HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

WP 19/24

Reviewing the Existing Evidence of the Conditional Cash Transfer in India through the Partial Identification Approach

Toshiaki Aizawa

October 2019

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

Reviewing the Existing Evidence of the Conditional Cash Transfer in India through the Partial Identification Approach

Toshiaki Aizawa

Abstract

This paper re-estimates the causal impacts of a conditional cash transfer programme in India, the Janani Suraksha Yojana (JSY), on maternal and child healthcare use. The main goal is to provide new evidence and assess the validity of the identification assumptions employed in previous studies on the JSY, through the conservative partial identification approach. We find that the average treatment effects estimated under the conditional independence assumption lie outside the bound of the treatment effect estimated under weaker but more credible assumptions, thereby suggesting that the selection bias could not have been fully controlled for by the observable characteristics and that the average treatment effects estimated in the previous studies may have been over- or under-estimated.

JEL codes: I12, I15, I18

Keywords: Conditional cash transfer; Partial identification; Conditional independence; India

1 Introduction

1.1 Background

The adequate use of healthcare services is an essential factor in successful maternal and child health outcomes (Campbell and Graham, 2006; Chou et al., 2015; Darmstadt et al., 2009; Scott and Ronsmans, 2009). However, women in developing countries are often faced with multiple barriers to accessing health services, and improvements in maternal healthcare access, especially by poor women, remain inadequate. In India, only 51.2 per cent of women access antenatal care, and in 2014, only 38.7 per cent gave birth in a health institution (UNICEF, 2018a). In addition, India has the highest neonatal mortality rate in the world – as of 2015, around 20 per cent (=1.2 million) of global under-five deaths occurred in the country (UNICEF, 2015, 2018b). Promoting institutional delivery and adequate use of antenatal care is a key method to improve infant and neonatal mortalities (Lahariya, 2009; Pathak and Mohanty, 2010; Freedman et al., 2007). However, women's uptake of maternal care in India has been associated strongly with wealth, and facilitating the adequate use of maternal care among poor and marginalised women is still a major challenge (Pathak et al., 2010; Kesterton et al., 2010).

The success of conditional cash transfer (CCT) programmes in Latin America has led to an enthusiastically mirrored response in Africa and Asia, and CCT has become one of the most adopted demand-side programmes to enhance healthcare use (Handa and Davis, 2006; Rawlings and Rubio, 2005). CCTs influence health-seeking behaviours through financial incentives by transferring money to households contingent on investments in human capital, such as health and education. Interest in using financial incentives to promote maternal and child healthcare utilisation has also spread to Southern Asia, which has low maternal healthcare use and high neonatal mortality rates. CCT programmes intending to encourage maternity and child healthcare use have been introduced in India, Nepal and Bangladesh (Devadasan et al., 2008).[1]

On 12 April 2005, the Indian government launched a nationwide CCT programme called Janani Suraksha Yojana (JSY), or Safe Motherhood Programme, which promotes institutional delivery and receiving timely antenatal and postnatal care to reduce infant and neonatal mortalities (Horton, 2010; Bhutta et al., 2010). The JSY has become one of the largest cash transfer programmes in the world in terms of the number of beneficiaries. The JSY provides a cash incentive to mothers that give birth in government-approved

[1] Pakistan and Cambodia also introduced a voucher-type programme with the same aims as the CCTs in South Asia (Jehan et al., 2012; Van de Poel et al., 2014).

health facilities[2], and in this sense, it has a much narrower aim than most of the CCTs in Latin America, which have multiple goals beyond maternal and child healthcare. Funded by central government, the JSY is administered through the National Rural Health Mission. In 2014-2015, 10.4 million women benefited from the scheme, representing a third of all women giving birth in the country annually (Ministry of Health and Family Welfare, 2015).

Evidence in most previous studies has largely relied on inappropriate data or potentially untenable identification assumptions in order to yield a definitive causal impact. Until now, no study has ever tried to explore the validity of the identification assumptions employed in previous studies on the JSY. The main goal of this paper is to re-estimate the causal impact of the JSY participation on maternal and child healthcare use, with weaker but more credible identification assumptions than used in previous studies. Specifically, we non-parametrically identify the average treatment effect (ATE) by bounds. The partial identification approach abandons point identification with strong assumptions and instead seeks causal effects with much more credible assumptions (Manski, 1990, 1997; Manski and Pepper, 2000). In this sense, the partial identification approach yields more conservative results.

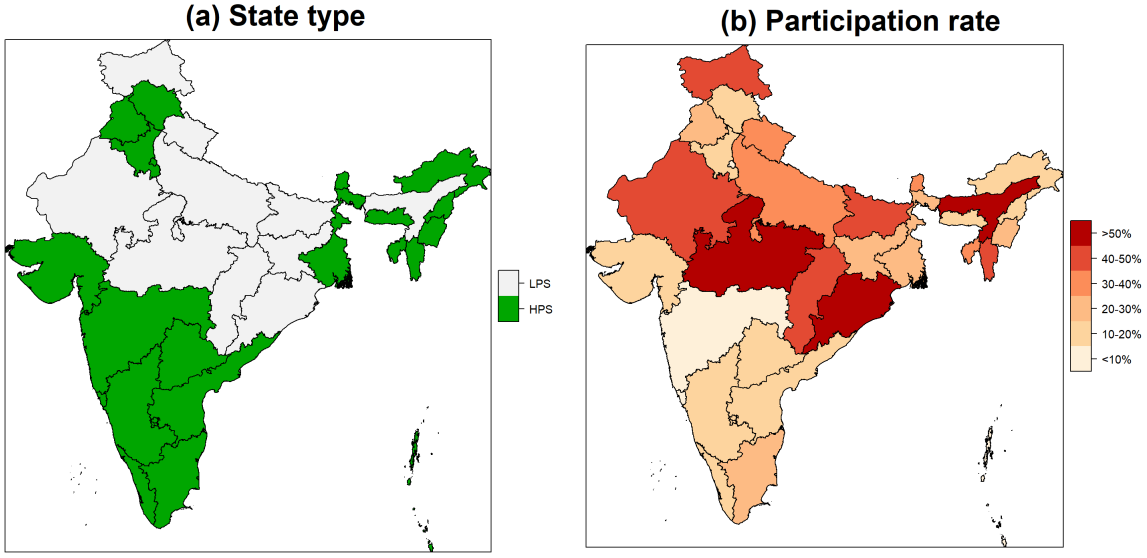
In this paper, we perform inference under a spectrum of assumptions of varying identification power, following the approach taken by Gundersen et al. (2012, 2017); Gonzalez (2005); Gerfin and Schellhorn (2006) and Kreider et al. (2012). Any point-estimate obtained under the conventionally imposed strong assumptions should lie within the bounds if these assumptions are valid. In this study, we show that the causal impacts based on the assumptions employed in previous studies are larger/smaller than the upper/lower limits of the ATE bounds, suggesting that the existing evidence could have been over-estimated/under-estimated.

1.2 Janani Suraksha Yojana (JSY)

In the JSY scheme, women of 19 years of age and above are eligible for a cash benefit if they have a below-the-poverty-line card issued by the government or belong to a scheduled caste or tribe. They can receive 600 Indian rupees in urban areas and 700 rupees in rural areas after delivery in a public health facility for their first two live births. The JSY does not cover the actual cost of institutional delivery and maternal healthcare. The basic JSY scheme used to be the same across the country, but after November

[2]This includes public hospitals and accredited private institutions.

Figure 1: State types and JSY participation rate in 2010-16



Source: Author's calculation from the National Family Health Survey.
Note: LPS=Low performing state; and HPS=High performing state.

2006, different eligibility criteria and cash transfer amounts began to be applied in ten states with high levels of maternal mortality and low levels of institutional delivery (i.e., Uttar Pradesh, Uttaranchal, Bihar, Jharkhand, Madhya Pradesh, Chhattisgarh, Assam, Rajasthan, Odisha and Jammu and Kashmir) (Bredenkamp, 2009). In these ten low-performing states (LPSs), the JSY provides a cash incentive to all women regardless of age, numbers of children or socio-economic status. In other words, every woman who gives birth in a public facility is eligible to receive a cash benefit. Moreover, in the LPSs, a higher cash incentive is provided, namely 1,000 rupees in urban areas and 1,400 rupees in rural areas.[3] In other states that are classified as high-performing states (HPSs), the original eligibility criteria, as well as the same cash incentives for pregnant mothers, continue to be applied. The left panel of Figure 1 illustrates the HPSs and LPSs in the country and the right panel of Figure 1 shows the participation rates of the JSY across the states.

The JSY addresses both the demand and the supply constraints of maternal healthcare services, i.e., it also has a supply-side component, in which community-level health workers are given incentive payments for encouraging pregnant women to give birth at health facilities. Community-level health workers, known as 'accredited social health activists', are the first and most important point of contact for pregnant women. These accredited social health activists identify and register all pregnant women, assist them in developing

[3]This is around 8-12 paid days off from minimum wage manual labour (Joshi and Sivaram, 2014).

their birth plans and arrange the JSY for them. Accredited social health activists are given additional cash incentives to encourage mothers to complete postnatal care. Cash incentives given to health workers are intended to reduce absenteeism and enhance their overall performance.

2 Related literature

2.1 Evidence of the impacts of the JSY on maternal healthcare use

The positive impacts of CCTs in Latin America on maternal healthcare utilisation are reported by numerous studies (Ranganathan and Lagarde, 2012; Barber and Gertler, 2010; Morris et al., 2004). For systematic literature reviews of the CCT programme, see Glassman et al. (2013); Lagarde et al. (2007); Bassett (2008) and Gaarder et al. (2010). In Mexico and El Salvador, for example, birth attendance by skilled personnel increased by 11.4 percentage points and 12.3 percentage points respectively (Urquieta et al., 2009; de Brauw and Peterman, 2011). However, different from CCTs which have much broader aims beyond maternal healthcare, studies on those with narrower aims, such as the JSY, are relatively sparse. Most of the existing literature explored the descriptive associations between JSY receipt and healthcare use and health (Gupta et al., 2012, 2011; Thongkong et al., 2017; Randive et al., 2014, 2013; Mukherjee and Singh, 2018; Purohit et al., 2014; Ng et al., 2014; Gopalan and Varatharajan, 2012), while the studies estimating the causal impacts of the JSY are rather limited. This is mainly because the rigorous randomised controlled trial (RCT) designed for the evaluation of the programme was not conducted in India (Joshi and Sivaram, 2014).

Lim et al. (2010) is the first seminal study estimating the causal impact of JSY from the observational data. Specifically, Lim et al. (2010) estimate the causal impacts on institutional delivery, use of antenatal care with the individual-level and district-level data. According to their estimates with matching, the JSY increases institutional deliveries by 43.5 percentage points and skilled birth attendance by 36.6 percentage points. In addition, they report a significant increase in the use of antenatal care by 10.7 percentage points.

After the study by Lim et al. (2010), a number of studies extend analyses; they explore heterogeneity in average treatment effect (ATE) across the population and estimate the impacts on the use of various healthcare services (Carvalho et al., 2014). As well as the increase in maternal healthcare utilisation, reductions in neonatal mortality and increases in the use of child healthcare are reported by Sengupta and

Sinha (2018), in which authors estimate the causal impacts with inverse probability weighting regression approach. In addition, improvements in the mental health of beneficiary mothers (Powell-Jackson et al., 2016), in breastfeeding and more pregnancies (Powell-Jackson et al., 2015; Nandi and Laxminarayan, 2016) are reported as additional – unintended – benefits.

Most studies exploit the second and the third waves of the District Level Household and Facility Survey (DLHS-2 and DLHS-3), of which data were collected in 2002-2004 and 2007-2008. DLHS-2 is the survey conducted before the JSY was launched, and DLHS-3 was conducted in the initial stage of the JSY implementation. Despite the introduction of JSY in 2005, its proper implementation actually started in 2007 (Das et al., 2011). Das et al. (2011) argues that in DLHS-3, many women who gave birth before 2007 were misclassified as JSY beneficiaries and only a smaller proportion of women knew about the programme itself at that time. Also, during the first few years of the implementation, many institutions were not adequately prepared for maternal and child healthcare (Gopichandran and Chetlapalli, 2012). It turned out that the DLHS-3 was not appropriate for causal analysis, and we had to wait for new data to become available in order to estimate the causal effect of the JSY programme.

Recently, Rahman and Pallikadavath (2017) and Rahman and Pallikadavath (2018) have re-estimated the impact of the JSY with the latest wave of the DLHS (DLHS-4), which was implemented in 2013-2014, when JSY had matured enough to be known by almost all pregnant women. Rahman and Pallikadavath (2017) estimate the impacts of the JSY on various healthcare utilisations with the propensity score matching (PSM). Rahman and Pallikadavath (2018) additionally estimate the impacts with the fuzzy regression discontinuity design (RDD), which exploits the changes of eligibility for the JSY programme with birth orders in HPSs. Major research into the causal effects of the JSY are listed in Table 1.

2.2 Identification assumptions in previous studies

Most of the existing literature which uses individual level data rests on strong assumptions in order to point-identify the ATE. Previous studies on the JSY imposed the conditional independence assumption (CIA), requiring that conditional on specific observable characteristics, the selection of the treatment is random (Imbens and Wooldridge, 2009).[4] Methods based on the CIA do not assume the possibility that unobserved differences between participants and non-participants are associated with the difference

[4] A notable exception is Powell-Jackson et al. (2016), who collect their own data and exploit the fact that some women did not receive the cash due to administrative problems in its disbursement.

Table 1: Major studies estimating the causal effects of the JSY

Literature	Data	Methods/Identification source	Treatment indicator	Identified Effect	Main outcomes
Study on targeted outcomes					
Lim et al. (2010)	DLHS 2-3	Matching, Regression District-level difference-in-differences	JSY participation JSY coverage rate	ATE ATE	Delivery, birth attendance, mortality
Carvalho et al. (2014)	DLHS 3	Propensity score matching	JSY participation	ATE	Immunisation, PNC, breastfeeding
Sengupta and Sinha (2018)	DLHS 3	OLS adjusted with the Inverse probability weighting	JSY participation	ATE	Mortality, stillbirth, PNC, immunisation
Rahman and Pallikadavath (2017)	DLHS 4	OLS, Propensity score matching	JSY participation	ATT	Delivery, ANC, PNC, immunisation
Rahman and Pallikadavath (2018)	DLHS 4	Propensity score matching, Fuzzy regression discontinuity design	JSY participation JSY participation	ATT, LATE	Delivery, ANC, PNC
Joshi and Sivaram (2014)	DLHS 2-3	Individual-level difference-in-differences	JSY eligibility	ITT	ANC, delivery, PNC, birth attendance
Study on additional unintended benefits					
Powell-Jackson et al. (2015)	DLHS 2-3	District-level difference-in-differences	JSY coverage rate	ATE	Mortality, delivery, ANC, pregnancy, breastfeeding
Powell-Jackson et al. (2016)	Own data, census	Quasi-experimental design, exploiting the administrative problems in the disbursement	JSY participation	ATE	Maternal mental health
Nandi and Laxminarayan (2016)	DLHS 2-3	Propensity score matching + difference-in-differences, exploiting the difference in the incentive structure	JSY participation	ATT	Fertility

Note: ATE=Average treatment effect; ATT=Average trreatment effect on the treated; ITT=Intention-to-treat effect; LATE=Local average treatment effect; ANC=Antenatal care; and PNC=Postnatal care.

in outcomes. However, participation in the JSY programme is likely to be dependent on individual unobservable characteristics that could also affect healthcare use, such as the degree of being risk-averse. Also, participants in the JSY may be more aware of the importance of maternal healthcare use and the potential risk of delivery at home. If the receipt of the JSY is related to such unobservable characteristics, the approaches based on the CIA still suffer selection bias and fail to estimate the causal impact.

Joshi and Sivaram (2014) estimated the intention-to-treat effect to deal with the selection problems that had not been properly addressed, using the eligibility criteria, but we should note that the intention-to-treat effect (ITT) is conceptually different from the treatment effects explored in other literature (Table 1). The district-level difference-in-differences (DID) is employed in Lim et al. (2010) and Powell-Jackson et al. (2015), but the effect identified with the district-level DID is also conceptually different from the ATEs identified in the other studies that rely on individual-level data (Table 1). As the district-level DID exploits the variation in the roll-out of the JSY, it estimates the effect of percent point change in the JSY coverage rate on the probability of healthcare use, not the effect of the JSY participation itself. Hence in this study, we do not consider the ITT or the district-level DID.

The fuzzy RDD can deal with the individual unobservable characteristics, but the estimate is local in the sense that it is the impact among those who are having a value of the running variable near the cut-off value (Imbens and Angrist, 1994). Hence it is often difficult to extrapolate the impact for the entire population from the estimated results among the sub-population. Rahman and Pallikadavath (2018) use the birth order (parity) as a running variable and exploit the gap in the probability to participate in the programme between mothers having one and two children and mothers having three to seven children. However, the parity itself is highly likely to be endogenous. Hence, the observed discontinuity in the probability of programme participation could be invalid as an exogenous source of identification.

In contrast to the existing literature, this study takes a totally different approach. Rather than relying on strong yet questionable assumptions to achieve point-identification of the causal impact, we impose just weak but credible identification assumptions to get the bound-estimation of the causal effect (Manski, 1990). The partial identification approach highlights what may be learned from the data without invoking potentially untenable assumptions (Manski, 2003, 2013). By gradually adding assumptions, we explore how much we can narrow the estimated bounds. The partial identification approach is very useful in the

situation in which a rigorous RCT was not implemented and hence valid instrumental variables to deal with possible endogeneity are not available to researchers. The partial identification approach has been applied to various topics in microeconomics, such as labour (Lee, 2009; Gonzalez, 2005; Blundell et al., 2007), crime (Manski, 2013; Manski and Nagin, 1998), education (Ginther, 2000; Huber et al., 2017) and health (Gerfin and Schellhorn, 2006; Gundersen et al., 2017, 2012; Kreider et al., 2012).

The identification assumptions employed in the existing literature have not been sufficiently justified or assessed. In this study, we first re-estimate the causal impacts of the JSY with the DLHS-4, closely following the approaches taken by Rahman and Pallikadavath (2018), Lim et al. (2010), and Sengupta and Sinha (2018), which all rely on the CIA. Then we estimate the bounds of ATE through the partial identification approach. We compare the point-identified impacts with the bound-identified impacts in order to assess the validity of the assumptions used by their respective studies.

3 Data

3.1 District Level Household and Facility Survey (DLHS)

We use the latest wave of the District Level Household and Facility Survey (DLHS-4), which was conducted in 2013-2014, where the JSY had been rolled out across the country and had been matured (Rahman and Pallikadavath, 2017, 2018). Hence the DLHS-4 does not suffer the data contamination problem that is found in DLHS-3 (Das et al., 2011). The DLHS-4 covers the 18 HPSs[5] and three high-performing union states.[6] The DLHS-4 is a representative survey only at the district level, and it collects the data only in HPSs. So this paper only focuses on the HPSs. We complement our finding in HPSs by proving additional evidence in LPSs using another latest survey data, the National Family Health Survey (NFHS-4), in which data have been collected also in LPSs though sample sizes are much smaller. Results for LPSs are discussed in section A.1 in Appendix A.

3.2 Outcomes

This study estimates the impacts on the following eight maternal and child healthcare utilisations. First, we estimate the impact on (1) giving birth at health institutions[7], which is the primary target outcome

[5]They are Andhra Pradesh, Arunachal Pradesh, Goa, Haryana, Himachal, Pradesh, Karnataka, Kerala, Maharashtra, Manipur, Meghalaya, Mizoram, Nagaland, Punjab, Sikkim, Tamil Nadu, Telangana, Tripura and West Bengal.

[6]They are Andaman and Nicobar Islands, Chandigarh, and Puducherry.

[7]Health institution herein includes both public and private health sectors.

of the JSY programme and (2) skilled birth attendance. We also estimate the impacts on the uses of (3) antenatal care (ANC) at least once, (4) ANC three times or more, (5) postnatal care (PNC) for mothers, (6) PNC for babies, (7) iron and folic acid (IFA) tablets/syrup during pregnancy and (8) tetanus toxoid (TT) injections to prevent babies from getting tetanus after birth.

3.3 Covariates for the point-identification

In order to make our results comparable with those in the existing literature as much as possible, we control for the individual- and household-level confounding factors, closely following Rahman and Pallikadavath (2018). As factors reflecting demographic and socio-economic characteristics, we control for below-the-poverty-line card ownership, maternal age, a residential location (urban/rural), and a birth order (parity). A Hindu dummy variable is used to reflect individual socio-cultural backgrounds. Other religions (e.g., Muslim, Christianity, Sikhism, Buddhism) are benchmark groups. In addition, we use dummy variables for the scheduled castes and tribes. Scheduled castes/tribes are the most socially disadvantaged groups, members of which have suffered the greatest burden of social and economic segregation and deprivation (Chitnis, 1997). We control for parental educational levels that are measured by the number of education years. The household wealth is captured through a composite index of relative standards of living.[8]

3.4 Sample selection

In this study, we focus only on women aged between 15 and 49 who gave birth in the five years prior to the survey, so that all of those analysed in this study gave birth after the proper implementation of the JSY. We restrict our attention to the most recent birth of each woman and exclude women who participated in programmes for childbirth other than the JSY. The DLHS-4 has 84,266 observations. After dropping observations with missing information[9], our sample size has become 67,595. Descriptive statistics are shown in Table 2.

[8]Following Rahman and Pallikadavath (2018), we derive the wealth index by applying the principal component analysis over various household characteristics. They are cooking fuel, house type, number of dwelling rooms, electricity, house ownership, landholding, radio, television, computer, internet, telephone, mobile phone, washing machine, refrigerator, sewing machine, watch, bicycle, motorcycle, car, tractor, tube well, cart, and air cooler.

[9]We dropped 27 observations because of the missing information about the below-the-poverty-line card ownership. We dropped 82 and two observations that do not have information about the caste and religions respectively. 2,916 observations have been dropped because of the missing information about parity. We dropped 79 observations and 13,171 observations because they do not have information about maternal and parental educational backgrounds. We dropped eight observations with missing information about household wealth. Lastly, we dropped 386 observations with missing information about outcomes.

Table 2: Descriptive statistics of the DLHS-4

	Non-participants mean	Participants mean
Institutional delivery	0.82	0.95
Skilled birth attendance	0.89	0.97
Antenatal care at least once	0.87	0.96
Antenatal care ($\geq$ 3 times)	0.70	0.81
Postnatal care for mother	0.64	0.72
Postnatal care for baby	0.78	0.83
Iron and folic acid (IFA) supplement	0.68	0.82
Tetanus toxoid (TT) injection	0.83	0.93
Below the poverty line card	0.30	0.46
Scheduled caste	0.77	0.75
Scheduled tribe	0.16	0.19
Rural	0.56	0.67
Birth order (parity)	2.01	1.78
Hindu	0.67	0.69
Maternal age	27.23	26.05
Maternal education years	9.81	9.23
Paternal education years	10.92	9.98
Wealth	0.54	-0.26
Observations	52732	14863

Source: DLHS-4.

4 Methods

4.1 Notations

Let $Y = \{0, 1\}$ be an indicator for observed health care utilisation. Y becomes 1 if a woman uses a health care service. $D = \{0, 1\}$ is an indicator for the participation in the JSY. It equals 1 for those who participated in the JSY programme and 0 otherwise. Following Rubin (1974), we assume that each individual has two potential outcomes, namely Y_0 in the absence of the treatment and Y_1 in the presence of the treatment. They are latent, in the sense that we can only observe one of them for each person, but never both. What is actually observed for researchers is

$$Y = DY_1 + (1 - D)Y_0. \tag{1}$$

The causal impact of the JSY for individual $j \in J$ is $Y_{1j} - Y_{0j}$, and we are interested in the average treatment effect (ATE), which is defined by $ATE \equiv P(Y_1 = 1) - P(Y_0 = 1)$. Theoretically, the ATE ranges from -1 to +1 and has a width of 2. Let X be observed individual characteristics, which determine Y and D.

Using the law of total probability, we can express $P(Y_1 = 1)$ and $P(Y_0 = 1)$ as

$$P(Y_1 = 1) = P(Y_1 = 1|D = 1)P(D = 1) + \underbrace{P(Y_1 = 1|D = 0)}_{Unobservable} P(D = 0) \quad (2)$$

$$P(Y_0 = 1) = \underbrace{P(Y_0 = 1|D = 1)}_{Unobservable} P(D = 1) + P(Y_0 = 1|D = 0)P(D = 0), \quad (3)$$

where $P(Y_1 = 1|D = 0)$ is the counterfactual probability that untreated women would use health care service under the treatment, and $P(Y_0 = 1|D = 1)$ is the counterfactual probability that treated women would use health care service if they had not been treated. $P(Y_1 = 1|D = 1)$, $P(Y_0 = 1|D = 0)$, $P(D = 1)$ and $P(D = 0)$ are immediately identifiable from the distribution of the observed data, but $P(Y_1 = 1|D = 0)$ and $P(Y_0 = 1|D = 1)$ are unobservable, which makes the ATE unidentifiable without further assumptions.

4.2 Point-identification with the independence assumption

Researchers must impose assumptions regarding $P(Y_1 = 1|D = 0)$ and $P(Y_0 = 1|D = 1)$ in order to identify the ATE. For example, if we assume that participation in the treatment is random, we can point-identify the ATE from the observed distribution. This assumption is called the independence assumption and can be expressed as

$$P(Y_1 = 1|D = 0) = P(Y_1 = 1|D = 1) \quad (4)$$

$$P(Y_0 = 1|D = 0) = P(Y_0 = 1|D = 1). \quad (5)$$

The ATE is point-identified by $ATE = P(Y_1 = 1|D = 1) - P(Y_0 = 1|D = 0)$. This independence assumption is valid under the rigorous RCTs, without which it is usually too stringent. Another identification assumption that is conventionally imposed in the policy evaluation literature is the conditional independence assumption (CIA) in which participation in the treatment is assumed to be random, conditional on the observable characteristics X. The CIA is formalised by

$$P(Y_1 = 1|D = 0, X) = P(Y_1 = 1|D = 1, X) \quad (6)$$

$$P(Y_0 = 1|D = 1, X) = P(Y_0 = 1|D = 1, X). \quad (7)$$

Table 3: Point-identification approaches and assumptions

Method	**Short name**	**Assumption**
Mean comparison	Independence	Exogenous/independent treatment assignment
Multivariate linear regression	OLS	Conditional independence; Outcome linear functional-from; Additive linearity of the treatment status and the unobservable; Same treatment effect for all individuals
Propensity score matching	PSM	Conditional independence; Propensity score functional-form
Nearest neighbour matching	NNM	Conditional independence
OLS adjusted with the inverse probability weighting	IPW+OLS	Conditional independence
Entropy balance weighting	EBW	Conditional independence
Regression discontinuity design	RDD	Exogenous discontinuity in the probability of the JSY participation at birth order two

For example, evaluations with the multivariate regression approach and the propensity score matching approach are based on this assumption (Imbens and Wooldridge, 2009; Abadie and Cattaneo, 2018). Although the CIA itself cannot be tested by the data, this assumption can be untenable in many cases. If the decision to participate in the programme is dependent on unobservable characteristics that can affect outcomes, the CIA fails, resulting in a biased causal effect estimate.

In this study, we point-estimate the ATEs by (1) mean comparison between the participants and non-participants, which we call Independence, (2) multivariate or ordinary least squares regression (OLS), (3) propensity score matching (PSM), (4) OLS regression with inverse probability weighting adjustment (IPW+OLS), (5) nearest neighbourhood matching (NNM), (6) entropy balance weighting (EBW) and (7) fuzzy regression discontinuity design (RDD). Their key assumptions are summarised in Table 3.

The mean comparison between the participants and non-participants rests on the independence assumption, and the OLS, PSM, IPW+OLS, NNM and EBW approach rest on the CIA. Some also rely on functional form assumptions. For example, the OLS assumes that the treatment variable and the error term are additively separated and the treatment effect is identical for everyone. The PSM and IPW+OLS rely on the functional form assumption of the propensity score. The EBW non-parametrically balances the moments of the covariate distributions, the algorithm of which is explained in the Appendix B.1.

For fuzzy RDD, we also closely follow the approach taken by Rahman and Pallikadavath (2018), which exploits the JSY eligibility, i.e. mothers are eligible for the JSY up to their second live birth. The eligibility change of JSY at birth order two was exploited as a source of identification. Following Rahman and Pallikadavath (2018), we first run OLS regression of the JSY dummy variable on $z_1 = 1\{parity \leq 2\}$, $z_2 = z_1 * (parity - 2)$, $z_3 = (parity - 2)$ and the other covariates as the first stage regression, where z_1 and z_2 are excluded instruments. Using the predicted value of the JSY dummy variable, we estimate the causal impact of JSY in the second stage OLS regression.

The DLHS is repeated cross-sectional data. If the longitudinal data were available, we could possibly employ the fixed effect (FE) model, which can deal with the situation where individual unobservable and time-invariant characteristics influence the selection of the treatment. However, the FE model is less appealing for the study on maternal health care, because it is not realistic to assume that the degree of risk-aversion is time-invariable; it can change as mothers experience more childbearing. For example, the motivation to join the JSY programme for the primiparity can be different from the one for the second or third childbirth. In this paper, we do not consider the individual-level DID approach either, mainly for two reasons. First, the JSY is the nationwide programme, which makes it difficult to set up the control group. Second, we have gaps of more than 10 years between the data collection before and after the JSY implementation; the DLHS-2 was conducted in 2002-2004. The long periods between the waves could substantially deteriorate the credibility of common-trend assumption that is essential for the DID.

4.3 Partial identification assumptions

In essence, partial identification first estimates sharp bounds for $P(Y_1 = 1)$ and $P(Y_0 = 1)$ and then constructs a sharp bound of ATE. A sharp bound is defined as the narrowest bound that can be obtained under the maintained assumptions regarding the unobservable distribution. When the bound of $P(Y_t = 1)$ is (LB_t, UB_t), the ATE bound is defined as

$$(LB_1 - UB_0) \leq ATE \leq (UB_1 - LB_0). \tag{8}$$

Note that first the ATE bound is conceptually different from the confidence interval of the point-estimated ATE. The width of ATE bound reflects the identification power of the imposed assumptions. The bound width indicates the tension between the strength of assumptions and their credibility (Manski, 2007),

and the bound width does not change with the change in sample sizes. On the other hand, the width of confidence interval for the point-identified ATE reflects the uncertainty of sampling variability, which does vary with the change in sample sizes. We discuss the uncertainty of sampling variability for the ATE bound in section 4.5. In this paper, unless otherwise indicated, we use the term 'bound' to denote an ATE bound of the partial identification and use the term 'interval' to indicate a confidence interval.

Second, as the partial identification approach does not assume that participation in the treatment is random conditional on the covariates, there is no specific need to condition on a long list of covariates. Hence the partial identification approach is not susceptible to criticisms of omitted variable bias (Manski and Nagin, 1998). Splitting the sample by a covariate to explore heterogeneity across the population is also possible for the partial identification as well, but in this paper we do not implement sample splitting.

4.3.1 No assumptions (Worst-case)

First, as a benchmark, we specify a range of $P(Y_1 = 1|D = 0)$ and $P(Y_0 = 1|D = 1)$ to construct a bound for $P(Y_1 = 1)$ and $P(Y_0 = 1)$ without imposing any assumption regarding the counterfactual probabilities. Since $P(Y_1 = 1|D = 0)$ and $P(Y_0 = 1|D = 1)$ are probabilities, they necessarily belong to [0,1]. Hence the bounds of $P(Y_1 = 1)$ and $P(Y_0 = 1)$ are given from equations (2) and (3) by

$$\underbrace{P(Y_1 = 1|D = 1)P(D = 1)}_{LB_1} \leq \quad \boldsymbol{P(Y_1 = 1)} \quad \leq \underbrace{P(Y_1 = 1|D = 1)P(D = 1) + P(D = 0)}_{UB_1} \tag{9}$$

$$\underbrace{P(Y_0 = 1|D = 0)P(D = 0)}_{LB_0} \leq \quad \boldsymbol{P(Y_0 = 1)} \quad \leq \underbrace{P(Y_0 = 1|D = 0)P(D = 0) + P(D = 1)}_{UB_0}. \tag{10}$$

In equation (9), the lower bound of $P(Y_1 = 1)$ is attained if all non-participants would have used the healthcare had they participated in the JSY. The upper bound of $P(Y_1 = 1)$ is attained if all non-participants would have had no health care had they participated in the JSY. The bound width of $P(Y_1 = 1)$ is $P(D = 0)$. The bounds of $P(Y_0 = 1)$ in equation (10) can be interpreted in the same fashion and its bound width is $P(D = 1)$. The sharp bound of ATE can be obtained via equation (8) (Manski,

Table 4: Partial-identification assumptions

Assumption	**Implication**
Monotone treatment response (MTR)	Individuals do not participate in the programme that makes them worse-off with respect to the outcome.
Positive/negative monotone treatment selection (MTS)	The treated are more/less likely to have the healthcare than the non-treated both in the presence and absence of the treatment.
Monotone instrumental variable (MIV)	Eligible people are less likely to have healthcare in the presence and absence of the treatment conditional on the treatment.

1990):

$$\begin{aligned} & \underbrace{P(Y_1=1|D=1)P(D=1)}_{LB_1} - \underbrace{\{P(Y_0=1|D=0)P(D=0)+P(D=1)\}}_{UB_0} \\ \leq \; & ATE \\ \leq \; & \underbrace{\{P(Y_1=1|D=1)P(D=1)+P(D=0)\}}_{UB_1} - \underbrace{P(Y_0=1|D=0)P(D=0)}_{LB_0}. \end{aligned} \tag{11}$$

This bound is the narrowest sharp band that can be inferred from the data alone, which is often called a worst-case bound. Note that even without any assumption about the unobserved probabilities, the data alone tighten the theoretical width of ATE to half if the support of the outcome variable is bounded (Manski, 1989). It is known that without any identification assumption, the bound on the ATE in equation (11) always has a width of 1 and includes 0. Hence without additional assumptions, we cannot evaluate the sign of the treatment effect on healthcare use. We consider three assumptions: (i) Monotone treatment response (MTR), (ii) Monotone treatment selection (MTS) and (iii) Monotone instrumental variable (MIV). The key implications of each assumption is summarised in Table 4.

4.3.2 Monotone treatment response

Second, we assume that individuals do not select a treatment that would make them worse off with respect to maternal and child healthcare use. This assumption is called the monotone treatment response (MTR) assumption and it implies that, *ceteris paribus*, response varies monotonically with treatment (Manski, 1997).

Formally, the MTR assumes that for all individual $j \in J$, $Y_{0j} \leq Y_{1j}$. We obtain the new bounds for $P(Y_1 = 1|D = 0)$ and $P(Y_0 = 1|D = 1)$ as follows:

$$\begin{aligned} P(Y_0 = 1|D = 0) &\leq \boldsymbol{P(Y_1 = 1|D = 0)} \leq 1 \quad (12) \\ 0 &\leq \boldsymbol{P(Y_0 = 1|D = 1)} \leq P(Y_1 = 1|D = 1). \quad (13) \end{aligned}$$

We observe the shrink in the counterfactual probability bounds. For $P(Y_1 = 1|D = 0)$, while the bound of this counterfactual probability in the worst-case ranges from 0 to 1, its new range under the MTR assumption ranges from $P(Y_0 = 1|D = 0)$ to 1. Also, the new range of $P(Y_0 = 1|D = 1)$ under the MTR assumption is from 0 to $P(Y_1 = 1|D = 1)$. The following bounds of $P(Y_1 = 1)$ and $P(Y_0 = 1)$ can be obtained from equations (1) and (3):

$$\begin{aligned} & P(Y_1 = 1|D = 1)P(D = 1) + P(Y_0 = 1|D = 0)P(D = 0) \\ = \; & \underbrace{P(Y = 1)}_{Updated\, LB_1} \\ \leq \; & \boldsymbol{P(Y_1 = 1)} \quad (14) \\ \leq \; & P(Y_1 = 1|D = 1)P(D = 1) + P(D = 0) \end{aligned}$$

and

$$\begin{aligned} & P(Y_0 = 1|D = 0)P(D = 0) \\ \leq \; & \boldsymbol{P(Y_0 = 1)} \quad (15) \\ \leq \; & P(Y_1 = 1|D = 1)P(D = 1) + P(Y_0 = 1|D = 0)P(D = 0) \\ = \; & \underbrace{P(Y = 1)}_{Updated\, UB_0}, \end{aligned}$$

where the lower bound of $P(Y_1 = 1)$ and the upper bound of $P(Y_0 = 1)$ in the worst-case are replaced both by $P(Y = 1)$. By equation (8), it follows

$$0 \leq ATE \leq \{P(Y_1 = 1|D = 1)P(D = 1) + P(D = 0)\} - P(Y_0 = 1|D = 0)P(D = 0). \quad (16)$$

Under the MTR assumption, the lower bound of ATE becomes 0. Hence, the MTR assumption precludes a non-negative ATE and it assumes away the possibility of the detrimental impact of the JSY. In

other words, the MTR assumes that the JSY programme would never decrease the likelihood of receiving healthcare. The MTR assumption, however, leaves open the question of whether the programme has a strong beneficial effect, mild beneficial effect, or no effect (Gundersen et al., 2017).

In general, the credibility of the MTR assumption depends on the type of policy that we are analysing. As the MTR assumes the sign of the treatment effect, the MTR assumption can be admittedly too stringent in the case where we have no clue at all regarding the sign of the ATE a priori. However, in the context of the JSY programme, it is unlikely that a mother does not use the healthcare service when she is treated and does use the service when she is untreated, because the accredited social health activists, who support pregnant mothers in the JSY programme, encourage women to receive timely maternal and child healthcare. Furthermore, they encourage women to undergo three antenatal cares and to give birth in a health institution. They also assist pregnant women to obtain tetanus toxoid injections and iron/folic acid supplements. The Ministry of Health and Family Welfare also states that all pregnant mothers should receive at least three ANCs in their guidelines for the implementation of JSY. Therefore, it is reasonable to presume that the JSY programme will not decrease healthcare use.

4.3.3 Monotone treatment selection

Third, we make an assumption on the selection mechanism through which women participate in the JSY. Specifically, we make an assumption as to whether individuals participating in the JSY programme are more likely to use maternal and child healthcare services on average or not, conditional on treatment assignment. This assumption is called the monotone treatment selection (MTS) (Manski and Pepper, 2000). The MTS assumption sounds similar to the MTR assumption, but they are different, although they are not mutually exclusive. In contrast to the MTR assumption, which is on individual-level potential behaviours, the MTS assumption is on the expected probability. The MTS can be regarded as the weaker version of the independence assumption, where the equalities in equations (4) and (5) are weakened to inequalities. There are two types of MTS assumption, namely positive MTS and negative MTS. They make different assumptions regarding the direction of the selection bias.

(i) Positive MTS: When we assume that mothers participating in the JSY are likely to use no less healthcare services on average than non-participants, conditional on treatment assignment, this assumption is called the positive MTS assumption. This assumption is plausible in the situations where risk-averse

women are more likely to participate in the JSY and use healthcare services than women with lower degrees thereof and/or where women in socially and economically segregated ethnic groups know less about the JSY. The positive MTS assumption is formalised as follows:

$$0 \leq \quad \boldsymbol{P(Y_1 = 1|D = 0)} \quad \leq P(Y_1 = 1|D = 1) \tag{17}$$

$$P(Y_0 = 1|D = 0) \leq \quad \boldsymbol{P(Y_0 = 1|D = 1)} \quad \leq 1. \tag{18}$$

Note that in equations (17) and (18), the MTS assumes only the direction of the selection bias, and it does not assume its strength. The positive MTS assumption lowers the upper bound of $P(Y_1 = 1|D = 0)$ in the worst-case from 1 to $P(Y_1 = 1|D = 1)$ and raises the lower bound of $P(Y_0 = 1|D = 1)$ in the worst-case from 0 to $P(Y_0 = 1|D = 0)$.

Then the new sharp bounds of $P(Y_1 = 1)$ and $P(Y_0 = 1)$ can be obtained from equations (2) and (3) as follows (Manski and Pepper, 2000),

$$P(Y_1 = 1|D = 1)P(D = 1) \leq \quad \boldsymbol{P(Y_1 = 1)} \quad \leq \underbrace{P(Y_1 = 1|D = 1)}_{Updated\, UB_1} \tag{19}$$

$$\underbrace{P(Y_0 = 1|D = 0)}_{Updated\, LB_0} \leq \quad \boldsymbol{P(Y_0 = 1)} \quad \leq P(Y_0 = 1|D = 0)P(D = 0) + P(D = 1). \tag{20}$$

Compared with their bounds in the worst-case, the upper bound of $P(Y_1 = 1)$ has become smaller and the lower bound of $P(Y_0 = 1)$ has become larger. The sharp bound of ATE is given via the equation (8) and it is

$$\begin{aligned} & P(Y_1 = 1|D = 1)P(D = 1) - \{P(Y_0 = 1|D = 0)P(D = 0) + P(D = 1)\} \\ \leq \quad & ATE \\ \leq \quad & P(Y_1 = 1|D = 1) - P(Y_0 = 1|D = 0). \end{aligned} \tag{21}$$

We find that, compared with the worst-case ATE bound in equation (11), the upper bound in equation (21) has become smaller. It is easy to show that the upper bound of ATE under the MTS assumption corresponds to the ATE estimated under the independence assumption (see the note in the Appendix B.2), which implies that the upper bound of ATE is achieved when there exists no selection bias regarding the participation in the JSY.

(ii) Negative MTS: If we anticipate that mothers who had the greater difficulty in using healthcare services have higher motivation to participate in the JSY or they are more strongly encouraged by the accredited social health activists to participate in the JSY, it may be more appropriate to assume that mothers participating in the JSY are likely to use no more healthcare services on average than non-participants. This assumption is called the negative MTS assumption. The negative MTS assumption is formalised as follows:

$$P(Y_1 = 1|D = 1) \leq \quad \boldsymbol{P(Y_1 = 1|D = 0)} \quad \leq 1 \tag{22}$$

$$0 \leq \quad \boldsymbol{P(Y_0 = 1|D = 1)} \quad \leq P(Y_0 = 1|D = 0). \tag{23}$$

The negative MTS assumption raises the worst-case upper bound of $P(Y_1 = 1|D = 0)$ from 0 to $P(Y_1 = 1|D = 1)$ and lowers the worst-case lower bound of $P(Y_0 = 1|D = 1)$ from 1 to $P(Y_0 = 1|D = 0)$. The bounds of $P(Y_1 = 1)$ and $P(Y_0 = 1)$ become

$$\underbrace{P(Y_1 = 1|D = 1)}_{Updated\, LB_1} \leq \quad \boldsymbol{P(Y_1 = 1)} \quad \leq P(Y_1 = 1|D = 1)P(D = 1) + P(D = 0) \tag{24}$$

$$P(Y_0 = 1|D = 0)P(D = 0) \leq \quad \boldsymbol{P(Y_0 = 1)} \quad \leq \underbrace{P(Y_0 = 1|D = 0)}_{Updated\, UB_0}. \tag{25}$$

The sharp bound of ATE is given via the equation (8) and it is

$$\begin{aligned} & P(Y_1 = 1|D = 1) - P(Y_0 = 1|D = 0) \\ \leq \quad & ATE \\ \leq \quad & \{P(Y_1 = 1|D = 1)P(D = 1) + P(D = 0)\} - P(Y_0 = 1|D = 0)P(D = 0). \end{aligned} \tag{26}$$

We find that, compared with the worst-case bound in equation (11), the lower bound of the ATE in equation (26) has become larger. Also, this bound shows that the lower bound of ATE under the negative MTS assumption corresponds to the ATE estimated under the independence assumption, which implies that the lower bound of ATE is achieved only when there exists no selection bias regarding the participation in the JSY (see the note in the Appendix B.2).

In general sources of selection bias are multiple. The ultimate direction of the selection mechanism is expected to vary across the outcomes. Moreover, in essence, there is no definitive way to examine the

sources of selection bias and the direction of the bias. However, we can obtain a clue by comparing the size of ATE estimated under the independence assumption and that estimated under the CIA. When the ATE under the independence assumption is larger than that under the CIA, it implies the existence of the positive selection bias, part of which is removed by controlling for covariates. On the other hand, when the difference of the ATEs is negative, it suggests the existence of the negative selection bias. We infer the sign of selection bias for each outcome, and impose either positive or negative MTS assumption accordingly.

4.3.4 Monotone instrumental variable

The last assumption we consider is the monotone instrumental variable (MIV) assumption (Manski and Pepper, 2000), which states that the latent probability of having healthcare varies weakly and monotonically with an observed instrument, v. Different from the standard instrumental variable approach (Imbens and Angrist, 1994), the MIV assumption does not impose the mean independence assumption requiring that the latent outcomes are mean independent of the instrument, namely $P(Y_t = 1|v = 1) = P(Y_t = 1|v = 0)$. The instrument v is allowed to be dependent on the mean of potential outcome as long as the direction of its effect is monotone. Hence, in this sense, the MIV is weaker than the mean independence assumption. Also the MIV assumption itself does not impose any assumption regarding the association between the instrument and the treatment status. Thus, the MIV assumption is not susceptible to criticisms of weak instruments.

Following Gundersen et al. (2017, 2012)[10], we use the non-eligibility status of the JSY as an instrument. We assume that ineligible women, who tend to be richer, are more likely to use healthcare services conditional on the treatment than eligible women. This assumption is supported by the observations in developing countries that poorer women are less likely to use maternal healthcare services (Pathak et al., 2010; Pathak and Mohanty, 2010; Kesterton et al., 2010; Balarajan et al., 2011). We discuss the plausibility of our MIV assumption further in section 6.2.

Formally, let v be a binary instrument, and $v = 1$ indicates that a woman is not eligible for the JSY. The

[10]Gundersen et al. (2017, 2012) estimate the ATE bounds of the food assistance programmes targeting poor households in the US on nutritional status.

MIV assumption imposed herein is expressed as

$$P(Y_1 = 1|v = 0) \leq P(Y_1 = 1|v = 1) \tag{27}$$

$$\underbrace{P(Y_0 = 1|v = 0)}_{Eligible\ women} \leq \underbrace{P(Y_0 = 1|v = 1)}_{Ineligible\ women}. \tag{28}$$

Note that equations (27) and (28) are the assumptions on the latent outcomes. The MTS assumption is a special case of the MIV assumption, in which the participation status itself is being used an instrument, i.e $v = D$.

The MIV assumption gives the following sharp bounds of $P(Y_1 = 1)$ and $P(Y_0 = 1)$ in equations (29) and (30). Their derivations are provided in the Appendix B.3.

$$\begin{aligned}
& P(v=0)P(Y_1=1|D=1,v=0)P(D=1|v=0) \\
& +P(v=1)\max\{P(Y_1=1|D=1,v=0)P(D=1|v=0), P(Y_1=1|D=1,v=1)P(D=1|v=1)\} \\
\leq\ & \boldsymbol{P(Y_1=1)} \qquad (29) \\
\leq\ & P(v=0)\min\{P(Y_1=1|D=1,v=0)P(D=1|v=0)+P(D=0|v=0), \\
& P(Y_1=1|D=1,v=1)P(D=1|v=1)+P(D=0|v=1)\} \\
& +P(v=1)[P(Y_1=1|D=1,v=1)P(D=1|v=1)+P(D=0|v=1)]
\end{aligned}$$

and

$$\begin{aligned}
& P(v=0)P(Y_0=1|D=0,v=0)P(D=0|v=0) \\
& +P(v=1)\max\{P(Y_0=1|D=0,v=0)P(D=0|v=0), P(Y_0=1|D=0,v=1)P(D=0|v=1)\} \\
\leq\ & \boldsymbol{P(Y_0=1)} \qquad (30) \\
\leq\ & P(v=0)\min\{P(Y_0=1|D=0,v=0)P(D=0|v=0)+P(D=1|v=0), \\
& P(Y_0=1|D=0,v=1)P(D=0|v=1)+P(D=1|v=1)\} \\
& +P(v=1)[P(Y_0=1|D=0,v=1)P(D=0|v=1)+P(D=1|v=1)].
\end{aligned}$$

The sharp bound of the ATE can be obtained by equation (8).[11] The ATE bounds under the MIV assumption and those in the worst-case scenario coincide if the worst-case lower and upper bounds of $P(Y_t = 1|v = u)$ weakly increase with u; in such cases, the MIV assumption has no identifying power and the MIV assumption does not make the ATE bounds narrower (Manski and Pepper, 2000; Richey, 2016).

4.4 Joint imposition of partial identification assumptions

We consider the case where we jointly impose the assumptions introduced so far. We will add mild assumptions one by one and see how much each additional assumption can make the ATE bound narrower. In the context of the JSY programme, the MTR assumption is the most plausible, followed by the MTS assumption. Lastly we add the MIV assumption which seems the most strongest of all three assumptions. Hence, in this paper we mainly consider the following two cases: MTR+MTS and MTR+MTS+MIV.

When we jointly impose different assumptions, we derive the bounds of $P(Y_1 = 1)$ and $P(Y_0 = 1)$ under the jointly imposed assumptions in order to get the ATE bound. ATE bounds estimators and their full derivations are in the Appendix C. For the other combinations, i.e, MTR+MIV and MTS+MIV, their estimators are also provided in the Appendix D.

4.5 Inference of partial identification

Estimating the ATE bounds requires us to consider sampling variabilities. The ATE bounds introduced so far concern the conclusions that could be drawn under different assumptions if we could observe the JSY participation status and outcomes experienced by everyone in the population. Sampling variability, however, arises when these data are available for only a sample of the population. We consider the sample variability by constructing the confidence intervals for the bound estimates.

A confidence interval for the bound estimate is the area that contains the parameter of interest. However,

[11] It is known that when we estimate the ATE bound under the MIV assumption with the sample analogue, the estimate can suffer finite-sample bias which could potentially lead the bound to be narrower than the true bound (Manski and Pepper, 2009). By Jensen's inequality, the estimated lower bounds are potentially biased upwards because of their maxima operators, and the estimated upper bounds are potentially biased downwards because of their minima operators. Details are provided in the Appendix B.4. To address this bias, we implement a bootstrap-based correction proposed by Kreider and Pepper (2007). This method estimates the bias by using the bootstrap distribution and adjusts the sample analogue estimate in accordance with the estimated bias. For example, when we have a random sample of size N and let LB_N be the sample analogue estimate of the lower bound in question. We denote $E^b(LB_N)$ as the mean of the estimate from the bootstrap distribution of size b. The bias is estimated as $E^b(LB_N) - LB_N$. The bias-corrected estimate is given by $LB_N - \{E^b(LB_N) - LB_N\} = 2LB_N - E^b(LB_N)$. We correct the bias of the upper bound in the same way. In this paper we estimate the bias with $b = 100$ bootstrap repetitions. The performance of this method is confirmed by Monte Carlo simulations in Manski and Pepper (2009).

there is no consensus on its definition and its research is still ongoing. When we consider a confidence interval in the partial identification, the question arises as to whether to construct a interval over the region of identification (Chernozhukov et al., 2007) or over the actual parameter of interest (Imbens and Manski, 2004).[12]

We obtain the confidence intervals for the estimated bounds by the method developed by Imbens and Manski (2004). The confidence intervals defined by Imbens and Manski (2004) are the ones that asymptotically cover the true parameter, namely ATE, with a fixed probability, rather than covering the entire identification region with fixed probability.

Formally, the confidence intervals for parameter θ at the $1-\alpha$ level are defined as the sets of the parameters where we cannot reject the null hypothesis that θ is an element of the identification set, Θ, at the α level. When the estimated lower and upper bounds are $[\widehat{LB}, \widehat{UB}]$ and their standard errors are $\widehat{\sigma_{LB}}$ and $\widehat{\sigma_{UB}}$, $1-\alpha$ level confidence intervals of Imbens and Manski (2004) are constructed as $CI_{1-\alpha} \in [\widehat{LB} - c\widehat{\sigma_{LB}}, \widehat{UB} + c\widehat{\sigma_{UB}}]$, where the parameter c is the one that solves $\Phi(c + \frac{(\widehat{\sigma_{UB}} - \widehat{\sigma_{LB}})}{\max\{\widehat{\sigma_{LB}}, \widehat{\sigma_{UB}}\}}) - \Phi(-c) = 1-\alpha$ with the Newton-Raphson method. For more details, see Imbens and Manski (2004). The confidence intervals are obtained by bootstrap with 200 repetitions.

5 Results

5.1 Point-identification

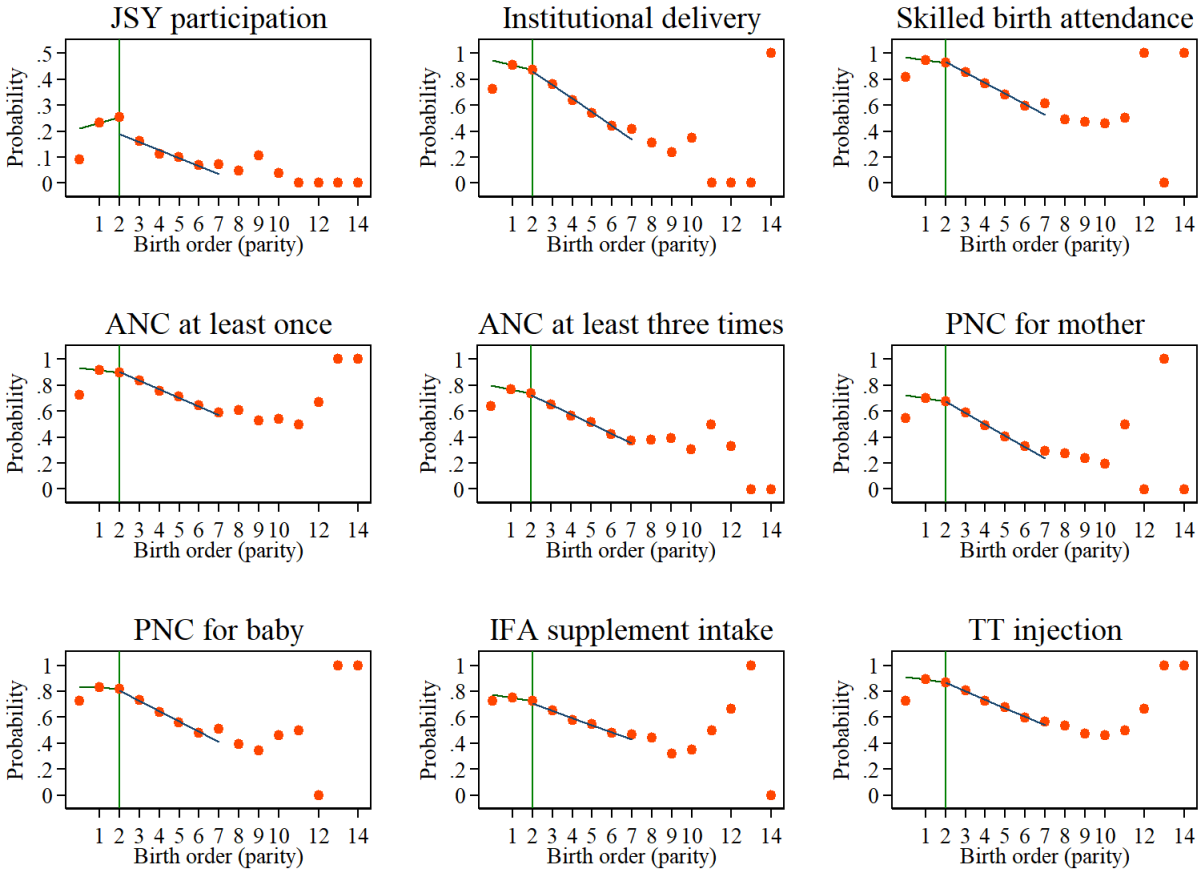
We point-estimate the ATEs by: (1) Independence, (2) OLS, (3) PSM, (4) IPW+OLS, (5) NNM, (6) EBW and (7) RDD. For these approaches (2)-(7), we use the set of covariates used in Rahman and Pallikadavath (2018). The logistic regression is used to estimate the propensity score.[13] Table 5 shows the estimation results, where we observe that the JSY has positive impacts on all outcomes ($p < 0.01$). When we relax the independence assumption and instead impose the CIA, we find larger effects on institutional delivery, skilled birth attendance, PNC for mothers and PNC for babies, and smaller effects on ANC at least once, ANC at least three times, iron and folic acid supplement intakes and tetanus toxoid injections.

[12] Instead of them, some studies provide confidence internals for the upper and lower ATE bounds respectively (Ginther, 2000).

[13] Its result is available from the author upon request.

The results of the fuzzy RDD approach are shown in the last two columns of Table 5. They show the largest ATEs on all outcomes except PNC for mothers. As is well known, the RDD estimates the ATE among the sub-population that are induced to participate in the treatment with the change in instrumental variables, and these people are often called "compliers". Hence, an ATE estimated by RDD can be different from ATEs estimated by other approaches if compliers are very much different from the entire sample population (Imbens and Angrist, 1994). The jumps and kinks in the probability of participating in the programme and the probability of using the healthcare are shown in Figure 2.

Figure 2: Discontinuity at parity more than two



Source: DLHS-4. Note: ANC=Antenatal care; PNC=Postnatal care; IFA=Iron and folic acid; and TT=tetanus toxoid.

5.2 Partial identification

Table 6 is the main result of this study, which reports the change in the bound estimates for each outcome as we incrementally tighten the ATE bounds by adding stronger assumptions one by one. Herein we are considering the following four cases: (1) No assumption, (2) MTR, (3) MTR+MTS, and (4)

Table 5: Point-estimated ATE

	Estimates	SEs	Confidence intervals
Institutional delivery			
Independence	0.126***	0.003	(0.12, 0.132)
OLS	0.129***	0.003	(0.123, 0.135)
PSM	0.129***	0.003	(0.123, 0.135)
IPW+OLS	0.132***	0.003	(0.126, 0.138)
NNM	0.122***	0.003	(0.116, 0.128)
EBW	0.131***	0.003	(0.125, 0.137)
RDD	0.19***	0.037	(0.117, 0.263)
Skilled delivery attendance			
Independence	0.081***	0.002	(0.077, 0.085)
OLS	0.088***	0.002	(0.084, 0.092)
PSM	0.088***	0.002	(0.084, 0.092)
IPW+OLS	0.09***	0.002	(0.086, 0.094)
NNM	0.082***	0.002	(0.078, 0.086)
EBW	0.094***	0.002	(0.09, 0.098)
RDD	0.134***	0.030	(0.075, 0.193)
Antenatal care +1			
Independence	0.091***	0.002	(0.087, 0.095)
OLS	0.079***	0.003	(0.073, 0.085)
PSM	0.078***	0.003	(0.072, 0.084)
IPW+OLS	0.08***	0.003	(0.074, 0.086)
NNM	0.072***	0.003	(0.066, 0.078)
EBW	0.079***	0.003	(0.073, 0.085)
RDD	0.139***	0.035	(0.07, 0.208)
Antenatal care +3			
Independence	0.108***	0.004	(0.1, 0.116)
OLS	0.082***	0.005	(0.072, 0.092)
PSM	0.07***	0.006	(0.058, 0.082)
IPW+OLS	0.079***	0.005	(0.069, 0.089)
NNM	0.075***	0.006	(0.063, 0.087)
EBW	0.091***	0.004	(0.083, 0.099)
RDD	0.129***	0.050	(0.031, 0.227)
Postnatal care for mother			
Independence	0.086***	0.004	(0.078, 0.094)
OLS	0.105***	0.005	(0.095, 0.115)
PSM	0.094***	0.006	(0.082, 0.106)
IPW+OLS	0.106***	0.005	(0.096, 0.116)
NNM	0.102***	0.006	(0.09, 0.114)
EBW	0.106***	0.004	(0.098, 0.114)
RDD	0.071	0.051	(-0.029, 0.171)
Postnatal care for baby			
Independence	0.054***	0.004	(0.046, 0.062)
OLS	0.076***	0.004	(0.068, 0.084)
PSM	0.068***	0.005	(0.058, 0.078)
IPW+OLS	0.077***	0.004	(0.069, 0.085)
NNM	0.073***	0.004	(0.065, 0.081)
EBW	0.071***	0.004	(0.063, 0.079)
RDD	0.113***	0.044	(0.027, 0.199)
Iron folic acid supplement intakes			
Independence	0.14***	0.004	(0.132, 0.148)
OLS	0.099***	0.005	(0.089, 0.109)
PSM	0.091***	0.006	(0.079, 0.103)
IPW+OLS	0.096***	0.005	(0.086, 0.106)
NNM	0.085***	0.006	(0.073, 0.097)
EBW	0.105***	0.004	(0.097, 0.113)
RDD	0.145***	0.051	(0.045, 0.245)
Tetanus toxoid injections			
Independence	0.099***	0.003	(0.093, 0.105)
OLS	0.088***	0.003	(0.082, 0.094)
PSM	0.088***	0.004	(0.08, 0.096)
IPW+OLS	0.088***	0.003	(0.082, 0.094)
NNM	0.078***	0.004	(0.07, 0.086)
EBW	0.09***	0.003	(0.084, 0.096)
RDD	0.17***	0.039	(0.094, 0.246)

Source: DLHS-4. Note: Number of observations is 67,595. The choice of covariates follows Rahman and Pallikadavath (2018). 95% confidence intervals are shown. PSM=Propensity score matching; IPW+OLS=OLS adjusted with the propensity score inverse probability weighting; NNM=Nearest neighbour matching; and EBW=Entropy balance weighting; and RDD=Regression discontinuity design. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

MTR+MTS+MIV. Table 6 also reports the confidence intervals for the respective ATE bound estimates. We observe that the widths of the confidence intervals are just slightly wider than the widths of the estimated ATE bounds. Compared with the widths of the bound estimates, the additional widths due to the sampling variability are very small, which is thanks to the large sample sizes. This suggests that identification problem is the dominant concern in inference on the impacts and sampling variance poses a much less serious problem for inference (Manski and Nagin, 1998).

Institutional delivery

First, for institutional delivery, in the worst-case scenario where we do not impose any assumption, the lower bound is -0.650 and the upper bound is 0.350. In the worst-case, the bound width is always 1 and the bound always includes 0. This bound just ensures that the true ATE can never be outside of this bound as long as the ATE is estimated from this data. Although this bound has a wider negative range, it does not necessarily mean that the ATE is more likely to be negative. Also, the centre of the bound (-0.15 in this case) is not necessarily the most probable ATE.

When we impose the MTR assumption, it truncates the lower bound of ATE at 0 because the MTR assumes that the ATE is non-negative. Then we add the MTS assumption to the MTR assumption. As the ATE estimated by OLS is larger than the ATE estimated by mean comparisons, we assume the negative MTS assumption here. The negative selection bias assumes that individual observable and unobservable characteristics are correlated with the participation in the programme and the institutional delivery in the opposite directions. One possible case would be that richer mothers who seek high quality healthcare at private hospitals may have lower motivation to participate in the JSY. Imposing the MTS assumption, we find that the lower bound becomes even larger – from 0 to 0.126. Further, adding the MIV assumption makes the ATE bound even narrower. The lower bound has become 0.262 and the upper bound has become 0.331.

After adding the MIV assumption, we find that all point-identified ATE estimates are outside of the ATE bound. The ATEs on institutional delivery estimated under the CIA are smaller than the lower ATE bound estimated under the MTR, MTS and MIV assumptions, thereby suggesting that their maintained assumptions are not compatible with the MTR, MTS and MIV assumptions used in the partial

Table 6: ATE bound estimates

	Estimates		Confidence intervals	
	Lower bound	Upper bound	Lower bound	Upper bound
Institutional delivery				
No assumption	−0.652	0.348	−0.654	0.351
MTR	0.000	0.348	0.000	0.351
MTR+MTS	0.126	0.348	0.122	0.351
MTR+MTS+MIV	0.262	0.331	0.255	0.335
Skilled delivery attendance				
No assumption	−0.702	0.298	−0.705	0.301
MTR	0.000	0.298	0.000	0.301
MTR+MTS	0.081	0.298	0.078	0.301
MTR+MTS+MIV	0.176	0.294	0.170	0.297
Antenatal care +1				
No assumption	−0.685	0.315	−0.688	0.319
MTR	0.000	0.315	0.000	0.319
MTR+MTS	0.000	0.091	0.000	0.095
MTR+MTS+MIV	0.000	0.042	0.000	0.050
Antenatal care +3				
No assumption	−0.587	0.413	−0.590	0.417
MTR	0.000	0.413	0.000	0.417
MTR+MTS	0.000	0.108	0.000	0.115
MTR+MTS+MIV	0.000	0.014	0.000	0.029
Postnatal care for mother				
No assumption	−0.558	0.442	−0.561	0.445
MTR	0.000	0.442	0.000	0.445
MTR+MTS	0.086	0.442	0.078	0.445
MTR+MTS+MIV	0.196	0.430	0.187	0.434
Postnatal care for baby				
No assumption	−0.644	0.356	−0.647	0.358
MTR	0.000	0.356	0.000	0.358
MTR+MTS	0.054	0.356	0.048	0.358
MTR+MTS+MIV	0.158	0.349	0.150	0.352
Iron folic acid supplement intakes				
No assumption	−0.570	0.430	−0.573	0.433
MTR	0.000	0.430	0.000	0.433
MTR+MTS	0.000	0.140	0.000	0.146
MTR+MTS+MIV	0.000	0.065	0.000	0.081
Tetanus toxoid injections				
No assumption	−0.664	0.336	−0.668	0.339
MTR	0.000	0.336	0.000	0.339
MTR+MTS	0.000	0.099	0.000	0.104
MTR+MTS+MIV	0.000	0.046	0.000	0.056

Source: DLHS-4. Note: Number of observations is 67,595. Note: 95% confidence intervals are calculated following Imbens and Manski (2004) by bootstrap with 200 repetitions.

identification approach.[14] This incompatibility is because: (1) the CIA is not valid; (2) functional-form assumptions are not correct; and/or (3) the MTR, MTS, and MIV assumptions are not valid. The finding that the non-parametric EBW estimate is outside of the bound implies that regardless of functional-form assumptions, the point-identified ATE is outside of the ATE bound.

Moreover, as we discussed in the previous section, the identification assumptions used for the bound estimation are supported by empirical and observational evidence. Given that the CIA is hard to be justified for the case of the JSY, it would be reasonable to think that the incompatibility between assumptions for the point-identification and those for the partial-identification mainly comes from the possibly unacceptable CIA and/or functional form specification errors.[15] Also, we found that the ATE estimated by the fuzzy RDD is also smaller than the lower bound of ATE, which implies that the compliers are very much different from the other populations and/or the assumption of the discontinuity is not valid.

Skilled birth attendance

For skilled birth attendance, in the worst-case, the ATE bound is from -0.702 to 0.298. The MTR assumption truncates the lower bound at 0. As the ATE estimated by OLS is larger than the ATE estimated by mean comparisons, we assume the negative MTS assumption. We speculate that most of the rich mothers who can afford to give birth in the presence of skilled birth attendants may not have been eligible for the JYS. Adding the negative MTS assumption raises the lower bound up to 0.081. Further adding the MIV assumption makes the band narrower. Under the three assumptions, the lower and upper bounds have become 0.176 and 0.294 respectively (Figure 4).

All point-estimated ATEs are all smaller than the lower limit of this ATE bound, which suggests that they could be under-estimated under the maintained assumptions. The confidence interval of the ATE bound under the MTR, MTS and MIV assumptions is between 0.170 and 0.297, which does not include the point-identified ATEs.

[14] Strictly speaking, even without the MTR assumption, all the point-identified ATEs are found to be outside of the bound estimates (see Figure 4).

[15] We estimated the propensity score differently with the probit and linear probability model, but we did obtain similar results.

ANC use at least once

For ANC use at least once, in the worst-case, the ATE bound is from -0.685 to 0.315. The MTR assumption makes the bound non-negative. As the ATE estimated by OLS is smaller than the ATE estimated by mean comparisons, we add the positive MTS assumption to the MTR. The positive MTS assumption implies that mothers with higher awareness of the importance of maternal healthcare may be more likely to participate in the JSY and receive ANC. Adding the MTS assumption substantially narrows the upper bound of ATE to 0.091, implying that the assumption regarding the direction of selection bias has large identifying power. The MIV assumption further narrows the upper bound to 0.042.

We observe that the ATEs under the independence assumption and the CIA lie beyond the upper limit of this ATE bound (Figure 4). We also observe that the ATE estimated by the RDD is outside of this bound, thereby implying that the compliers are very much different from the entire populations and/or the assumption of the discontinuity is not valid. We do not observe the overlapping of confidence intervals between point-estimated ATEs and bound-estimated ATE.

ANC use at least three times

We also observe similar shrinks in the ATE bound for ANC use three times and more. Its worst-case bound ranges from -0.587 to 0.413. The MTR assumption truncates the lower bound at 0. As the ATE estimated by OLS is smaller than the ATE estimated by mean comparisons, we assume the positive MTS assumption. Adding the positive MTS assumption narrows the upper bound from 0.413 to 0.108. The MIV assumption further narrows it up to 0.014.

After adding the MIV assumption, the point-identified ATEs are again beyond this upper bound (Figure 4). We do not observe the overlapping of confidence intervals between point-estimated ATEs and bound-estimated ATE.

PNC for mothers

For PNC for mothers, the worst-case bound is -0.558 to 0.442 and the MTR assumption makes it non-negative. As the ATE estimated by OLS is larger than the ATE estimated by mean comparisons, we assume the negative MTS assumption. We speculate that the observed negative selection bias suggests that those mothers who have greater difficulty in having healthcare may have been more strongly en-

couraged by the accredited social health activists to join the JSY programme or/and that mothers who can afford to receive PNC may not have been eligible for the JSY. We find that the negative MTS has large identifying power and succeeds in increasing the lower bound up to 0.086. The MIV assumption contributes to further narrowing the bounds of the ATE.

Under the three assumptions, we observe that the ATE is no less than 0.193 and no larger than 0.430. Under the three assumptions, we observe that the point-identified ATEs are all smaller than the lower limit of ATE bound, suggesting the possibility of underestimation of ATEs under the CIA (Figure 5). The ATE under the RDD assumption lies also outside the bound. We do not observe the overlapping of confidence intervals between point-estimated ATEs and bound-estimated ATE.

PNC for babies

For PNC for babies, the worst-case bound is from -0.644 to 0.356. Along with PNC for mothers, as the ATE estimated by OLS is larger than the ATE estimated by mean comparisons, we assume the negative MTS assumption. The MTR and negative MTS assumptions raise the lower limit of the ATE bound up to 0.054. The MIV assumption further narrows the lower bound to 0.158 and the upper bound to 0.349. We find that the ATEs under the CIA and the RDD assumption lie below this lower bound, suggesting that they could be under-estimated (Figure 5). Even when considering the sampling variability, the confidence interval of the ATE bound does not include the point-estimated ATEs.

Iron and folic acid supplement intakes

For iron and folic acid supplement intakes, we find that in the worst-case, the ATE ranges from -0.570 to 0.430. As the ATE estimated by OLS is smaller than the ATE estimated by mean comparisons, we impose the positive MTS assumption, implying that mothers who understand the importance of maternal healthcare more are more likely to participate in the programme and receive the supplement. The MTR assumption makes the bound non-negative and the positive MTS assumption substantially narrows the upper bound to 0.140. Adding the MIV assumption further narrows the upper bound up to 0.065. We find that all point-estimated ATEs are greater than the upper bound, thereby suggesting their over-estimation (Figure 5). Although the confidence interval of the ATE bound does not include the point-identified ATEs, but we observe the overlapped confidence intervals of the ATEs point-estimated by PSM, NNM and RDD with that of the bound-estimated ATE.

Tetanus toxoid injections

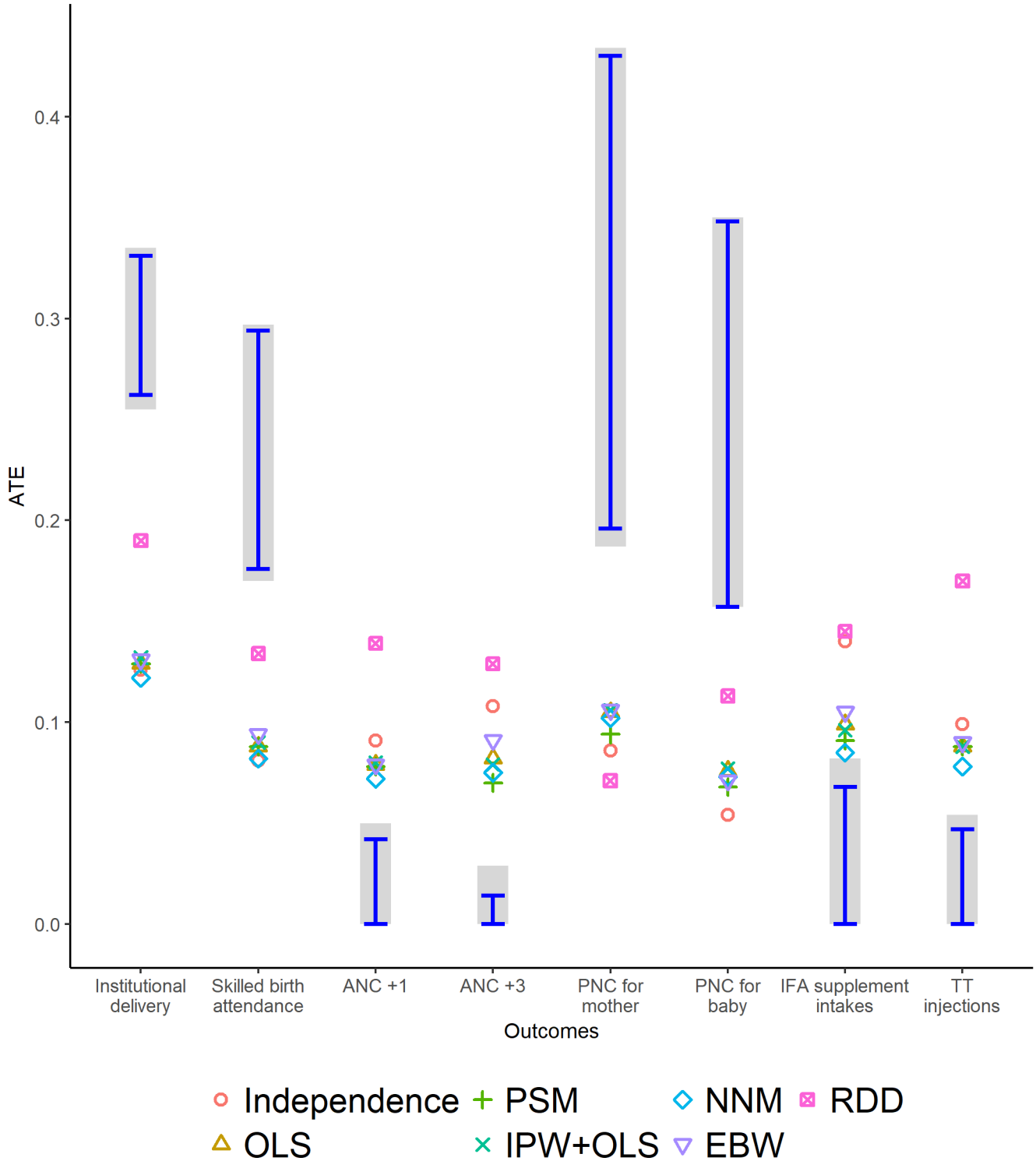
Finally, for the tetanus toxoid injections, the worst-case bound is from -0.664 to 0.336. The MTR assumption truncates the lower bound at 0. As with the case with iron and folic acid supplement, as the ATE estimated by OLS is smaller than the ATE estimated by mean comparisons, we assume the positive MTS assumption. We conjecture that more risk averse mothers are more likely to join the JSY and get the treatment. We observe that the positive MTS assumption substantially contributes to narrowing the upper bound up to 0.099, and adding the MIV assumption contributes to further narrowing it up to 0.046. We find that the point-estimated ATEs are all outside of the bound again, indicating the evidence of over-estimation (Figure 5). We do not observe the overlapping of confidence intervals between point-estimated ATEs and bound-estimated ATE.

5.3 Result summary

Figure 3 visually summarises the results presented above. The bounds shown in Figure 3 are the sharp bounds under the MTR, MTS and MIV assumptions. We find that the point-identified ATEs are below the lower ATE bound of institutional delivery, skilled birth attendance, PNC for mothers and children. On the other hand, for ANC at least once and ANC at least three times, the point-identified ATEs are over the upper limit of the ATE bounds. In particular, a larger deviation from the lower bound is observed for institutional delivery. Even considering the uncertainty of sampling variability, we do not observe the overlapping of the confidence intervals between point-identified and partially-identified ATEs for institutional delivery, ANC at least once, PNC for mothers and tetanus toxoid injections.

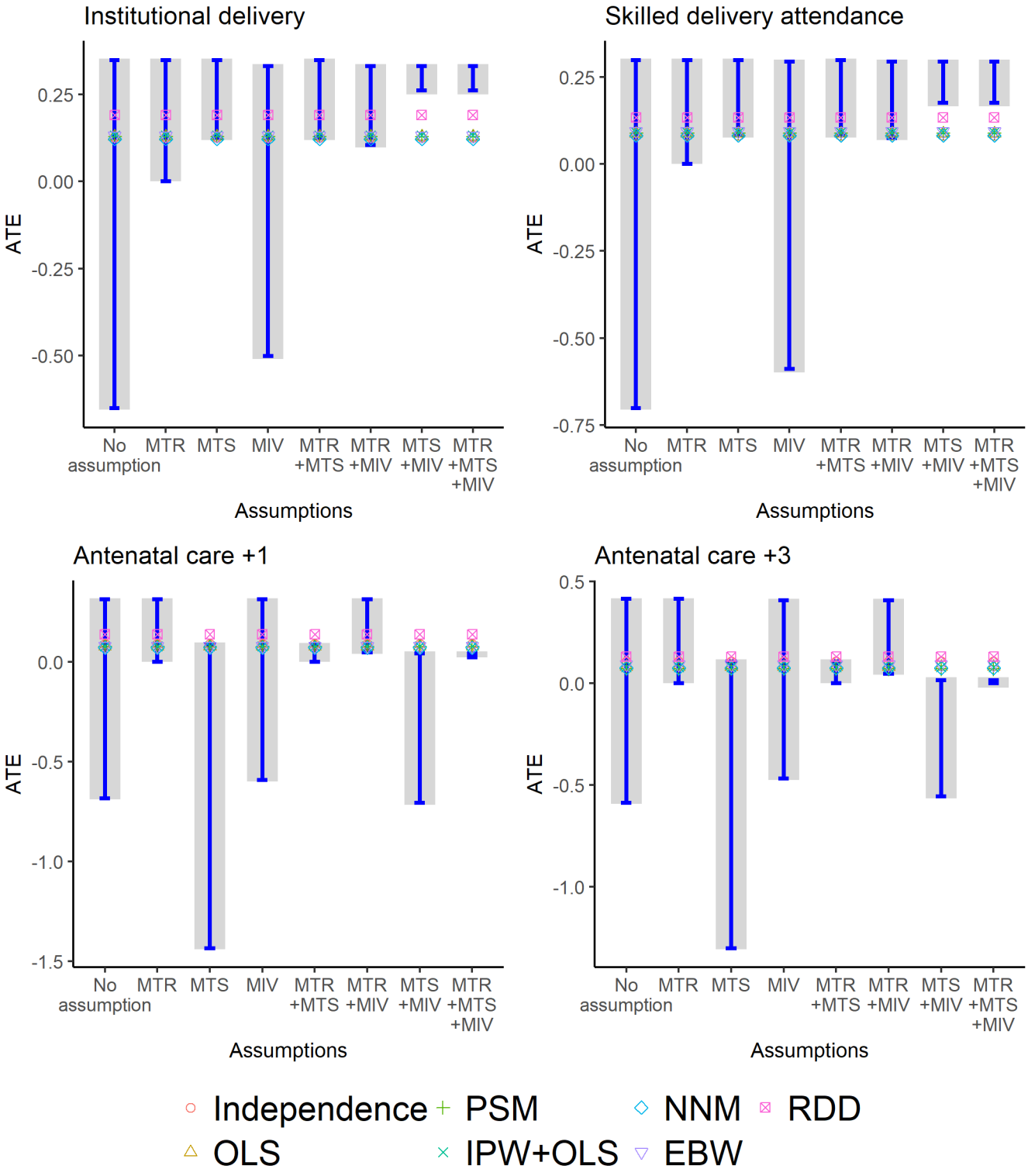
Next we explore how each of the three assumptions contributes to estimating the ATE bounds for each outcome. Figures 4 and 5 illustrate the sharp ATE bounds for each outcome estimated under the eight different combinations of assumptions. By comparing the bound widths, we can infer the identification power of each assumption. The joint imposition of the MTS and MIV assumptions leads to our main finding that point-estimated ATEs are outside the bound-estimated ATEs. Interestingly, even without the MTR assumption, we could have reached the same finding although the estimated ATE bound can have a larger width. Another interesting finding is that the MTR assumption is not binding in the presence of the negative MTS assumption. This is because the negative MTS assumption itself made the lower limits of the ATE bounds positive. Nevertheless, the MTR assumption significantly contributes to narrowing the ATE bounds unless the negative MTS assumption is imposed.

Figure 3: Point-estimated ATE and bound-estimated ATE under the MTR, MTS and MIV assumptions



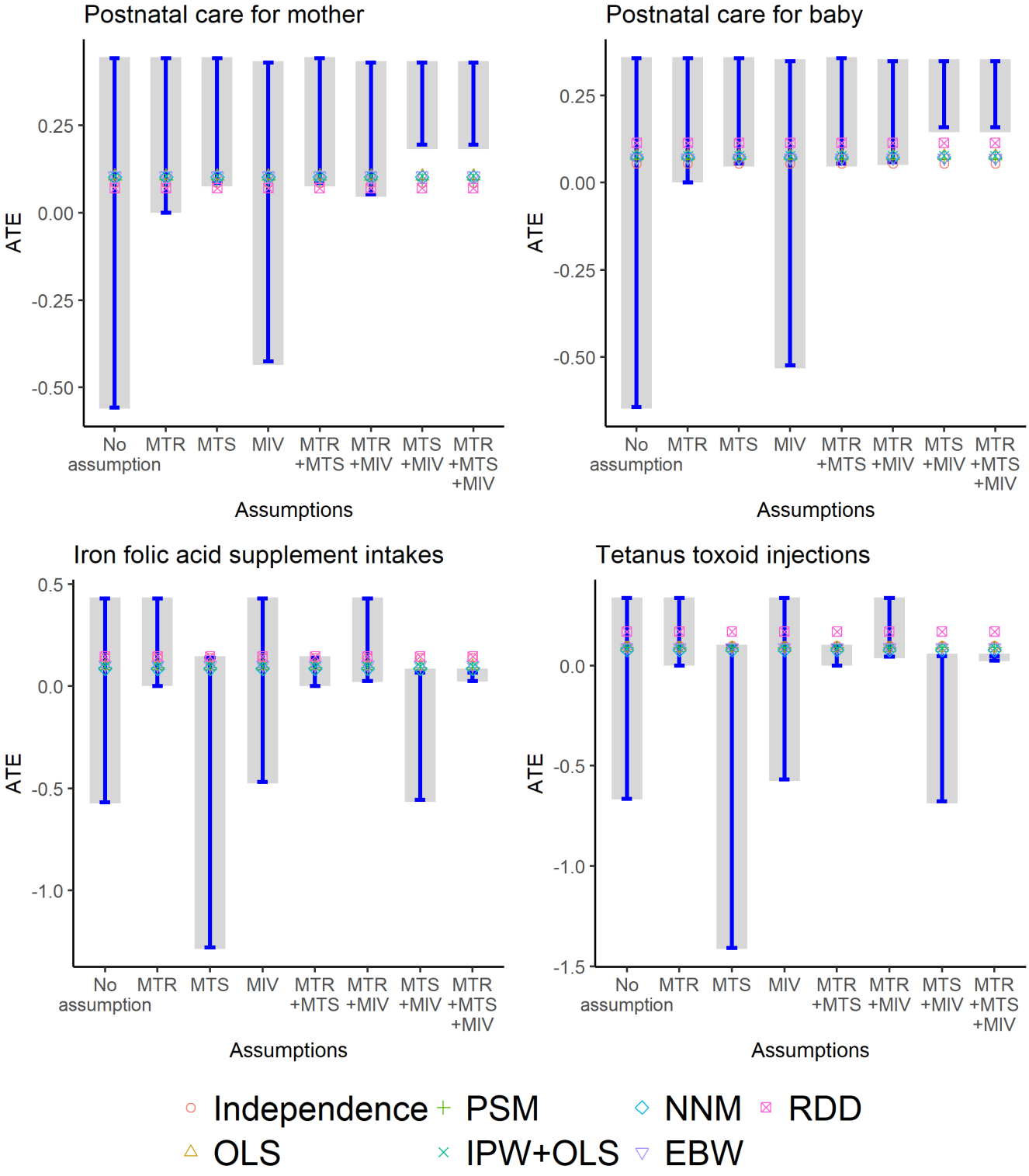
Source: DLHS-4. Note: The choice of covariates follows Rahman and Pallikadavath (2018). PSM=Propensity score matching; IPW+OLS=OLS adjusted with the propensity score inverse probability weighting; NNM=Nearest neighbour matching; EBW=Entropy balance weighting; RDD=Regression discontinuity design; ANC=Antenatal care; PNC=Postnatal care; IFA=Iron and folic acid; and TT=Tetanus toxoid. The bounds shown are sharp bounds estimated under the MTR, MTS and MIV assumptions. The shaded areas show the 95% level confidence interval of the ATE bound under the MTR, MTS and MIV assumptions.

Figure 4: ATE bounds for institutional delivery, Skilled birth attendance and ANC use



Source: DLHS-4. Note: The choice of covariates follows Rahman and Pallikadavath (2018). PSM=Propensity score matching; IPW+OLS=OLS adjusted with the propensity score inverse probability weighting; NNM=Nearest neighbour matching; EBW=Entropy balance weighting; and RDD=Regression discontinuity design. The bounds shown are sharp bounds under the MTR, MTS and MIV assumptions. The shaded areas show the 95% level confidence interval of the ATE bound under the MTR, MTS and MIV assumptions.

Figure 5: ATE bounds for PNC, iron and folic acid supplement intakes and tetanus toxoid injections



Source: DLHS-4. Note: The choice of covariates follows Rahman and Pallikadavath (2018). PSM=Propensity score matching; IPW+OLS=OLS adjusted with the propensity score inverse probability weighting; NNM=Nearest neighbour matching; EBW=Entropy balance weighting; and RDD=Regression discontinuity design. The bounds shown are sharp bounds under the MTR, MTS and MIV assumptions. The shaded areas show the 95% level confidence interval of the ATE bound under the MTR, MTS and MIV assumptions.

6 Robustness checks and further analysis

6.1 Robustness check: choice of covariates

As robustness checks with respect to the choice of covariates, we point-identify ATEs again with different covariate sets used in the previous studies and the full covariate sets in order to see whether the choice of covariates substantially affects our findings. Note that as an ATE bound is unconditional on covariates, it is invariant across the covariate sets. First, we closely follow Lim et al. (2010) and use the following covariate sets: state of residence, urban residence, below-the-poverty-line card ownership, wealth quintile, scheduled caste, education, parity, and maternal age.

Next, we closely follow Sengupta and Sinha (2018), in which the authors used the following variables as determinants of the JSY participation and healthcare use: residence in North-Western states, urban residence, below-the-poverty-line card ownership, number of live births, scheduled caste, scheduled tribe, whether pregnancy was known within three months, any previous stillbirths, any previous miscarriages, any children who died, age at which living with husband, age at time of birth, ratio of male to female members in household, being Hindu, wealth quintiles, living in a *kachcha* (made of natural materials) house, maternal and paternal education dummies.[16] As well as them, following Sengupta and Sinha (2018), we use the following variables that affect the healthcare use only: sex of baby, number of brothers and sisters, whether any household member is covered by health insurance, having no toilet facility, and any water treatment for drinking.

Third, we control for more variables to fully take into account of the selection bias as much as we could. We control for the following covariates: birth-year fixed effects, birth-month fixed effects, state-fixed effects, urban residence, below-the-poverty-line card ownership, number of live births, scheduled caste, scheduled tribe, whether pregnancy was known within three months, any previous stillbirths, any previous miscarriages, any children who died, age at which living with husband, age at time of birth, ratio of male to female members in household, wealth quintiles, maternal and paternal education levels and being Hindu. This complete set of covariates is expected to more effectively remove the selection bias attributable to measurable characteristics. If the point-estimated ATEs estimated with this full covariate set are outside of the ATE bounds, then it would support our finding that the CIA would be implausible

[16] We defined the following three educational levels for mothers and fathers respectively: 1-6 years of education, 7-12 years of education, and more than 12 years of education.

and the estimated ATEs could be biased.

The results estimated with these three sets of covariates are shown in Table 7 and summarised visually in Figure 6. We find that all point-identified ATEs lie outside of ATE bound in three sets of covariates. We reconfirm that the point-identified ATEs are likely to be over-estimated for ANC, iron and folic acid supplement intakes and tetanus toxoid injections and to be under-estimated for institutional delivery, skilled birth attendance and PNC. Overall, these findings are consistent with our main results in Figure 3, implying that regardless of the choice of covariate sets, the conditional independence and functional form assumptions are likely to be invalid.

6.2 Robustness check: validity of the MIV assumption

Manski (2003) argues that the credibility of estimates decreases with the strength of assumptions maintained (*The Law of Decreasing Credibility*). We observe that, in both HPSs and LPSs, the MIV assumption made a key role in our finding that point-identified ATEs lie outside the bound-estimated ATEs. Hence it is important to examine the credibility of the MIV assumption. Although a large volume of literature reports that socio-economic status is negatively associated with maternal and child healthcare use in developing countries (Pathak et al., 2010; Pathak and Mohanty, 2010; Kesterton et al., 2010; Balarajan et al., 2011), we cannot directly test the validity of MIV assumption formalised in equations (27) and (28) in principle, because these are assumptions on latent probabilities. Nevertheless, we can still explore the validity of equation (27), using the National Family Health Survey data, which was surveyed in 2005-2006. By focusing only on women who delivered before 12 April 2005, we can observe both $P(Y_0 = 1|v = 0)$ and $P(Y_0 = 1|v = 1)$ among women who gave birth before the introduction of the JSY. Figure 7 shows that $P(Y_0 = 1|v = 0)$ is larger than $P(Y_0 = 1|v = 1)$ for all outcomes ($p < 0.01$), supporting the validity of equation (27) and providing confidence in our results.

7 Conclusion

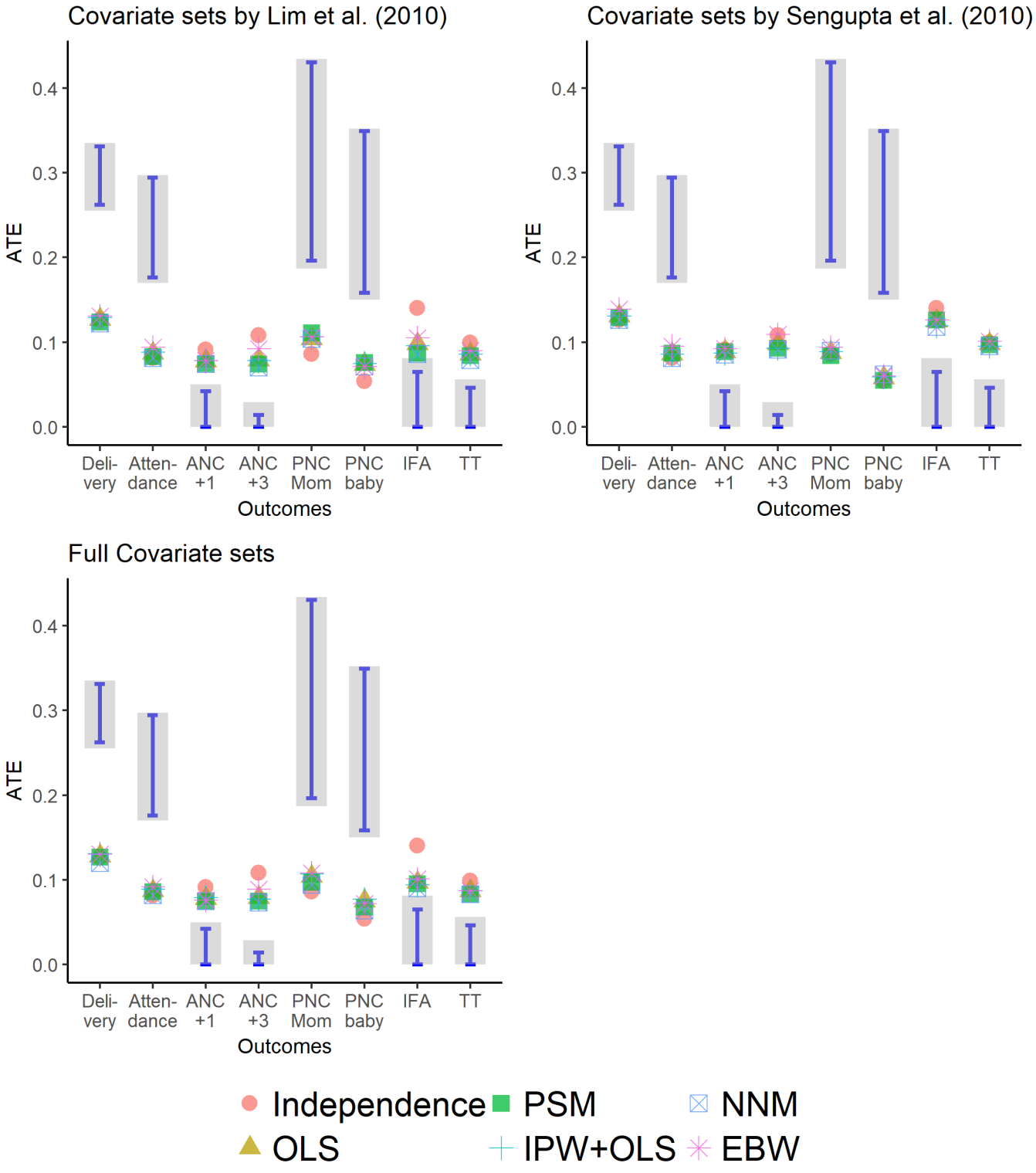
In India, the CCT programme, the JSY, was introduced in 2005, which promotes the use of maternal and child healthcare with cash incentives in order to reduce infant and neonatal mortalities. In contrast to other countries with CCT programmes, a rigorous RCT was not conducted in India for the JSY (Joshi and Sivaram, 2014), and hence a valid instrumental variable has not been available for researchers. Lim et al.

Table 7: Point-estimated ATE with different covariate sets

	Lim et al. (2010)		Sengupta et al. (2018)		Full sets	
	Estimates	SEs	Estimates	SEs	Estimates	SEs
Institutional delivery						
Independence	0.432***	0.005	0.432***	0.005	0.432***	0.005
OLS	0.127***	0.003	0.131***	0.002	0.128***	0.003
PSM	0.124***	0.003	0.129***	0.003	0.127***	0.003
IPW+OLS	0.129***	0.003	0.131***	0.003	0.131***	0.003
NNM	0.122***	0.003	0.126***	0.003	0.12***	0.003
EBW	0.131***	0.003	0.139***	0.003	0.13***	0.003
Skilled delivery attendance						
Independence	0.349***	0.005	0.349***	0.005	0.349***	0.005
OLS	0.087***	0.002	0.086***	0.002	0.087***	0.002
PSM	0.083***	0.002	0.087***	0.002	0.086***	0.002
IPW+OLS	0.088***	0.002	0.086***	0.002	0.089***	0.002
NNM	0.081***	0.002	0.081***	0.002	0.082***	0.002
EBW	0.094***	0.002	0.095***	0.002	0.092***	0.002
Antenatal care +1						
Independence	0.105***	0.005	0.105***	0.005	0.105***	0.005
OLS	0.078***	0.003	0.089***	0.002	0.078***	0.003
PSM	0.075***	0.004	0.089***	0.003	0.074***	0.004
IPW+OLS	0.078***	0.003	0.087***	0.003	0.079***	0.003
NNM	0.074***	0.004	0.085***	0.003	0.075***	0.004
EBW	0.078***	0.003	0.092***	0.002	0.076***	0.003
Antenatal care +3						
Independence	0.101***	0.007	0.101***	0.007	0.101***	0.007
OLS	0.079***	0.005	0.098***	0.004	0.079***	0.005
PSM	0.074***	0.006	0.093***	0.006	0.075***	0.006
IPW+OLS	0.078***	0.005	0.093***	0.004	0.077***	0.005
NNM	0.07***	0.006	0.091***	0.005	0.073***	0.005
EBW	0.092***	0.004	0.109***	0.004	0.089***	0.004
Postnatal care for mother						
Independence	0.178***	0.007	0.178***	0.007	0.178***	0.007
OLS	0.104***	0.005	0.088***	0.005	0.104***	0.005
PSM	0.111***	0.006	0.084***	0.006	0.097***	0.006
IPW+OLS	0.106***	0.005	0.089***	0.005	0.107***	0.005
NNM	0.104***	0.006	0.09***	0.006	0.094***	0.006
EBW	0.106***	0.004	0.094***	0.004	0.108***	0.004
Postnatal care for baby						
Independence	0.065***	0.007	0.065***	0.007	0.065***	0.007
OLS	0.074***	0.004	0.058***	0.004	0.075***	0.004
PSM	0.076***	0.005	0.055***	0.005	0.068***	0.005
IPW+OLS	0.075***	0.004	0.059***	0.004	0.077***	0.004
NNM	0.071***	0.004	0.062***	0.004	0.064***	0.004
EBW	0.071***	0.004	0.06***	0.004	0.072***	0.004
Iron folic acid supplement intakes						
Independence	0.104***	0.006	0.104***	0.006	0.104***	0.006
OLS	0.098***	0.005	0.125***	0.004	0.097***	0.005
PSM	0.087***	0.006	0.126***	0.005	0.095***	0.006
IPW+OLS	0.096***	0.005	0.119***	0.004	0.094***	0.005
NNM	0.086***	0.006	0.118***	0.005	0.09***	0.006
EBW	0.105***	0.004	0.126***	0.004	0.101***	0.004
Tetanus toxoid injections						
Independence	0.056***	0.004	0.056***	0.004	0.056***	0.004
OLS	0.086***	0.003	0.098***	0.003	0.087***	0.003
PSM	0.084***	0.004	0.097***	0.004	0.083***	0.004
IPW+OLS	0.086***	0.003	0.095***	0.003	0.087***	0.004
NNM	0.079***	0.004	0.095***	0.004	0.083***	0.004
EBW	0.09***	0.003	0.101***	0.003	0.087***	0.003

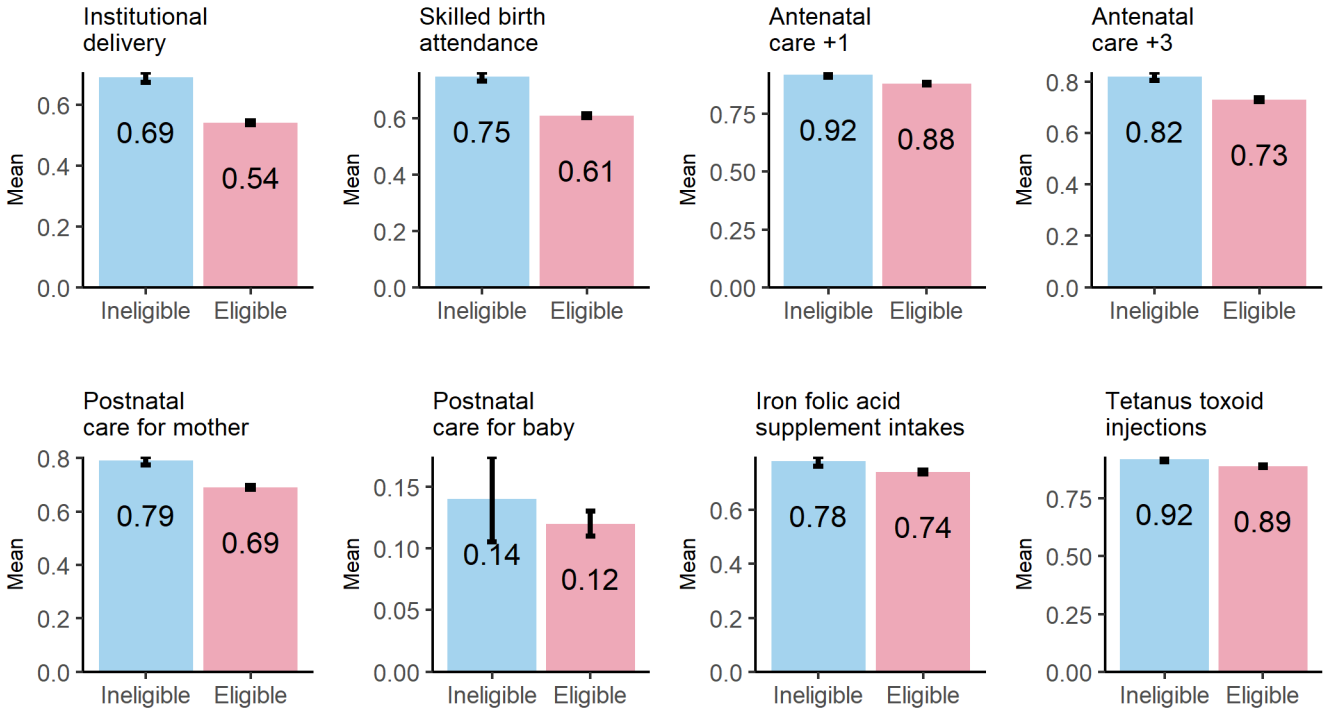
Source: DLHS-4. Note: 95% confidence intervals are shown. PSM=Propensity score matching; IPW+OLS=OLS adjusted with the propensity score inverse probability weighting; NNM=Nearest neighbour matching; and EBW=Entropy balance weighting. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 6: ATE bound and point-estimated ATE with different covariate sets



Source: DLHS-4. Note: PSM=Propensity score matching; IPW+OLS=OLS adjusted with the propensity score inverse probability weighting; NNM=Nearest neighbour matching; EBW=Entropy balance weighting; ANC=Antenatal care; PNC=Postnatal care; IFA=Iron and folic acid; and TT=tetanus toxoid. The bounds shown are sharp bounds estimated under the MTR, MTS and MIV assumptions. The shaded areas show the 95% level confidence interval of the ATE bound under the MTR, MTS and MIV assumptions.

Figure 7: Comparison in the mean outcomes between eligible and non-eligible mothers without the treatment



Source: National Family Health Survey. Note: 95% confidence intervals are shown.

(2010) conducted a first estimate of the causal impacts of the JSY, which has attracted much attention for years. Up to now, several studies have attempted to estimate its causal impacts on various healthcare services and health outcomes. However, according to Das et al. (2011), the JSY was immature until 2007 and the data set analysed by Lim et al. (2010) and others (i.e., DLHS-3) was not reliable enough to estimate the causal impact. The latest wave of the DLHS (DLHS-4) has become available only recently, and Rahman and Pallikadavath (2017) and Rahman and Pallikadavath (2018) re-estimated the causal impacts, acknowledging the potential serious errors that had occurred in DLHS-3.

Up until now, however, no study has ever tried to assess the validity of the identification assumptions employed in previous studies. For example, Lim et al. (2010), Rahman and Pallikadavath (2018) and Sengupta and Sinha (2018) all relied on the CIA that participation in the JSY programme can be regarded as random if we control for observable household and individual characteristics. However, the CIA is extremely hard to justify for the case of JSY. The utilisation of maternal and child healthcare use and the participation in the JSY are likely to be dependent on individual unobservable characteristics such as the degree of being risk-averse and the awareness of the importance of maternal and child healthcare. If

there exist some selection mechanisms that cannot be controlled for by the observable characteristics, the estimated ATE under the CIA suffers selection bias, potentially resulting in flawed and conflicting conclusions (Manski, 2013). This study assessed the validity of the identification assumptions used in previous studies and provided new evidence through a partial identification approach. The partial identification taken herein yielded an honest and credible bound of ATE, which is valuable when we are not confident of the conventionally imposed identification assumptions.

If the imposed assumptions are valid, any point-estimate of ATE should lie within the ATE bound. However, we find that the ATEs under the CIA lie beyond the ATE bounds in HPSs, suggesting the invalidity of identification assumptions and the possibility of over- or under-estimation. Especially, we find that the point-identified ATEs are below the lower ATE bounds of institutional delivery, skilled birth attendance, PNC for mothers and children. On the other hand, for ANC at least once and ANC at least three times, the point-identified ATEs are over the upper limits of the ATE bounds. For institutional delivery, the largest deviations of the point-estimated ATEs from the lower limit of its ATE bound is observed.[17] We find consistent results even when we used different covariate sets that have been used in two other previous studies.

Overall, this study provided sufficiently strong evidence that the point-estimated ATEs could have been biased in previous studies. This study re-estimated the causal impacts through a partial identification approach and shows the conservative bounds of ATEs. Also, this study quantified how much at least the point-identified ATEs under questionable identification assumptions could be far from the true ATE. Certainly, the ATE bounds themselves do not give definite values of the causal impacts. We believe, however, that the honest estimation of the bounds based on the credible assumptions could be more useful and valuable in policy making than definitive ATE estimates that heavily rely on unpalatable or untenable assumptions. We hope the results of this study contribute to evaluating the JSY thoroughly.

[17] However, it does not necessarily mean that the actual size of under-estimation for the institutional delivery is the largest. It just means that its lower limit of potential underestimation is the largest.

References

Abadie, A. and M. D. Cattaneo (2018). Econometric Methods for Program Evaluation. *Annual Review of Economics 10*(1), 465–503.

Balarajan, Y., S. Selvaraj, and S. V. Subramanian (2011). Health Care and Equity in India. *The Lancet 377*(9764), 505 – 515.

Barber, S. L. and P. J. Gertler (2010). Empowering Women: How Mexico's Conditional Cash Transfer Programme Raised Prenatal Care Quality and Birth Weight. *Journal of Development Effectiveness 2*(1), 51–73.

Bassett, L. (2008). Can Conditional Cash Transfer Programs Play a Greater Role in Reducing Child Undernutrition. Social Protection & Labor Discussion Paper No. 0835.

Bhutta, Z. A., M. Chopra, H. Axelson, P. Berman, T. Boerma, J. Bryce, F. Bustreo, E. Cavagnero, G. Cometto, B. Daelmans, A. de Francisco, H. Fogstad, N. Gupta, L. Laski, J. Lawn, B. Maliqi, E. Mason, C. Pitt, J. Requejo, A. Starrs, C. G. Victora, and T. Wardlaw (2010). Countdown to 2015 Decade Report (2000–10): Taking Stock of Maternal, Newborn, and Child Survival. *The Lancet 375*(9730), 2032 – 2044.

Blundell, R., A. Gosling, H. Ichimura, and C. Meghir (2007). Changes in the Distribution of Male and Female Wages Accounting for Employment Composition Using Bounds. *Econometrica 75*(2), 323–363.

Bredenkamp, C. (2009). India - The Janani Suraksha Yojana (JSY) Program. Results-based Financing (RBF) for Health: Country Snapshot. Technical report, World Bank, Washington DC.

Campbell, O. M. and W. J. Graham (2006). Strategies for Reducing Maternal Mortality: Getting on with What Works. *The Lancet 368*(9543), 1284 – 1299.

Carvalho, N., N. Thacker, S. S. Gupta, and J. A. Salomon (2014). More Evidence on the Impact of India's Conditional Cash Transfer Program, Janani Suraksha Yojana: Quasi-Experimental Evaluation of the Effects on Childhood Immunization and Other Reproductive and Child Health Outcomes. *PLOS ONE 9*(10), 1–13.

Chernozhukov, V., H. Hong, and E. Tamer (2007). Estimation and Confidence Regions for Parameter Sets in Econometric Models. *Econometrica 75*(5), 1243–1284.

Chitnis, S. (1997). Definition of the Terms Scheduled Castes and Scheduled Tribes: A Crisis of Ambivalence. In V. A. Pai-Panandiker (Ed.), *The Politics of Backwardness: Reservation Policy in India*, pp. 88–107. Delhi, India: Kornark Publishers.

Chou, D., B. Daelmans, R. R. Jolivet, M. Kinney, and L. Say (2015). Ending Preventable Maternal and Newborn Mortality and Stillbirths. *BMJ 351*, 19–22.

Corsi, D. J., M. Neuman, J. E. Finlay, and S. Subramanian (2012). Demographic and Health Surveys: A Profile. *International Journal of Epidemiology 41*(6), 1602–1613.

Darmstadt, G. L., A. C. Lee, S. Cousens, L. Sibley, Z. A. Bhutta, F. Donnay, D. Osrin, A. Bang, V. Kumar, S. N. Wall, A. Baqui, and J. E. Lawn (2009). 60 Million Non-facility Births: Who Can Deliver in Community Settings to Reduce Intrapartum-related Deaths? *International Journal of Gynecology & Obstetrics 107*(Supplement), S89–S112.

Das, A., D. Rao, and A. Hagopian (2011). India's Janani Suraksha Yojana: Further Review Needed. *The Lancet 377*(9762), 295 – 296.

de Brauw, A. and A. Peterman (2011). Can Conditional Cash Transfers Improve Maternal Health and Birth Outcomes? Evidence from El Salvador's Comunidades Solidarias Rurales. IFPRI discussion papers 1080, International Food Policy Research Institute (IFPRI).

Devadasan, N., M. A. Elias, D. John, S. Grahacharya, and L. Ralte (2008). A Conditional Cash Sssistance Programme for Promoting Institutional Deliveries among the Poor in India: Process Evaluation Results. *Studies in Health Service Organ Policy 24*, 257–273.

Dongre, A. (2012). Can Conditional Cash Transfers Impact Institutional Deliveries? Evidence from Janani Suraksha Yojana in India. *ZBW Working Paper, No.707*.

Freedman, L. P., W. J. Graham, E. Brazier, J. M. Smith, T. Ensor, V. Fauveau, E. Themmen, S. Currie, and K. Agarwal (2007). Practical Lessons from Global Safe Motherhood Initiatives: Time for a New Focus on Implementation. *The Lancet 370*(9595), 1383 – 1391.

Gaarder, M. M., A. Glassman, and J. E. Todd (2010). Conditional Cash Transfers and Health: Unpacking the Causal Chain. *Journal of Development Effectiveness 2*(1), 6–50.

Gerfin, M. and M. Schellhorn (2006). Nonparametric Bounds on the Effect of Deductibles in Health Care Insurance on Doctor Visits – Swiss Evidence. *Health Economics 15*(9), 1011–1020.

Ginther, D. K. (2000). Alternative Estimates of the Effect of Schooling on Earnings. *The Review of Economics and Statistics 82*(1), 103–116.

Glassman, A., D. Duran, L. Fleisher, and et al. (2013). Impact of Conditional Cash Transfers on Maternal and Newborn Health. *Journal of Health Population 4-2*(31), S48–S66.

Gonzalez, L. (2005). Nonparametric Bounds on the Returns to Language Skills. *Journal of Applied Econometrics 20*(6), 771–795.

Gopalan, S. S. and D. Varatharajan (2012). Addressing Maternal Healthcare through Demand Side Financial Incentives: Experience of Janani Suraksha Yojana Program in India. *BMC Health Services Research 12*(319), 1–10.

Gopichandran, V. and S. K. Chetlapalli (2012). Conditional Cash Transfer to Promote Institutional Deliveries in India: Toward a Sustainable Ethical Model to Achieve MDG 5A. *Public Health Ethics 5*(2), 173–180.

Gundersen, C., B. Kreider, and J. Pepper (2012). The Impact of the National School Lunch Program on Child Health: A Nonparametric Bounds Analysis. *Journal of Econometrics 166*(1), 79 – 91.

Gundersen, C., B. Kreider, and J. V. Pepper (2017). Partial Identification Methods for Evaluating Food Assistance Programs: A Case Study of the Causal Impact of SNAP on Food Insecurity. *American Journal of Agricultural Economics 99*(4), 875–893.

Gupta, S. K., D. K. Pal, R. Tiwari, R. Garg, R. Sarawagi, A. K. Shrivastava, and P. Gupta (2011). Assessment of Janani Suraksha Yojana (JSY) in Jabalpur, Madhya Pradesh: Knowledge, Attitude and Utilization Pattern of Beneficiaries: A Descriptive Study. *International Journal of Current Biological and Medical Science 1*(2), 6–11.

Gupta, S. K., D. K. Pal, R. Tiwari, R. Garg, A. K. Shrivastava, R. Sarawagi, R. Patil, L. Agarwal, P. Gupta, and C. Lahariya (2012). Impact of Janani Suraksha Yojana on Institutional Delivery Rate and Maternal Morbidity and Mortality: An Observational Study in India. *Journal of Health, Population and Nutrition 30*(4), 464–471.

Hainmueller, J. (2012). Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. *Political Analysis 20*(1), 25–46.

Handa, S. and B. Davis (2006). The Experience of Conditional Cash Transfers in Latin America and the Caribbean. *Development Policy Review 24*(5), 513–536.

Horton, R. (2010). The Continuing Invisibility of Women. *The Lancet 375*(9730), 1941–1943.

Huber, M., L. Laffers, and G. Mellace (2017). Sharp IV Bounds on Average Treatment Effects on the Treated and Other Populations Under Endogeneity and Noncompliance. *Journal of Applied Econometrics 32*(1), 56–79.

Iacus, S. M., G. King, and G. Porro (2012). Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis 20*(1), 1–24.

Imbens, G. W. and J. D. Angrist (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica 62*(2), 467–475.

Imbens, G. W. and C. F. Manski (2004). Confidence Intervals for Partially Identified Parameters. *Econometrica 72*(6), 1845–1857.

Imbens, G. W. and J. M. Wooldridge (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature 47*(1), 5–86.

Jehan, K., K. Sidney, H. Smith, and A. de Costa (2012). Improving Access to Maternity Services: An Overview of Cash Transfer and Voucher Schemes in South Asia. *Reproductive Health Matters 20*(39), 142–154.

Joshi, S. and A. Sivaram (2014). Does it Pay to Deliver? An Evaluation of India's Safe Motherhood Program. *World Development 64*, 434 – 447.

Kesterton, A. J., J. Cleland, A. Sloggett, and C. Ronsmans (2010). Institutional Delivery in Rural India: The Relative Importance of Accessibility and Economic Status. *BMC Pregnancy and Childbirth 10*(30), 1–9.

Kreider, B. and J. V. Pepper (2007). Disability and Employment. *Journal of the American Statistical Association 102*(478), 432–441.

Kreider, B., J. V. Pepper, C. Gundersen, and D. Jolliffe (2012). Identifying the Effects of SNAP (Food Stamps) on Child Health Outcomes When Participation Is Endogenous and Misreported. *Journal of the American Statistical Association 107*(499), 958–975.

Lagarde, M., A. Haines, and N. Palmer (2007). Conditional Cash Transfers for Improving Uptake of Health Interventions in Low- and Middle-income Countries: A Systematic Review. *The Journal of the American Medical Association 298*(16), 1900–1910.

Lahariya, C. (2009). Cash Incentives for Institutional Delivery: Linking with Antenatal and Post Natal Care May Ensure 'Continuum of Care' in India. *Indian Journal of Community Medicine 34*(1), 15–18.

Lee, D. S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *The Review of Economic Studies 76*(3), 1071–1102.

Lim, S. S., L. Dandona, J. A. Hoisington, S. L. James, M. C. Hogan, and E. Gakidou (2010). India's Janani Suraksha Yojana, A Conditional Cash Transfer Programme to Increase Births in Health Facilities: An Impact Evaluation. *The Lancet 375*(9730), 2009 – 2023.

Manski, C. F. (1989). Anatomy of the Selection Problem. *The Journal of Human Resources 24*(3), 343–360.

Manski, C. F. (1990). Nonparametric Bounds on Treatment Effects. *The American Economic Review 80*(2), 319–323.

Manski, C. F. (1997). Monotone Treatment Response. *Econometrica 65*(6), 1311–1334.

Manski, C. F. (2003). *Partial Identification of Probability Distributions*. Springer series in statistics. London, New York: Springer.

Manski, C. F. (2007). *Identification for Prediction and Decision*. MA: Harvard University Press.

Manski, C. F. (2013). *Public Policy in an Uncertain World: Analysis and Decisions*. Cambridge, MA: Harvard University Press.

Manski, C. F. and D. S. Nagin (1998). Bounding Disagreements About Treatment Effects: A Case Study of Sentencing and Recidivism. *Sociological Methodology 28*(1), 99–137.

Manski, C. F. and J. V. Pepper (2000). Monotone Instrumental Variables: With an Application to the Returns to Schooling. *Econometrica 68*(4), 997–1010.

Manski, C. F. and J. V. Pepper (2009). More on Monotone Instrumental Variables. *The Econometrics Journal 12*(1), S200–S216.

McCarthy, I., D. L. Millimet, and M. Roy (2015). Bounding Treatment Effects: A Command for the Partial Identification of the Average Treatment Effect with Endogenous and Misreported Treatment Assignment. *The Stata Journal 15*(2), 411–436.

Ministry of Health and Family Welfare (2015). Beneficiaries under Janani Suraksha Yojana. http://pib.nic.in/newsite/PrintRelease.aspx?relid=133709, Accessed in 20/06/2018.

Montgomery, M. R., M. Gragnolati, K. A. Burke, and E. Paredes (2000). Measuring Living Standards with Proxy Variables. *Demography 37*(2), 155–174.

Morris, S. S., R. Flores, P. Olinto, and J. M. Medina (2004). Monetary Incentives in Primary Health Care and Effects on Use and Coverage of Preventive Health Care Interventions in Rural Honduras: Cluster Randomised Trial. *The Lancet 364*(9450), 2030 – 2037.

Mukherjee, S. and A. Singh (2018). Has the Janani Suraksha Yojana (A Conditional Maternity Benefit Transfer Scheme) Succeeded in Reducing the Economic Burden of Maternity in Rural India? Evidence from the Varanasi District of Uttar Pradesh. *Journal of Public Health Research 7*(957), 1–8.

Nandi, A. and R. Laxminarayan (2016). The Unintended Effects of Cash Transfers on Fertility: Evidence from the Safe Motherhood Scheme in India. *Journal of Population Economics 29*(2), 457–491.

Ng, M., A. Misra, V. Diwan, M. Agnani, A. Levin-Rector, and A. D. Costa (2014). An Assessment of the Impact of the JSY Cash Transfer Program on Maternal Mortality Reduction in Madhya Pradesh, India. *Global Health Action 7*(1), 1–10.

Pathak, P. K. and S. K. Mohanty (2010). Does the Safe-motherhood Programme Reach the Poor in Uttar Pradesh, India? *Asian Population Studies 6*(2), 173–191.

Pathak, P. K., A. Singh, and S. V. Subramanian (2010). Economic Inequalities in Maternal Health Care: Prenatal Care and Skilled Birth Attendance in India, 1992–2006. *PLOS ONE 5*(10), 1–17.

Powell-Jackson, T., S. Mazumdar, and A. Mills (2015). Financial Incentives in Health: New Evidence from India's Janani Suraksha Yojana. *Journal of Health Economics 43*, 154 – 169.

Powell-Jackson, T., S. K. Pereira, V. Dutt, S. Tougher, K. Haldar, and P. Kumar (2016). Cash Transfers, Maternal Depression and Emotional Well-being: Quasi-experimental Evidence from India's Janani Suraksha Yojana Programme. *Social Science & Medicine 162*, 210 – 218.

Purohit, N., V. Mehra, and D. Govil (2014). Tracking Benefits of Janani Suraksha Yojana–A Maternal Cash Transfer Scheme: Evidence from Rajasthan, India. *Journal of Health Management 16*(2), 289–302.

Rahman, M. M. and S. Pallikadavath (2017). Intended and Unintended Effects of Janani Suraksha Yojana on Maternal and Child Health Care Services. *Working paper No.2*, 1–47.

Rahman, M. M. and S. Pallikadavath (2018). How much Do Conditional Cash Transfers Increase the Utilization of Maternal and Child Health Care Services? New Evidence from Janani Suraksha Yojana in India. *Economics & Human Biology 31*, 164 – 183.

Randive, B., V. Diwan, and A. De Costa (2013). India's Conditional Cash Transfer Programme (the JSY) to Promote Institutional Birth: Is There an Association between Institutional Birth Proportion and Maternal Mortality? *PLOS ONE 8*(6), 1–8.

Randive, B., M. S. Sebastian, A. D. Costa, and L. Lindholm (2014). Inequalities in Institutional Delivery Uptake and Maternal Mortality Reduction in the Context of Cash Incentive Program, Janani Suraksha Yojana: Results from Nine States in India. *Social Science & Medicine 123*, 1 – 6.

Ranganathan, M. and M. Lagarde (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.

Rawlings, L. B. and G. M. Rubio (2005). Evaluating the Impact of Conditional Cash Transfer Programs. *The World Bank Research Observer 20*(1), 29–55.

Richey, J. A. (2016). An Odd Couple: Monotone Instrumental Variables and Binary Treatments. *Econometric Reviews 35*(6), 1099–1110.

Rubin, D. B. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology 66*(5), 688–701.

Rutstein, S. and S. Staveteig (2014). Making the Demographic and Health Surveys Wealth Index Comparable. DHS Methodological Reports No. 9. Rockville, Maryland, USA: ICF International.

Scott, S. and C. Ronsmans (2009). The Relationship between Birth with a Health Professional and Maternal Mortality in Observational Studies: A Review of the Literature. *Tropical Medicine & International Health 14*(12), 1523–1533.

Sengupta, N. and A. Sinha (2018). Is India's Safe Motherhood Scheme Leading to Better Child Health Care Practices? *Global Social Welfare 5*(1), 49–58.

Thongkong, N., E. van de Poel, S. S. Roy, S. Rath, and T. A. J. Houweling (2017). How Equitable is the Uptake of Conditional Cash Transfers for Maternity Care in India? Evidence from the Janani Suraksha Yojana Scheme in Odisha and Jharkhand. *International Journal for Equity in Health 16*(48), 1–9.

UNICEF (2015). Committing to Child Survival: A Promise Renewed, Progress Report. New York, USA.

UNICEF (2018a). UNICEF Data: Monitoring the Situation of Children and Women. https://data.unicef.org/topic/maternal-health/antenatal-care/, Accessed in 20/06/2018.

UNICEF (2018b). UNICEF Data: Monitoring the Situation of Children and Women. https://data.unicef.org/topic/child-survival/under-five-mortality/, Accessed in 20/06/2018.

Urquieta, J., G. Angeles, T. Mroz, H. Lamadrid-Figueroa, and B. Hernández (2009). Impact of Oportunidades on Skilled Attendance at Delivery in Rural Areas. *Economic Development and Cultural Change 57*(3), 539–558.

Van de Poel, E., G. Flores, P. Ir, O. O'Donnell, and E. Van Doorslaer (2014). Can Vouchers Deliver? An Evaluation of Subsidies for Maternal Health Care in Cambodia. *Bulletin of World Health Organization 92*, 331–339.

Appendix

A Appendix – Additional results

A.1 Additional evidence in LPSs from the National Family Health Survey (NFHS)

We complement the analysis with the latest National Family Household Survey (NFHS-4) conducted in 2015-2016. In contrast to the DLHS-4, the NFHS is nationally representative, which allows us to estimate the impacts in LPSs. The NFHS is a nationwide household survey that provides information on health, health-related behaviours and household socio-economic status; it is the Indian version of the Demographic Health Survey conducted in more than 85 low- and middle-income countries (Corsi et al., 2012). The NFHS-4 has 19,906 female observations in LPSs.[18] After dropping the observations with missing information[19], our final sample size has become 13,112. Descriptive statistics for the NFHS-4 are shown in Table A.1.

Table A.1: Descriptive statistics in LPSs of the NFHS-4

	Non-participants LPSs mean	Participants LPSs mean
Institutional delivery	0.80	1.00
Skilled birth attendance	0.82	0.99
Antenatal care at least once	0.86	0.95
Antenatal care (≥3 times)	0.76	0.87
Postnatal care for mother	0.72	0.87
Postnatal care for baby	0.35	0.48
Iron and folic acid (IFA) supplement	0.83	0.92
Tetanus toxoid (TT) injection	0.89	0.94
Below the poverty line card	0.31	0.38
Scheduled caste	0.16	0.24
Scheduled tribe	0.25	0.30
Rural	0.65	0.69
Birth order (parity)	2.18	1.90
Hindu	0.64	0.65
Maternal age	27.62	26.88
Maternal education years	3.55	3.72
Paternal education years	3.73	3.74
Wealth	3.38	2.03
Observations	10783	2338

Source: NFHS-4.

A.1.1 Point-identification in LPSs

In this subsection, we show the results for the LPSs obtained from the NFHS-4 data. In contrast to the DLHS-4, the NFHS-4 collected data in LPSs as well as HPSs (Figure 1). To make the results comparable to the ones obtained from the DLHS-4 data as much as possible, we use the same covariate sets used for

[18] In the NFHS data, not all women in the samples were interviewed about their husband's characteristics. In this study, we use the nationally representative female samples in which information about a woman's husband is also collected.

[19] We dropped 67 observations with missing information about the JSY status and 68 observations that do not have information about paternal educational background.

the DLHS-4, i.e., the covariate set used in Rahman and Pallikadavath (2018).[20] The only difference in terms of covariates is the component of household wealth. In both data sets, wealth is created from the principal component analysis from indicators of asset ownership, housing characteristics and water and sanitation facilities, but their components are different. In the NLHS-4, we use the wealth which was officially derived in collaboration with the World Bank, and it has been shown to be a consistent proxy for household income and expenditure (Rutstein and Staveteig, 2014; Montgomery et al., 2000).[21] Although in LPSs, all women are eligible for the JSY regardless of their age, socio-economic status, or number of children, we continue to use the same eligibility criteria that are applied in HPSs as an instrument, which makes the results comparable with the previous results in HPSs.

Table A.2 reports the ATEs estimated by: (1) Independence, (2) OLS, (3) PSM, (4) IPW+OLS, (5) NNM, and (6) EBW. All the point-estimated ATEs are positive and significant ($p < 0.01$). We find much larger ATEs in LPSs on institutional delivery and skilled birth attendance than ATEs in HPSs estimated from the DLHS-4 data, which could be attributed to the observation that in LPSs the initial outcome levels are far lower than those in HPSs. The lower initial level and wider coverage of the programme may have led to larger ATEs (Dongre, 2012).

A.1.2 Partial identification in LPSs

First, we surmise the direction of the selection bias. Comparing the ATE under the independence assumption and the ATE estimated by OLS, we find that the ATEs under the independence assumption are larger than those estimated by OLS for ANC at least once, PNC for babies, and iron and folic acid supplement intakes and tetanus toxoid injections. Hence for these outcomes, we impose the positive MTS assumption. For the other outcomes, namely institutional delivery, skilled birth attendance, ANC three times and more, and PNC for mothers, the negative MTS assumption is imposed.

Table A.3 shows the ATE bounds in LPSs estimated from the NFHS-4 data. Figure A.1 visually summarises the bound-estimated ATEs and the point-estimated ATEs. Overall, we find that the point-identified ATEs are outside of the ATE bounds for institutional delivery, skilled birth attendance, ANC at least once, ANC three times and more and PNC for mothers. Compared with the results in HPSs, the deviations of the point-identified ATEs from the upper/lower limits of ATE bounds are substantially smaller in LPSs, suggesting the lower limits of the selection bias in LPSs. For PNC for babies, iron and folic acid supplement intakes and tetanus toxoid injections, some of the ATEs estimated under the CIA are included within the ATE bounds. In particular, for the iron and folic acid supplement intakes, all the point-identified ATEs are within the bound, which is attributable to the weak identification power of the MIV assumption. Additional imposition of the MIV little contributes to tightening the ATE bound.

Overall, we find that most of the ATEs under the CIA are outside the ATE bound. The exception is the effect on the iron and folic acid supplement intakes in which all the point-identified ATEs are inside the partially-identified ATE bound. Comparing the results in both HPSs and LPSs, we observe that the deviations from the ATE bounds are larger in HPSs. The contrast between HPSs and LPSs suggests that we are more likely to suffer larger selection bias when the eligibility of the JSY is restricted to marginalised mothers.

[20] We also estimate the ATEs with the covariate sets in Lim et al. (2010) and Sengupta and Sinha (2018), and we obtained very similar results.

[21] The list of indicators can be downloaded from the official website (https://www.dhsprogram.com/programming/wealth%20index/India%20DHS%202015-16/IndiaNFHS4.pdf)–Accessed in 14/02/2019–.

Table A.2: Point-estimated ATE in LPSs from the NFHS-4

	Estimates	SEs	Confidence intervals
Institutional delivery			
Independence	0.432***	0.005	(0.422, 0.442)
OLS	0.433***	0.004	(0.425, 0.441)
PSM	0.435***	0.005	(0.425, 0.445)
IPW+OLS	0.433***	0.004	(0.425, 0.441)
NNM	0.428***	0.005	(0.418, 0.438)
EBW	0.429***	0.005	(0.419, 0.439)
Skilled delivery attendance			
Independence	0.349***	0.005	(0.339, 0.359)
OLS	0.35***	0.005	(0.34, 0.36)
PSM	0.35***	0.005	(0.34, 0.36)
IPW+OLS	0.35***	0.005	(0.34, 0.36)
NNM	0.346***	0.005	(0.336, 0.356)
EBW	0.345***	0.005	(0.335, 0.355)
Antenatal care +1			
Independence	0.105***	0.005	(0.095, 0.115)
OLS	0.105***	0.005	(0.095, 0.115)
PSM	0.104***	0.006	(0.092, 0.116)
IPW+OLS	0.105***	0.005	(0.095, 0.115)
NNM	0.097***	0.006	(0.085, 0.109)
EBW	0.095***	0.005	(0.085, 0.105)
Antenatal care +3			
Independence	0.101***	0.007	(0.087, 0.115)
OLS	0.104***	0.007	(0.09, 0.118)
PSM	0.105***	0.008	(0.089, 0.121)
IPW+OLS	0.104***	0.007	(0.09, 0.118)
NNM	0.093***	0.008	(0.077, 0.109)
EBW	0.099***	0.007	(0.085, 0.113)
Postnatal care for mother			
Independence	0.178***	0.007	(0.164, 0.192)
OLS	0.18***	0.007	(0.166, 0.194)
PSM	0.176***	0.008	(0.16, 0.192)
IPW+OLS	0.18***	0.007	(0.166, 0.194)
NNM	0.178***	0.008	(0.162, 0.194)
EBW	0.178***	0.007	(0.164, 0.192)
Postnatal care for baby			
Independence	0.065***	0.007	(0.051, 0.079)
OLS	0.061***	0.007	(0.047, 0.075)
PSM	0.055***	0.008	(0.039, 0.071)
IPW+OLS	0.06***	0.007	(0.046, 0.074)
NNM	0.058***	0.008	(0.042, 0.074)
EBW	0.064***	0.007	(0.05, 0.078)
Iron folic acid supplement intakes			
Independence	0.104***	0.006	(0.092, 0.116)
OLS	0.095***	0.006	(0.083, 0.107)
PSM	0.093***	0.007	(0.079, 0.107)
IPW+OLS	0.095***	0.006	(0.083, 0.107)
NNM	0.086***	0.007	(0.072, 0.1)
EBW	0.091***	0.006	(0.079, 0.103)
Tetanus toxoid injections			
Independence	0.056***	0.004	(0.048, 0.064)
OLS	0.052***	0.004	(0.044, 0.06)
PSM	0.052***	0.004	(0.044, 0.06)
IPW+OLS	0.052***	0.004	(0.044, 0.06)
NNM	0.05***	0.004	(0.042, 0.058)
EBW	0.047***	0.003	(0.041, 0.053)

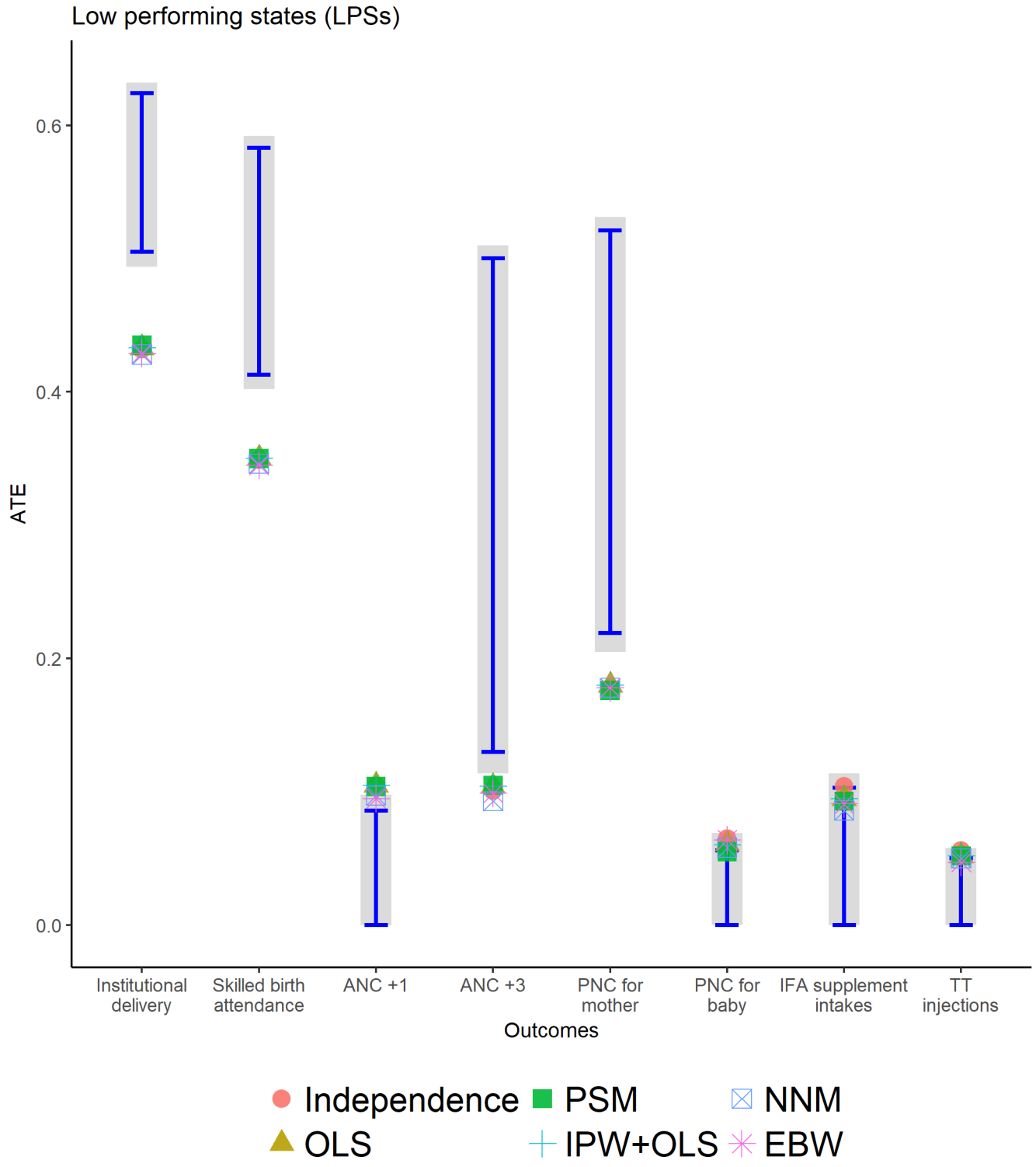
Source: NFHS-4. Note: Number of observations is 13,112. The choice of covariates follows Rahman and Pallikadavath (2018). PSM=Propensity score matching; IPW+OLS=OLS adjusted with the propensity score inverse probability weighting; NNM=Nearest neighbour matching; and EBW=Entropy balance weighting. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: ATE bounds in LPSs estimated from the NFHS-4

	Estimates		**Confidence intervals**	
	Lower bound	Upper bound	Lower bound	Upper bound
Institutional delivery				
No assumption	−0.319	0.681	−0.325	0.686
MTR	0.000	0.681	0.000	0.686
MTR+MTS	0.432	0.681	0.425	0.686
MTR+MTS+MIV	0.505	0.624	0.494	0.632
Skilled delivery attendance				
No assumption	−0.364	0.636	−0.369	0.642
MTR	0.000	0.636	0.000	0.642
MTR+MTS	0.349	0.636	0.341	0.642
MTR+MTS+MIV	0.413	0.583	0.402	0.592
Antenatal care +1				
No assumption	−0.487	0.513	−0.493	0.519
MTR	0.000	0.513	0.000	0.519
MTR+MTS	0.000	0.105	0.000	0.114
MTR+MTS+MIV	0.000	0.086	0.000	0.098
Antenatal care +3				
No assumption	−0.460	0.540	−0.466	0.547
MTR	0.000	0.540	0.000	0.547
MTR+MTS	0.101	0.540	0.089	0.547
MTR+MTS+MIV	0.130	0.500	0.114	0.510
Postnatal care for mother				
No assumption	−0.432	0.568	−0.437	0.574
MTR	0.000	0.568	0.000	0.574
MTR+MTS	0.178	0.568	0.167	0.574
MTR+MTS+MIV	0.219	0.521	0.205	0.531
Postnatal care for baby				
No assumption	−0.449	0.551	−0.455	0.557
MTR	0.000	0.551	0.000	0.557
MTR+MTS	0.000	0.065	0.000	0.075
MTR+MTS+MIV	0.000	0.056	0.000	0.069
Iron folic acid supplement intakes				
No assumption	−0.477	0.523	−0.483	0.529
MTR	0.000	0.523	0.000	0.529
MTR+MTS	0.000	0.104	0.000	0.114
MTR+MTS+MIV	0.000	0.103	0.000	0.114
Tetanus toxoid injections				
No assumption	−0.526	0.474	−0.532	0.480
MTR	0.000	0.474	0.000	0.480
MTR+MTS	0.000	0.056	0.000	0.061
MTR+MTS+MIV	0.000	0.050	0.000	0.058

Source: NFHS-4. Note: Number of observations is 13,112. 95% confidence intervals are calculated following Imbens and Manski (2004) by bootstrap with 200 repetitions.

Figure A.1: ATE bound and point-estimated ATE in LPSs from the NFHS-4



Source: NFHS-4. Note: The choice of covariates follows Rahman and Pallikadavath (2018). PSM=Propensity score matching; IPW+OLS=OLS adjusted with the propensity score inverse probability weighting; NNM=Nearest neighbour matching; EBW=Entropy balance weighting; ANC=Antenatal care; PNC=Postnatal care; IFA=Iron and folic acid; and TT=tetanus toxoid. The bounds shown are sharp bounds estimated under the MTR, MTS and MIV assumptions. The shaded areas show the 95% level confidence interval of the ATE bound under the MTR, MTS and MIV assumptions.

B Appendix – Note on methodology

B.1 Entropy balance weighting

In the non-experimental study, wherein participation in treatment is not randomised, the pure comparison between treated people and untreated people does not generally provide the causal treatment effects because those who are treated can be fundamentally different from the untreated people with regards to their characteristics, such as the educational background. An entropy balance weighting, developed by Hainmueller (2012), perfectly balances the moments of the covariate distributions.

The entropy balance weighting method assigns weights such that the covariate distributions in the weighted data satisfy a set of moment conditions set by a researcher. Specifically, weights are assigned to the control group so that the moments of the weighted covariate distributions in the control group match those in the treatment group. The entropy balance weighting method has a few advantages over the propensity score weighting. Firstly, the entropy balance weighting can "perfectly" balance a set of moments of the covariate distributions between two groups, while the propensity score methods often fail to jointly balance all the covariate distributions in practice partly because of the misspecification of the propensity score. In order to improve the balance of some covariates, the propensity score method usually needs to sacrifice the balance of the other covariates (Iacus et al., 2012). Thanks to the perfect balancing obtained with the entropy balancing method, manual balance checking is no longer necessary. Hence, we don't have to go back and forth between estimating the propensity score and manually checking the balance of the covariates (Hainmueller, 2012).

Each observation in the untreated group gets a weight satisfying a set of balance constraints set by the researcher a priori. Entropy balancing estimates the weights directly from a set of balance constraints and the normalisation and non-negativity constraints. The optimal weights for the untreated group are chosen in such a way that they minimise the following entropy distance metric:

$$H(w) = \sum_{\{i|Untreated\}} w_i ln(w_i/q_i), \tag{B.1}$$

subject to the following balance and normalisation constraints:

$$\sum_{\{i|Untreated\}} w_i c_{ri}(X_i) = m_r \; with \; r \in \{1, ..., R\}, \tag{B.2}$$

$$\sum_{\{i|Untreated\}} w_i = 1, \; and \tag{B.3}$$

$$w_i \geq 0, \; \forall i \in (Untreated). \tag{B.4}$$

Equation (B.2) is a set of balance constraints. Equations (B.3) and (B.4) are the normalisation constraint and the non-negativity constraint respectively. The part $w_i c_{ri}(X_i) = m_r$ in equation (B.2) describes a set of R balance constraints imposed on the moments of the weighted distributions of the covariates in the control group. Each balance constraint equates a certain order moment of the weighted covariate distributions in the control group to the corresponding moment of the covariate distributions in the treated group. As the entropy balancing can balance the higher order moments of the distributions, the moment constraints can be not only the mean (first moment), but also the variance (second moment), the skewness (third moment), and beyond them. For example, if we are interested in estimating the weights for balancing the rth order moment of a specific covariate, say X_p, we set $c_{ri}(X_{ip}) = (X_{ip})^r$ or $c_{ri}(X_{ip}) = (X_{ip} - \mu_p)^r$ with mean μ_p. The choice of the moment order can vary across the covariates and largely depends on the researcher's a priori knowledge about their distribution types. For example,

if the covariates are binary variables, the first order moment balancing is sufficient, whereas in the case of the variables that are normally distributed, balancing the first and second order moments is sufficient to balance their entire distributions. In this study, we will balance the covariate distributions up to their second moments. Using the estimated weight, we non-parametrically estimate the means and the quantiles of the distributions of Y_0 and Y_1.

B.2 MTS sharp bound and independence

This note shows why the upper bound of ATE under the positive MTS assumption corresponds to the ATE under the independence assumption. This note is based on the explanation made by McCarthy et al. (2015). Since the upper bound of ATE is, by definition, the difference between the upper bound of $P(Y_1 = 1)$ and the lower bound of $P(Y_0 = 1)$, we firstly discuss in what circumstance the upper bound of $P(Y_1 = 1)$ and the lower bound of $P(Y_0 = 1)$ are attained.

Applying the law of total probability for $P(Y_1 = 1)$,

$$\begin{aligned} P(Y_1 = 1) &= P(Y_1 = 1|D = 1)P(D = 1) + P(Y_1 = 1|D = 0)P(D = 0) \\ &= P(Y_1 = 1|D = 1)\{1 - P(D = 0)\} + P(Y_1 = 1|D = 0)P(D = 0) \\ &= \{P(Y_1 = 1|D = 0) - P(Y_1 = 1|D = 1)\}P(D = 0) + P(Y_1 = 1|D = 1). \end{aligned} \tag{B.5}$$

Under the MTS assumption, the term in the bracket is non-positive. The upper bound of $P(Y_1 = 1)$, which is $P(Y_1 = 1|D = 1)$, is achieved when the term in the bracket becomes 0, where the independence assumption, $P(Y_1 = 1|D = 0) = P(Y_1 = 1|D = 1)$, is satisfied.

Next, applying the law of total probability for $P(Y_0 = 1)$,

$$\begin{aligned} P(Y_0 = 1) &= P(Y_0 = 1|D = 1)P(D = 1) + P(Y_0 = 1|D = 0)P(D = 0) \\ &= P(Y_0 = 1|D = 1)P(D = 1) + P(Y_0 = 1|D = 0)\{1 - P(D = 1)\} \\ &= \{P(Y_0 = 1|D = 1) - P(Y_0 = 1|D = 0)\}P(D = 1) + P(Y_0 = 1|D = 0), \end{aligned} \tag{B.6}$$

Under the MIS assumption, the term in the bracket is non-negative. The lower bound of $P(Y_0 = 1)$, which is $P(Y_0 = 1|D = 0)$, is achieved when the term in the bracket becomes 0, where the independence assumption, $P(Y_0 = 1|D = 1) = P(Y_0 = 1|D = 0)$, is satisfied.

Combined them together, the upper limit of ATE, which is $P(Y_1 = 1|D = 1) - P(Y_0 = 1|D = 0)$, is achieved when $P(Y_t = 1|D = 1) = P(Y_t = 1|D = 0) = P(Y_t = 1)$ holds for $t \in \{0, 1\}$. This condition is indeed the independence assumption. In the same way, we can show that the lower bound of ATE under the negative MTS assumption corresponds to the ATE estimated under the independence assumption.

B.3 Derivation of the MIV bounds in equations (29) and (30)

First, for $t \in \{0, 1\}$ and $t' \neq t$, the bound of the conditional probability, $P(Y_t = 1|v = u)$ is given by

$$\begin{aligned} & P(Y_t = 1|D = t, v = u)P(D = t|v = u) \\ \leq\ & \boldsymbol{P(Y_t = 1|v = u)} \\ \leq\ & \{P(Y_t = 1|D = t, v = u)P(D = t|v = u) + P(D = t'|v = u)\}. \end{aligned} \tag{B.7}$$

Under the MIV assumption, for $u_1 \leq u \leq u_2$,

$$\begin{aligned} & \max_{u_1 \leq u} P(Y_t = 1|D = t, v = u_1)P(D = t|v = u_1) \\ \leq \; & \boldsymbol{P(Y_t = 1|v = u)} \\ \leq \; & \min_{u \leq u_2} \{P(Y_t = 1|D = t, v = u_2)P(D = t|v = u_2) + P(D = t'|v = u_2)\}. \end{aligned} \tag{B.8}$$

Applying the law of total probability, we obtain the bounds of $P(Y_t = 1)$ as follows:

$$\begin{aligned} & \sum_{u \in V} P(v = u) \max_{u_1 \leq u} P(Y_t = 1|D = t, v = u_1)P(D = t|v = u_1) \\ \leq \; & \boldsymbol{P(Y_t = 1)} \\ \leq \; & \sum_{u \in V} P(v = u) \min_{u \leq u_2} \{P(Y_t = 1|D = t, v = u_2)P(D = t|v = u_2) + P(D = t'|v = u_2)\}. \end{aligned} \tag{B.9}$$

When an instrument is binary, we can simplify this. As u, u_0 and u_1 take the values of 0 or 1,

$$\begin{aligned} & P(v = 0)P(Y_t = 1|D = t, v = 0)P(D = t|v = 0) \\ & +P(v = 1) \max\{P(Y_t = 1|D = t, v = 0)P(D = t|v = 0), P(Y_t = 1|D = t, v = 1)P(D = t|v = 1)\} \\ \leq \; & \boldsymbol{P(Y_t = 1)} \\ \leq \; & P(v = 0) \min\{P(Y_t = 1|D = t, v = 0)P(D = t|v = 0) + P(D = t'|v = 0), \\ & P(Y_t = 1|D = t, v = 1)P(D = t|v = 1) + P(D = t'|v = 1)\} + \\ & P(v = 1)[P(Y_t = 1|D = t, v = 1)P(D = t|v = 1) + P(D = t'|v = 1)]. \end{aligned} \tag{B.10}$$

This leads to equations (29) and (30).

B.4 Finite-sample bias in the presence of maxima and minima operators

First, for $t \in \{0, 1\}$ and $t' \neq t$, we show the estimate of the lower bound of $P(Y_t = 1|v = u)$ has an upward bias.

When the estimate of $P(.)$ is denoted by $\widehat{P}(.)$, the estimated lower bound is expressed as

$$\max_{u_1 \leq u} \widehat{P}(Y_t = 1|D = t, v = u_1)\widehat{P}(D = t|v = u_1). \tag{B.11}$$

Since a maximum operator is a convex function, by the Jensen's inequality,

$$\begin{aligned} & E[\max_{u_1 \leq u} \widehat{P}(Y_t = 1|D = t, v = u_1)\widehat{P}(D = t|v = u_1)] \\ \geq \; & \max_{u_1 \leq u} E[\widehat{P}(Y_t = 1|D = t, v = u_1)\widehat{P}(D = t|v = u_1)] \\ = \; & \max_{u_1 \leq u} \widehat{P}(Y_t = 1|D = t, v = u_1)\widehat{P}(D = t|v = u_1). \end{aligned} \tag{B.12}$$

Equality in the last holds due to the unbiasedness of $\widehat{P}(.)$. Hence equation (B.12) shows that the estimate of the lower bound can have a upward bias.

Next, we show the estimate of the upper bound of $P(Y_t = 1|v = u)$ has an upward bias. The estimated upper bound is expressed as

$$\min_{u \leq u_0} \{\widehat{P}(Y_t = 1|D = t, v = u_0)\widehat{P}(D = t|v = u_0) + \widehat{P}(D = t'|v = u_0)\}. \tag{B.13}$$

Since a maximum operator is a concave function, by the Jensen's inequality,

$$\begin{aligned} & E[\min_{u\leq u_0}\{\widehat{P}(Y_t=1|D=t,v=u_0)\widehat{P}(D=t|v=u_0)+\widehat{P}(D=t'|v=u_0)\}] \\ \leq & \min_{u\leq u_0} E[\{\widehat{P}(Y_t=1|D=t,v=u_0)\widehat{P}(D=t|v=u_0)+\widehat{P}(D=t'|v=u_0)\}] \\ = & \min_{u\leq u_0}\{P(Y_t=1|D=t,v=u_0)P(D=t|v=u_0)+P(D=t'|v=u_0)\}. \end{aligned} \tag{B.14}$$

Equation (B.14) shows that the estimate of the upper bound can have a downward bias.

C Appendix – Derivations of ATE bounds under joint assumptions

C.1 MTR+MTS

We consider the case in which the MTR and MTS assumptions are jointly imposed. Imposing the MTR assumption and the MTS assumption jointly means that we are assuming the intersection of the bounds of $P(Y_1=1)$ and $P(Y_0=1)$ that are derived above under each assumption.

(i) MTR+ Positive MTS: If we combine the MTR assumption and the positive MTS assumption, the bounds of $P(Y_1=1)$ and $P(Y_0=1)$ become

$$\underbrace{P(Y=1)}_{MTR} \leq \quad \boldsymbol{P(Y_1=1)} \quad \leq \underbrace{P(Y_1=1|D=1)}_{Positive\ MTS} \tag{C.15}$$

$$\underbrace{P(Y_0=1|D=0)}_{Positive\ MTS} \leq \quad \boldsymbol{P(Y_0=1)} \quad \leq \underbrace{P(Y=1)}_{MTR}, \tag{C.16}$$

where for $t\in\{0,1\}$ the bound of $P(Y_t=1)$ is an intersection of the corresponding bounds under the MTR assumption and the positive MTS assumption. The new bound of ATE is given by

$$0 \leq ATE \leq P(Y_1=1|D=1)-P(Y_0=1|D=0). \tag{C.17}$$

The upper bound of ATE is given by the upper bound under the positive MTS assumption, while the lower bound is 0 thanks to the MTR assumption.

(ii) MTR+ Negative MTS: The combination of the MTR assumption and the negative MTS assumption can be obtained in the same fashion.

$$\max\{\underbrace{P(Y=1)}_{MTR},\underbrace{P(Y_1=1|D=1)}_{Negative\ MTS}\} \leq \quad \boldsymbol{P(Y_1=1)} \quad \leq \underbrace{P(Y_1=1|D=1)P(D=1)+P(D=0)}_{Worst\ case} \tag{C.18}$$

$$\underbrace{P(Y_0=1|D=0)P(D=0)}_{Worst\ case} \leq \quad \boldsymbol{P(Y_0=1)} \quad \leq \min\{\underbrace{P(Y=1)}_{MTR},\underbrace{P(Y_0=1|D=0)}_{Negative\ MTS}\} \tag{C.19}$$

and the ATE is given by

$$\begin{aligned} & \max\{P(Y=1),P(Y_1=1|D=1)\}-\min\{P(Y=1),P(Y_0=1|D=0)\} \\ \leq & \ ATE \\ \leq & \ \{P(Y_1=1|D=1)P(D=1)+P(D=0)\}-P(Y_0=1|D=0)P(D=0), \end{aligned} \tag{C.20}$$

where the upper bound corresponds to that in the worst-case.

C.2 MTR+MTS+MIV

We estimate the bounds under the MTR, MTS and MIV assumptions. When we impose the MIV assumption, the sharp bounds of $P(Y_1 = 1)$ and $P(Y_0 = 1)$ cannot be expressed simply by the intersection of the bounds estimated under each of assumption. We firstly derive the bounds for conditional probabilities $P(Y_1 = 1|v = u)$ and $P(Y_0 = 1|v = u)$.

(i) MTR+ positive MTS+ MIV: Under the MTR and positive MTS assumptions, the bounds of the conditional probabilities $P(Y_1 = 1|v = u)$ and $P(Y_0 = 1|v = u)$ are give by

$$\underbrace{P(Y = 1|v = u)}_{MTR} \leq \boldsymbol{P(Y_1 = 1|v = u)} \leq \underbrace{P(Y_1 = 1|D = 1, v = u)}_{Positive\ MTS} \tag{C.21}$$

$$\underbrace{P(Y_0 = 1|D = 0, v = u)}_{Positive\ MTS} \leq \boldsymbol{P(Y_0 = 1|v = u)} \leq \underbrace{P(Y = 1|v = u)}_{MTR}. \tag{C.22}$$

Under the MIV assumption, for $u_1 \leq u \leq u_2$,

$$\max_{u_1 \leq u} P(Y = 1|v = u_1) \leq \boldsymbol{P(Y_1 = 1|v = u)} \leq \min_{u \leq u_2} P(Y_1 = 1|D = 1, v = u_2) \tag{C.23}$$

$$\max_{u_1 \leq u} P(Y_0 = 1|D = 0, v = u_1) \leq \boldsymbol{P(Y_0 = 1|v = u)} \leq \min_{u \leq u_2} P(Y = 1|v = u_2). \tag{C.24}$$

Applying the law of total probability, we obtain the bounds of $P(Y_1 = 1)$ and $P(Y_0 = 1)$.

$$\begin{aligned} & \sum_{u \in V} P(v = u) \max_{u_1 \leq u} P(Y = 1|v = u_1) \\ \leq\ & \boldsymbol{P(Y_1 = 1)} \\ \leq\ & \sum_{u \in V} P(v = u) \min_{u \leq u_2} P(Y_1 = 1|D = 1, v = u_2) \end{aligned} \tag{C.25}$$

and

$$\begin{aligned} & \sum_{u \in V} P(v = u) \max_{u_1 \leq u} P(Y_0 = 1|D = 0, v = u_1) \\ \leq\ & \boldsymbol{P(Y_0 = 1)} \\ \leq\ & \sum_{u \in V} P(v = u) \min_{u \leq u_2} P(Y = 1|v = u_2). \end{aligned} \tag{C.26}$$

When an instrument is binary, we can simplify them. As u, u_0 and u_1 take the values of 0 or 1, arranging these equations lead to equations (C.27) and (C.28).

$$\begin{aligned} & P(v = 0)P(Y = 1|v = 0) + P(v = 1) \max\{P(Y = 1|v = 0), P(Y = 1|v = 1)\} \\ \leq\ & \boldsymbol{P(Y_1 = 1)} \\ \leq\ & P(v = 0) \min\{P(Y_1 = 1|D = 1, v = 0), P(Y_1 = 1|D = 1, v = 1)\} + P(v = 1)P(Y_1 = 1|D = 1, v = 1) \end{aligned} \tag{C.27}$$

and

$$\begin{aligned} & P(v = 0)P(Y_0 = 1|D = 0, v = 0) + P(v = 1) \max\{P(Y_0 = 1|D = 0, v = 0), P(Y_0 = 1|D = 0, v = 1)\} \\ \leq\ & \boldsymbol{P(Y_0 = 1)} \\ \leq\ & P(v = 0) \min\{P(Y = 1|v = 0), P(Y = 1|v = 1)\} + P(v = 1)P(Y = 1|v = 1). \end{aligned} \tag{C.28}$$

The sharp bound of the ATE can be obtained by equation (8).

(i) MTR+ negative MTS+ MIV: Under the MTR, negative MTS and MIV assumptions, the bounds of the conditional probabilities $P(Y_1 = 1|v = u)$ and $P(Y_0 = 1|v = u)$ are

$$\begin{aligned} & \max\{\underbrace{P(Y = 1|v = u)}_{MTR}, \underbrace{P(Y_1 = 1|D = 1, v = u)}_{Negative\ MTS}\} \\ \leq\ & \boldsymbol{P(Y_1 = 1|v = u)} \\ \leq\ & P(Y_1 = 1|D = 1, v = u)P(D = 1|v = u) + P(D = 0|v = u) \end{aligned} \tag{C.29}$$

and

$$\begin{aligned} & \max_{u_1 \leq u} P(Y_0 = 1|D = 0, v = u)P(D = 0|v = u) \\ \leq\ & \boldsymbol{P(Y_0 = 1|v = u)} \\ \leq\ & \min_{u \leq u_2}\{\underbrace{P(Y = 1|v = u)}_{MTR}, \underbrace{P(Y_0 = 1|D = 0, v = u)}_{Negative\ MTS}\}. \end{aligned} \tag{C.30}$$

Under the MIV assumption, for $u_1 \leq u \leq u_2$,

$$\begin{aligned} & \max_{u_1 \leq u} \max\{P(Y = 1|v = u_1), P(Y_1 = 1|D = 1, v = u_1)\} \\ \leq\ & \boldsymbol{P(Y_1 = 1|v = u)} \\ \leq\ & \min_{u \leq u_2}[P(Y_1 = 1|D = 1, v = u_2)P(D = 1|v = u_2) + P(D = 0|v = u_2)] \end{aligned} \tag{C.31}$$

and

$$\begin{aligned} & \max_{u_1 \leq u} P(Y_0 = 1|D = 0, v = u_1)P(D = 0|v = u_1) \\ \leq\ & \boldsymbol{P(Y_0 = 1|v = u)} \\ \leq\ & \min_{u \leq u_2} \min\{P(Y = 1|v = u_2), P(Y_0 = 1|D = 0, v = u_2)\}. \end{aligned} \tag{C.32}$$

Applying the law of total probability,

$$\begin{aligned} & \sum_{u \in V} P(v = u) \max_{u_1 \leq u} \max\{P(Y = 1|v = u_1), P(Y_1 = 1|D = 1, v = u_1)\} \\ \leq\ & \boldsymbol{P(Y_1 = 1)} \\ \leq\ & \sum_{u \in V} P(v = u) \min_{u \leq u_2}\{P(Y_1 = 1|D = 1, v = u_2)P(D = 1|v = u_2) + P(D = 0|v = u_2)\} \end{aligned} \tag{C.33}$$

and

$$\begin{aligned} & \sum_{u \in V} P(v = u) \max_{u_1 \leq u} P(Y_0 = 1|D = 0, v = u_1)P(D = 0|v = u_1) \\ \leq\ & \boldsymbol{P(Y_0 = 1)} \\ \leq\ & \sum_{u \in V} P(v = u) \min_{u \leq u_2} \min\{P(Y = 1|v = u_2), P(Y_0 = 1|D = 0, v = u_2)\}. \end{aligned} \tag{C.34}$$

As u, u_1 and u_2 take the values of 0 or 1, arranging them leads to equations (C.35) and (C.36).

$$
\begin{aligned}
& P(v=0)\max\{P(Y=1|v=0), P(Y_1=1|D=1,v=0)\} + \\
& P(v=1)\max\{P(Y=1|v=0), P(Y_1=1|D=1,v=0), P(Y=1|v=1), P(Y_1=1|D=1,v=1)\} \\
\leq\ & \boldsymbol{P(Y_1=1)} \qquad \text{(C.35)} \\
\leq\ & P(v=0)\min\{P(Y_1=1|D=1,v=0)P(D=1|v=0)+P(D=0|v=0), \\
& P(Y_1=1|D=1,v=1)P(D=1|v=1)+P(D=0|v=1)\} \\
& +P(v=1)\{P(Y_1=1|D=1,v=1)P(D=1|v=1)+P(D=0|v=1)\}
\end{aligned}
$$

and

$$
\begin{aligned}
& P(v=0)P(Y_0=1|D=0,v=0)P(D=0|v=0) \\
& +P(v=1)\max\{P(Y_0=1|D=0,v=0)P(D=0|v=0), P(Y_0=1|D=0,v=1)P(D=0|v=1)\} \\
\leq\ & \boldsymbol{P(Y_0=1)} \qquad \text{(C.36)} \\
\leq\ & P(v=0)\min\{P(Y=1|v=0), P(Y_0=1|D=0,v=0), P(Y=1|v=1), P(Y_0=1|D=0,v=1)\} \\
& +P(v=1)\min\{P(Y=1|v=1), P(Y_0=1|D=0,v=1)\}.
\end{aligned}
$$

The sharp bound of the ATE can be obtained by equation (8).

D Appendix – Derivations of ATE under the other combinations of assumptions

D.1 MTS+MIV

(i) Positive MTS+MIV: Under the positive MTS assumption, the bounds of the conditional probabilities $P(Y_1=1|v=u)$ and $P(Y_0=1|v=u)$ are give by

$$
\begin{aligned}
& P(Y_1=1|D=1,v=u)P(D=1|v=u) \\
\leq\ & \boldsymbol{P(Y_1=1|v=u)} \qquad \text{(D.37)} \\
\leq\ & P(Y_1=1|D=1,v=u)
\end{aligned}
$$

and

$$
\begin{aligned}
& P(Y_0=1|D=0,v=u) \\
\leq\ & \boldsymbol{P(Y_0=1|v=u)} \qquad \text{(D.38)} \\
\leq\ & P(Y_0=1|D=0,v=u)+P(D=1|v=u).
\end{aligned}
$$

Applying the law of total probability, we obtain the bounds of $P(Y_1=1)$ and $P(Y_0=1)$ as follows:

$$
\begin{aligned}
& \sum_{u\in V} P(v=u)\max_{u_1\leq u} P(Y_1=1|D=1,v=u_1)P(D=1|v=u_1) \\
\leq\ & \boldsymbol{P(Y_1=1)} \qquad \text{(D.39)} \\
\leq\ & \sum_{u\in V} P(v=u)\min_{u\leq u_2} P(Y_1=1|D=1,v=u_2)
\end{aligned}
$$

and

$$\begin{aligned} & \sum_{u\in V} P(v=u) \max_{u_1 \le u} P(Y_0=1|D=0,v=u_1) \\ \le\ & \boldsymbol{P(Y_0=1)} \\ \le\ & \sum_{u\in V} P(v=u) \min_{u \le u_2} \{P(Y_0=1|D=0,v=u_2) + P(D=1|v=u_2)\}. \end{aligned} \tag{D.40}$$

When an instrument is binary, we can simplify them. As u, u_0 and u_1 take the values of 0 or 1,

$$\begin{aligned} & P(v=0)P(Y_1=1|D=1,v=0)P(D=1|v=0) + \\ & P(v=1)\max\{P(Y_1=1|D=1,v=0)P(D=1|v=0), P(Y_1=1|D=1,v=1)P(D=1|v=1)\} \\ \le\ & \boldsymbol{P(Y_1=1)} \\ \le\ & P(v=0)\min\{P(Y_1=1|D=1,v=0), P(Y_1=1|D=1,v=1)\} + \\ & P(v=1)P(Y_1=1|D=1,v=1) \end{aligned} \tag{D.41}$$

and

$$\begin{aligned} & P(v=0)P(Y_0=1|D=0,v=0) + \\ & P(v=1)\max\{P(Y_0=1|D=0,v=0), P(Y_0=1|D=0,v=1)\} \\ \le\ & \boldsymbol{P(Y_0=1)} \\ \le\ & P(v=0)\min\{P(Y_0=1|D=0,v=0)+P(D=1|v=0), P(Y_0=1|D=0,v=1)+P(D=1|v=1)\} + \\ & P(v=1)\{P(Y_0=1|D=0,v=1)+P(D=1|v=1)\}. \end{aligned} \tag{D.42}$$

(ii) Negative MTS+MIV: Under the negative MTS assumption, the bounds of the conditional probabilities $P(Y_1=1|v=u)$ and $P(Y_0=1|v=u)$ are give by

$$\begin{aligned} & P(Y_1=1|D=1,v=u) \\ \le\ & \boldsymbol{P(Y_1=1|v=u)} \\ \le\ & P(Y_1=1|D=1,v=u)P(D=1|v=u) + P(D=0|v=u) \end{aligned} \tag{D.43}$$

and

$$\begin{aligned} & P(Y_0=1|D=0,v=u)P(D=0|v=u) \\ \le\ & \boldsymbol{P(Y_0=1|v=u)} \\ \le\ & P(Y_0=1|D=0,v=u). \end{aligned} \tag{D.44}$$

Applying the law of total probability, we obtain the bounds of $P(Y_1=1)$ and $P(Y_0=1)$ as follows:

$$\begin{aligned} & \sum_{u\in V} P(v=u) \max_{u_1 \le u} P(Y_1=1|D=1,v=u_1) \\ \le\ & \boldsymbol{P(Y_1=1)} \\ \le\ & \sum_{u\in V} P(v=u) \min_{u \le u_2} \{P(Y_1=1|D=1,v=u_2)P(D=1|v=u_2) + P(D=0|v=u_2)\} \end{aligned} \tag{D.45}$$

and

$$\begin{aligned} & \sum_{u\in V} P(v=u)\max_{u_1\le u} P(Y_0=1|D=0,v=u_1)P(D=0|v=u_1) \\ \le\ & \boldsymbol{P(Y_0=1)} \\ \le\ & \sum_{u\in V} P(v=u)\min_{u\le u_2} P(Y_0=1|D=0,v=u_2). \end{aligned} \tag{D.46}$$

When an instrument is binary, we can simplify them. As u, u_0 and u_1 take the values of 0 or 1,

$$\begin{aligned} & P(v=0)P(Y_1=1|D=1,v=0)+ \\ & P(v=1)\max\{P(Y_1=1|D=1,v=0),P(Y_1=1|D=1,v=1)\} \\ \le\ & \boldsymbol{P(Y_1=1)} \\ \le\ & P(v=0)\min\{P(Y_1=1|D=1,v=0)P(D=1|v=0)+P(D=0|v=0), \\ & P(Y_1=1|D=1,v=1)P(D=1|v=1)+P(D=0|v=1)\}+ \\ & P(v=1)\{P(Y_1=1|D=1,v=1)P(D=1|v=1)+P(D=0|v=1)\} \end{aligned} \tag{D.47}$$

and

$$\begin{aligned} & P(v=0)P(Y_0=1|D=0,v=0)P(D=0|v=0)+ \\ & P(v=1)\max\{P(Y_0=1|D=0,v=0)P(D=0|v=0),P(Y_0=1|D=0,v=1)P(D=0|v=1)\} \\ \le\ & \boldsymbol{P(Y_0=1)} \\ \le\ & P(v=0)\min\{P(Y_0=1|D=0,v=0),P(Y_0=1|D=0,v=1)\}+ \\ & P(v=1)P(Y_0=1|D=0,v=1). \end{aligned} \tag{D.48}$$

D.2 MTR+MIV

Under the MTR assumption, the bounds of the conditional probabilities $P(Y_1=1|v=u)$ and $P(Y_0=1|v=u)$ are give by

$$\begin{aligned} & P(Y=1|v=u) \\ \le\ & \boldsymbol{P(Y_1=1|v=u)} \\ \le\ & P(Y_1=1|D=1,v=u)P(D=1|v=u)+P(D=0|v=u) \end{aligned} \tag{D.49}$$

and

$$\begin{aligned} & P(Y_0=1|D=0,v=u)P(D=0|v=u) \\ \le\ & \boldsymbol{P(Y_0=1|v=u)} \\ \le\ & P(Y=1|v=u). \end{aligned} \tag{D.50}$$

Applying the law of total probability, we obtain the bounds of $P(Y_1=1)$ and $P(Y_0=1)$ as follows:

$$\begin{aligned} & \sum_{u\in V} P(v=u)\max_{u_1\le u} P(Y=1|v=u_1) \\ \le\ & \boldsymbol{P(Y_1=1)} \\ \le\ & \sum_{u\in V} P(v=u)\min_{u\le u_2}[P(Y_1=1|D=1,v=u_2)P(D=1|v=u_2)+P(D=0|v=u_2)] \end{aligned} \tag{D.51}$$

and

$$\begin{aligned} & \sum_{u\in V} P(v=u)\max_{u_1\le u} P(Y_0=1|D=0,v=u_1)P(D=0|v=u_1) \\ \le\ & \boldsymbol{P(Y_0=1)} \\ \le\ & \sum_{u\in V} P(v=u)\min_{u\le u_2} P(Y=1|v=u_2). \end{aligned} \tag{D.52}$$

When an instrument is binary, we can simplify them. As u, u_0 and u_1 take the values of 0 or 1,

$$\begin{aligned} & P(v=0)P(Y=1|v=0)+P(v=1)\max\{P(Y=1|v=0),P(Y=1|v=1)\} \\ \le\ & \boldsymbol{P(Y_1=1)} \\ \le\ & P(v=0)\min\{P(Y_1=1|D=1,v=0)P(D=1|v=0)+P(D=0|v=0), \\ & P(Y_1=1|D=1,v=1)P(D=1|v=1)+P(D=0|v=1)\} \\ & +P(v=1)\{P(Y_1=1|D=1,v=1)P(D=1|v=1)+P(D=0|v=1)\} \end{aligned} \tag{D.53}$$

and

$$\begin{aligned} & P(v=0)P(Y_0=1|D=0,v=0)P(D=0|v=0)+ \\ & P(v=1)\max\{P(Y_0=1|D=0,v=0)P(D=0|v=0),P(Y_0=1|D=0,v=1)P(D=0|v=1)\} \\ \le\ & \boldsymbol{P(Y_0=1)} \\ \le\ & P(v=0)\min\{P(Y=1|v=0),P(Y=1|v=1)\}+P(v=1)P(Y=1|v=1). \end{aligned} \tag{D.54}$$