3 Spontaneous termination and induced abortion according to contraceptive use at the time of pregnancy[8][9]

3.1 Resumen

Motivación: El uso de anticonceptivos afecta a la fecundidad no solo reduciendo la probabilidad de quedar embarazada sino también porque disminuye la probabilidad que un embarazo termine en nacimiento. Las terminaciones de un embarazo incluyen el aborto inducido y las terminaciones espontáneas. Para una distinción adecuada se requiere tomar en cuenta el riesgo en competencia entre los posibles resultados de un embarazo. No se han realizado estudios comparativos tomando en cuenta la información de los calendarios de vida reproductiva.

Metodología: Usamos 52,616 embarazos de mujeres con edades entre 15 y 49 años a partir de 14 encuestas DHS recogidas entre 2003 y 2017, las cuales incluyen calendarios de historia reproductiva. Estimamos la probabilidad diferencial entre terminación espontánea y aborto inducido de acuerdo con si la mujer usó o no anticonceptivos al momento de quedarse embarazada, controlando por variables demográficas y socioeconómicas. Utilizamos modelos logísticos multinomiales para tomar en cuenta el riesgo en competencia. También, exponemos las limitaciones en el uso de los datos.

Resultados: El uso de anticonceptivos al momento del embarazo está asociado con una mayor probabilidad de aborto inducido y también con un riesgo mayor de terminación espontánea. Si no se toma en cuenta el riesgo en competencia se obtienen estimaciones sesgadas hacia abajo mostrando un menor riesgo de pérdida espontánea. Los gradientes por edad son importantes, pero fuertemente influenciados por la inclusión de las características de

[8] Una versión previa de este capítulo fue presentada en la XIII edition of the Population Days 2019 en Milán, Italia.

[9] Este capítulo se encuentra actualmente bajo revisión por pares en una revista internacional de alto impacto.

la historia reproductiva de la mujer, como paridad, lo que indica el uso del aborto inducido para limitar el tamaño de la familia.

Discusión: Embarazos luego de un fallo en el uso de anticonceptivos tienen una mayor probabilidad de no terminar en un nacimiento debido al mayor riesgo de aborto inducido y de terminación espontánea. Los modelos agregados sobre el impacto de la planificación familiar deberían reflejar que el uso de anticonceptivos y el aborto inducido conforman estrategias interdependientes, mientras que la terminación espontánea es un riesgo en competencia del aborto inducido.

3.2 Abstract

Background: Contraceptive use affects fertility not only by reducing the chances of getting pregnant but also by lowering the probability of a pregnancy ending in a live birth. Pregnancy terminations include both induced abortion and spontaneous terminations. Proper separation requires accounting for the competing risk among pregnancy outcomes. No previous comparative studies of pregnancy outcomes are based on the rich information available in contraceptive calendars.

Methods: Using 52,616 pregnancies of women aged 15-49 from 14 DHS surveys collected between 2003 and 2017 with reproductive history calendars, we estimate the differential odds of spontaneous termination and induced abortion according to contraceptive use at the time of pregnancy, controlling for demographic and socioeconomic covariates and addressing potential data limitations. Multinomial logistic models account for competing risks.

Results: Contraceptive use at the time of pregnancy is associated with much higher odds of induced abortion but also moderately higher risk of spontaneous termination. Not accounting for competing risks biases estimates downwards often indicating a lower risk of spontaneous terminations. Age-gradients are important, but strongly influenced by the inclusion of reproductive history characteristics such as parity suggesting the use of induced abortion to

limit family size.

Discussion: Pregnancies after contraceptive failure are much more likely not to end in a live-birth, because of increased risk of induced abortion but also changing risk of spontaneous termination. Aggregate models of the impact of family planning should reflect that contraceptive use and induced abortion conform interdependent strategies, and that spontaneous termination is a competing risk of induced abortion.

3.3 Background

Fertility levels depend on the probability of pregnancies ending in a live birth. A comparative study shows proportions of pregnancy terminations ranging from 4.9% and 52.0% in 20 countries, with induced abortion explaining the higher values (Bradley, Croft, and Rutstein 2011). Pregnancy terminations include both spontaneous termination (ST) and induced abortion (IA), which are potentially associated with very different factors (Ahmed and Ray 2014). The goal of this article is to analyze the association of contraceptive use at pregnancy with the type of pregnancy outcome.

Global estimates of the regional prevalence of IA range between 12% and 39% of pregnancies in the period 2010-2014 (Sedgh et al. 2016). These are consensus estimates based on relatively good quality data for countries with high incidence but very scanty data from heterogeneous sources for countries where IA laws are restrictive. The prevalence of IA is known to be associated with institutional factors, such as abortion laws and the functioning of health systems, and to characteristics of the woman. At the personal level, IA is a behavioral choice associated with parity, marital status, age, and socioeconomic variables like level of education or wealth (Dickson, Adde, and Ahinkorah 2018; Chae et al. 2017; Maharana 2017; Souza e Silva et al. 2012). In this respect, an unplanned or unwanted pregnancy is the most commonly reported reason behind an IA (Bankole, Singh, and Haas 1998). Current estimates of the potential impact of contraceptive use are based on the probability of IA in unintended pregnancies (Askew et al. 2017; Bearak et al. 2018). However, contraceptive use at pregnancy indicates a more intense desire to avoid pregnancy since an action is already in place. Therefore, we expect that pregnancies resulting from contraceptive failure are more likely to end in IA.

In contrast to IA, ST, including both miscarriages and stillbirths, should not be perceived as a choice. This does not preclude an association with personal and institutional variables: In addition to the role of health systems, there can be differential biological risk and behavioral differences. Empirical studies on ST confirm the relevance of demographic (e.g., age, parity), health (e.g., illness, antenatal care), and socioeconomic determinants (e.g., education, wealth) (Mosley and Chen 2003; Cai and Feng 2005; Norsker et al. 2012; Zheng et al. 2017; Nfii 2017). Country-case studies show estimates of ST ranging from 4% to 20% of pregnancies with incidence rising with age (Carlson, Hoem, and Rychtarikova 1999; Nybo-Andersen et al. 2000; Cai and Feng 2005; Akker 2012), but evidence on global patterns of ST, and in particular its determinants, is still feeble (Lawn et al. 2016; Askew et al. 2017). In most global models of reproductive health, miscarriages are estimated ad-hoc as a fixed proportion of births and IAs (Stover and Winfrey 2017; Darroch 2018), while evidence starts to accumulate regarding global patterns of stillbirths (Blencowe et al. 2016).

The main role of contraception is to avoid pregnancy. Therefore, pregnancies occurring while using contraceptives are labeled as contraceptive failures and are mostly associated with the use of traditional methods, discontinuation or misuse (Polis et al. 2016). The use of contraceptives denotes not only an interest in avoiding pregnancy but also the determination to do something about it. In terms of the `ready, willing, able` framework (Lesthaeghe and Vanderhoeft 2001) we can reasonably expect women experiencing a contraceptive failure to be more likely to engage in an IA to stop their pregnancy. Evidence of this can be found in high-abortion countries (Westoff 2005).

On the other hand, abortion and the use of modern contraceptives can be considered as substitutes regarding birth prevention. This is consistent with the observed reduction in the prevalence of modern contraception in Nepal in areas where new abortion centers opened (Miller and Valente 2016) or a higher likelihood of IA among users of traditional methods compared to users of modern methods in several post-soviet countries (Westoff 2005). Non-contraceptive users are more heterogeneous: They include women that want to get pregnant together with women not wanting to give birth who have an unmet need for contraception (Sedgh, Ashford, and Hussain 2016). Non-users with unmet need for contraception are, in general, less likely than users to abort, but much more likely to do so than women seeking pregnancy (Westoff 2005).

In contrast to IA, less is known about the relationship between contraceptive use at the time of pregnancy and ST. Medical studies agree in the absence of a causal effect of contraceptives on ST (Jellesen et al. 2008; Waller et al. 2010). However, since pregnancies resulting from contraceptive failure are undesired, they are linked to behavioral differences in prenatal care that can result in higher rates of ST (Marston and Cleland 2004; Cheng et al. 2009): Women whose pregnancy is unwanted or mistimed are less likely to seek prenatal care in a 5-country study (Marston and Cleland 2003) and a study from 32 low-income countries (Guliani, Sepehri, and Serieux 2013). However, another study based on seven countries did not find a relevant link between unwantedness and antenatal care (Saad-Haddad et al. 2016). A recent systematic review (Hall et al. 2017) showed increased odds of low birth weight and neonatal mortality for unintended pregnancies; however, it could not locate studies from developing countries looking at the relationship between ST and pregnancy intentions.[10]

A major methodological challenge is that we can view live-births, IA, and ST as competing outcomes. An early ST might make a subsequent IA not necessary, and some pregnancies ending in IA would have otherwise ended in an ST. Moreover, a live-birth requires that the pregnancy did not terminate earlier (Potter, Ford, and Moots 1975; Meister and Schaefer

[10]The review points to one article from Ethiopia, but upon closer inspection, the article looks at the odds of all types of termination including IA.

2008). Naïve estimators based on the proportion of pregnancies ending in an outcome with no control for competing risks are biased, and different alternative indicators have been proposed (Susser 1983; Figà-Talamanca and Repetto 1988; Hammerslough 1992). In terms of statistical modeling, different strategies have been used in the literature to account for the competing risk: A trivariate probit model treating IA, ST, and live-birth as related separate outcomes (Ahmed and Ray 2014); a multinomial logit to differentiate among IA decided by medical persons, IA decided by others, and ST, conditional on pregnancy termination (Maharana 2017); or a multinomial logit considering ST, IA, and ectopic pregnancy conditional on termination (Schwandt et al. 2011). In our case, there are three reasons to model conditional on pregnancy. First, only in this way it is possible to include pregnancy level covariates such as contraceptive use at pregnancy. Second, pregnancy termination only occurs in the context of a previous pregnancy, and third, contraceptive use at pregnancy carries with it a meaning of contraceptive failure that would not be present, for instance, in the trivariate logit model of unconditional risk: Contraceptive use has two different simultaneous effects. It reduces terminations by lowering the risk of pregnancy while increasing the probability of IA conditional on pregnancy since the pregnancy is unintended. Our interest in this research is not on the net effect, but rather on the second effect on the probability of pregnancy outcomes.

Our goal is, therefore, to measure the differential odds of ST and IA according to use of contraceptives at the time of pregnancy while accounting for the competing risk of pregnancy outcomes, and controlling for the demographic and socioeconomic variables indicated as relevant in the literature. We use contraceptive calendar data from Demographic and Health Surveys (DHS) that meet quality checks. The relationship is not necessarily causal: We expect it mostly to be associated with differential behavior.

This research has policy implications regarding aggregate models of the effects of family planning on births, abortions, miscarriages, stillbirths, and maternal mortality (Darroch 2018;

Stover and Winfrey 2017; Askew et al. 2017). While a variety of methods exist to estimate the impact of contraceptive use, an emerging consensus is building around the reference concept of unintended pregnancies in order to estimate IAs (Askew et al. 2017; Bearak et al. 2018). This is a much simpler perspective than the Westoff approach (Westoff 2005) that also subsumes our proposal based on contraceptive use at the time of pregnancy. A significant advantage of our approach is the admission from the outset that contraceptive use and abortion are dependent on each other. This is, for instance, absent in the proximate determinants of fertility framework (Bongaarts and Potter 1983) in which the Spectrum model is based (Stover and Winfrey 2017), and which is also used to estimate IA indirectly by the `residual method` (Rossier 2003). To our knowledge, this is the first time that the risks of ST and IA have been jointly modeled based on DHS calendar data in an international comparison. We are also providing evidence on the determinants of ST. While more evidence is becoming available on stillbirths (Blencowe et al. 2016), very little is known regarding its relationship with contraceptive failure. Miscarriages are a weak point in current aggregate models. Given the lack of reliable statistics (Askew et al. 2017), they are imputed based on ad-hoc assumptions such as a constant rate of miscarriages for all surveys (Stover and Winfrey 2017; Darroch 2018). Our results provide instructive evidence that can help in refining aggregate models. Aggregate models are essential in informing policy since they are used to measure key Family Planning 2020 indicators such as unsafe abortions averted due to modern contraceptive use (Askew et al. 2017).

3.4 Data and methods

3.4.1 Data

DHS surveys include, in most cases, a contraceptive calendar going back up to 72 months before the interview (The DHS Program 2017). In this monthly calendar, women report pregnancies, the outcome of those pregnancies (live-birth or termination), and contraceptive methods used. We use all possible surveys meeting the following requirements: Having a contraceptive calendar, identifying the type of termination (miscarriage, stillbirth, or IA), interviewing women not-in-union, and including all our covariates of interest, in particular education and wealth quintiles. Unfortunately, most DHS surveys do not report the type of pregnancy termination (Christou, Dibley, and Raynes-Greenow 2017). Hence, we can only employ data from 14 DHS surveys collected between 2003 and 2017. Figure 5 shows the countries included in a map and the proportion of pregnancies ending in ST and IA in the different surveys. Our sample includes individual-level information for 52,616 pregnancies of women aged 15-49 at the time of interview that started in the period of 48 to 9 months before the pregnancy (table 3). We exclude pregnancies starting in the eight months preceding the survey to avoid right censoring.

Since we are interested in risk factors at the pregnancy level, we use the available information to infer the values of covariates at the time of pregnancy. Our main variable of interest, contraceptive use at the time of pregnancy, is directly available in the contraceptive calendar. The pregnancy history makes it possible to calculate the woman's parity and the number of previous terminations for each pregnancy. About half of the surveys (Turkey 2003, Philippines 2003, Moldova 2005, Kazakhstan 1999, Indonesia 2012, and Armenia 2000 and 2005) include a specific calendar on union status that provides information on union status at pregnancy. In the rest of surveys, union status has been imputed based on current union status and the moment and duration of the first union. Regarding women's age, we use age-groups defined according to the imputed age at birth. The imputed age at birth is equal to the mother's age at birth in pregnancies carried to term, and age at pregnancy plus nine months for terminated pregnancies. These age groups are then comparable to those standard in fertility analysis. Based on calendar data it is not possible to infer intention status for all pregnancies since those questions are only asked for live births and ongoing pregnancies.

Socioeconomic variables used as controls include the level of education, wealth quintile, employment, and place of residence. They are only observed at the time of interview.

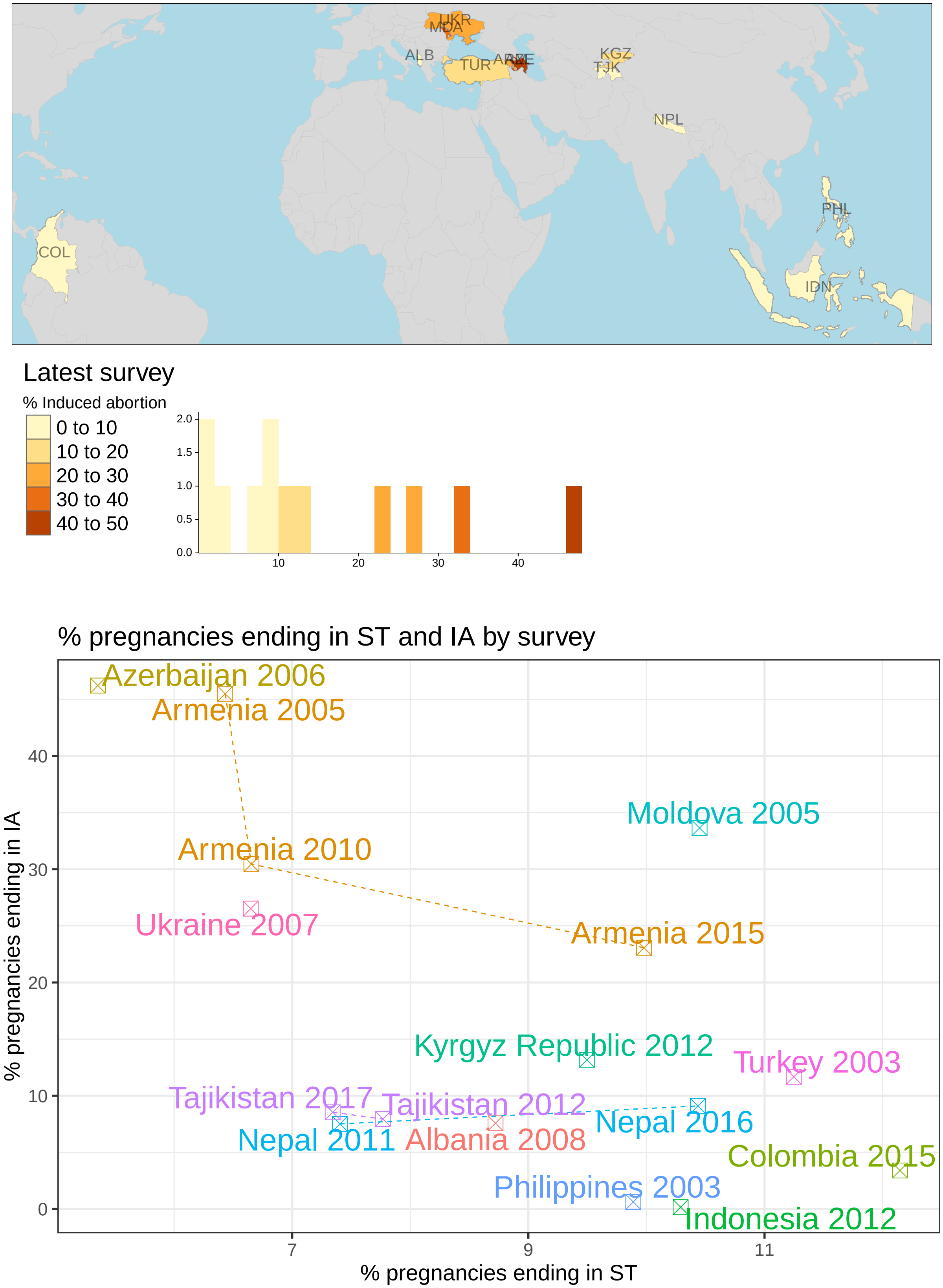

Figure 5: Percentage of pregnancies ending in termination by survey and country.

3.4.2 Methods

A pregnancy ends in birth only if the competing risks of ST and IA do not cause a premature termination. Ignoring the competing nature of risks leads to biased estimates of all risks (Potter, Ford, and Moots 1975, Meister and Schaefer (2008)). In such multiple outcome situations, the multinomial logit model (MNL) provides consistent and efficient estimates when the assumption of independence of irrelevant alternatives (IIA) is met (Cheng and Long 2007). IIA implies that removing an alternative does not alter the relative odds of the rest of alternatives. As a result, removing one outcome would increase the chances of the rest of outcomes. IIA is met in our specific context to the extent that in the absence of IA the biological and behavioral risks for ST would still be there, and that the decision of IA is taken irrespective of the possibility of an ST. STs are generally classified as miscarriages and stillbirths according to the week of pregnancy.[11] While this distinction makes sense from a medical and public health point of view, miscarriages and stillbirths should not be included as separate outcomes in an MNL since they are sequential instead of competing risks. Its inclusion would lead to a violation of the IIA assumption since eliminating stillbirths can only lead to increased likelihood of a birth outcome, leaving IA and miscarriages unaltered. There can also be boundary problems in separating stillbirths from miscarriages (Carlson, Hoem, and Rychtarikova 1999). In our sample miscarriages are much more common than stillbirths: They represent 91.3% of STs. Therefore, the results on ST should be interpreted as referring to miscarriage.

In our analysis, we first look at the conditional probabilities of pregnancy outcomes and binomial logistic regressions for ST and IA using the rest of pregnancy outcomes as the complementary category. Such estimates fail to incorporate competing risks and are biased. We compare them to the consistent and efficient MNL estimates. We estimate survey-specific models that are summarily discussed in the results section and pooled models with data from

[11] WHO recommends 28 weeks as the threshold, https://www.who.int/maternal_child_adolescent/epidemiology/stillbirth/en/, DHS defines stillbirths as occurring in month 7 or later (MacQuarrie et al. 2018).

all the surveys assessed to have no problems of misclassification (see subsection 3.4.3).

For each of the binomial and MNL regressions, a baseline model is estimated. This model only includes our main variable of interest: Contraceptive use at the time of pregnancy. Comparing binomial and MNL estimates provide an idea of biases arising when ignoring the competing risk of pregnancy outcomes. We then estimate four different MNL models. The first model adds age-group to the baseline model. The second model adds the interaction of contraceptive use and age-group to capture the different age-gradients according to contraceptive use. Models 1 and 2 are in line with studies concluding that age is a significant predictor for both IA and ST (Santow and Bracher 1989; Koo et al. 2012). The third model adds women's demographic characteristics like marital status at pregnancy and reproductive history summarized by parity (number of previous live-births at the moment of pregnancy) and the number of previous terminations (the difference between gravidity and the number of deliveries before the current pregnancy). This last variable captures differential risk for women that previously experienced terminations. Since there is little evidence supporting the causality of IA on subsequent ST, our interpretation in the case of ST is a biological predisposition while for IA it would signal the acceptance of IA as a method to avoid unwanted births (Thorp, Hartmann, and Shadigan 2005; Bhattacharya et al. 2012). The fourth model adds socioeconomic variables at the time of the interview. All models include survey-level fixed-effects and controls for recall error (see subsection 3.4.3). All estimations use women weights rescaled to an average of 1 at the survey level. The research is carried out in R (R Core Team 2019) using `multinom` from the `nnet` package for the estimation of MNL models (Venables and Ripley 2002).

Note that MNL estimates both equations simultaneously. Likelihood-ratio tests of significance would only indicate the relevance of a variable without identifying a particular outcome. To test the relevance of a variable for a particular outcome, we use asymptotic Wald tests of joint significance for groups of variables of interest. Note also that in models with interactions, a test of significance for a variable requires a joint test of all terms including that variable.

3.4.3 Addressing data limitations

Contraceptive calendar data includes data for all pregnancies occurring to a woman in the selected period before the interview. It is subject, however, to certain limitations including recall error, omission error, and misclassification of outcomes. We address the potential role of these effects based on current knowledge and devise methods to limit their impact.

Recall error can be present in all retrospective reporting. Still, it is not clear that alternative methods provide more reliable estimates. A comparison of a retrospective survey and continuous population monitoring showed the retrospective survey missing fewer births and fewer terminations than monitoring (Kadobera et al. 2017), and miscarriage rates reported in a recent large-scale prospective cohort study (Ahmed et al. 2018) seem much lower than retrospective survey estimates. Recall error can be identified by a systematic pattern of decline in events registered when going back in time. Different degrees of recall error in DHS terminations have been found using that approach (MacQuarrie et al. 2018). We have taken two actions to mitigate and measure the impact of recall error: First, we do not use all the information in the calendar data, limiting our analysis to pregnancies starting in the 48 months before the interview. This avoids the data with worse deterioration problems together with displacement problems around the cutoff year (Schoumaker 2014). Second, we include a recall error covariate in all models (MacQuarrie et al. 2018). The recall error variable is defined as the distance in years between the month when the pregnancy started and the baseline month of 9 months before the interview. Since this variable is included both in ST and IA regressions, we allow for differential recall error according to outcome.

In our study, there can be omissions regarding pregnancies, their outcomes, or contraceptive use at the time of pregnancy. Regarding pregnancies, there can be different degrees of omission according to the outcome. Miscarriages, particularly those happening early in the pregnancy, can be missing due to ignorance of being pregnant, forgetting, or cultural differences (Cai and Feng 2005). We have found exploratory evidence of this in finding larger differences according to education in first-trimester miscarriages. On the other hand, IA could be either reported as miscarriage (misclassification) or omitted. Intentional omission can be high in contexts where IA is not legal or is not socially accepted (Barreto et al. 1992). There can also be an unintentional omission of medication abortion not being considered as an IA (Jilozian and Agadjanian 2016).

Regarding stillbirths, a DHS-based study evaluated the consistency of DHS calendar data (Bradley, Winfrey, and Croft 2015): Stillbirth rates in some surveys seemed underestimated since they were lower than expected based on levels of early neonatal mortality. Underreporting could be about 50% for countries including full pregnancy history such as those in our sample, and larger for other surveys. The same study evaluated omissions in contraceptive use. In many surveys average contraceptive prevalence estimated from calendar data is lower than rates based on current use from earlier surveys, suggesting underreporting of contraceptive use. Again, in problematic surveys, the discrepancy tends to increase with time since the interview, so that limiting ourselves to pregnancies in the previous four years might help in that respect. Underreporting of contraceptive prevalence could lead to a bias towards zero in our estimates of the effect of contraceptive use at pregnancy since some women reporting no use might have been using.

Concerning misclassification, the biggest concern is misreporting of IA as ST. While this can certainly happen, one advantage of our data is that misclassification can be detected by an abnormal increase of reported ST among contraceptive users. Since they are at higher risk of IA, this would suggest misreporting. Note that many of the surveys in our sample belong to countries with less restrictive laws on IA, where the problem is less likely. The effect seems absent in most surveys, but it can be identified in two particular surveys, Colombia 2015 and the Kyrgyz Republic 2012. We remove those surveys from the pooled sample: While the percentage of pregnancies misclassified is not a large percent of all pregnancies, it can severely bias the estimates regarding ST since a large proportion of reported STs could be

IAs. Estimated effects in the ST equation would be contaminated and resemble the patterns for IA. It does not necessarily affect IA estimates as much since misclassification or omission are captured in a lower survey fixed-effect. The problem would be systematic misclassification or omission according to contraceptive use at pregnancy. While we cannot know if that is the case, it is more likely that misreporting is correlated with socioeconomic and cultural factors that might be captured in the model by wealth, education, and employment variables. In this respect, including these variables in the model makes the estimates of differences according to contraceptive use more robust, while blurring the interpretation of the coefficients of socioeconomic variables.

3.5 Results

Table 3 provides the descriptive characteristics of the pregnancies in our sample, together with their classification according to pregnancy outcome. We provide both unweighted and weighted counts. Sample weights have been defined to have a mean of 1 at the survey level. The outcomes are distributed in 79.1% (n=41,636) live-births, 11.9% (n=6,274) IA, and 8.9% (n=4,706) ST. Pregnancies resulting from contraceptive failure represent 11.7% (n=6,174) of the sample and only 52.3% (n=3,231) of them end in live-birth. The proportion of pregnancies ending in IA was 39.9% (n=2,465) for users compared to only 9.1% (n=4,228) for non-users. In contrast, the proportion of pregnancies ending in ST is smaller for users than non-users.

Table 3: Characteristics of pregnancies and conditional probabilities of outcomes. Probabilities estimated from the weighted sample.

	Total number		Percentage ending in			
				Termination		
	Unweighted	Weighted	Birth	Spontaneous	Induced	p-value
Sample (only included surveys)						
Pregnancies	52,616	52,616	79.1	8.9	11.9	
Surveys						
Included						
Albania 2008	1,150	1,150	83.7	8.7	7.6	< 1e-10

Table 3: Characteristics of pregnancies and conditional probabilities of outcomes. Probabilities estimated from the weighted sample. *(continued)*

	Total number		Percentage ending in			
				Termination		
	Unweighted	Weighted	Birth	Spontaneous	Induced	p-value
Armenia 2005	1,993	1,993	48.1	6.4	45.5	
Armenia 2010	1,634	1,634	62.9	6.7	30.5	
Armenia 2015	1,708	1,708	66.9	10.0	23.1	
Azerbaijan 2006	3,193	3,193	48.4	5.4	46.2	
Indonesia 2012	13,353	13,353	89.6	10.3	0.2	
Moldova 2005	1,974	1,974	55.9	10.5	33.7	
Nepal 2011	4,134	4,134	85.1	7.4	7.5	
Nepal 2016	4,130	4,130	80.5	10.4	9.1	
Philippines 2003	5,276	5,276	89.5	9.9	0.6	
Tajikistan 2012	4,220	4,220	84.3	7.8	7.9	
Tajikistan 2017	4,929	4,929	84.1	7.3	8.5	
Turkey 2003	3,714	3,714	77.1	11.2	11.7	
Ukraine 2007	1,208	1,208	66.8	6.6	26.5	
Excluded due to misclassification						
Colombia 2015	9,013	9,013	84.5	12.1	3.4	
Kyrgyz Republic 2012	3,866	3,866	77.3	9.5	13.2	
Contraceptive use						
Non-users	46,570	46,442	82.7	9.1	8.2	$< 1e\text{-}10$
Users	6,046	6,174	52.3	7.7	39.9	
Union status						
In-union	50,883	50,944	79.0	9.0	12.0	0.001
Not-in-union	1,733	1,672	84.6	7.1	8.3	
Age-group						
< 20	5,780	5,750	86.5	10.1	3.5	$< 1e\text{-}10$
20-24	17,005	17,146	85.0	7.5	7.5	
25-29	14,470	14,606	78.7	8.1	13.3	
30-34	8,780	8,684	73.7	9.0	17.3	
35-39	4,887	4,771	67.3	12.5	20.2	
40-49	1,694	1,659	60.0	16.5	23.5	
Parity						
0	17,005	17,502	88.4	9.8	1.8	$< 1e\text{-}10$
1	14,400	14,752	83.9	8.2	7.9	
2	10,229	10,115	65.9	8.2	25.8	
3	5,201	4,915	66.2	7.6	26.2	
4	2,587	2,384	69.3	9.0	21.7	
5	1,348	1,256	72.3	11.1	16.6	
6+	1,846	1,693	77.1	13.1	9.8	
Previous terminations						
0	38,741	38,796	84.5	9.3	6.2	$< 1e\text{-}10$
1	7,637	7,654	70.3	7.4	22.3	
2	2,979	2,976	59.6	9.6	30.8	
3	1,534	1,507	56.5	9.1	34.3	
4+	1,725	1,683	49.7	7.7	42.7	
Level of education						

Table 3: Characteristics of pregnancies and conditional probabilities of outcomes. Probabilities estimated from the weighted sample. *(continued)*

	Total number		Percentage ending in			
				Termination		
	Unweighted	Weighted	Birth	Spontaneous	Induced	p-value
No education	4,718	4,452	85.4	8.4	6.2	< 1e-10
Primary	10,577	11,018	84.1	10.6	5.3	
Secondary	27,835	28,001	77.0	8.3	14.7	
Higher	9,486	9,145	76.6	9.3	14.1	
Place of residence						
Urban	24,614	23,289	76.2	9.5	14.4	< 1e-10
Rural	28,002	29,327	81.5	8.5	10.0	
Currently working						
No	34,031	34,593	79.9	8.1	12.0	< 1e-10
Yes	18,585	18,023	77.7	10.6	11.7	
Wealth quintile						
Quintile 1	12,880	11,324	80.8	8.4	10.9	< 1e-10
Quintile 2	10,692	10,859	80.6	9.2	10.2	
Quintile 3	10,145	10,690	80.1	8.5	11.4	
Quintile 4	9,845	10,448	78.6	9.3	12.1	
Quintile 5	9,054	9,294	75.0	9.5	15.6	

3.5.1 Detection of misclassification

Before proceeding with the analysis, we assessed potential misclassification problems by looking at changes in the conditional probabilities of outcomes according to use. Misclassification would produce that part (or all) of the increased probability of IA for users would shift to an increased probability of ST. Figure 6 shows that the proportion of pregnancies ending in IA is higher among contraceptive users in all surveys, while in almost all surveys the probability of ST is lower for users as would be expected due to competing risks. Two exceptions stand out: Kyrgyz Republic 2012 and Colombia 2015, where the proportion of pregnancies ending in ST is 5.5 and 2.9 percentage points higher for users, respectively. In these two countries, an explanation in terms of IA reported as SA makes more sense than a significant increase in the risk of ST. As explained in section 3.4.3, we exclude these two surveys from the pooled analysis.

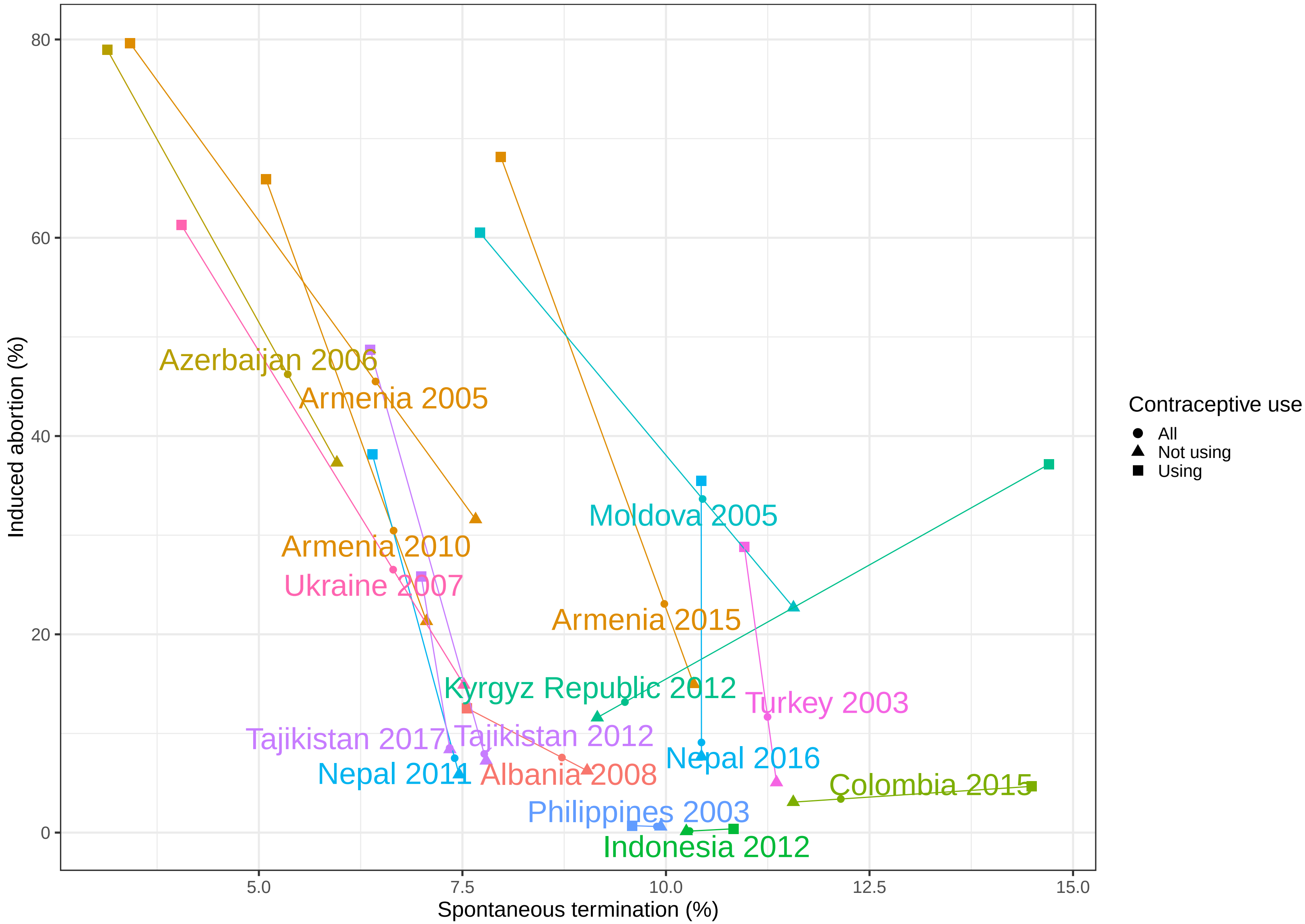

Figure 6: Percentage of pregnancies ending in termination by contraceptive use and survey.

3.5.2 Biases in binomial logistic regressions

Biases in logistic regressions not accounting for competing risks can be detected by comparing the MNL and logit coefficients for contraceptive use at pregnancy in the baseline model that only controls for recall error and survey fixed-effects. Table 4 shows the differences in the pooled sample. MNL estimates that take competing risks into account show users slightly more likely to experience ST than non-users (OR=1.24, p-value=6e-05). In the case of IA, users are much more likely to recur to IA (OR=7.2, p-value<1e-10).

Binomial logistic regression estimates that do not incorporate competing risks are biased downwards. This is particularly extreme in the case of ST, where the OR becomes lower than one (0.83, p-value=3e-04). This means that the competing risks can explain all the observed decline in the conditional probability of ST for contraceptive users. In the case of IA, the estimate is also biased downwards (OR=7.02, p-value<1e-10), but given the intensity of the effects, the order of magnitude is still similar.

Figure 7 explores the survey-specific patterns in individual survey regressions of the baseline model. Solid and hollow points indicate the OR from the MNL and binomial models, respectively. The dotted lines defining a cross around the MNL estimates indicate 95% confidence intervals. We can see that in most countries, particularly those with a higher incidence of IA, the estimates from the logistic regression have a negative bias both for IA and ST detected in the negative slope of the solid line connecting both estimates. In the case of ST, most binomial estimates are lower than one, consistent with the lower proportion of STs among contraceptive users. However, the MNL estimates are mostly higher than 1 indicating a slightly higher risk of ST for users, consistent with lower levels of care. Regarding IA, in all surveys the ORs are higher than 1 for users, corroborating the link between contraceptive failure and IA.

Table 4: Odds ratios (OR) from baseline model.

	Binomial logistic						Multinomial logistic					
	Spontaneous			Induced			Spontaneous			Induced		
	OR	95% CI	p-value	OR	95% CI	p-value	OR	95% CI	p-value	OR	95% CI	p-value
Contraceptive use												
Non-users	1.00			1.00			1.00			1.00		
Users	0.83	0.74 - 0.91	3e-04	7.02	6.49 - 7.60	<1e-10	1.24	1.12 - 1.37	6e-05	7.20	6.65 - 7.81	<1e-10
Recall error												
Per year	0.93	0.90 - 0.96	9e-06	0.94	0.91 - 0.97	2e-04	0.93	0.90 - 0.96	2e-06	0.93	0.90 - 0.96	3e-05
Fixed-effect												
Survey			<1e-10			<1e-10			<1e-10			<1e-10

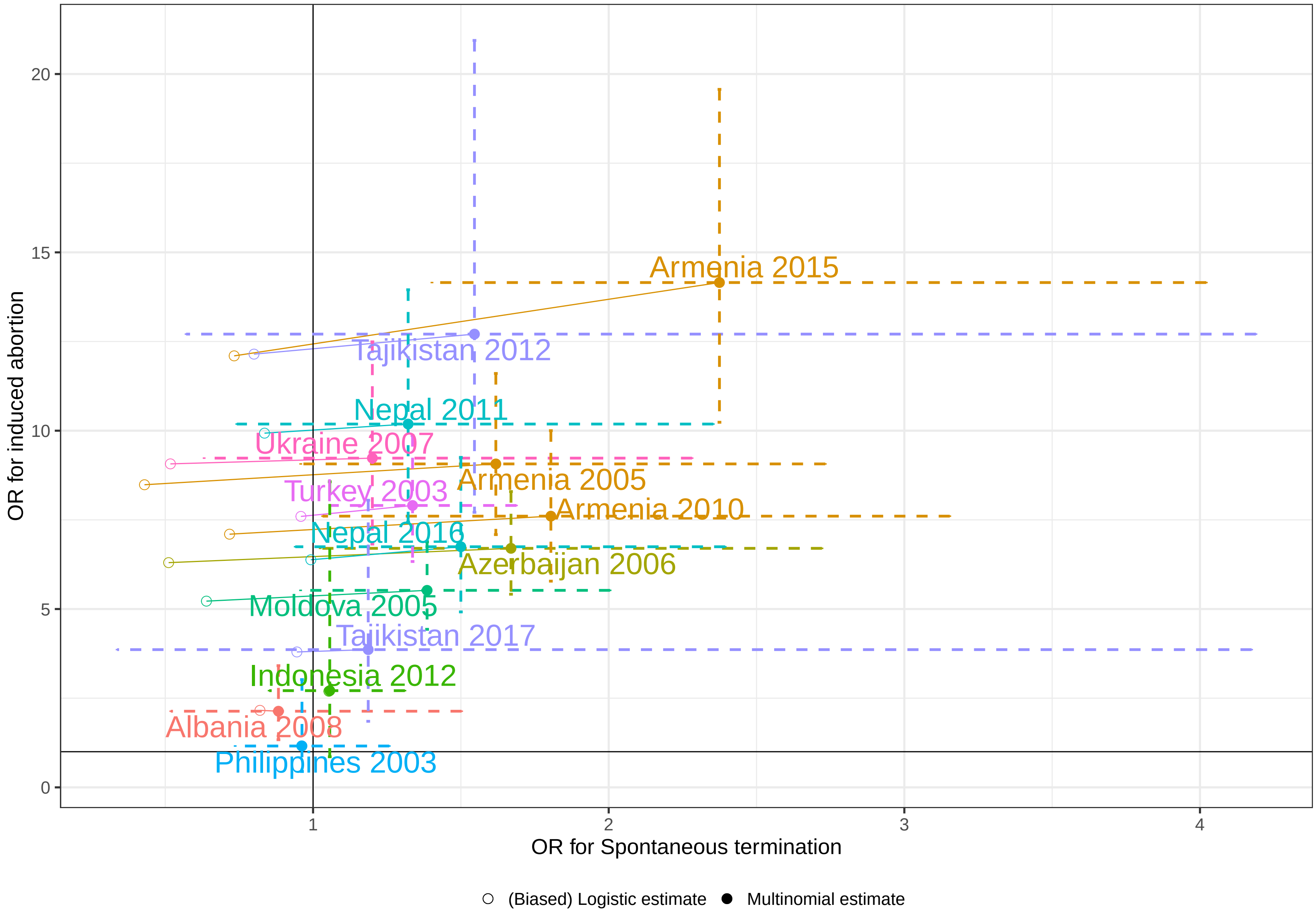

Figure 7: Survey-specific Odds Ratios (OR) for contraceptive use in the multinomial model and logistic estimates. 95% confidence intervals.

3.5.3 Multivariate models accounting for competing risk

Table 5 shows the MNL estimates for models that progressively introduce controls for demographic and socioeconomic variables into the baseline model. Given the general dependence of IA and ST on age, age-group is added in model 1. Models 2 to 4 progressively include interactions between contraceptive use at pregnancy and age, and additional controls for demographic and socioeconomic characteristics. Table 6 provides asymptotic Wald tests of joint significance for the variables on contraceptive use, age, interactions of age and use, and survey-fixed effects. All the tests show a significant effect of both contraceptive use at the time of pregnancy and age on both ST and IA. The interaction of age and use is highly significant for IA and not significant or borderline significant for ST (table 6). This is consistent with combined strategies of contraceptive use and IA by age, whereas the age-gradient for ST might be more connected to biological risk which would have liitle connection with pregnancy intentions. Note that significance tests for individual coefficients are of little interest in interaction models. They only measure the difference to the reference category, in this case `women less than 20 not using contraceptives`.

Table 5: Adjusted odds ratios (AOR) from multinomial logistic regression accounting for competing risk (Birth is the reference).

	Model 1				Model 2				Model 3				Model 4			
	Spontaneous		Induced		Spontaneous		Induced		Spontaneous		Induced		Spontaneous		Induced	
	AOR	p-value	AOR	p-value	AOR	p-value	AOR	p-value	AOR	p-value	AOR	p-value	AOR	p-value	AOR	p-value
Contraceptive use																
Non-users	1.00		1.00		1.00		1.00		1.00		1.00		1.00		1.00	
Users	1.15	0.0085	5.94	<1e-10	1.50	0.069	13.48	<1e-10	1.61	0.032	6.65	<1e-10	1.56	0.045	5.86	<1e-10
Age group																
< 20	1.00		1.00		1.00		1.00		1.00		1.00		1.00		1.00	
20-24	0.76	3.6e-07	1.57	6e-08	0.76	9e-07	1.74	3e-09	0.79	2e-05	0.68	2e-04	0.76	2e-06	0.59	5e-07
25-29	0.89	0.0248	3.39	<1e-10	0.89	0.043	3.80	<1e-10	0.97	0.568	0.65	3e-05	0.90	0.110	0.49	<1e-10
30-34	1.06	0.2962	5.65	<1e-10	1.07	0.239	6.31	<1e-10	1.20	0.010	0.70	0.001	1.09	0.219	0.46	<1e-10
35-39	1.63	< 1e-10	8.56	<1e-10	1.65	<1e-10	10.72	<1e-10	1.87	<1e-10	1.04	0.734	1.68	1e-10	0.64	3e-04
40-49	2.44	< 1e-10	13.91	<1e-10	2.65	<1e-10	18.89	<1e-10	3.08	<1e-10	1.94	7e-06	2.67	<1e-10	1.23	0.16
Use x Age-group																
Users x < 20					1.00		1.00		1.00		1.00		1.00		1.00	
Users x 20-24					0.84	0.482	0.49	1e-03	0.84	0.493	0.60	0.028	0.86	0.554	0.62	0.04
Users x 25-29					0.77	0.289	0.46	3e-04	0.75	0.244	0.68	0.092	0.77	0.282	0.72	0.16
Users x 30-34					0.76	0.278	0.46	5e-04	0.72	0.184	0.77	0.266	0.73	0.205	0.83	0.44
Users x 35-39					0.75	0.250	0.30	2e-07	0.71	0.171	0.49	0.005	0.73	0.209	0.55	0.02
Users x 40-49					0.46	0.009	0.20	8e-09	0.42	0.004	0.29	3e-05	0.44	0.006	0.31	9e-05
Union status																
Not-in-union									1.00		1.00		1.00		1.00	
In-union									1.34	0.003	0.22	<1e-10	1.34	0.003	0.22	<1e-10
Parity																
0									1.00		1.00		1.00		1.00	
1									0.77	1e-09	4.59	<1e-10	0.77	6e-09	5.17	<1e-10
2									0.89	0.023	18.57	<1e-10	0.92	0.109	25.53	<1e-10
3									0.75	4e-05	23.73	<1e-10	0.79	0.001	39.94	<1e-10
4									0.77	0.003	25.58	<1e-10	0.82	0.029	50.26	<1e-10
5									0.81	0.048	18.58	<1e-10	0.89	0.291	43.80	<1e-10
6+									0.75	0.002	13.58	<1e-10	0.84	0.074	39.78	<1e-10
Previous terminations																
0									1.00		1.00		1.00		1.00	

Table 5: Adjusted odds ratios (AOR) from multinomial logistic regression accounting for competing risk (Birth is the reference). *(continued)*

	Model 1				Model 2				Model 3				Model 4			
	Spontaneous		Induced		Spontaneous		Induced		Spontaneous		Induced		Spontaneous		Induced	
	AOR	p-value	AOR	p-value	AOR	p-value	AOR	p-value	AOR	p-value	AOR	p-value	AOR	p-value	AOR	p-value
1									0.85	0.001	2.21	<1e-10	0.84	8e-04	2.11	<1e-10
2									1.35	1e-04	2.29	<1e-10	1.34	1e-04	2.08	<1e-10
3									1.28	0.020	2.18	<1e-10	1.31	0.010	2.13	<1e-10
4+									1.11	0.374	2.30	<1e-10	1.16	0.197	2.46	<1e-10
Level of education																
No education													1.00		1.00	
Primary													1.33	4e-05	2.45	<1e-10
Secondary													1.20	0.011	4.36	<1e-10
Higher													1.14	0.126	4.10	<1e-10
Place of residence																
Urban													1.00		1.00	
Rural													0.96	0.246	0.77	1e-08
Wealth quintile																
Quintile 1													1.00		1.00	
Quintile 2													1.13	0.014	1.13	0.03
Quintile 3													1.08	0.154	1.42	4e-09
Quintile 4													1.19	0.001	1.63	<1e-10
Quintile 5													1.22	0.001	2.02	<1e-10
Currently working																
No													1.00		1.00	
Yes													1.36	<1e-10	1.64	<1e-10
Recall error																
Per year	0.93	3.9e-06	0.94	6e-04	0.93	3e-06	0.94	9e-04	0.93	4e-06	0.95	0.003	0.91	4e-08	0.93	2e-04
Fixed-effects																
Survey		< 1e-10		<1e-10		<1e-10		<1e-10		<1e-10		<1e-10		<1e-10		<1e-10

Table 6: Wald test from multinomial logistic regressions.

		Spontaneous		Induced	
Test	df	χ^2	p-value	χ^2	p-value
No contraceptive use					
Baseline model	1	16.1	6e-05	2340	< 1e-10
Model 1	1	6.9	0.008	1773	< 1e-10
Model 2	6	15.0	0.020	1845	< 1e-10
Model 3	6	21.4	0.002	1026	< 1e-10
Model 4	6	19.5	0.003	919	< 1e-10
No age					
Model 1	5	371.0	<1e-10	1642	< 1e-10
Model 2	10	381.0	<1e-10	1680	< 1e-10
Model 3	10	311.7	<1e-10	157	< 1e-10
Model 4	10	265.5	<1e-10	147	< 1e-10
No interaction use - age					
Model 2	5	8.7	0.121	49	2.8e-09
Model 3	5	11.2	0.048	30	1.5e-05
Model 4	5	10.7	0.058	28	3.7e-05
No survey fixed-effects					
Baseline model	13	130.3	<1e-10	3889	< 1e-10
Model 1	13	151.1	<1e-10	3833	< 1e-10
Model 2	13	149.0	<1e-10	3815	< 1e-10
Model 3	13	109.6	<1e-10	3503	< 1e-10
Model 4	13	117.4	<1e-10	2674	< 1e-10

Note:
df = degrees of freedom.

Controlling for age as in model 1 leads to lower estimates of contraceptive use compared to the baseline model for both ST (AOR=1.15, p-value=0.008) and IA (AOR=5.94, p-value<1e-10). Such reduction is consistent with older women being more likely to use contraceptives and more at risk of IA and ST. Age-gradients are highly significant (table 6): U-shaped in the case of ST with a minimum risk at ages 20-24, and increasing with age for IA.

Since IA and contraceptive use provide elements of a combined strategy of fertility control, the age-gradients can be different for contraceptive users and not users. Models 2 to 4 include interaction of age-groups with use. Coefficients for interacted variables in models 2 to 4 are best interpreted collectively as in figure 8 displaying the estimated age-gradients for users and not-users respectively. Model estimates (respect to 0) are shown in the main axis, with AOR in the secondary axis. Only models 2 and 4 are shown due to the similarity of models 3 and

4. Regarding IA, age-profiles are very different for users and not-users, even after controlling for demographic and socioeconomic characteristics. Interactions are highly significant in all models (table 6).

Model 2, with no controls, shows that odds increase sharply with age. In contrast, model 4, with controls, shows relatively flat U-shaped patterns with maximum levels for younger women in the case of users, and sharp U-shaped pattern with minimum levels at ages 30-34 for non-users. The large reduction in the coefficients is mostly connected to the very high effects of parity and union-status. This indicates that the large effect of age without control is due to birth-avoidance of women having 2 children or more (AOR>25 at parities 2 and above compared to nulliparous women in model 4) or not being in union (AOR=0.22 for women in-union compared to those not-in-union). The peak at ages less than 20 in model 4 indicates that these are the women likely to recur to IA at lower parities. The results for individual countries are generally consistent with this idea.

Regarding ST, while the age-patterns are much less marked than for IA, they are always highly significant. They remain as important after controlling for demographic and socio-economic characteristics. In all cases, risks of ST are minimum for women 20-24 and maximum for older women. The interaction between age and use is of borderline significance (p-value=0.058 in model 4) but suggests that the increased risk of ST for users is higher for women in the younger age groups. The effect of parity is not so important for SA, with nulliparous women having the highest risk.

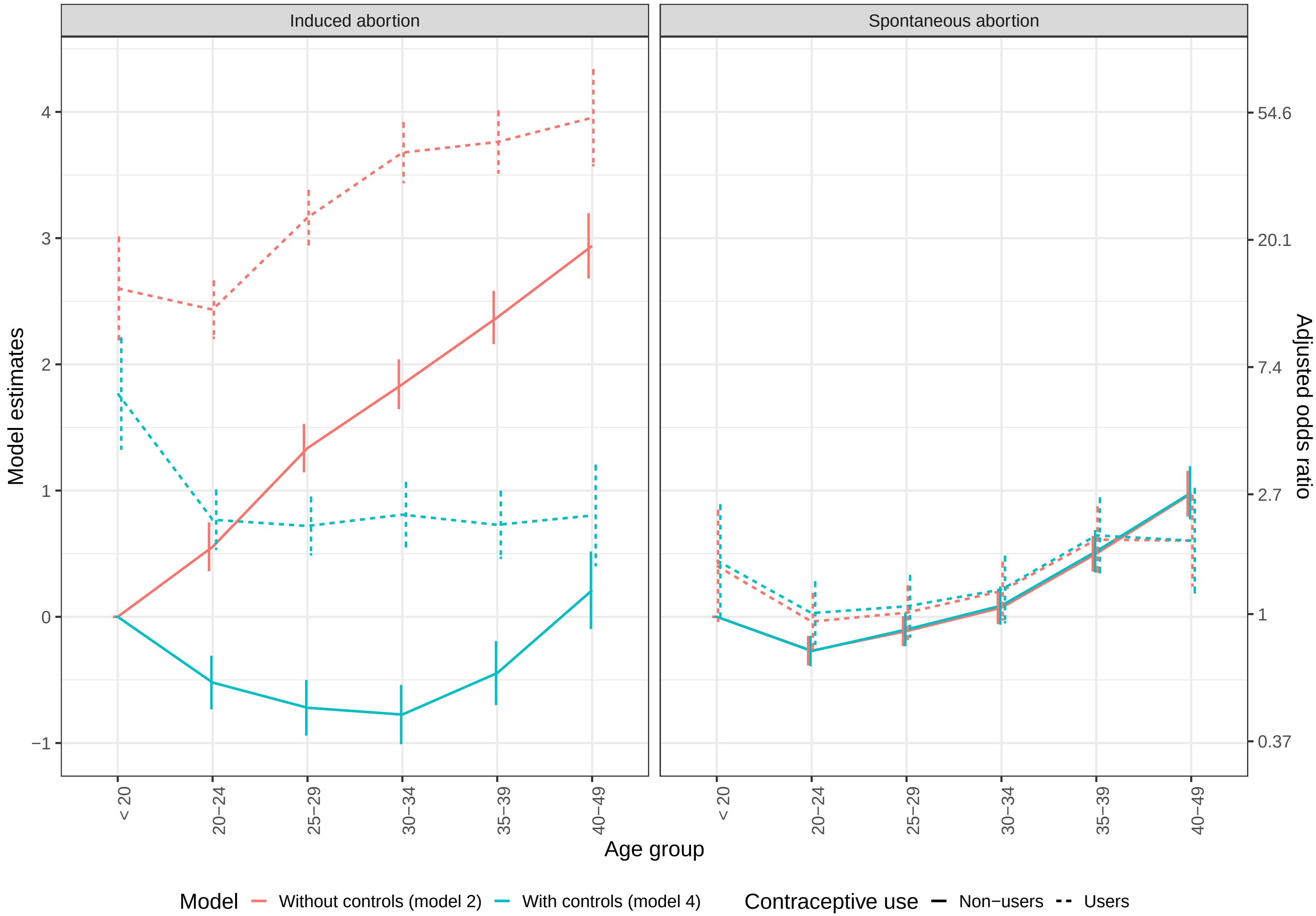

Figure 8: Estimated age-profiles of pregnancy outcomes according to contraceptive use without (model 2) and with (model 4) control for demographic and socioeconomic characteristics. Asymptotic 95% confidence intervals are displayed in vertical.

Regarding the rest of variables in table 5, previous terminations are connected to highly-significant increasing odds for IA in line with the `ready-willing-able` framework, with no apparent pattern for ST. In-union women have higher odds of ST (AOR=1.34, p-value=0.003) in contrast to the lower likelihood of IA. Regarding socioeconomic variables, the odds of ST are surprisingly lower for women with no education compared to the rest. Differences are stronger for IA, with increasing odds with education ranging from AOR of 2.45 (p-value<1e-10) for primary education to an AOR of 4.10 (p-value<1e-10) for women with higher education. Regarding the place of residence, there are no significant differences in ST (p-value=0.25), with a lower likelihood of IA in rural areas (AOR=0.77, p-value=1.5e-08). Wealth quintile patterns indicate that poorest women are more likely to report the lowest levels of both ST and IA. Again, the effects are more prominent in IA, where the AOR at quintile 5 rises to 2.02 (p-value<1e-10) than for ST, with a maximum AOR=1.22 (p-value=0.001). Women currently working are also more likely to experience both IA and ST with stronger effects for IA.

Recall error seems to be present in all cases, with estimates of comparable magnitude for ST and IA.

3.6 Discussion

This study presents original estimates of the differential odds of ST and IA according to contraceptive use at the time of pregnancy while accounting for the competing risk of pregnancy outcomes. It is the first such comparative study making use of the rich information contained in DHS calendar data.

The share of pregnancies not ending in live-birth in our sample is within the ranges reported in the literature. At the survey level, terminated pregnancies range between 10.4% and 51.9%, with the incidence of IA explaining most of the differences (Bradley, Croft, and Rutstein 2011). The incidence of ST ranges from 5.4% to 12.1% of total pregnancies, in line with previous findings (Cai and Feng 2005; Nfii 2017; Nybo-Andersen et al. 2000). In the case of IA, estimates are between 0.2% and 46.2% of total pregnancies (Sedgh et al. 2016).

Our estimates are consistent with previous findings regarding the bond between contraceptive failure and IA using a more extense empirical base. Pregnancies resulting from contraceptive failure are much more likely to end in termination, particularly an IA, but also increase the risk of ST (Bankole, Singh, and Haas 1998; Bradley, Croft, and Rutstein 2011; Marston and Cleland 2004). Even though contraceptive use has been increasing in the last decades, there is still a significant share of pregnancies considered as unintended, mostly due to contraceptive failure (Polis et al. 2016). Both access to and use of contraception are therefore first steps to avoid unwanted pregnancies. Contraceptive failure can be reduced by more efficient contraceptive use and the use of more effective contraceptive methods leading to less unintended pregnancies and fewer abortions (Bongaarts and Westoff 2000).

Regarding ST, since medical studies do not find a causal effect of contraceptive use on miscarriages (Jellesen et al. 2008; Waller et al. 2010), the higher odds of ST could be explained by women being less careful with their pregnancies when resulting from contraceptive failure (Cheng et al. 2009; Guliani, Sepehri, and Serieux 2013; Saad-Haddad et al. 2016). This includes both prenatal care and behavioral factors such as smoking, drinking, or eating patterns. These results are robust when controlling for demographic and socioeconomic variables and are also found in most individual surveys. However, as discussed in the methods section, this result is subject to bias if IAs are misclassified as STs. Excluding suspicious surveys from the pooled sample makes the estimates more robust, but we cannot exclude misclassification in the rest of surveys.

The introduction of demographic and socioeconomic variables, and in particular, separate age-gradients for users and non-users have allowed us to identify combined strategies of contraceptive use and IA in birth prevention. Whereas the age-gradient without further controls shows older women at higher risk of IA, controlling for parity and union-status suggests that parity is more important than age, with women at parities 2 and above or not-in-union, and particularly those using contraceptives, being most likely to abort. In this context, women below the age of 20 are most likely to abort conditional on parity and union status. Working women would also incur a higher opportunity cost from birth and be more likely to abort.

In the case of ST, age-gradients seem more connected to biological factors than to behavioral factors. Age-patterns remain after controlling for other variables: Women 20-24 have the lowest risk of ST. At younger ages, women using contraceptives at pregnancy are more likely to experience ST probably due to lower levels of care, but the present evidence is tentative given the possible contamination of ST coefficients in the presence of misclassification.

Regarding differentials in IA according to the rest of demographic and socioeconomic variables, they are consistent with the `ready, willing, able` framework (Lesthaeghe and Vanderhoeft 2001): more educated women living in urban areas with a higher wealth status could be more likely to recur to IA since they might have better access and be more knowledgeable regarding available options. Such results are in line with others in the literature (Westoff 2005). There is also a possibility that these variables capture differential reporting according to socioeconomic status, with more disadvantaged women less likely to admit an IA. If this is the case, controlling for these variables makes the estimates for contraceptive use more robust.

In the case of ST, while some patterns are consistent with the literature such as the higher risk for nulliparous women or women experiencing previous terminations, the patterns suggest that women in disadvantage (less educated or less wealthy) are less likely to experience ST. While studies based on good quality data show less likelihood of ST with higher socioeconomic status (Carlson, Hoem, and Rychtarikova 1999; Norsker et al. 2012; Zheng et al. 2017), a negative gradient has also been found in other retrospective surveys (Cai and Feng 2005). We have already hinted at alternative explanations: A first one, differential reporting of

ST according to socioeconomic status, with more disadvantaged women being less able to identify or remember a previous ST. A second possibility is contamination from the estimated equation for IA due to misclassification.

On the methodological side, our results confirm the importance of adequately capturing the competing risks between IA and ST (Potter, Ford, and Moots 1975). Not accounting for the competing outcomes would lead to wrongly conclude that the risk of ST is lower for women using contraceptives in contrast to the MNL estimates suggesting a higher risk. It would also underestimate the association between contraceptive failure and IA. The multinomial logit model of pregnancy outcomes conditional on pregnancy proposed here is new in the literature. It is a simple way to adequately control for the competing risks while keeping the results interpretable. It requires the assumption of independence of irrelevant alternatives (Cheng and Long 2007), something that can be defended on a-priori grounds to the extent that miscarriages and stillbirths are grouped together. Analysis of the separate determinants for miscarriages and stillbirths would require taking into account the sequential (not competing) nature of those terminations. Note also that we are focusing on the understudied topic of pregnancy outcomes according to contraceptive use at the time of pregnancy. This is only one part of the overall impact of contraceptive use on birth outcomes: since the main reason why contraceptives are used is that they make unintended pregnancies much less likely, this reduction in the probability of conception leads to a lower number of conceptions and fewer abortions. We focus on what happens in the event of a contraceptive failure. Note also that we have not explored differential patterns according to the contraceptive method used. We believe that the main impact of using more effective methods is avoiding pregnancy. Once a contraceptive failure happens, it is not so relevant why it happened but what is done about it.

Regarding data issues, our sample is biased towards countries where laws regarding IA are less restrictive. This makes recourse to IA more likely. It would be interesting to carry out a similar analysis for countries where IA is illegal or heavily restricted. That study cannot be done using the DHS as a source to the extent that only the surveys in our sample report the nature of the termination, whether it is IA or ST. For other DHS surveys, terminations are registered but not classified. Furthermore, misreporting and omission are more likely in those contexts. We have found problems of misclassification in two surveys identified by an abnormal increase of reported ST among contraceptive users. We have excluded those surveys from the pooled estimates. We have also addressed recall error and omission error by limiting our sample to the most recent 48 months of contraceptive calendar data and including a recall error covariate defined by time since the baseline period.

On the policy side, our research has implications regarding methods for estimating the impact of contraceptive use on abortion and pregnancy outcomes. A first implication is that contraceptive use and IA are dependent strategies. Therefore, methods based on independence such as the residual methods of estimating IA or the Spectrum model are not realistic (Rossier 2003; Stover and Winfrey 2017). Second, since IA and ST are competing risks, scenarios that change one probability while keeping the other constant are not realistic. That is the case of many aggregate models partly due to little evidence on ST. Our research fills a gap in that sense suggesting that ST is much less dependent on contraceptive use than IA, but that still, due to competing risks, there will be a lower probability of ST in high abortion contexts.