of complementarity would require randomly varying both vaccination and birth intervals, however, with the data I have, I can look at the correlational relationship between vaccination and birth intervals. The idea is to see whether parents who value their children more, are more likely to vaccinate their children *and* to also have longer birth intervals after those children. If this relationship does not exist in the data, it would suggest that either there is no complementarity between vaccination and birth intervals or parents are not aware of it, which could help in understanding why there is no effect on birth intervals.

To test for evidence of this relation, I use all years of data that I have and I also use the sample of all births in Haryana. The results are presented in table 12. The first column shows the results for the full sample. The hazard regressions also control for various socio-economic indicators such as mother's education, father's education, mother's age, household assets and amenities. As we can see from the table, there is a negative relationship between vaccination and the probability of having a child in any given month (or a positive relationship between vaccination and birth intervals). The odds of the next child being born in any given month after a child who has a vaccination card are 14% (odds ratio=$exp(-0.15) = 0.86$) less than the odds of the next child being born after a child who does not have a vaccination card. This relationship is statistically different from 0 at 1% significance level. It might be the case that this relationship holds only for boys but the girls who are vaccinated may not be followed by longer birth intervals. To test for that, I run the same specification for females separately in column 2. The same result holds for females as well, though the coefficient is slightly smaller in magnitude.

Since I find stronger evidence of increase in vaccination in rural areas only, I test for the presence of relationship between vaccination and birth intervals in rural areas separately. The last two columns show the results for the rural sample only. For rural areas also, we can see that there is evidence that children who are vaccinated, have longer subsequent birth intervals as well. So, there seems to be some suggestive evidence that parents who vaccinate their children, are also likely to have longer birth intervals after these children[41].

10.1 Limitation

An important limitation of my analysis is that I am able to look at only one form of non-incentivized investment, which is birth intervals following the child[42]. I cannot rule out that the program had an effect on other forms of investments in the child. For example, since I do not have weight or height data of children, I cannot say whether the program had any effect on the nutrition of the girl child. Future work should look at these other forms of investments to see whether the program had any effect.

11 Conclusion

In this paper, I study the impact of *Ladli*, a conditional cash transfer program in Haryana in India, on the investments parents make in their daughters. Using difference-in-difference strategy and District Level Household Survey (DLHS) data, I find evidence that the program increases the vaccination rates (one of the conditions of the transfer) of eligible girls by 11 percentage points relative to the control group of boys, which is almost 100% of the gap between the two groups in the pre-program period. I also study whether the program is able to reach its goal of increasing the perceived value of daughters by measuring value as the birth intervals following the eligible girls. However, I do not find any evidence of a differential change in the birth intervals after the treatment group of girls. I write down a model in which I show that the program may not have an effect on the non-incentivized investments if there is a lack of complementarity between vaccination and other investments in the health production function of the child.

My results on vaccination imply that the conditional cash transfers can be very useful in reducing the male-female gap in specific investments. However, the results on birth interval suggest that such programs may not be successful in increasing the perceived value of daughters as parents do not seem to be making investments which are not incentivized.

---

[41] This relationship also gives evidence of the fact that parents seem to know about the benefits of longer birth intervals.

[42] I also used the outcomes of whether the child had diarrhea in the past 2 weeks and whether the child was taken to the doctor when ill, but I did not find any evidence of a gap between males and females in the pre-program period or any evidence of a change after the program in these outcomes.

References

Aizer, A. & Cunha, F. (2012), The production of human capital: Endowments, investments and fertility, Technical report, National Bureau of Economic Research.

Arnold, F., Choe, M. K. & Roy, T. K. (1998), 'Son preference, the family-building process and child mortality in india', *Population studies* **52**(3), 301–315.

Bank, W. (2011), *World development report 2012: gender equality and development*, World Bank Publications.

Buckles, K. S. & Munnich, E. L. (2012), 'Birth spacing and sibling outcomes', *Journal of Human Resources* **47**(3), 613–642.

Caldwell, P. & Caldwell, J. C. (1990), 'Gender implications for survival in south asia.'.

Chakraborty, T. & Kim, S. (2010), 'Kinship institutions and sex ratios in india', *Demography* **47**(4), 989–1012.

Cunha, F. & Heckman, J. (2007), The technology of skill formation, Technical report, National Bureau of Economic Research.

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

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

Debnath, S. (2013), Improving maternal health with incentives to mothers vs. health workers: Evidence from india, Technical report, Technical report, Dissertation, University of Virginia.

Department of Planning, G. o. P. (2009), 'New schemes for girl child'.

Filmer, D., King, E. M. & Pritchett, L. (1998), *Gender disparity in South Asia: comparisons between and within countries*, Vol. 1867, World Bank Publications.

Foster, A., Menken, J., Chowdhury, A. & Trussell, J. (1986), 'Female reproductive development: A hazards model analysis', *Social Biology* **33**(3-4), 183–198.

Gertler, P. (2004), 'Do conditional cash transfers improve child health? evidence from progresa's control randomized experiment', *The American Economic Review* **94**(2), 336–341.

Harris, M. (1993), 'The evolution of human gender hierarchies: A trial formulation', *Sex and gender hierarchies* pp. 57–79.

Jayachandran, S. (2015), 'The roots of gender inequality in developing countries', *economics* **7**(1), 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.

Lagarde, M., Haines, A. & Palmer, N. (2007), 'Conditional cash transfers for improving uptake of health interventions in low-and middle-income countries: a systematic review', *Jama* **298**(16), 1900–1910.

Maitra, P. & Pal, S. (2001), 'Birth spacing and child survival: Comparative evidence from india and pakistan'.

Mazumdar, C. T. (2012), 'Can incentives increase preference for daughters? evidence from a cash transfer scheme in haryana, india'.

Miller, B. D. (1981), *The endangered sex*.

Muhuri, P. K. & Menken, J. (1997), 'Adverse effects of next birth, gender, and family composition on child survival in rural bangladesh', *Population Studies* **51**(3), 279–294.

Omran, A. (1981), 'Review of the evidence–an update.'.

Organization, W. H. et al. (2007), 'Report of a who technical consultation on birth spacing: Geneva, switzerland 13-15 june 2005'.

Palloni, A. & Tienda, M. (1986), 'The effects of breastfeeding and pace of childbearing on mortality at early ages', *Demography* **23**(1), 31–52.

Pande, R. P. (2003), 'Selective gender differences in childhood nutrition and immunization in rural india: the role of siblings', *Demography* **40**(3), 395–418.

Pande, R. P. & Yazbeck, A. S. (2003), 'What's in a country average? wealth, gender, and regional inequalities in immunization in india', *Social Science & Medicine* **57**(11), 2075–2088.

Ranganathan, M. & Lagarde, M. (2012), 'Promoting healthy behaviours and improving health outcomes in low and middle income countries: a review of the impact of conditional cash transfer programmes', *Preventive medicine* **55**, S95–S105.

Retherford, R. D. & Roy, T. (2003), 'Factors affecting sex-selective abortion in india and 17 major states'.

Rindfuss, R. R., Bumpass, L. L., Palmore, J. A. & Han, D. W. (1982), 'The transformation of korean child-spacing practices', *Population Studies* **36**(1), 87–104.

Sekher, T. (2012), Skewed sex ratios, conditional cash transfers and promotion of girl children: The indian experience, *in* 'Annual Meeting of the Population Association of America, San Francisco'.

Siddhanta, S. & Nandy, D. (2004), 'Disaggregated pattern of gender gap in education in indian population: A fresh exploration'.

Singh, P. K. & Parasuraman, S. (2014), "looking beyond the male–female dichotomy'–sibling composition and child immunization in india, 1992–2006', *Social Science & Medicine* **107**, 145–153.

Sinha, N. & Yoong, J. (2009), 'Long-term financial incentives and investment in daughters', *Gender and Development* .

Thompson, O. (2014), 'Genetic mechanisms in the intergenerational transmission of health', *Journal of health economics* **35**, 132–146.

Trussell, J., Martin, L. G., Feldman, R., Palmore, J. A., Concepcion, M. & Abu Bakar, D. N. L. B. D. (1985), 'Determinants of birth-interval length in the philippines, malaysia, and indonesia: A hazard-model analysis', *Demography* **22**(2), 145–168.

Verma, J., Sachar, R., Prakash, V., Sehgal, R. & Aggarwal, S. (1990), 'Effect of sex of preceding child on birth interval.', *Indian Journal of Maternal and Child Health* **1**(2), 50–1.

Figure 1: Probability of vaccination card in rural areas

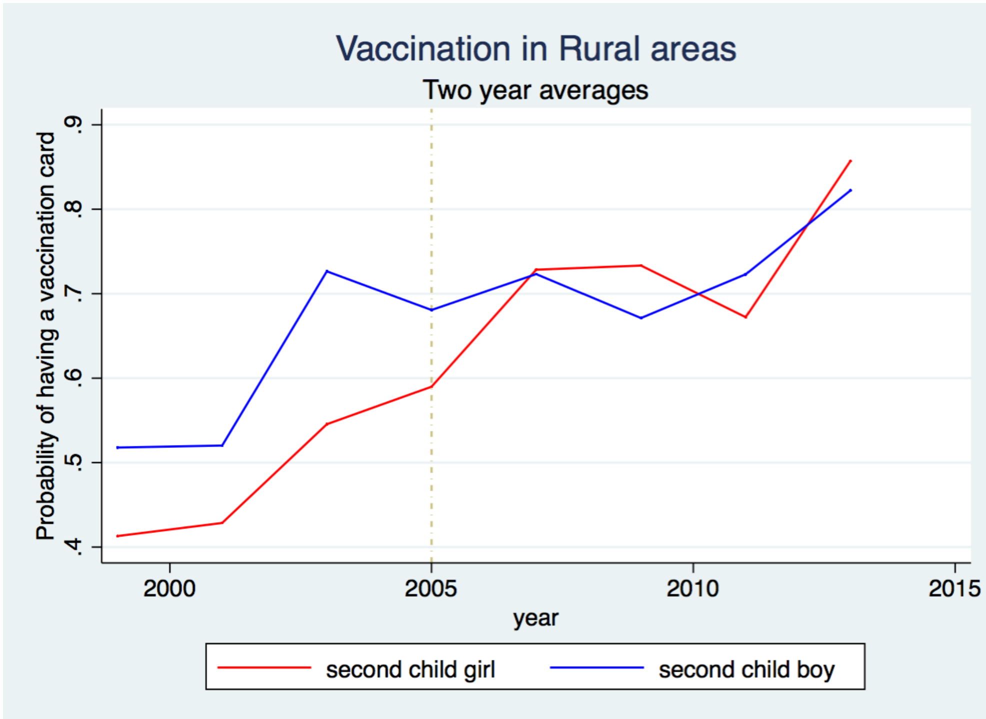

Figure 2: Hazard rates for the birth of third child

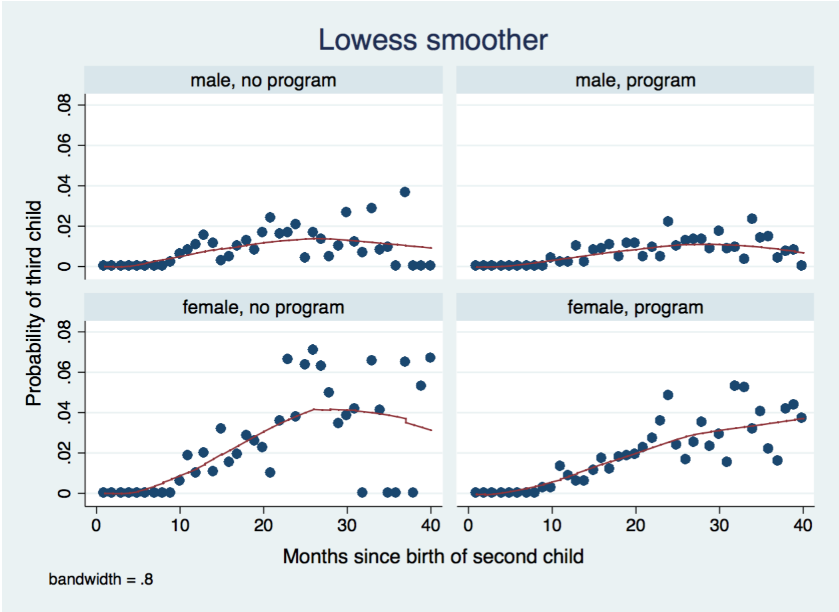

Figure 3: Hazard ratio

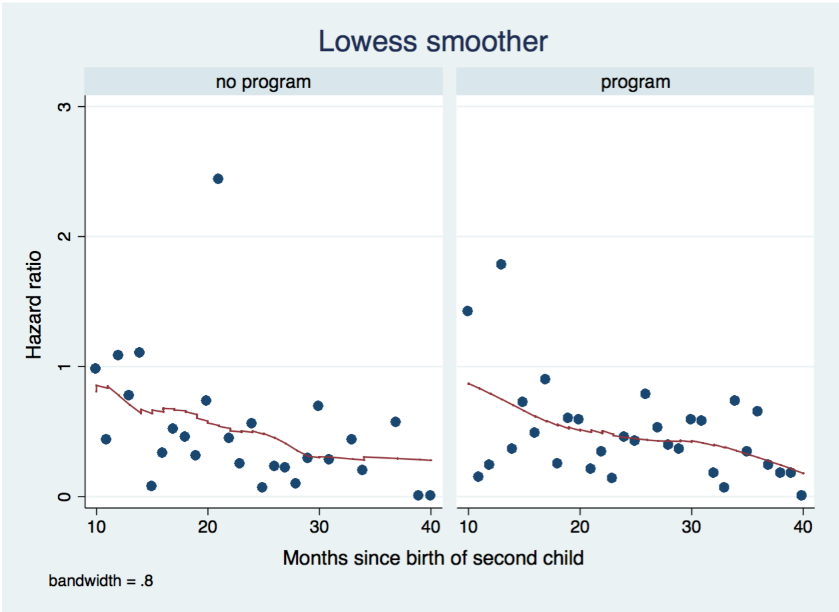

Figure 4: Kaplan-Meier survival curves for the birth of the third child

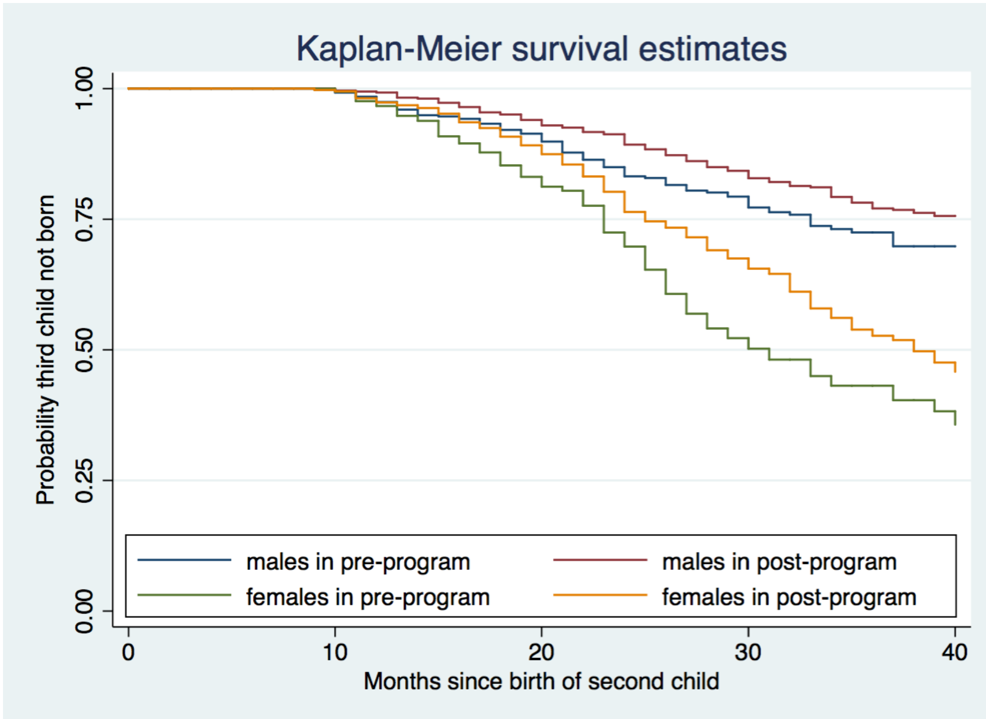

Table 1: Summary Statistics

	Pre-program		Post-program		
	Males	Females	Males	Females	**Diff-in-diff**
Dependent Variables					
Probability Vaccination card	0.62	0.54	0.74	0.73	
Number of vaccines	4.15	3.70	5.10	5.10	
	(1.65)	(1.79)	(1.29)	(1.22)	
Probability birth intervals < 24 months	0.11	0.20	0.07	0.18	
Had third birth	98	132	107	171	
Covariates					
Mother's age	25.86	25.65	29.02	28.34	-0.229
	(3.293)	(3.527)	(3.770)	(3.526)	(0.334)
Mother's age at marriage	18.66	18.57	19.87	19.48	-0.197
	(2.493)	(2.726)	(3.262)	(3.044)	(0.272)
Mother's education	8.703	8.427	10.51	10.20	-0.12
	(3.905)	(3.659)	(4.005)	(4.116)	(0.42)
Father's education	10.38	9.921	10.91	10.87	0.388
	(3.636)	(3.287)	(3.533)	(3.739)	(0.35)
Pucca house	0.642	0.544	0.780	0.799	0.10**
	(0.480)	(0.499)	(0.415)	(0.402)	(0.043)
Type of fuel	2.248	2.411	1.943	2.024	-0.0280
	(1.127)	(1.144)	(1.203)	(1.188)	(0.107)
Source of lighting	1.013	1.041	1.024	1.038	-0.0251
	(0.111)	(0.220)	(0.233)	(0.268)	(0.0243)
Source of water	2.092	2.166	1.677	1.799	0.0555
	(1.372)	(1.353)	(1.225)	(1.339)	(0.130)
Type of toilet	2.737	2.917	1.290	1.528	-0.144
	(1.755)	(1.759)	(0.986)	(1.328)	(0.146)
Radio	0.383	0.373	0.0484	0.0208	-0.00172
	(0.487)	(0.485)	(0.215)	(0.143)	(0.0330)
TV	0.840	0.817	0.949	0.934	0.0393
	(0.367)	(0.387)	(0.219)	(0.249)	(0.0359)
Telephone	0.343	0.290	0.0747	0.0764	0.0320
	(0.475)	(0.455)	(0.263)	(0.266)	(0.0322)
Bicycle	0.575	0.606	0.378	0.427	0.00709
	(0.495)	(0.490)	(0.485)	(0.496)	(0.0461)
Motor cycle	0.362	0.382	0.593	0.524	-0.0446
	(0.481)	(0.487)	(0.492)	(0.500)	(0.0443)
Car/Jeep	0.0575	0.0539	0.185	0.163	-0.0209
	(0.233)	(0.226)	(0.388)	(0.370)	(0.0262)
Tractor	0.100	0.0871	0.0615	0.0382	0.00152
	(0.300)	(0.283)	(0.241)	(0.192)	(0.0226)
Sewing machine	0.797	0.755	0.789	0.726	-0.0122
	(0.402)	(0.431)	(0.408)	(0.447)	(0.0438)
Observations	580	377	563	393	

Table 2: Probability of vaccination in Rural areas

	Vaccination Card		
	(1)	(2)	(3)
Female	-0.118***	-0.0997**	-0.101***
	(0.0408)	(0.0396)	(0.0390)
Post	0.122***	0.0878	0.147**
	(0.0345)	(0.0615)	(0.0626)
Female * post	0.116**	0.111**	0.115**
	(0.0543)	(0.0538)	(0.0531)
Mother's education		0.00786	0.0110*
		(0.00663)	(0.00645)
Father's education		0.00221	0.00441
		(0.00507)	(0.00507)
Observations	1283	1276	1276
District dummies	No	No	Yes
Controls for household assets	No	Yes	Yes

Robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the difference-in-difference results for having a vaccination card for rural areas only. The first column controls for age of the child only (in months), the second column also controls for mother's education, father's education, mother's age, household assets and amenities and the third column additionally also controls for district dummies.

Table 3: Probability of vaccination- Entire sample

	Vaccination card		
	(1)	(2)	(3)
Female	-0.0798**	-0.0553*	-0.0511
	(0.0342)	(0.0327)	(0.0329)
Post	0.116***	0.105**	0.129***
	(0.0267)	(0.0468)	(0.0473)
Female * post	0.0696	0.0562	0.0521
	(0.0429)	(0.0417)	(0.0416)
Mother's education		0.0117**	0.0141***
		(0.00495)	(0.00492)
Father's education		0.00488	0.00611
		(0.00415)	(0.00420)
Observations	2072	2062	2062
District dummies	No	No	Yes
Controls for household assets and amenities	No	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the difference-in-difference results for having a vaccination card for entire sample. The first column controls for age of the child only (in months), the second column also controls for mother's education, father's education, mother's age, household assets and amenities and the third column additionally also controls for district dummies.

Table 4: Number of vaccines

	Number of vaccines received					
	Total sample			Rural sample		
	(1)	(2)	(3)	(1)	(2)	(3)
Female	-0.458***	-0.369***	-0.328***	-0.466***	-0.410***	-0.403***
	(0.122)	(0.114)	(0.114)	(0.139)	(0.132)	(0.130)
Post	0.642***	0.407**	0.610***	0.670***	0.438**	0.713***
	(0.102)	(0.163)	(0.169)	(0.129)	(0.201)	(0.210)
Female*post	0.458***	0.415***	0.367***	0.502***	0.492***	0.472***
	(0.146)	(0.139)	(0.138)	(0.176)	(0.170)	(0.167)
Father's education		0.0361**	0.0391***		0.0585***	0.0719***
		(0.0147)	(0.0151)		(0.0173)	(0.0178)
Mothers's education			0.0614***			0.0587***
			(0.0171)			(0.0211)
Observations	1832	1823	1823	1137	1130	1130
District dummies	No	No	Yes	No	No	Yes
Controls for household assets	No	Yes	Yes	No	Yes	Yes

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the difference-in-difference results for number of vaccines received by the child in both rural areas as well as the total sample. The first 3 columns show the results for total sample and the last 3 for the rural sample. In both the samples, first column controls for age of the child only (in months), the second column also controls for mother's education, father's education, mother's age, household assets and amenities and the third column additionally also controls for district dummies.

Table 5: Vaccination results- Falsification checks

	Punjab				First child boy sample	
	Vaccination Card		Number vaccines		Vaccination card	Number vaccines
	Full	Rural	Full	Rural	Rural	Rural
Female	-0.0501	-0.0584	-0.205	-0.154	-0.0209	0.179
	(0.0349)	(0.0424)	(0.135)	(0.163)	(0.0394)	(0.124)
Post	0.103***	0.131***	0.809***	1.006***	0.168***	0.632***
	(0.0250)	(0.0318)	(0.0970)	(0.124)	(0.0403)	(0.132)
Female*Post	0.0206	0.00838	0.203	0.0437	0.0761	-0.173
	(0.0413)	(0.0511)	(0.153)	(0.187)	(0.0530)	(0.166)
Observations	1822	1166	1599	1030	1290	1130

Robust standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the various falsification checks. The first 4 columns show the results for Punjab which did not have the scheme and the last 2 columns show the results for the second child sample in Haryana where the first child was a boy. In this sample, girls are the first daughters in the family and hence are not eligible for the incentive. The only control included is the age of the child (in months).

Table 6: Hazard regression

	Third birth dummy			
	Logit	Logit	Logit	Marginal effects
Female	0.910***	0.762***	0.764***	0.0105***
	(0.137)	(0.154)	(0.155)	(0.00216)
Post	-0.375***	-0.501**	-0.552**	-0.00761**
	(0.143)	(0.229)	(0.238)	(0.00330)
Female*post	0.0473	0.116	0.0882	0.00122
	(0.184)	(0.197)	(0.198)	(0.00273)
Mother's age		-0.0396**	-0.0412**	-0.000567**
		(0.0167)	(0.0170)	(0.000236)
Mother's education		-0.118***	-0.125***	-0.00173***
		(0.0262)	(0.0274)	(0.000385)
Father's education		-0.0404**	-0.0448**	-0.000618**
		(0.0177)	(0.0183)	(0.000254)
Mother's age at marriage		0.00789	0.0219	0.000301
		(0.0210)	(0.0218)	(0.000301)
Observations	35376	33097	33097	33097
District dummies	No	No	Yes	Yes
Controls for household assets	No	Yes	Yes	Yes

Standard errors clustered by mother ID in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the hazard regression results for the entire sample. The dependent variable here is a dummy which takes the value 1 if the mother had a third birth and 0 if she did not have a third birth or was censored. The first 3 columns show the logit coefficients while the final column shows the marginal effects. The dummies for each month since the birth of the second child are also added as controls to allow for non-parametric baseline hazard.

Table 7: Hazard regression- Rural areas

	Third birth dummy			
	Logit	Logit	Logit	Marginal effects
Female	0.855***	0.790***	0.836***	0.0140***
	(0.160)	(0.177)	(0.178)	(0.00304)
Post	-0.237	-0.391	-0.467	-0.00784
	(0.169)	(0.272)	(0.286)	(0.00482)
Female*post	0.239	0.224	0.165	0.00277
	(0.217)	(0.231)	(0.232)	(0.00389)
Mother's age		-0.0396**	-0.0431**	-0.000723**
		(0.0195)	(0.0208)	(0.000352)
Mother's education		-0.139***	-0.158***	-0.00265***
		(0.0329)	(0.0353)	(0.000600)
Father's education		-0.0340	-0.0358*	-0.000601*
		(0.0208)	(0.0215)	(0.000363)
Mother's age at marriage		0.0166	0.0366	0.000613
		(0.0268)	(0.0276)	(0.000464)
Observations	21016	19415	19415	19415
District dummies	No	No	Yes	Yes
Controls for household assets	No	Yes	Yes	Yes

Standard errors clustered by mother ID in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the hazard regression results for rural areas only. The dependent variable here is a dummy which takes the value 1 if the mother had a third birth and 0 if she did not have a third birth or was censored. The first 3 columns show the logit coefficients while the final column shows the marginal effects. The dummies for each month since the birth of the second child are also added as controls to allow for non-parametric baseline hazard.

Table 8: Hazard regression

	Third birth dummy			
	Punjab		First child boy sample	
	Full	Rural	Full	Rural
Female	0.943***	0.967***	0.610***	0.682***
	(0.163)	(0.188)	(0.173)	(0.201)
Post	-0.833***	-0.723***	-0.686***	-0.562**
	(0.180)	(0.214)	(0.199)	(0.237)
Female*Post	0.216	0.189	0.116	0.113
	(0.225)	(0.262)	(0.252)	(0.298)
Observations	32651	20780	38748	24379

Standard errors clustered by mother ID in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the hazard regression for the samples which were not eligible for Ladli. The first 2 columns show the results for Punjab- full sample and rural sample respectively and the last 2 columns show the results for the sample in Haryana where the first child was a boy. The third column shows the results for full sample and the fourth one for rural sample only. No other controls are included. The dummies for each month since the birth of the second child are also added as controls to allow for non-parametric baseline hazard.

Table 9: Probability of vaccination in Rural areas

	Vaccination Card			
	Post-2005	Common trend	Differential trends	Pre-period trends
Female	-0.0860***	-0.0857***	-0.0732	0.0231
	(0.0291)	(0.0289)	(0.0481)	(0.0289)
Post Ladli period	0.0370	-0.0666	-0.0760	-0.117***
	(0.0338)	(0.0438)	(0.0515)	(0.0340)
Female*Post	0.108**	0.106**	0.129	0.113***
	(0.0424)	(0.0423)	(0.0784)	(0.0424)
Year		0.0243***	0.0258***	
		(0.00670)	(0.00805)	
Female*year			-0.00361	
			(0.0107)	
Observations	1927	1927	1927	1927

Robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the difference-in-difference results for the probability of having a vaccination card. The first column shows the results when taking all years after 2005 as the post period, without controlling for trends. The second column controls for common trend for males and females by including variable 'year', the third column controls for 'year' as well as 'female*year'. The final column estimates the residuals by allowing for differential trends for males and females only in the pre-program data and uses those residuals as the dependent variable. All the specifications control for mother's education, father's education, mother's age, household assets and amenities, and district dummies.

Table 10: Hazard regression for different durations

	Third birth dummy		
	Less than 20 months	21-30 months	31-40 months
Female	0.507**	1.071***	0.835*
	(0.210)	(0.218)	(0.426)
Post	-0.455	-0.433	-0.326
	(0.336)	(0.359)	(0.585)
Female*post	0.0635	-0.199	0.390
	(0.305)	(0.290)	(0.488)
Observations	18734	10375	5763

Standard errors clustered by mother ID in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the logit coefficients from the hazard regression for different durations. The first column shows the regression with less than or equal to 20 months since the birth of the second child, the second column shows the regression with duration since birth of second birth between 21-30 months, the final column shows the results for the duration since birth of second child between 31-40 months. The regressions also control for mother's education, father's education, mother's age, household assets and amenities. The dummies for each month since the birth of the second child are also included as controls to allow for non-parametric baseline hazard.

Table 11: Hazard regression- Linear interaction with time

	Third birth dummy	
	Full	Rural
Female	0.136	0.401
	(0.307)	(0.358)
Post	-0.823**	-1.066**
	(0.365)	(0.444)
Female*post	0.236	0.446
	(0.419)	(0.502)
Female*Duration	0.0503**	0.0344
	(0.0223)	(0.0266)
Post*Duration	0.0256	0.0462*
	(0.0218)	(0.0265)
Female*Post*Duration	-0.0203	-0.0240
	(0.0276)	(0.0332)
Observations	33097	19415

Standard errors clustered by mother ID in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the logit coefficients from hazard regression, while interacting the variables with the duration since the birth of the second child. The first column shows the results for the entire sample and the second column shows the results for the rural sample only. The regressions also control for mother's education, father's education, mother's age, household assets and amenities, and district dummies. The dummies for each month since the birth of the second child are also added as controls to allow for non-parametric baseline hazard.

Table 12: Hazard regression- relation between vaccination and birth intervals

	Next birth dummy			
	Full sample		Rural sample	
	All children	Females only	All children	Females only
Vaccination card	-0.152***	-0.117***	-0.162***	-0.115**
	(0.0294)	(0.0399)	(0.0338)	(0.0457)
Mother's education	-0.0376***	-0.0341***	-0.0324***	-0.0270***
	(0.00610)	(0.00816)	(0.00719)	(0.00954)
Father's education	-0.00627	-0.00242	-0.00458	-0.00156
	(0.00390)	(0.00477)	(0.00427)	(0.00517)
Observations	357830	152201	241640	102166

Standard errors clustered by mother ID in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The above regression shows the logit coefficients from the hazard regression. The total sample here includes all the children who were born first in the reference period in DLHS-2, DLHS-3 and DLHS-4 data in Haryana. Here, I am not limiting the data by second children only. The dependent variable takes the value 1 if the next child was born after these children and 0 otherwise. Vaccination card takes the value 1 if the child has a vaccination card and 0 otherwise. The first 2 columns show the results for the full sample and the last two show the results for the rural sample only. The second and fourth columns show the results for females only. The regressions also control for mother's education, father's education, mother's age, number of brothers and number of sisters of the child, household assets and amenities, and district dummies. The dummies for each month since the birth of the second child are also included as controls to allow for non-parametric baseline hazard.

Appendix

Figure 5: Probability of vaccination card in rural areas- yearly averages

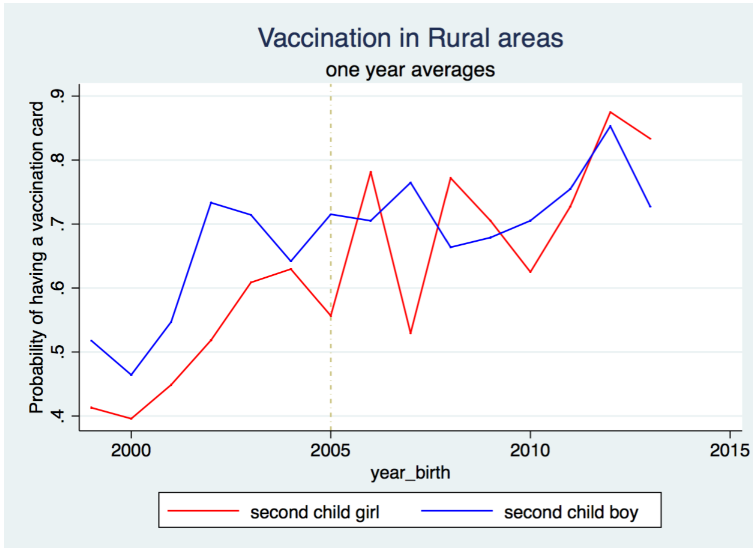

Figure 6: Probability of vaccination card- entire sample

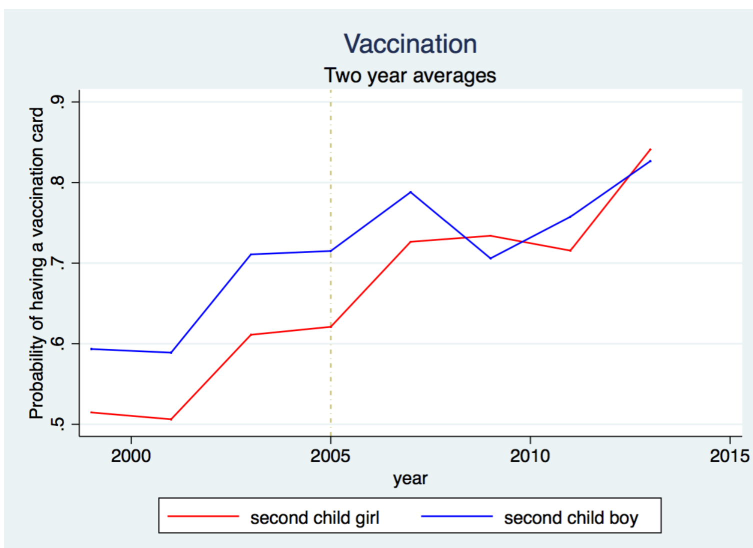

Table 13: Wealth heterogeneity

	Vaccination Card	Third birth dummy	
	Rural	Full	Rural
Female	-0.0979**	0.753***	0.835***
	(0.0458)	(0.169)	(0.185)
Post	0.150**	-0.297	-0.233
	(0.0683)	(0.241)	(0.288)
Female * post	0.0776	-0.212	-0.187
	(0.0748)	(0.231)	(0.256)
High SES	-0.0159	-0.249	-0.413
	(0.0730)	(0.363)	(0.563)
Female*High SES	-0.00426	0.149	0.0889
	(0.0892)	(0.399)	(0.654)
Post*High SES	-0.119	-0.866**	-0.766
	(0.0763)	(0.393)	(0.623)
Female*Post*High SES	0.0572	0.704	1.122
	(0.117)	(0.482)	(0.745)
Observations	1276	33097	19415

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the test for wealth heterogeneity for vaccination and birth intervals. I create wealth index based on the assets of the household. High SES are defined as those who are above the median for the index. Robust standard errors for vaccination and standard errors clustered by mother ID in parentheses. The first column shows the results for vaccination card and the second and third columns show the logit coefficients from the hazard regressions. The regressions also control for mother's education, father's education, mother's age, number of brothers and number of sisters of the child, household assets and amenities, and district dummies. The dummies for each month since the birth of the second child are also included as controls to allow for non-parametric baseline hazard.

Table 14: Probability of vaccination- Urban sample

	(1) Vaccination card
Female	0.0602
	(0.0569)
Post	0.0836^{*}
	(0.0445)
Female * post	-0.0834
	(0.0679)
Observations	789

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the difference-in-difference results for having a vaccination card for urban sample. Only additional control included is age in months of the child.

Table 15: Probability of vaccination in Rural areas- One brother sample

	(1) Vaccination card
female	0.0135
	(0.0520)
Post	0.331^{***}
	(0.0568)
Female * Post	-0.0784
	(0.0780)
Observations	608

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table shows the difference-in-difference results for having a vaccination card for rural areas for the sample where both girls and boys have one elder brother and one elder sister. Only additional control included is age in months of the child.

Mathematical Appendix

$$U(H(V,I),C)$$

where

$$H(V,I) = (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}}$$

and

$$p_v * V + p_i * I + C = Y$$

Setting up the lagrangean, we have:

$$L = U(H(V,I),C) + \lambda * (p_v * V + p_i * I + C - Y)$$

$$\frac{\partial L}{\partial V} = 0 \Rightarrow \frac{\partial U}{\partial H} * \frac{\sigma}{\sigma-1} * (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{1}{\sigma-1}} * \frac{\sigma-1}{\sigma} * V^{\frac{-1}{\sigma}} = -\lambda * p_v \tag{10}$$

$$\frac{\partial L}{\partial I} = 0 \Rightarrow \frac{\partial U}{\partial H} * \frac{\sigma}{\sigma-1} * (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{1}{\sigma-1}} * \frac{\sigma-1}{\sigma} * I^{\frac{-1}{\sigma}} = -\lambda * p_i \tag{11}$$

$$\frac{\partial L}{\partial C} = 0 \Rightarrow \frac{\partial U}{\partial C} = -\lambda \tag{12}$$

From (10) and (11),

$$\frac{V^{\frac{-1}{\sigma}}}{I^{\frac{-1}{\sigma}}} = \frac{p_v}{p_i}$$

$$\frac{V}{I} = \Big(\frac{p_v}{p_i}\Big)^{-\sigma}$$

$$V = \frac{p_i^\sigma}{p_v^\sigma} * I \tag{13}$$

From (11) and (12), we have:

$$\frac{\partial U}{\partial H} * \frac{\partial H(V,I)}{\partial I} = \frac{\partial U}{\partial C} * p_i \tag{14}$$

And the budget constraint is:

$$p_v * V + p_i * I + C = Y \tag{15}$$

We are interested in the partial derivative of I with respect to p_v, so I take the total differentiation of equations (13), (14) and (15) with respect to p_v.

Differentiating (13):

$$\frac{\partial V}{\partial p_v} = \frac{p_i^\sigma}{p_v^\sigma} * \frac{\partial I}{\partial p_v} - \frac{\sigma * I * p_i^\sigma}{p_v^{\sigma+1}} \tag{16}$$

Differentiating (14):

$$\begin{aligned}&\Big(\frac{\partial^2 U}{\partial H^2} * \frac{\partial H}{\partial V} * \frac{\partial V}{\partial p_v} + \frac{\partial^2 U}{\partial H^2} * \frac{\partial H}{\partial I} * \frac{\partial I}{\partial p_v}\Big) * \frac{\partial H}{\partial I} \\ &+ \frac{\partial U}{\partial H} * \Big(\frac{\partial^2 H}{\partial I^2} * \frac{\partial I}{\partial p_v} + \frac{\partial^2 H}{\partial I \partial V} * \frac{\partial V}{\partial p_v}\Big) \\ &= \frac{\partial^2 U}{\partial C^2} * \frac{\partial C}{\partial p_v} * p_i\end{aligned} \tag{17}$$

Differentiating (15):

$$\begin{aligned}&p_v * \frac{\partial V}{\partial p_v} + V + p_i * \frac{\partial I}{\partial p_v} + \frac{\partial C}{\partial p_v} = 0 \\ \Rightarrow &\frac{\partial C}{\partial p_v} = -p_v * \frac{\partial V}{\partial p_v} - V - p_i * \frac{\partial I}{\partial p_v}\end{aligned} \tag{18}$$

Using values of $\frac{\partial V}{\partial p_v}$ and $\frac{\partial C}{\partial p_v}$ from (16) and (18) respectively and substituting them in (17), we have:

$$\begin{aligned}
&\Big(\frac{\partial H}{\partial I} * \frac{\partial^2 U}{\partial H^2} * \frac{\partial H}{\partial V} + \frac{\partial U}{\partial H} * \frac{\partial^2 H}{\partial I \partial V}\Big) * \Big(- \sigma I \frac{p_i^\sigma}{p_v^{\sigma+1}} + \frac{\partial I}{\partial p_v} * \frac{p_i^\sigma}{p_v^\sigma}\Big) \\
+&\Big(- p_i * \frac{\partial^2 U}{\partial C^2}\Big) * \Big(- V + \sigma I \frac{p_i^\sigma}{p_v^\sigma} - p_i * (1 + \frac{p_i^{\sigma-1}}{p_v^{\sigma-1}}) * \frac{\partial I}{\partial p_v}\Big) \\
+&\Big((\frac{\partial H}{\partial I})^2 * \frac{\partial^2 U}{\partial H^2} + \frac{\partial U}{\partial H} * \frac{\partial^2 H}{\partial I^2}\Big) * \frac{\partial I}{\partial p_v} = 0
\end{aligned} \tag{19}$$

Using the functional form of $H(I, V)$ to substitute $\frac{\partial H}{\partial I}, \frac{\partial H}{\partial V}, \frac{\partial^2 H}{\partial I^2}, \frac{\partial^2 H}{\partial I \partial V}$ in (19) and solving for $\frac{\partial I}{\partial p_v}$, we have:

$$\begin{aligned}
&\Big[- V^{\frac{-1}{\sigma}} \frac{p_i^\sigma}{p_v^{\sigma+1}} I^{\frac{\sigma-1}{\sigma}} * (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{2-\sigma}{\sigma-1}} * U_H \\
&+ U_{HH} * \Big(- \sigma V^{\frac{-1}{\sigma}} \frac{p_i^\sigma}{p_v^{\sigma+1}} I^{\frac{\sigma-1}{\sigma}}\Big) * (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{2}{\sigma-1}} \\
&- \frac{p_i}{\sigma} * U_{CC} * \Big(- V + \sigma I \frac{p_i^\sigma}{p_v^\sigma}\Big)\Big] \\
=&\frac{\partial I}{\partial p_v} * \Big[- p_i^2 U_{CC} \Big(1 + \frac{p_i^{\sigma-1}}{p_v^{\sigma-1}}\Big) \\
&- \frac{1}{\sigma} * U_H * (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{2-\sigma}{\sigma-1}} * (p_v^{-\sigma} V^{\frac{-1}{\sigma}} I^{\frac{-1}{\sigma}} p_i^\sigma + I^{\frac{-2}{\sigma}}) \\
&- U_{HH} * (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{2}{\sigma-1}} * (p_v^{-\sigma} V^{\frac{-1}{\sigma}} I^{\frac{-1}{\sigma}} p_i^\sigma + I^{\frac{-2}{\sigma}}) \\
&+ \frac{1}{\sigma} * (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{1}{\sigma-1}} * I^{\frac{-\sigma-1}{\sigma}} * U_H\Big]
\end{aligned}$$

Using $H = (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{\sigma}{\sigma-1}}$ and $V = I \frac{p_i^\sigma}{p_v^\sigma}$:

$$\begin{aligned}
&\Rightarrow \Bigg(- V^{\frac{-1}{\sigma}} \frac{p_i^\sigma}{p_v^{\sigma+1}} I^{\frac{\sigma-1}{\sigma}}\Bigg)\Bigg(V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}}\Bigg)^{\frac{2}{\sigma-1}} \Bigg(\frac{U_H}{H} + U_{HH}\sigma\Bigg) - \Bigg(p_i * U_{CC} I \frac{p_i^\sigma}{p_v^\sigma}(\sigma - 1)\Bigg) \\
=&\frac{\partial I}{\partial p_v} * \Bigg[- \Bigg(p_v^{-\sigma} V^{-\frac{1}{\sigma}} I^{-\frac{1}{\sigma}} P_i^\sigma + I^{-1\frac{2}{\sigma}}\Bigg)\Bigg(V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}}|\Bigg)^{\frac{2}{\sigma-1}} \Bigg(\frac{U_H}{\sigma_H} + U_{HH}\Bigg) \\
&+ \Bigg(\frac{1}{\sigma} * (V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}})^{\frac{1+\sigma}{\sigma-1}} * I^{\frac{-\sigma-1}{\sigma}} * \frac{U_H}{H}\Bigg) - \Bigg(P_i^2 U_{CC} (1 + \frac{p_i^{\sigma-1}}{p_v^{\sigma-1}})\Bigg)\Bigg]
\end{aligned}$$

$$\begin{aligned}
&\Rightarrow \Bigg(- V^{\frac{-1}{\sigma}} \frac{p_i^\sigma}{p_v^{\sigma+1}} I^{\frac{\sigma-1}{\sigma}}\Bigg)\Bigg(V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}}\Bigg)^{\frac{2}{\sigma-1}} \Bigg(\frac{U_H}{H} + U_{HH}\sigma\Bigg) - \Bigg(p_i * U_{CC} I \frac{p_i^\sigma}{p_v^\sigma}(\sigma - 1)\Bigg) \\
=&\frac{\partial I}{\partial p_v} * \Bigg[I^{\frac{-\sigma-1}{\sigma}} \Bigg(V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}}\Bigg)^{\frac{1+\sigma}{\sigma-1}} \Bigg[\Bigg(- I \frac{p_v^{-\sigma} p_i^\sigma V^{\frac{-1}{\sigma}} + I^{\frac{-1}{\sigma}}}{V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}}}\Bigg)\Bigg(\frac{U_H}{\sigma_H} + U_{HH}\Bigg) + \frac{U_H}{\sigma_H}\Bigg] \\
&- P_i^2 U_{CC} \Bigg(1 + \frac{p_i^{\sigma-1}}{p_v^{\sigma-1}})\Bigg)\Bigg]
\end{aligned}$$

Using $V = \frac{p_i^\sigma}{p_v^\sigma} I$:

$$\begin{aligned}
&\Rightarrow \Bigg(- V^{\frac{-1}{\sigma}} \frac{p_i^\sigma}{p_v^{\sigma+1}} I^{\frac{\sigma-1}{\sigma}}\Bigg)\Bigg(V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}}\Bigg)^{\frac{2}{\sigma-1}} \Bigg(\frac{U_H}{H} + U_{HH}\sigma\Bigg) - \Bigg(p_i * U_{CC} I \frac{p_i^\sigma}{p_v^\sigma}(\sigma - 1)\Bigg) \\
=&\frac{\partial I}{\partial p_v} * \Bigg[I^{\frac{-\sigma-1}{\sigma}} \Bigg(V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}}\Bigg)^{\frac{1+\sigma}{\sigma-1}} \Bigg(- \frac{U_H}{\sigma_H} + \frac{U_H}{\sigma_H} - U_{HH}\Bigg) - P_i^2 U_{CC} \Bigg(1 + \frac{p_i^{\sigma-1}}{p_v^{\sigma-1}})\Bigg)\Bigg]
\end{aligned}$$

$$\Rightarrow \frac{\partial I}{\partial p_v} = \frac{\left(- V^{\frac{-1}{\sigma}} \frac{p_i^{\sigma}}{p_v^{\sigma+1}} I^{\frac{\sigma-1}{\sigma}} \right) \left(V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}} \right)^{\frac{2}{\sigma-1}} \left(\frac{U_H}{H} + U_{HH}\sigma \right) - \left(p_i * U_{CC} I \frac{p_i^{\sigma}}{p_v^{\sigma}} (\sigma - 1) \right)}{\left[- U_{HH} I^{\frac{-\sigma-1}{\sigma}} \left(V^{\frac{\sigma-1}{\sigma}} + I^{\frac{\sigma-1}{\sigma}} \right)^{\frac{1+\sigma}{\sigma-1}} - P_i^2 U_{CC} \left(1 + \frac{p_i^{\sigma-1}}{p_v^{\sigma-1}}) \right) \right]}$$