Positive Treatment Effects and High Heterogeneity of Hormonal Contraceptive Use on Women's Sexuality

Registered Report Stage 2 Based on a Programmatic Registered Report Stage 1

Stage 2 Submitted to *Peer Community In Registered Reports* (*PCI RR*) on March 15~~th~~th, 2024

Stage 2 First Revised Version Submitted to *PCI RR* on November 13th, 2024

Stage 1 Submitted to *PCI RR* on November 30th, 2022

Stage 1 First Revised Version Submitted to *PCI RR* on July 13th, 2023

Stage 1 Second Revised Version Submitted to *PCI RR* on September 20th, 2023

Laura J. Botzet[a,b,*], Julia M. Rohrer[c], Lars Penke[a,b], and Ruben C. Arslan[c,d]

[a]Biological Personality Psychology, University of Goettingen, Goettingen, Germany

[b]Leibniz ScienceCampus Primate Cognition, Goettingen, Germany

[c]Department of Psychology, University of Leipzig, Leipzig, Germany

[d]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

This version of the manuscript includes tracked changes compared to the clean version of the first submission of the manuscript (Registered Report Stage 2). Text in black font is from the clean version of the first submission of the manuscript, text in blue font has been added to the manuscript, and text in ~~red strikethrough font~~ has been deleted. A clean version of the first revised version of the Registered Report Stage 2 without tracked changes can be found on the Open Science Framework: https://osf.io/u8ntf/.
~~This is a clean version of the the manuscript without tracked changes compared to the Programmatic Registered Report Stage 1. A version of the manuscript with tracked changes can be found on the Open Science Framework: https://osf.io/u8ntf/.~~

*Corresponding author:

Laura J. Botzet; Biological Personality Psychology, Institute of Psychology; Goßlerstraße 14, 37073 Goettingen, Germany; Email: botzet@uni-goettingen.de

Abstract

Different women experience hormonal contraceptives differently, reporting side effects on their sexuality that range from negative to positive. But research on such causal effects of hormonal contraceptives on psychological outcomes struggles both to identify average causal effects and capture the high heterogeneity in women's treatment responses. In this study, we leveraged longitudinal data to improve our ability to separate the causal effects of hormonal contraceptives from other sources of association, including observed and unobserved confounding, reverse causality, and attrition. In this programmatic registered report (programmatic registered stage 1 protocol: https://osf.io/kj3h2; date of in-principle acceptance: 28/09/2023), we analyzed data from up to 5,041 women (23,130 observations~~current waves~~), who participated in PAIRFAM, a German longitudinal panel dataset consisting of 14 waves, using Bayesian multilevel regressions. To deal with confounding and probe the robustness of findings, we implemented two analysis approaches: adjusted regression analysis and inverse probability of treatment weighting approach. We found evidence for positive average treatment effects of hormonal contraceptives on sexual frequency and sexual satisfaction, but no robust evidence for effects on desired sexual frequency. Furthermore, to move beyond average treatment effects, we analyzed heterogeneity in treatment responses. We found relatively high heterogeneity in individual treatment effects on sexual frequency and sexual satisfaction. Interindividual differences were not systematically related to individual treatment effects, and those treatment effects did not predict women's decisions about which contraceptive method to use in the long run. Our results contribute to understanding the effects of hormonal contraceptives on sexuality in a naturalistic setting, where women adapt their choice of contraceptive method to their own experiences.

Keywords: causal inference, contraception, hormones, longitudinal analyses, sexuality

This manuscript is a stage 2 article based on a programmatic registered report stage 1 (Botzet et al., 2023). The registry for the programmatic registered report stage 1 including in-principal acceptance can be found here: https://osf.io/kj3h2. Based on the programmatic registered report stage 1 two separate stage 2 articles were planned. The current article focuses on sexuality (including desired sexual frequency, reported sexual frequency, and sexual satisfaction as outcomes) and the other stage 2 article focuses on well-being (including depressiveness, life satisfaction, and self-esteem as outcomes). All other parts of the stage 1 registered report apply for both stage 2 articles.

This manuscript contains supporting information including the supplement as well as rmd files and html files for the blind code, the simulation code, planned analyses, data wrangling, and conducted analyses online at https://osf.io/u8ntf/.

Positive Treatment Effects and High Heterogeneity of Hormonal Contraceptive Use on Women's Sexuality

Registered Report Stage 2 Based on a Programmatic Registered Report Stage 1

The impact of hormonal contraceptives on women's sexuality has been discussed since their approval in 1960. Before their invention, only so-called barrier methods existed, which prevent fertilization by blocking the union of egg and sperm (e.g., condoms, diaphragms, cervical caps, and chemical spermicides). In contrast, hormonal contraceptives (including oral hormonal contraceptives, but also hormonal implants, hormonal shots, skin patches, and vaginal rings) include synthetic hormones (progestins and sometimes synthetic estrogens) that enter the bloodstream and, in most cases, prevent ovulation (Watkins, 2012).

By altering the endocrine system, hormonal contraceptives can have effects on other aspects of the female body and brain—including negative medical and psychological side effects. For instance, two randomized controlled trials reported small negative effects of oral hormonal contraceptives on sexual desire, arousal, and pleasure (Zethraeus et al., 2016) as well as sexual interest (Lundin et al., 2018). But a recent review by Both et al. (2019) found that only a minority of women reported changes in sexual functioning and concluded that the effects of hormonal contraceptives on sexual functioning – and sexual desire in particular – are understudied and therefore poorly understood.

Experiments are considered the gold standard to answer causal research questions, such as the effects of hormonal contraceptives on sexuality. However, experimental evidence can only partly tell us how these effects affect women's everyday lives. As Graham (2019) points out, women's experiences with hormonal contraceptives are highly heterogeneous – ranging from negative side effects to no effects to positive effects. These heterogeneous responses to hormonal contraceptive use might be due to varying sensitivity to hormones (Kiesner, 2017). Such differences in sensitivity are also supported by evidence that ovulatory cycle shifts with average increases in sexual desire and self-perceived attractiveness during the fertile phase vary between women (Arslan et al., 2021; Schleifenbaum et al., 2021). Hormonal contraceptives inhibit ovulation, and so hormonal

contraceptive users no longer experience the same ovulatory cycle shifts. Heterogeneous effects of hormonal contraceptives might therefore be due to varying sensitivity to ovulatory cycle shifts before starting hormonal contraceptive use, with sensitive women showing stronger effects and insensitive women showing smaller effects on sexuality.

Such differences in the effects of hormonal contraceptives can be studied in an experimental context, as suggested by Hill and Mengelkoch (2022) who propose a precision medicine approach. They suggest researchers collect detailed information about contraceptive methods, duration of contraceptive use, mental health history, as well as sexual activity and relationship status as important potential moderators of the relationships between hormonal contraceptive use and psychological outcomes (see Box 3 and 4 in Hill & Mengelkoch, 2022).

Carefully isolated experimental settings are valuable to establish the effects of (individual) hormonal contraceptives on women's sexuality. In contrast, in everyday life, women actively choose between different non-hormonal and hormonal contraceptive methods and often try multiple methods during their lifespan. As women try to find a balance between efficacy, ease-of-use, as well as desirable and undesirable side effects, the causal effects of synthetic hormones are interwoven with confounding, attrition effects, and reverse causality. This poses unique causal inference challenges, but also allows one to investigate additional research questions such as whether side effects determine which contraceptive women eventually choose. Furthermore, the different requirements of observational data collection (as opposed to randomized clinical trials) make it easier to include a broad range of variables such as personality, thus making it possible to more thoroughly investigate potential predictors of women’s heterogeneous responses to hormonal contraceptives.

The current study aims to close the gap between the available experimental and correlational evidence about the relationship between hormonal contraceptives and women’s sexuality. By analyzing the effects of starting and discontinuing hormonal contraceptives on sexuality in a longitudinal dataset with around 5,000 women, observed over up to 14 yearly

waves[1] (years of data collection: 2008–2022), we aim to answer questions about potentially heterogeneous average treatment effects of hormonal contraceptives in real world settings while accounting for (un)observed confounders as well as attrition effects.

Empirical Evidence of Positive and Negative Effects of Hormonal Contraceptives on Sexuality

Hormonal contraceptives contain synthetic versions of progesterone (also called progestin) and sometimes estrogen, which inhibit the natural production of progesterone and estrogens as well as the natural production of pituitary hormones (luteinizing hormone and follicle-stimulating hormone). This reduction of natural hormonal fluctuation across the menstrual cycle prevents the maturation of the ovarian follicle and therefore hinders ovulation (Frye, 2006). In general, women who are using hormonal contraceptives have lower levels of estradiol, progesterone, follicle stimulating hormones, luteinizing hormones, and total and free testosterone as well as higher levels of sex-binding globulins (Gaspard et al., 1983; Zethraeus et al., 2017; Zimmerman et al., 2014). Their endogenous hormone levels remain constantly similar to those found in the early follicular phase of normally cycling women (Mishell et al., 1972).

This intervention into the endocrine system (Stomati et al., 1998) has been hypothesized to negatively affect women's sexuality (Both et al., 2019). Some empirical evidence supports these hypotheses regarding sexual functioning (e.g., Læssøe et al., 2014) and libido (e.g., Lee et al., 2017; Lundin et al., 2018; Zethraeus et al., 2016), as well as sexual activity, arousal, pleasure, orgasm, and lubrication (Smith et al., 2014). While hypotheses and evidence for negative side effects of hormonal contraception exist, the use of hormonal contraception has also been hypothesized to positively affect women's sexuality through several mechanisms, including, for example, overcoming the fear of unwanted pregnancy during sexual activity (Blumenstock & Barber, 2022) and the resolution of painful

[1] In stage 1 of this registered report we only mentioned 13 waves. The 14th wave was released on July 31st, 2023, and we decided to include all available information up to date in our analyses.

or troublesome gynecologic disorders (Both et al., 2019). Empirical evidence in support of positive effects on sexuality has been reported concerning sexual functioning (e.g., Oranratanaphan & Taneepanichskul, 2006), libido (McCoy & Matyas, 1996), and, most strongly, sexual frequency (e.g., Caruso et al., 2005; McCoy & Matyas, 1996). In addition, women using hormonal contraceptives reported higher sexual satisfaction (e.g., Caruso et al., 2005) and higher relationship satisfaction (e.g., Taggart et al., 2018).

Obstacles to Estimating Psychological Effects of Hormonal Contraceptives

Taken together, evidence concerning potential psychological effects of hormonal contraceptives remains inconclusive. While randomized-controlled trials provide somewhat consistent evidence of small negative average treatment effects on various aspects of women's sexuality (e.g., Graham et al., 1995; Lundin et al., 2018; Sabatini & Cagiano, 2006; Zethraeus et al., 2016; but see Oranratanaphan & Taneepanichskul (2006) and Strufaldi et al. (2010) for evidence of positive causal effects of certain methods of hormonal contraception), evidence based on correlational data often shows no or even positive relationships between the use of hormonal contraceptives and sexuality (e.g., Caruso et al., 2005; McCoy & Matyas, 1996; but see Wallwiener et al. (2010, 2015) for evidence of a negative relationship). Some reviews about potential effects of hormonal contraceptives conclude that there are negative effects of hormonal contraceptives (Lee et al., 2017) or no effects of hormonal contraceptives (Pastor et al., 2013). However, most reviews conclude that the effects of hormonal contraceptives on sexuality have not been well studied and remain controversial (Both et al., 2019; Burrows et al., 2012; Davis & Castaño, 2004; Schaffir, 2006).

Several explanations for this mixed and inconclusive body of evidence are plausible:

(1) *Contraceptive method and dosage effects*: differing psychological responses are due to differences between hormonal contraceptives (e.g., application methods or different dosages of synthetic progesterone and estrogen; for

supporting empirical evidence see e.g., Boozalis et al., 2016; Læssøe et al., 2014; Sabatini & Cagiano, 2006; Strufaldi et al., 2010)

(2) *Treatment heterogeneity*: differing psychological responses are due to interindividual differences between women (Graham, 2019) and studies systematically vary in sampling procedures (e.g., some only including women with a regular ovulatory cycle)

(3) *Treatment heterogeneity leading to selective attrition:* women who experience negative effects of hormonal contraceptives discontinue them, leaving only women who experience no effects or positive effects in the group of hormonal contraceptive users in correlational studies

(4) *Confounders*: pre-existing differences in women influence the decision what contraceptive method to use *and* affect psychological outcomes, leading to differences between the groups of hormonal contraceptive users and non-hormonal contraceptive users in correlational studies

(5) *Reverse causality*: in some cross-sectional studies, relationships between psychological outcomes and hormonal contraceptive use might occur because the outcome influences the contraceptive choice (e.g., higher frequency of vaginal intercourse might lead to the decision to start using hormonal contraceptives).

Randomized controlled trials with a placebo control group are regarded as the superior approach for estimating the average treatment effect of hormonal contraceptives and their contraceptive efficacy. They can also expand the knowledge about (1) *contraceptive method and dosage effects* and (2) *treatment heterogeneity*. While the estimated effects will not be biased through (4) *confounders* and (5) *reverse causality* as their impacts are nullified by randomization, this also means that the design cannot inform us about the extent to which these two affect correlations between contraceptive usage and outcomes in everyday life. Furthermore, this design is not optimized to inform us about how (3) *treatment heterogeneity might lead to selective attrition* in everyday life. A related

concern is sometimes termed *healthy user bias*: the women who volunteer for a randomized controlled trial will not include, for example, women who, based on previous experience, fear bouts of severe depression if they are assigned to hormonal contraception. By randomly assigning different forms of contraceptives to women, they remove the decision process to start or to discontinue using contraceptives that is inherent to real world settings. In addition, owing to their cost, randomized controlled trials usually have small sample sizes that preclude the rigorous investigation of subgroups, heterogeneity, and uncommon side effects. Finally, trials with a non-hormonal contraceptive control group are uncommon, in part because pharmaceutical trials tend to focus on comparing different formulations and in part because many non-hormonal methods are less efficacious, increasing the risk of unplanned pregnancies. For example, in the randomized trial with a non-hormonal contraceptive control group by Zethraeus et al. (2016, 2017) women were blinded and did not know whether they were using hormonal contraceptives. To avoid unwanted pregnancies, all women were instructed to use additional non-hormonal contraceptive methods during the study and received free condoms (Zethraeus et al., 2017). Therefore, any beneficial effects resulting from knowing that one is using a highly effective birth control method (Both et al., 2019) may be underestimated in such blinded randomized controlled trials.

Observational Cross-Sectional and Longitudinal Designs

In comparison to randomized-controlled trials, observational cross-sectional designs also capture any association induced by the decision process. Therefore, (3) *selective attrition*, (4) *confounders*, and (5) *reverse causality* will often bias the estimated effects. At the same time, they are usually based on larger sample sizes and include users of multiple contraceptive methods as well as those who use no contraceptive method at all. They operate like photographs of the real world. While they only show patterns at one specific time point, they still provide important pieces of the picture (such as the associations between demographic variables and contraceptive method) that could not be obtained based on randomized controlled trials alone. Going beyond mere associations, we can at least

attempt to infer causal effects from cross-sectional data, if we are willing to transparently discuss and defend the necessary strong assumptions and statistical adjustments (e.g., Botzet et al., 2021).

One way to reduce the number of assumptions necessary for causal identification in observational data is examining change over time within individuals, because many of the potential confounding factors that vary between individuals are held constant by design. Longitudinal designs can rule out between-subject confounders by allowing the use of within-subject analyses (Rohrer & Murayama, 2021). Therefore, time-invariant confounders can be ruled out when estimating causal effects based on appropriately specified longitudinal designs.

Such panel studies operate like a series of photos:[2] We can track change, but still have to be cautious not to confuse cause and effect, since multiple events can occur in the interim—a longitudinal design alone is no guarantee of appropriate causal inference. Still, given transparent assumptions and adequate statistical control, we can at least attempt to infer causal effects. Specific statistical models are needed to remove confounders (Hamaker et al., 2015) and all modeling decisions ultimately reflect assumptions about the underlying causal network (Rohrer & Lucas, 2020).

Given the correct modeling decisions, time-invariant confounders are automatically controlled for in longitudinal designs. As they do not vary within a woman, they will not induce spurious correlations between her time-varying predictor and her time-varying outcome. Time-varying confounders on the other hand are not automatically controlled by longitudinal designs, but instead need to be accounted for (Rohrer & Murayama, 2021). A time-varying confounder might affect a woman's choice of contraceptive method as well as the outcome of interest at a given time. For example, an ineffable or at least unmeasured shift from a casual to a steadier exclusive relationship may affect the decision to use hormonal contraceptives. In addition, this shift could cause more frequent sexual activity at a

[2] Going a step further, by analogy to movies, we could do even better by having more granular, potentially daily longitudinal data on contraception, which would, for example, allow us to explicitly model the effects of the menstrual cycle.

later time. In a longitudinal design that only measures hormonal contraceptive use and sexual activity but not this relationship shift, it will appear like there is a positive causal effect of hormonal contraceptives on sexual activity.

Some of these time-varying confounders might not have been observed in the available dataset or might even be completely unobservable – they thus cannot be accounted for in the statistical analysis. Such unobserved confounders bias the estimate no matter what analytic strategy is used, which we analyzed in our simulations reported in the supplement. However, additional sensitivity analyses can be conducted to estimate the influence unobserved confounders would need to have to fully account for the remaining observed relationship between treatment and outcome, thus providing at least the opportunity to make an educated guess about the internal validity of the results (for early work on sensitivity analysis for unobserved confounders see Rosenbaum & Rubin, 1983).

Longitudinal designs investigating potential medical effects of hormonal contraception are relatively common (e.g., Eng et al., 2008; Riggs et al., 2007; Wang et al., 2016), although all of these studies implement randomized treatment assignment rather than an observational approach. To our knowledge, only two studies investigated effects of hormonal contraception on sexuality with an observational longitudinal design. Blumenstock and Barber (2022) analyzed data from a weekly survey over 2.5 years from 893 women. They showed that women had a higher sexual frequency when they were using hormonal contraceptives. Frequency of sexual intercourse increased after starting using hormonal contraception, remained high for several months, and then slowly declined. Ott et al. (2008) showed in a 41-month long study with 328 participants that sexual interest based on daily diaries did not change when women started using oral contraceptives. But when women stopped using oral contraceptives, sexual interest decreased.

To summarize, causal inference from longitudinal data is only possible on the basis of assumptions. We strive to make our analysis goal (Lundberg et al., 2021) and the assumptions underlying our causal identification strategy as transparent as possible. In addition, we apply two different analytical approaches with different underlying assumptions.

Heterogeneity in Treatment Responses

While evidence for a negative average treatment effect on sexuality based on randomized controlled trials exist (Zethraeus et al., 2016), self-reports by women indicate that individual treatment effects on sexuality might vary widely (Malmborg et al., 2016). Heterogeneity in treatment responses might be caused by individual differences in responses to steroids (Kiesner, 2017). To our knowledge, treatment heterogeneity of hormonal contraceptives on sexuality has not been estimated quantitatively. Based on longitudinal data analyses, individual treatment effects on sexuality for each woman can be estimated and the distribution of individual treatment effects and their uncertainty can be visualized.

Estimating individual treatment effects allowed us to answer further questions about the underlying causal network connecting hormonal contraceptives and sexuality. Is there a large number of women who experience either positive or negative effects? Do women use their own experience with individual effects of hormonal contraceptives on sexuality to make a decision about their contraceptive method? For example, are women who experience adverse effects of hormonal contraceptives on sexuality more likely to stop using them during a specific time span? In addition, we want to answer the question whether interindividual differences like demography and personality predict individual treatment effects. Older women might be more likely to experience beneficial side effects of hormonal contraceptives on sexuality because they found the method that fits them best. In line with this reasoning, empirical findings suggest that higher age was associated with less negative side effects of hormonal contraceptive use on depression with particularly strong negative effects during adolescents (Skovlund et al., 2016). Nevertheless, these findings might be accounted for by other explanations, e.g., a possible decrease in sensitivity to steroid hormones with age or a specifically strong sensitivity to steroid hormones during puberty. Women with higher scores on openness might be more likely to experience beneficial side effects as well because they are more likely to try out different contraceptive methods until they find their perfect method. Other personality dimensions might be related to negative or

positive individual treatment effects. For example, women with higher scores on neuroticism may experience more positive psychological effects as their heightened worries about unwanted pregnancies are reduced.

Focusing on individual treatment effects of hormonal contraceptives on sexuality allowed us to broaden our understanding about the individual nature of potential effects of hormonal contraceptives as well as confounding and attrition effects.

The Current Study

In the current study we aimed to answer the questions whether hormonal contraceptive use influences women's sexuality (over and above attrition effects, accounting for observed and unobserved confounders) as well as whether and to which extent the effects of hormonal contraceptives on sexuality vary between users. Outcomes included desired sexual frequency in the last three months as a measure for libido, reported sexual frequency in the last three months, and sexual satisfaction. By using a longitudinal design, we can partly rule out alternative explanations such as reverse causality. Analyses were based on the German Family Panel (PAIRFAM), a panel dataset containing information about contraceptive use and women's sexuality from more than 5,000 women over 14 waves, starting in 2008 (Brüderl et al., 2021; Huinink et al., 2011).

Conceptual Design and Underlying Assumptions

The conceptual design of the study, including all underlying assumptions, is outlined in Figure 1. These two graphs correspond to the two analytical approaches that were used to estimate the causal effect of hormonal contraceptives on the three outcomes.

The graph in panel A shows the adjusted regression approach, which estimates the effect of contraceptive method on the outcome while controlling for the respective outcome in the previous wave, contraceptive method in the previous wave, and their interaction in~~at~~ the previous wave, as well as potential observed time-varying confounders (i.e.,

demography, relationship information). In addition, the potential influence of unobserved (and unobservable) confounders was estimated.

The graph in panel B corresponds to the conceptual design underlying the inverse probability of treatment weighting approach (IPTW; Thoemmes & Ong, 2016). For this approach, individuals were weighted by their probability to receive a specific treatment, in our case hormonal contraceptive use. This weight for each individual is modeled with effects of the outcome in the previous wave, contraceptive method in the previous wave, and their interaction in~~at~~ the previous wave, as well as potential observed time-varying confounders (i.e., demography, relationship information) on the treatment itself (i.e., hormonal contraceptive use). When estimating the effect of hormonal contraceptives on the respective outcome this weight was taken into account.

Why implement two approaches instead of only one line of analyses? According to Thoemmes and Ong (2016), the adjusted regression approach has several disadvantages: (1) regressions with different numbers of covariates can be estimated easily and therefore may introduce biases through cherry-picking (Rubin, 2001); (2) the adjusted regression approach relies on the untested key assumption that the relationships between the covariates and the outcome are modeled appropriately (more narrowly described as the linearity assumption, see Gutman & Rubin, 2017); (3) any comparisons between the treated and the untreated group might be due to extrapolation because there are no treated participants who are comparable to the untreated participants (King & Zeng, 2006).

While we agree that the IPTW approach outperforms adjusted regression analysis in estimating the causal effect of a treatment on an outcome in many possible scenarios (Fuentes et al., 2021), the first two disadvantages of adjusted regression mentioned above can also apply to the IPTW approach: (1) models estimating the individual weights *are* regression models that can be performed as easily with a different number of covariates and therefore potential bias through cherry-picking is not meaningfully precluded, and (2) the IPTW approach relies on the untested key assumption that the relationships between the

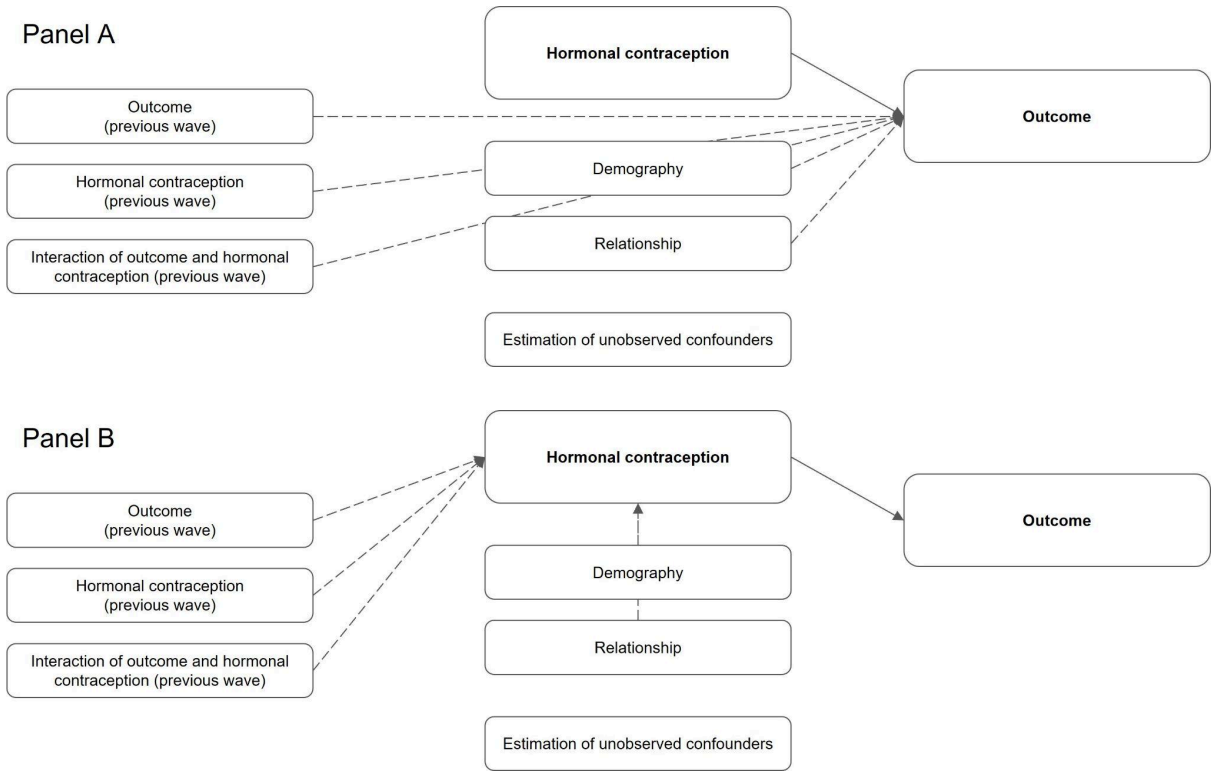
covariates and the treatment are modeled appropriately (as opposed to the relationships between covariates and outcome, see assumptions of adjusted regression approach).

We addressed the first concern of both approaches (introduction of bias through cherry-picking) by carefully laying out the assumed underlying causal network and preregistering our models in form of a registered report before having access to the data. To address the second concern (nonlinearity between covariates and outcome or treatment, respectively), we decided to perform and compare both approaches to estimate the causal effect of hormonal contraceptives on the outcome robustly under different sets of assumptions. Nevertheless, both approaches still rely on the assumptions of (1) no unobserved confounders; (2) positivity (i.e., every individual having a probability of receiving the treatment that is larger than 0 and smaller than 1); and (3) a correct specification of the underlying models (Thoemmes & Ong, 2016). To estimate the dependency of our analyses on these three underlying assumptions, we tested the proposed models with different specifications based on simulated data with varying data generating mechanisms. The models, simulations, and results are described in more detail in the supplement. Given our interest in the immediate effects of hormonal contraceptive use (rather than the lagged effects after one year), and to avoid adding superfluous complexity, we decided against a popular alternative modeling approach (RI-CLPM, Hamaker et al., 2015) which simultaneously attempts to estimate causal effects pointing into the opposite direction.

In addition, we estimated the potential influence of *unobserved* confounders on the average treatment effect. We ran additional sensitivity analysis to estimate how sensitive the results are to hidden bias. Although a sensitivity analysis does not compensate for unobserved confounding, it quantifies how large the hidden bias would need to be to change the conclusions substantially (see methods section).

Figure 1

Conceptual design of the analyses approaches.



Note. Panel A shows the conceptual design and assumptions underlying the adjusted regression model. Panel B shows the conceptual design and assumptions underlying the inverse probability of treatment weighted (IPTW) regression model.

Estimands

In the context of this study, we were not interested in assessing dichotomous hypotheses (i.e., whether an effect of contraception on sexuality does or does not exist), but rather in estimating the magnitude and heterogeneity of a range of effects of interest. Thus, instead of formulating hypotheses, we identified clear analysis goals and theoretical estimands, defined estimation strategies, and specified the corresponding empirical estimands (Lundberg et al., 2021). By precisely defining all target quantities, estimands connect theory with statistical evidence. The study design template in Table 1 ~~is~~ based on the template provided by *Peer Community In Registered Report*

(https://rr.peercommunityin.org/help/guide_for_authors), therefore includes theoretical estimands and empirical estimands instead of hypotheses.

First of all, we were interested in overall descriptive patterns, including the percentage of hormonal contraceptive users across observations~~waves~~ and common patterns in use and switches of hormonal contraceptives. Based on the full sample of all eligible women participating in PAIRFAM, we examined descriptives and general trends over the course of the study.

Second, we were interested in why women choose hormonal contraceptive methods. To get a better understanding of potential causes, we investigated whether time-varying covariates predicted the choice of contraceptive method. This was based on the IPTW model as this approach explicitly models how likely women are to use hormonal contraceptive methods. Our empirical estimands were quantified as percentage points based on marginal effects.

Third, we wanted to estimate the average treatment effect of hormonal contraceptive use on all three outcomes. Therefore, adjusted as well as IPTW regression models were performed to estimate the causal effect, taking into account observed confounders. In addition, the sensitivity of the models to unobserved confounders was estimated. Our empirical estimand was the unstandardized mean difference in the outcome between non-hormonal and hormonal contraceptive use. For the reported sexual frequency outcome, this difference can be seen as a very rough approximation of the percentage change in sexual frequency.[3]

In addition, we were interested in treatment heterogeneity. Therefore, we investigated individual treatment effects on the outcome based on the adjusted regression models (see section Analysis for an explanation why we did not investigate treatment effect heterogeneity in the context of IPTW regression models).To help interpret this quantity, we visualized the

[3] This is the case because the response scale of this item is very roughly a log-transformed version of frequency, e.g., on the response scale, the difference between 2 = *once per month and less* and 4 = *once per week* is as large as 4 = *once per week* and 6 = *more than three times a week*. For the full response scale see Table S1.

distribution and uncertainty of individual estimates and reported for how many women we estimated negative and positive effects.

Furthermore, we wanted to explore the correlation between individual treatment effects and age as well as the correlations between individual treatment effects and Big Five personality traits. While these analyses were less focused on causal identification, they might still provide tentative evidence for substantively plausible causal hypotheses.

In addition, we wanted to investigate whether women's individual treatment effects on sexuality informed their decision of which contraceptive method to use by investigating the correlation between estimated individual treatment effects and the number of years using hormonal contraceptives during the course of PAIRFAM. Ideally, we would have sufficient data to instead estimate individual treatment effects (e.g., using all but the last observation of each participant~~wave of data~~) to predict individual behavior (e.g., contraceptive method in the very last observation of each participant~~wave of data~~). However, in the context of the available data, this would result in very low statistical power, and we thus decided on a different approach which would only provide very rough evidence for potential assortment based on experiences with contraceptive methods. Such an assortment based on experiences would result in the type of selective attrition explained above and may provide a partial explanation for the mixed evidence concerning effects of hormonal contraceptives on sexuality.

Table 1

Study design

Theoretical estimand	Quantification of empirical estimand	Sampling plan	Analysis plan / Estimation	Interpretation given different outcomes
Descriptive patterns in hormonal contraceptive use	Percentages of hormonal contraceptive users Probability to switch between hormonal and non-hormonal contraceptive use Average number of switches	All available data from PAIRFAM across 14 waves	Descriptive analyses	—
"Confounding" effects on hormonal contraceptive use	Percentage points based on marginal effects	up to n = 5,041 women with a mean average of 4.59 observations~~waves~~ → 2,169 women reported using both hormonal contraceptives and non-hormonal contraceptives at some point while participating in PAIRFAM → approximately 2,716 switches between contraceptive methods	Linear binomial regression with hormonal contraceptive method as a dichotomous outcome and all treatment predictors as predictors (same model is used for the weights of the inverse probability of treatment weighting approach)	—
Average treatment effects of hormonal contraceptive use on sexuality	Unstandardized mean difference between non-hormonal and hormonal contraceptive use		Adjusted regression analysis Inverse probability of treatment weighting approach	If outcomes based on the two estimations differ, adjusted regression analysis was treated as the main analysis and the inverse probability of treatment weighting approach was treated as a robustness analysis for identifying the average treatment effect
Heterogeneity in treatment effects of hormonal contraceptive use on sexuality	Percentage of women with negative estimated effects and positive estimated effects		Extracted individual treatment effects from adjusted regression analysis	—
Link between individual treatment effects and predictors of individual treatment effects as well as contraceptive decision	Correlations between individual estimated treatment effects and age, personality traits, as well as years spent on hormonal contraceptives		Extracted individual treatment effects correlated with age, personality traits, as well as years spent on hormonal contraceptives ~~weighted by inverse standard error~~	—

Note. This table is adapted based on the study design template provided by *Peer Community In Registered Report* here:

https://rr.peercommunityin.org/help/guide_for_authors. PAIRFAM = German Family Panel (Brüderl et al., 2021; Huinink et al., 2011).

Methods

Data

Analyses were based on data from a German panel study called PAIRFAM (Brüderl et al., 2021; Huinink et al., 2011). It contains information about contraceptive use and women's sexuality from more than 5,000 women. The longitudinal design consists of annual waves with the first data collection in 2008 and the latest available data from 2022 (wave 14). The ethics committee of the Faculty of Management, Economics, and Social Sciences of the University of Cologne approved PAIRFAM. Huinink et al. (2011) provide a detailed description of the PAIRFAM dataset. In addition, the present manuscript contains supporting information including rmd files and html files for the blind code, the simulation code, planned analyses, data wrangling, and conducted analyses online at https://osf.io/u8ntf/.

The data on which our analyses are based were already available and can be used for scientific purposes; the Leibniz Institute for the Social Sciences (GESIS) grants access to the scientific community. Only one of the authors had previously accessed the PAIRFAM data; JMR was granted access to Release 7.0 (waves 1-7) in 2016 within the context of a different research project but never actually worked with the data beyond an initial screening of the included variables to determine suitability for her research question (birth order effects on personality). Thus, some of the data used to answer this research question had been previously downloaded by one of the authors, but we certify that we have not observed any part of the data relevant to the present research question (Level 3 bias control based on the categorization in Table 1 by *Peer Community In Registered Report;* https://rr.peercommunityin.org/help/guide_for_authors).

Exclusion Process and Participants

We excluded individuals who did not identify as female. Furthermore, once a woman crossed the age of 50 or reported to be (post-)menopausal, her data (including subsequent waves) were excluded, but previous waves of data collection remained in the analysis. In addition, we excluded all individual waves of data in which participants indicated being in a

homosexual relationship or only reported homosexual relationships in the past, were pregnant, trying to become pregnant, gave birth to a child in the last year, were currently breastfeeding, or indicated using the morning-after-pill or an unknown contraceptive method. Besides these exclusion criteria explicitly mentioned in stage 1 of this registered report, the implemented models for effects of hormonal contraception on sexuality only used information from one wave if the information from the previous wave was also available (because we used predictors from the previous waves in our models). In addition, we could only include information from participants if (1) hormonal contraception was available in the previous and the current wave (2) all additional predictors were available in the current wave and (3) one of the three outcomes was available in the previous and the current wave. To make these data exclusion steps following from modeling decisions more explicit, we now list them as additional exclusion criteria in stage 2 of this registered report.

In further separate robustness analyses, we additionally excluded waves in which participants indicated that they are sterilized, as well as all subsequent waves of those participants. We also excluded all waves in which participants indicated that their partner is sterilized and all waves in which women indicated using no contraceptive method[4], an intrauterine device as a contraceptive method, or hormonal methods other than the oral contraceptive pill. In addition, we excluded all waves in which women indicated that they had never been sexually active. All exclusion criteria, reasons for exclusion, and excluded observations~~unit(s)~~ are summarized in Table 2.

In addition to these robustness analyses, which focused on excluding specific women or waves that might bias the estimates of the originally registered main analysis, we would have liked to conduct exploratory subanalyses based *only* on women who reported being in a homosexual relationship or who have reported only homosexual relationships in the past

[4] In PAIRFAM, women who indicated that they had never been sexually active in their life were not asked about their contraceptive method. These women were coded as using no contraceptive method, i.e., a non-hormonal method (see the section about the variables for more information). This coding may introduce some errors as some women may use hormonal methods without being sexually active; we thus exclude them in an additional robustness analysis to ensure that this coding decision does not systematically affect results.

(otherwise using the same exclusion criteria as in the originally registered main analysis). While we hoped to gain some initial insight into the potential effects of hormonal contraceptives on sexuality of homosexual women, the sample size after applying our registered exclusion criteria was already too small to perform any meaningful analysis (*n* = 188, observations~~waves~~ = 539). We had the information from the previous and the current wave that was necessary to perform the proposed models for only 62 exclusively homosexual women (213 observations~~waves~~). Of these, only three women reported a switch from non-hormonal to hormonal contraception and no women reported a switch from hormonal to non-hormonal contraception (compared to our registered threshold of 200 homosexual women reporting a switch between hormonal and non-hormonal contraception at least once). As the threshold of a sufficient number of exclusively homosexual women was not met, we did not perform the additional exploratory subanalyses based on this sample.

Table 2

Exclusion criteria, reasons for exclusion, and excluded observations~~waves~~

Originally r~~R~~egistered main analysis				
Exclusion criteria	**Reasons for exclusion**	**Excluded wave(s)~~unit(s)~~**	$n_{excluded}$	**observations~~waves~~$_{e}$** $_{xcluded}$
identifying as non-female	potential hormonal influences	current and all subsequent waves	9,157	49,283
older than 50 years	potential hormonal influences	current and all subsequent waves	3	3
(post-)menopausal	potential hormonal influences	current and all subsequent waves	161	509
only homosexual relationships	no need to use contraceptives to prevent pregnancy	current wave	330	1,046
pregnant	potential hormonal influences	current wave	2,345	2,898
trying to become pregnant	no need to use contraceptives to prevent pregnancy	current wave	2,712	5,026
gave birth in the last year	potential hormonal influences	current wave	3,121	4,250
breastfeeding	potential hormonal influences	current wave	1,507	2,287

Table 2 (continued)

Exclusion criteria, reasons for exclusion, and excluded observations~~waves~~

Originally r~~R~~egistered main analysis (continued)				
Exclusion criteria	**Reasons for exclusion**	**Excluded wave(s)~~unit(s)~~**	**$n_{excluded}$**	**observations~~waves~~$_{e}$ $_{xcluded}$**
using the morning-after-pill as a contraceptive method	potential hormonal influences	current wave	559	687
using an unknown contraceptive method	not possible to classify method as hormonal or non-hormonal	current wave	607	953
previous wave completely missing	data required to fit model missing	current wave	18,912	23,032
hormonal contraception missing (previous or current wave)	data required to fit model missing	current wave	12,648	66,073
other predictors missing (current wave)	data required to fit model missing	current wave	17,204	27,448
all outcomes missing (previous or current wave)	data required to fit model missing	current wave	18,912	33,218
Further robustness analyses				
Exclusion criteria	**Reasons for exclusion**	**Excluded wave(s)~~unit(s)~~**	**$n_{excluded}$**	**observations~~waves~~$_{e}$ $_{xcluded}$**
sterilized	no need to use contraceptives to prevent pregnancy	current and all subsequent waves	327	1,762
partner sterilized	no need to use contraceptives to prevent pregnancy	current wave	430	1,594
using no contraceptive method	imprecise classification as non-hormonal in originally registered main analysis	current wave	8,375	24,398
using an intrauterine device as a contraceptive method	imprecise classification as non-hormonal in originally registered main analysis	current wave	2,019	6,712
using other hormonal methods	investigate effects of oral contraceptive pills only	current wave	1,021	2,243
never sexually active	potentially conditioning on sexual frequency as an outcome	current wave	4,248	10,274

Variables

All variables, including the predictor variable, potential time-varying confounders, outcome variables, and variables used to investigate treatment heterogeneity are listed in Table S1 in the supplement. The original German item wording can be found here: https://www.pairfam.de/dokumentation/fragebogen/.

The predictor hormonal contraception was based on the items about the contraceptive method; participants were able to report multiple contraceptive methods. Hormonal contraception was coded as 0 if participants indicated that they use no contraceptive method at all. The variable hormonal contraception was also coded as 0 if participants indicated that they use no hormonal contraceptive method and at least one of the following methods: *condom*; *intrauterine device*;[5] *diaphragm, foam, suppository, gel*; *natural birth control*; *female sterilization*; *male sterilization*; or *withdrawal method, coitus interruptus*. In addition, the variable hormonal contraception was coded as 0 if participants were never sexually active in their life, as these participants were not asked about their contraceptive method. The variable hormonal contraception was coded as 1 if participants indicated that they use a *birth control pill*, *mini pill*, or *other hormonal method (implant, patch, ring)*, even if they additionally use non-hormonal methods. Exclusion criteria for originally registered main as well as robustness analyses based on the contraceptive method are described above.

Simulation

In order to contrast our different analytical approaches, we compared the performance of our models (conceptually summarized in Figure 1) on data simulated under different data generating mechanisms. These simulations are discussed in more detail in the

[5] Participants were not asked whether they used a hormonal or copper intrauterine device. Therefore, we coded the choice *intrauterine device* as hormonal if participants had indicated earlier in the survey that they use *other hormonal method (implant, patch, ring)*, assuming that women who use a hormonal intrauterine device would classify this as another hormonal method after the option *birth control pill, mini pill*. If participants only indicated that they use an *intrauterine device* but no hormonal method, this was coded as non-hormonal contraception.

supplement and the implications are mentioned in the section Analysis. Based on these simulation results, to estimate the average treatment effects and treatment effect heterogeneity, we performed adjusted regression analysis without accounting for systematic missingness. In addition, we estimated the average treatment effects based on the IPTW approach accounting for systematic missingness.

Analysis

To answer the question whether hormonal contraceptive use influences women's sexuality, and to separate these potential causal effects from confounders and attrition effects, we used two different analytical approaches, as outlined in Figure 1. This decision was based on simulations contrasting our different analytical approaches and comparing the performance of our models on data simulated under different data generating mechanisms (for more details on the simulation see the supplement). Based on these simulation results, to estimate the average treatment effects and treatment effect heterogeneity, we performed adjusted regression analysis without accounting for systematic missingness. In addition, we estimated the average treatment effects based on the IPTW approach with stabilized, truncated weights at 1% (Thoemmes & Ong, 2016) accounting for systematic missingness. All planned analyses can be found in form of an rmd file and an html file: https://osf.io/u8ntf/.

All Bayesian models included a random intercept and a random slope for hormonal contraceptive use nested within participants. In addition, each model included information from the previous wave about the outcome, hormonal contraception, and their interaction as predictors.[6] In order to be able to estimate the causal relationship between the hormonal contraception and the outcome, we controlled for individual mean levels of hormonal contraceptive use across observations~~waves~~ (see Bafumi & Gelman, 2007, and Hamaker &

[6] We decided to include the interaction term in the IPTW approach to model the possibility that certain outcomes might have stronger effects in hormonal contraceptive users than in non-hormonal contraceptive users on the contraceptive choice (e.g., strong negative side effects on sexuality might be more likely to be attributed to the contraceptive choice in hormonal contraceptive users leading to the decision to stop using this method). To keep both approaches parallel, we also included this interaction term in the adjusted regression analysis.

Muthén, 2020, for further information); this approach effectively controls for stable confounding influences that work between women (time-invariant confounders). For both models, potential time-varying confounders included linear effects for log transformed net income, educational attainment, and fertility plans; a thin-plate spline effect (Wood, 2003) for age; and a categorical effect for number of children (no children, one child, two children, three or more children).

Furthermore, relationship duration was included as a nested variable. This allowed us to model a linear association with relationship duration which is only informed by women who are in a relationship, while simultaneously including those who are not in the analysis. Technically, we achieved this by including a dummy coded variable for current relationship status (single vs. non-single) and its interaction with log transformed relationship duration as a predictor. No main effect of relationship duration was included in the model. Relationship duration for singles was set to -1; this value is arbitrary and does not affect the resulting estimates because when multiplied with the relationship status dummy, relationship duration for singles is dropped from the analysis. In addition, we included two dummy coded variables: one indicating whether a woman started a relationship between the previous wave and the current wave and one indicating whether a woman became single between the previous wave and the current wave.

In the IPTW approach the outcome in the first model was the contraceptive method. The first model results in an estimated weight which was then included in the second model. In the IPTW approach, the effects were additionally weighted for systematic missingness based on weights provided by PAIRFAM.[7] Separate analysis for desired sexual frequency, reported sexual frequency, and sexual satisfaction as outcomes were performed. All included variables are listed in Table S1.

To answer the question whether interindividual differences predicted individual treatment effects, we extracted individual treatment effect estimates from the adjusted

[7] We used the calibration weights which adjust for differences between the population and the sample on the following characteristics: gender, federal state, education level, migration background, settlement structure, family status, number of children in household.

regression analysis and subsequently correlated them with age ~~(continuous)~~ and the Big Five personality traits at draw level. ~~These correlation analyses were weighted by the inverse of the standard error of the individual treatment effect estimates to propagate uncertainties in their estimation.~~ Finally, these correlations were averaged across draws (we call this approach "correlate, then average across draws"; Ly et al. (2017) for example use the term "plausible values").[8]

To answer the question whether women guide their contraceptive method choices by deciding against hormonal contraceptive methods after experiencing adverse effects, we again used individual treatment effect estimates from the adjusted regression analysis, this time correlating them with the proportion of years using hormonal contraceptives (observations~~waves~~ in which hormonal contraceptives were used divided by total number of observations~~waves participating in PAIRFAM~~) at draw level. ~~This correlation analysis was again weighted by the inverse of the standard error of the individual treatment effect estimates.~~ Finally, these correlations were averaged across draws (we call this approach "correlate, then average across draws"; Ly et al. (2017) for example use the term "plausible values").[9] This analysis can potentially provide tentative evidence for assortment based on experiences with contraceptive methods.

Additionally, given the possibility of unobserved confounding, we ran sensitivity analysis to estimate how sensitive our results are to hidden bias. We calculated E-values for the effect of hormonal contraception on all outcomes (VanderWeele & Ding, 2017). As VanderWeele and Ding (2017) write "The E-value is defined as the minimum strength of

---

[8] In stage 1 of this registered report, we planned to first average the individual treatment effects across draws, then compute correlations with age and the Big Five personality traits weighted by the inverse of the standard error of the individual treatment effect estimates. This planned approach, here referred to as "average across draws, then correlate", overestimates the shared variance of individual treatment effects and interindividual differences. Therefore, we instead decided to correlate, then average across draws (sometimes called "plausible values"). We report the originally planned approach in Table S6 in the supplement.

[9] In stage 1 of this registered report, we planned to first average the individual treatment effects across draws, then compute correlations with the proportion of years using hormonal contraceptives weighted by the inverse of the standard error of the individual treatment effect estimates. This planned approach, here referred to as "average across draws, then correlate", overestimates the shared variance of individual treatment effects and interindividual differences. Therefore, we instead decided to correlate, then average across draws (sometimes called "plausible values"). We report the originally planned approach in Table S6 in the supplement.

association, on the risk ratio scale, that an unobserved confounder would need to have with both the treatment and the outcome to fully explain away a specific treatment-outcome association, conditional on the measured covariates.“ A large E-value implies that unobserved confounding would need to be relatively substantial to explain away an effect. Conversely, a small E-value implies that even just a little unobserved confounding would be able to explain away the estimated effect. E-values are one of the few approaches to unobserved confounding that can be applied to longitudinal designs (VanderWeele et al., 2020).

All planned analyses can be found in the form of an rmd file and an html file: https://osf.io/u8ntf/. In addition, all models are outlined in the supplement using a simplified readable notation.

Model Convergence

Using the default settings for *brms*, a lot of the models had divergent transitions (between 6 and 240 diverging transitions). Therefore, we increased the sampler's target acceptance rate during Stan's adaptation period in these models from brms's default 0.80 to 0.99 (Stan Development Team, 2024). With this increased average proposal acceptance probability, none of the originally registered main models had divergent transitions, except for the adjusted regression analysis with sexual frequency as an outcome (4 divergent transitions), which we considered harmless. In all originally registered main models, the Rhats were at most 1.01, indicating that the 4 chains of the models converged.[10]

[10] For the robustness models with the increased average proposal acceptance probability, 4 of the 36 performed models had divergent transitions (between 1 and 3 divergent transitions). In all robustness models except one, the Rhats were at most 1.04, indicating that the 4 chains of the models converged. For the adjusted regression analysis with desired sexual frequency as an outcome in the main analysis~~robustness analysis 6~~, the~~one~~ Rhat for the~~a~~ standard deviation of the random effect of hormonal contraceptive use was 1.08, indicating potential convergence issues.

Deviations From Stage 1 Concerning the Analysis Sample

In contrast to our registered analysis plan, we decided to report the results from robustness analysis 6 (additional exclusion of women who had never been sexually active) as our *main analysis*. All results based on the original sample for the *originally registered main analysis* are additionally presented in the supplement. We did so because we realized that including women who had never been sexually active likely induces bias. This particularly applies to the IPTW approach and less so to the adjusted regression analysis, with the overall consequence that including these women makes results more dependent on modeling choices (i.e., means of covariate adjustment) that should not make a substantive difference.

In the originally registered main analysis (and all other robustness analyses), women who had never been sexually active were classified as non-hormonal contraceptive users (even though they were not asked about their contraceptive method) and their sexual frequency was 0 (*never*). This coding decision appears to systematically bias results, because women who are not sexually active over a longer period of time are incorporated in the analyses as non-hormonal contraceptive users (because by study design, they are not asked questions about contraception usage) with the lowest score for sexual frequency (nobody else was assigned this score). The adjusted regression analysis seems to partly account for this bias by controlling for hormonal contraception in~~at~~ the previous wave, the linear effect of sexual frequency from the previous wave (which was 0 = *never* as well), and their interaction on the sexual frequency in the current wave. In contrast, in the IPTW approach the control for hormonal contraception in~~at~~ the previous wave, the linear effect of sexual frequency from the previous wave, and their interaction on hormonal contraceptive use in the current wave seems to be unable to account for this bias, presumably because the effect of sexual frequency from the previous wave was modeled as a linear effect. Thus, when including these women, estimates from the IPTW approach were about twice as large as estimated based on the adjusted regression analysis (see the section Robustness of Results for more information). When excluding these women, results are much less

dependent on the choice of the data-analytic approach and thus more robust. The new exclusion implies that we should not generalize our conclusions to women who start being sexually active and/or start using hormonal contraception.

For women who are being sexually active for the first time, several factors may make their experience different from that of women who have been sexually active for a longer period of time. Initially, these women may be in the process of exploring and understanding their sexual preferences, desires, and boundaries, which may influence their sexual frequency and satisfaction. These factors may result in a unique pattern of sexual behavior that differs from those with more established sexual routines and experiences. Women who begin using hormonal contraceptives without being sexually active may experience different dynamics than those who are already sexually active. For these women, the decision to use hormonal contraception may be driven by reasons other than birth control, such as managing menstrual cycles, reducing menstrual pain, or addressing hormonal imbalances. Without the context of sexual activity, their experience with contraception may be primarily focused on managing these health concerns.

Results

Description of Samples

In PAIRFAM, 18,912 individuals provided a total of~~with information from~~ 104,661 observations (i.e., annual assessments of individuals)~~waves participated~~. After applying our predefined exclusion criteria and excluding observations~~waves~~ that could not be incorporated into the analyses because of modeling decisions, up to ~~*n* = 5,041 women and~~ 23,130 observations~~current waves~~ from *n* = 5,041 women were eligible for our main analysis (registered as robustness analysis 6) estimating effects of hormonal contraception on sexuality.[11] Due to further missingness in single outcomes, the sample sizes for our main

[11] The sample sizes for the models estimating the predictors of hormonal contraceptive use were slightly larger because the sexuality variables (sexual frequency, desired sexual frequency, sexual satisfaction) were only required in the current but not in the previous wave (leading to a total sample size of up to *n* = 5,133 women and 23,615 observations~~23,165 current waves~~; for the model including sexual frequency as a predictor: *n* = 4,564 women and 20,501 observations~~current waves~~; for ~~for~~ the

analysis vary between outcomes~~estimating effects of hormonal contraception on sexuality fluctuated between analyses focusing on different outcomes~~. In~~For~~ the main analysis estimating the effect of hormonal contraception on sexual frequency 4,403 women with~~and~~ 19,801 observations~~current waves~~ were included (sexual frequency was only measured in waves 2 to 14); in~~for~~ the main analysis estimating the effect of hormonal contraception on desired sexual frequency 3,057 women with~~and~~ 9,939 observations~~current waves~~ were included (desired sexual frequency was only measured in waves 7 to 14); in~~for~~ the main analysis estimating the effect of hormonal contraception on sexual satisfaction 4,983 women with~~and~~ 22,622 observations~~current waves~~ were included. The full exclusion process for the main analysis is outlined in Figure 2. The total eligible sample sizes for estimating effects of hormonal contraception on sexuality for all samples are displayed in Table 3.

---

model including desired sexual frequency as a predictor: *n* = 3,224 women and 10,468 observations~~current waves~~; for ~~for~~ the model including sexual satisfaction as a predictor: *n* = 5,101 women and 23,266 observations~~current waves~~).

Figure 2

Exclusion process for main analysis (registered as robustness analysis 6) estimating effects of hormonal contraception on sexuality

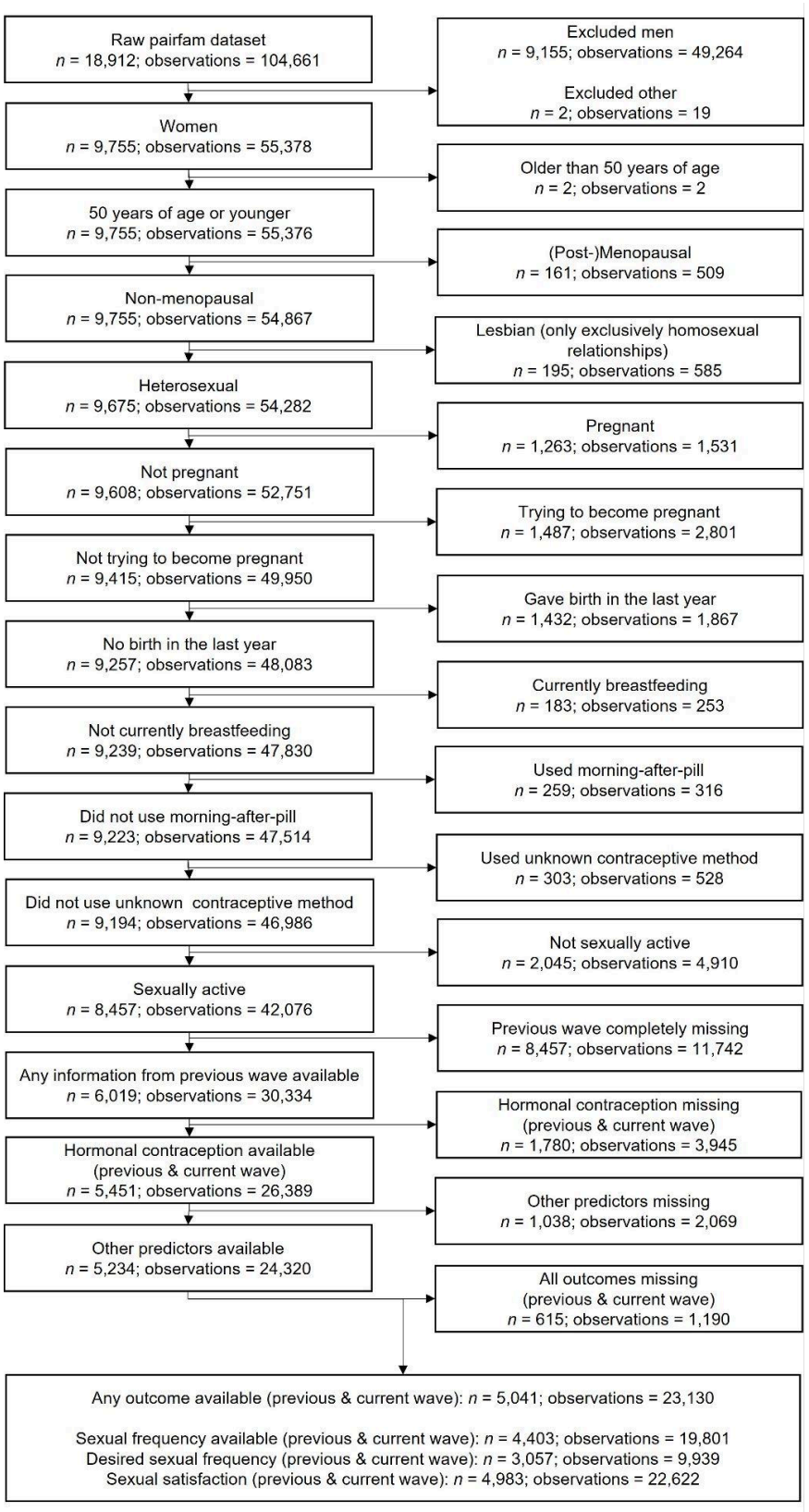


Table 3

Total eligible sample size for estimating effects of hormonal contraception on sexuality for all samples

Sample	Additional exclusion criteria	*n*	observationswaves
Main analysis (registered as robustness analysis 6)	all waves in which women indicated that they had never been sexually active	5,041	23,130
Originally rRegistered main analysis	—	5,684	25,891
Robustness analysis 1	all waves in which women indicated that they are sterilized (and all subsequent waves)	5,586	25,201
Robustness analysis 2	all waves in which women indicated that their partner is sterilized	5,618	25,159
Robustness analysis 3	all waves in which women indicated that they used no contraceptive method	5,114	20,870
Robustness analysis 4	all waves in which women indicated that they used an intrauterine device as contraceptive method	5,358	23,231
Robustness analysis 5	all waves in which women indicated that they used hormonal methods other than the oral contraceptive pill	5,601	25,023

Note. For all samples women and waves were excluded according to the section Exclusion Process and Participants.

Descriptives

Overall, after applying all exclusion criteria *n* = 5,041 women with~~and~~ 23,130 observations~~current waves~~ were eligible for our main analysis estimating effects of hormonal contraception on sexuality.[12] Each woman was observed between 1 and 13 times (the first wave was always excluded because no information from the previous wave was available); with an average of 4.59 observations (standard deviation of 3.46).[13] ~~We had up to 4.59 current eligible waves on average per woman (standard deviation of 3.46), ranging from 1 to 13 available waves (the first wave of each woman had to be excluded from all analyses because no information from the previous wave was available).~~ Table 4 displays the number of waves (for all categorical variables) or means, standard deviations, and ranges (for all continuous variables) for all included women and all observations~~current waves~~.

Table 5 shows zero-order correlations of all variables included in the same~~main~~ analysis ~~estimating effects of hormonal contraception on sexuality (except the categorical variable number of children)~~ averaged within women across all available observations~~current waves~~. Hormonal contraception measured in the current wave correlated positively with hormonal contraception measured in the previous wave (*r* = .86 [95% CI: .86,~~;~~ .87]). As is to be expected, b~~B~~oth measures (hormonal contraception ~~measured~~ in the current wave, ~~and~~ hormonal contraception ~~measured~~ in the previous wave) correlated positively with the average frequency of using hormonal contraception (*r* = .95 [.94,~~;~~ .95] and *r* = .95 [.94,~~;~~ .95], respectively).

Hormonal contraception ~~measured~~ in the current wave correlated negatively with income, years of education, relationship duration, completed fertility plans, and age (ordered from weakest to strongest negative correlation, *r*s ranging from -.11 [-.14,~~;~~ -.08] to -.50 [-.52,~~;~~ -.48]), as well as positively with the end of a relationship in the last year (*r* = .11 [.09,~~;~~ .14]) and the start of a relationship in the last year (*r* = .15 [.12,~~;~~ .17]).

---

[12] The sample sizes for the models estimating the predictors of hormonal contraceptive use were slightly larger because the sexuality outcomes (sexual frequency, desired sexual frequency, sexual satisfaction) were only necessary to be available in the current, but not in the previous wave (see section Exclusion Process and Participants for more information).

[13] These numbers vary slightly between analyses due to missing outcome data.

Each sexuality variable measured in the current wave correlated positively with its value in the previous wave (*r*s ranging from .71 [.70,~~;~~ .72] to .82 [.81,~~;~~ .83]). When measured in the current wave, sexual frequency correlated positively with sexual satisfaction (*r* = .57 [.55,~~;~~ .59]), while desired sexual frequency correlated negatively with sexual frequency (*r* = -.35 [-.38,~~;~~ -.32]) and sexual satisfaction (*r* = -.34 [-.37,~~;~~ -.31]). Hormonal contraception in the current wave correlated positively with sexual frequency in the current wave (*r* = .22 [.19,~~;~~ .24]) and sexual satisfaction in the current wave (*r* = .16 [.13,~~;~~ .19]), but did not correlate with desired sexual frequency in the current wave (*r* = -.02 [-.05,~~;~~ .01]). For the remaining covariates measured in the current wave, being in a relationship showed the strongest and most consistent correlations with sexual frequency (*r* = .44 [.42,~~;~~ .47]), ~~and~~ sexual satisfaction (*r* = .26 [.24,~~;~~ .29]), and desired sexual frequency (*r* = -.27 [-.30,~~;~~ -.24]).

Table 4

Number of observations~~waves~~ (categorical variables) or means, standard deviations, and ranges (continuous variables) for all included variables across all women and all observations~~current waves~~

Variable	Number of observations ~~waves~~	mean	sd	min	max
Hormonal contraception in current wave					
- 0 = no	12,321 (53%)				
- 1 = yes	10,809 (47%)				
Average frequency of using hormonal contraception (0 = only non-hormonal contraception, 1 = only hormonal contraception)		0.47	0.40	0	1
Hormonal contraception in previous wave					
- 0 = no	11,869 (51%)				
- 1 = yes	11,261 (49%)				
Age (in years)		32.90	9.08	16	50
Net income (in Euros)		1007.00	1174.00	0	80,000
Education (in years)		13.374~~0~~	2.77	1	20
Relationship status					
- 0 = single	5,024 (22%)				
- 1 = in a relationship	18,106 (78%)				
Relationship duration based on 4,441 women and 18,106 observations~~waves~~		10.374~~0~~	8.32	0	36.83~~0~~

(in years)	
Start of relationship in the last year	
- 0 = no	21,648 (94%)
- 1 = yes	1,482 (6%)
End of relationship in the last year	
- 0 = no	21,814 (94%)
- 1 = yes	1,316 (6%)

Table 4 (continued)

Number of observations~~waves~~ (categorical variables) or means, standard deviations, and ranges (continuous variables) for all included variables across all women and all observations~~current waves~~

Variable	Number of observations ~~waves~~	mean	sd	min	max
Number of Children					
- no children	11,017 (48%)				
- 1 child	3,654 (16%)				
- 2 children	5,745 (25%)				
- 3 or more children	2,714 (12%)				
Completed fertility plans					
- 0 = no	11,040 (48%)				
- 1 = yes	12,090 (52%)				
Sexual frequency in current wave based on 4,904 women and 22,346 observations~~waves~~ (on a scale from 0[a] to 7)		3.32	1.52	1	7
Sexual frequency in previous wave based on 4,498 women and 20,241 observations~~waves~~ (on a scale from 0[a] to 7)		3.32	1.52	1	7
Desired sexual frequency in current wave based on 3,396 women and 12,193 observations~~waves~~ (on a scale from 1 to 5)		3.58	0.81	1	5
Desired sexual frequency in previous wave based on 3,188 women and 10,379 observations~~waves~~ (on a scale from 1 to 5)		3.58	0.81	1	5
Sexual satisfaction in current wave based on 5,022 women and 22,855 observations~~waves~~ (on a scale from 0 to 10)		6.22	2.64	0	10
Sexual satisfaction in previous wave based on 5,007 women and 22,827 observations~~waves~~ (on a scale from 0 to 10)		6.29	2.65	0	10
Big Five personality extraversion based		3.63	0.81	1	5

on 3,426 women and 5,474 observations~~waves~~ (4 items on a scale from 1 to 5)				
Big Five personality agreeableness based on 3,425 women and 5,473 observations~~waves~~ (4 items on a scale from 1 to 5)	3.32	0.72	1	5
Big Five personality conscientiousness based on 3,426 women and 5,476 observations~~waves~~ (4 items on a scale from 1 to 5)	3.90	0.60	1	5
Big Five personality neuroticism based on 3,425 women and 5,474 observations~~waves~~ (4 items on a scale from 1 to 5)	2.84	0.82	1	5
Big Five personality openness based on 3,426 women and 5,475 observations~~waves~~ (4 items on a scale from 1 to 5)	3.68	0.69	1	5

Table 4 (continued)

Number of observations~~waves~~ (categorical variables) or means, standard deviations, and ranges (continuous variables) for all included variables across all women and all observations~~current waves~~

Note. [a] All waves in which women indicated that they had never been sexually active (i.e., sexual frequency was 0 = *never*) were excluded from the main analysis.

All variables were measured in all 14 waves, except for reported sexual frequency (measured in waves 2 to 14), desired sexual frequency (measured in waves 7 to 14), and the Big Five personality measures (measured in waves 2, 6, 10, and for the refreshment sample in wave 11). Where no numbers of women and observations~~waves~~ are given in the "Variable" column, the information is based on the full sample for the main analysis (registered as robustness analysis 6; *n* = 5,041 women and 23,130 observations~~waves~~).

Table 5

Zero-order correlations of all included variables averaged within women across all available observations~~current waves~~

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) HC (cw)																
(2) Average frequency of using HC	.95															
(3) HC (pw)	.86	.95														
(4) Age	-.50	-.51	-.49													
(5) Net income	-.11	-.11	-.10	.33												
(6) Education	-.12	-.12	-.11	.23	.43											
(7) Relationship status	-.02	-.03	-.03	.21	.07	.08										
(8) Relationship duration; 4,441 women; 18,106 observations~~waves~~	-.37	-.38	-.37	.70	.16	.12	.43									
(9) Start of relationship last year	.15	.12	.10	-.26	-.08	-.09	-.03	-.41								
(10) End of relationship last year	.11	.12	.13	-.24	-.11	-.09	-.44	-.35	.13							
(11) Completed fertility plans	-.41	-.43	-.42	.78	.16	.03	.20	.63	-.22	-.18						
(12) Sexual frequency (cw); 4,904 women; 22,346 observations~~waves~~	.22	.20	.18	-.21	-.08	-.10	.44	-.17	.18	-.07	-.14					
(13) Sexual frequency (pw); 4,498 women; 20,241 observations~~waves~~	.18	.18	.18	-.19	-.08	-.09	.39	-.12	-.03	.08	-.12	.82				
(14) Desired sexual frequency (cw); 3,396 women; 12,193 observations~~waves~~	-.02	-.02	-.005	-.06	.03	.06	-.27	-.14	.01	.09	-.11	-.35	-.29			
(15) Desired sexual frequency (pw); 3,188 women; 10,379 observations~~waves~~	-.02	-.02	-.02	-.05	.03	.07	-.25	-.14	.09	.07	-.09	-.29	-.35	.81		
(16) Sexual satisfaction (cw); 5,022 women; 22,855 observations~~waves~~	.16	.15	.14	-.19	-.09	-.08	.26	-.10	.12	-.02	-.13	.57	.49	-.34	-.30	
(17) Sexual satisfaction (pw); 5,007 women; 22,827 observations~~waves~~	.15	.16	.17	-.20	-.10	-.09	.24	-.09	-.02	.06	-.13	.49	.57	-.25	-.30	.71

Table 5 (continued)

Zero-order correlations of all included variables averaged within women across all available observations~~current waves~~

Note. HC = Hormonal contraception; cw = current wave; pw = previous wave. If no information about the wave is available, the variable was measured in the current wave. Continuous variables (average frequency of using hormonal contraception, age, net income, education, relationship duration, sexual frequency, desired sexual frequency, and sexual satisfaction) are averaged within women across observations~~waves~~. Binary variables (hormonal contraception, relationship status, start of relationship in the last year, end of relationship in the last year, and completed fertility) are coded as either 0 or 1 and then averaged within women across observations~~waves~~. The categorical predictor number of kids is not included in this table. Where no numbers of women and observations~~waves~~ are given in the "Variable" column, the information is based on the full sample for the main analysis (registered as robustness analysis 6; n = 5,041 women and 23,130 observations~~current waves~~). ~~If no information about the wave is available, the variable was measured in the current wave.~~

Descriptive Patterns in Hormonal Contraceptive Use

Of the 23,130 observations of~~current waves reported by reported by~~ 5,041 women included in our main analysis, 12,321 (53%) were observations~~waves~~ in which women reported using non-hormonal contraceptive methods and 10,809 (47%) were observations~~waves~~ in which women reported using hormonal contraceptive methods. Over the whole available information (including previous waves), 1,537 women (30%) reported using only non-hormonal contraceptives, 1,335 women (26%) reported using only hormonal contraceptives, and 2,169 (43%) women reported using non-hormonal and hormonal contraceptives at least once. The 2,169 women who used both non-hormonal and hormonal contraceptives reported 2,716 switches (total number of switches across all 5,041 women: *mean* = 0.54, *sd* = 0.93, *min* = 0, *max* = 7) with 1,584 switches from hormonal to non-hormonal and 1,132 switches from non-hormonal to hormonal contraceptives.

Predictors of Hormonal Contraceptive Use

Figure 3 displays the three separate models predicting~~estimating the effects on~~ hormonal contraceptive use in the current wave from, among other predictors,~~controlling (besides other predictors) for~~ a) sexual frequency in~~at~~ the previous wave and its interaction with hormonal contraceptive use in~~at~~ the previous wave; b) desired sexual frequency in~~at~~ the previous wave and its interaction with hormonal contraceptive use in~~at~~ the previous wave; and c) sexual satisfaction in~~at~~ the previous wave and its interaction with hormonal contraceptive use in~~at~~ the previous wave.

Stability of Contraceptive Method Use

In all three models the effect of the average frequency of using hormonal contraception was significant (model a): unstandardized regression coefficient *b* = 6.25 [6.02,~~;~~ 6.49]; model b): *b* = 7.18 [6.70,~~;~~ 7.68]; model c): *b* = 6.02 [5.82,~~;~~ 6.24]). Using average marginal effects, this corresponds to an increase around 4% in the probability to use hormonal contraception, when women's average frequency of using hormonal contraception

increases by 10% (model a): 4.16% [4.04%,; 4.28%]; model b): 4.30% [4.10%,; 4.51%]; model c): 4.11% [4.00%,; 4.22%]). In addition, hormonal contraception in~~at~~ the previous wave was a significant predictor in all three models (model a): *b* = 0.73 [0.48,; 0.98]; model b): *b* = 1.58 [0.89,; 2.30]; model c): *b* = 1.09 [0.84,; 1.34]). This corresponds to an increase of around 15% in the probability to use hormonal contraception for women who used hormonal contraception in~~at~~ the previous wave compared to women who used non-hormonal contraception (model a): 15.30% [13.80%,; 16.80%]; model b): 14.30% [12.20%,; 16.50%]; model c): 15.40% [14.00%,; 16.80%]).

Sexuality

The use of hormonal contraception in~~at~~ the current wave was negatively predicted by sexual frequency in~~at~~ the previous wave (model a): *b* = -0.13 [-0.19,; -0.07]) and its interaction with hormonal contraceptive use in~~at~~ the previous wave (model a): *b* = 0.27 [0.20,; 0.34]). If a woman used non-hormonal contraceptives in~~at~~ the previous wave, for each increase of 1 on the scale for sexual frequency this corresponds to a decrease of -0.65% [-0.95%,; -0.35%] in the probability to use hormonal contraception in the current wave. If a woman used hormonal contraceptives in~~at~~ the previous wave, for each increase of 1 on the scale for sexual frequency this corresponds to an increase of 1.17% [0.75%,; 1.59%] in the probability to use hormonal contraception in the current wave. Similarly, the use of hormonal contraception in~~at~~ the current wave was negatively predicted by sexual satisfaction in~~at~~ the previous wave (model c): *b* = -0.04 [-0.07,; -0.02]) and its interaction with hormonal contraceptive use in~~at~~ the previous wave (model c): *b* = 0.08 [0.04,; 0.12]). If a woman used non-hormonal contraceptives in~~at~~ the previous wave, each increase of 1 on the scale for sexual satisfaction corresponds to a decrease of -0.25% [-0.40%,; -0.09%] in the probability to use hormonal contraception in the current wave. If a woman used hormonal contraceptives in~~at~~ the previous wave, each increase of 1 on the scale for sexual satisfaction corresponds to an increase of 0.31% [0.10%,; 0.51%] in the probability to use hormonal contraception in the current wave. In other words, with increased sexual frequency and

increased sexual satisfaction, the model predicted that women were more likely to stick with the contraceptive method reported in the previous wave.

Neither the effect of desired sexual frequency in~~at~~ the previous wave (model b): *b* = -0.03 [-0.18,~~;~~ 0.13]) nor its interaction with hormonal contraceptive use in~~at~~ the previous was a significant predictor of hormonal contraceptive use in~~at~~ the current wave (model b): *b* = 0.03 [-0.17,~~;~~ 0.21]).

Relationship and Family Situation

The probability of using hormonal contraceptives was significantly predicted by relationship status in models a) and c) (model a): *b* = 0.32 [0.07,~~;~~ 0.58]; model c): *b* = 0.37 [0.15,~~;~~ 0.60]; but not in model b): *b* = 0.01 [-0.29,~~;~~ 0.34]) corresponding to an increase around 6% in the probability to use hormonal contraception for women being in a relationship compared to women being single (model a): 5.42% [1.19%,~~;~~ 9.44%]; model c): 6.40% [2.62%,~~;~~ 10.30%]). The probability of using hormonal contraceptives was significantly predicted by a relationship start in the last year in all three models (model a): *b* = 0.64 [0.41,~~;~~ 0.87]; model b): *b* = 0.50 [0.16,~~;~~ 0.83]; model c): *b* = 0.49 [0.28,~~;~~ 0.71]) corresponding to an increase around 3% to 4% in the probability to use hormonal contraception for women who started a relationship in the last year (model a): 4.24% [2.69%,~~;~~ 5.78%]; model b): 2.94% [0.94%,~~;~~ 4.90%]; model c): 3.31% [1.92%,~~;~~ 4.83%]). Ending a relationship was a significant negative predictor of hormonal contraceptive use in all three models (model a): *b* = -0.36 [-0.60,~~;~~ -0.12]; model b): *b* = -0.60 [-0.97,~~;~~ -0.23]; model c): *b* = -0.29 [-0.51,~~;~~ -0.06]) corresponding to a decrease around 3% in the probability to use hormonal contraception for women who ended a relationship in the last year (model a): -2.43% [-4.10%,~~;~~ -0.79%]; model b): -3.67% [-6.12%,~~;~~ -1.38%]; model c): -2.00% [-3.54%,~~;~~ -0.42%]). In addition, in all three models, the probability of using hormonal contraceptives was significantly predicted by relationship duration for women in a relationship (model a): *b* = -0.02 [-0.03,~~;~~ -0.01]; model b): *b* = -0.02 [-0.03,~~;~~ -0.001]; model c): *b* = -0.02 [-0.03,~~;~~ -0.01]) corresponding to a decrease around 0.12~~1~~% in the probability to use hormonal contraception for each additional year

women spent in a relationship (model a): -0.13% [-0.20%,; -0.06%]; model b): -0.1009% [-0.19%,; -0.003%]; model c): -0.13% [-0.19%,; -0.07%]).

The positive effect of having one child compared to no children was significant in all three models (model a): *b* = 0.23 [0.04,; 0.42]; model b): *b* = 0.30 [0.01,; 0.60]; model c): *b* = 0.21 [0.03,; 0.39]). This corresponds to an increase around 2% in the probability to use hormonal contraception for women having one child compared to women who have no children (model a): 1.46% [0.24%,; 2.74%]; model b): 1.79% [0.05%,; 3.45%]; model c): 1.38% [0.23%,; 2.5672%]). The positive effect of having two children compared to no children was significant in model a) and model c) (model a): *b* = 0.25 [0.05,; 0.46]; c): *b* = 0.22 [0.03,; 0.41]; but not in model b): *b* = 0.28 [-0.02,; 0.58]). This corresponds to an increase around 1.5% in the probability to use hormonal contraception for women having two children compared to women who have no children (model a): 1.66% [0.36%,; 2.94%]; model c): 1.49% [0.24%,; 2.72%]). The positive effect of completed fertility plans was significant in model b) (model b): *b* = 0.29 [0.02,; 0.57]; but not in model a): *b* = 0.12 [-0.07,; 0.30] and model c): *b* = 0.06 [-0.11,; 0.22]). This corresponds to an increase around 2% in the probability to use hormonal contraception for women having completed their fertility plans compared to women who have not completed their fertility plans (model b): 1.72% [0.14%,; 3.30%]).

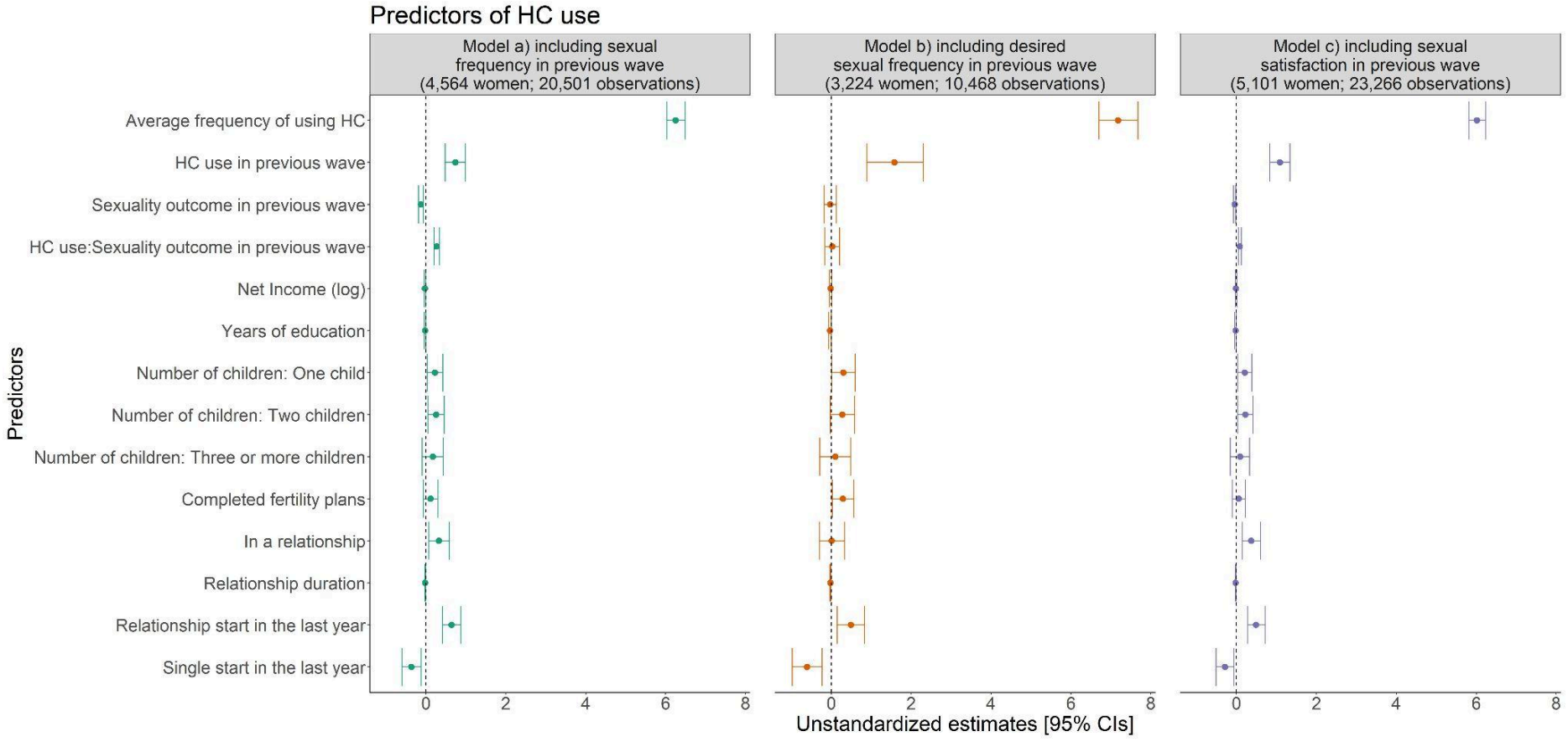
Income and Education

The negative effect of income was significant in model a) (model a): *b* = -0.02 [-0.04,; -0.002]; but not in model b): *b* = -0.01 [-0.04,; 0.01] and model c): *b* = -0.01 [-0.03,; 0.01]). This corresponds to a decrease around 0.15% in the probability to use hormonal contraception, for each increase of 1 in the log transformed net income per month (model a): -0.15% [-0.28%,; -0.02%]). The negative effect of education was significant in model c) (model c): *b* = -0.03 [-0.05,; -0.004]; but not in model a): *b* = -0.02 [-0.04,; 0.002] and model b): *b* = -0.03 [-0.06,; 0.003]). This corresponds to a decrease around 0.17% in the probability

to use hormonal contraception, for each increase of one year of education (model c): -0.17% [-0.31%,; -0.03%]). All other predictors were non-significant across all three models.

Figure 3

Predictors of HC use controlling separately for three sexuality outcomes and their interaction with HC use in~~at~~ the previous wave



Note. HC = hormonal contraception. Figure S8 shows the thin-plate spline effects of age on HC use. ~~The thin-plate spline effects of age only show the linear part of the effects on HC use (see Figure S8 for the full thin-plate spline effects of age on HC use).~~

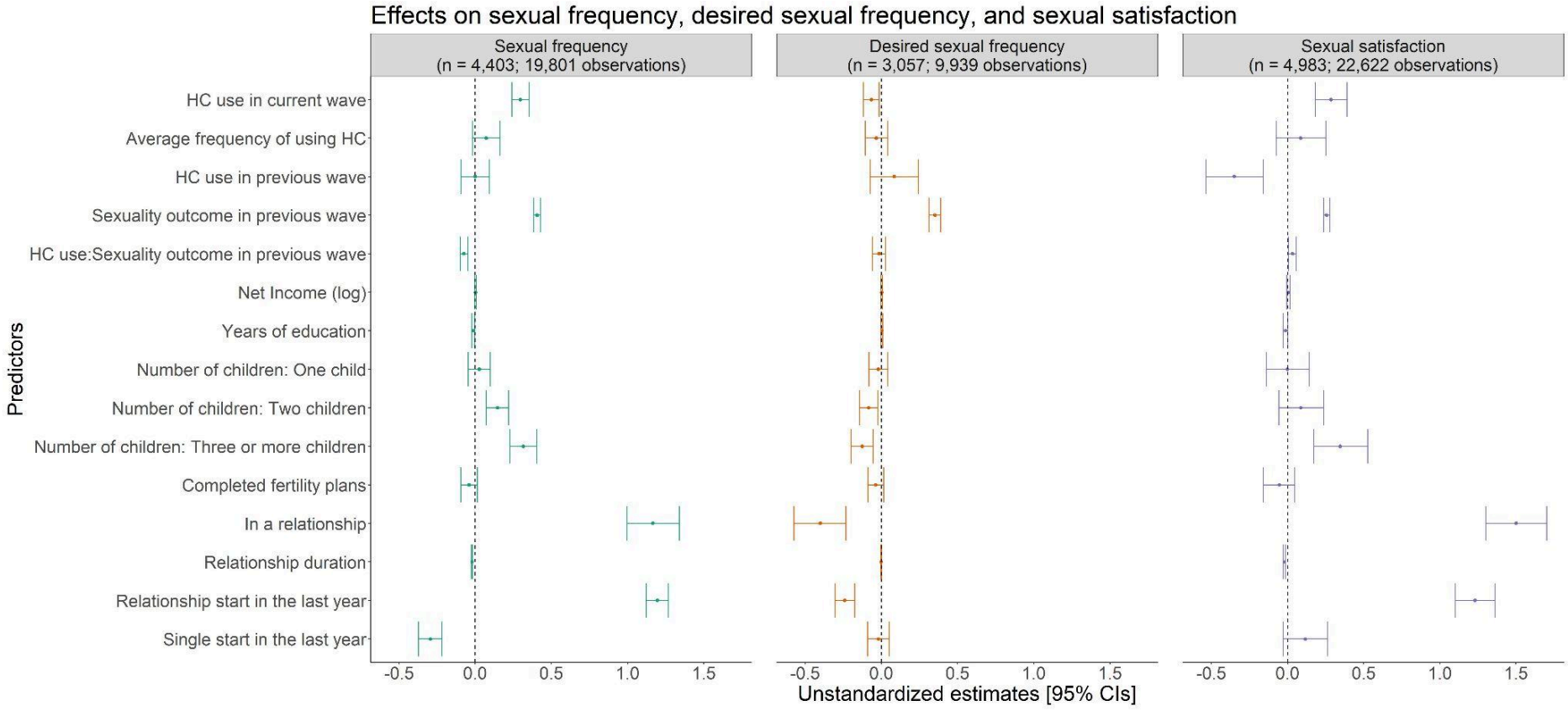
Average Treatment Effects of Hormonal Contraceptive Use on Sexuality

Adjusted Regression Analysis

Figure 4 displays the three separate adjusted regression analyse~~i~~s estimating the effects of hormonal contraceptive use and additional predictors on sexual frequency, desired sexual frequency, and sexual satisfaction. Hormonal contraception in~~at~~ the current wave positively predicted sexual frequency and sexual satisfaction. For sexual frequency, this corresponds to an unstandardized mean difference between non-hormonal contraceptive use and hormonal contraceptive use of 0.30 [0.24,~~;~~ 0.35] on a 7-point scale (ranging from 1 = *no sexual intercourse during the last three months* to 7 = *daily*). This unstandardized mean difference can be seen as a very rough approximation of the percentage change in sexual frequency (as the response scale of this item is very roughly a log-transformed version of frequency). For sexual satisfaction, the unstandardized mean difference between non-hormonal contraceptive use and hormonal contraceptive use was 0.29 [0.18,~~;~~ 0.39] on a 11-point scale (ranging from 0 = *very unsatisfied* to 10 = *very satisfied*). Hormonal contraception in~~at~~ the current wave negatively predicted desired sexual frequency indicating a small unstandardized mean difference between non-hormonal contraceptive use and hormonal contraceptive use of -0.06 [-0.12,~~;~~ -0.02] on a 5-point scale. As sexual frequency differed between the groups and desired sexual frequency was measured relative to the actual sexual frequency (“If it was only up to you, would you like to have less or more sexual intercourse compared to the last three months?”) on a scale from 1 = *a lot less*; 2 = *a little bit less*; 3 = *same amount*; 4 = *a little bit more*; to 5 = *a lot more*, this result cannot, on its own, be interpreted as ~~is not indicative of~~ an overall higher sexual desire for women using non-hormonal contraceptives. The distribution of desired sexual frequency for women included in this model using non-hormonal and hormonal methods is displayed in Figure S9 in the supplement.

Figure 4

Adjusted regression analysis estimating the effects of hormonal contraceptive use and other predictors on sexual frequency, desired sexual frequency, and sexual satisfaction



Note. HC = hormonal contraception. Figure S10 shows the thin-plate spline effects of age on the sexuality outcomes. ~~The thin-plate spline effects of age only show the linear part of the effects on the sexuality outcomes (see Figure S10 for the full thin-plate spline effects of age on the sexuality outcomes).~~

Inverse Probability of Treatment Weighting

We had registered (1) adjusted regression analyses that do not account for systematic missingness (already summarized in the previous section) and (2) inverse probability of treatment weighting models (IPTW) that do account for systematic missingness. The resulting effect estimates are contrasted in Figure 5. ~~Figure 5 displays the estimated effects of hormonal contraceptive use on sexual frequency, desired sexual frequency, and sexual satisfaction for the registered adjusted regression analysis without accounting for systematic missingness and the registered IPTW approach accounting for systematic missingness.~~

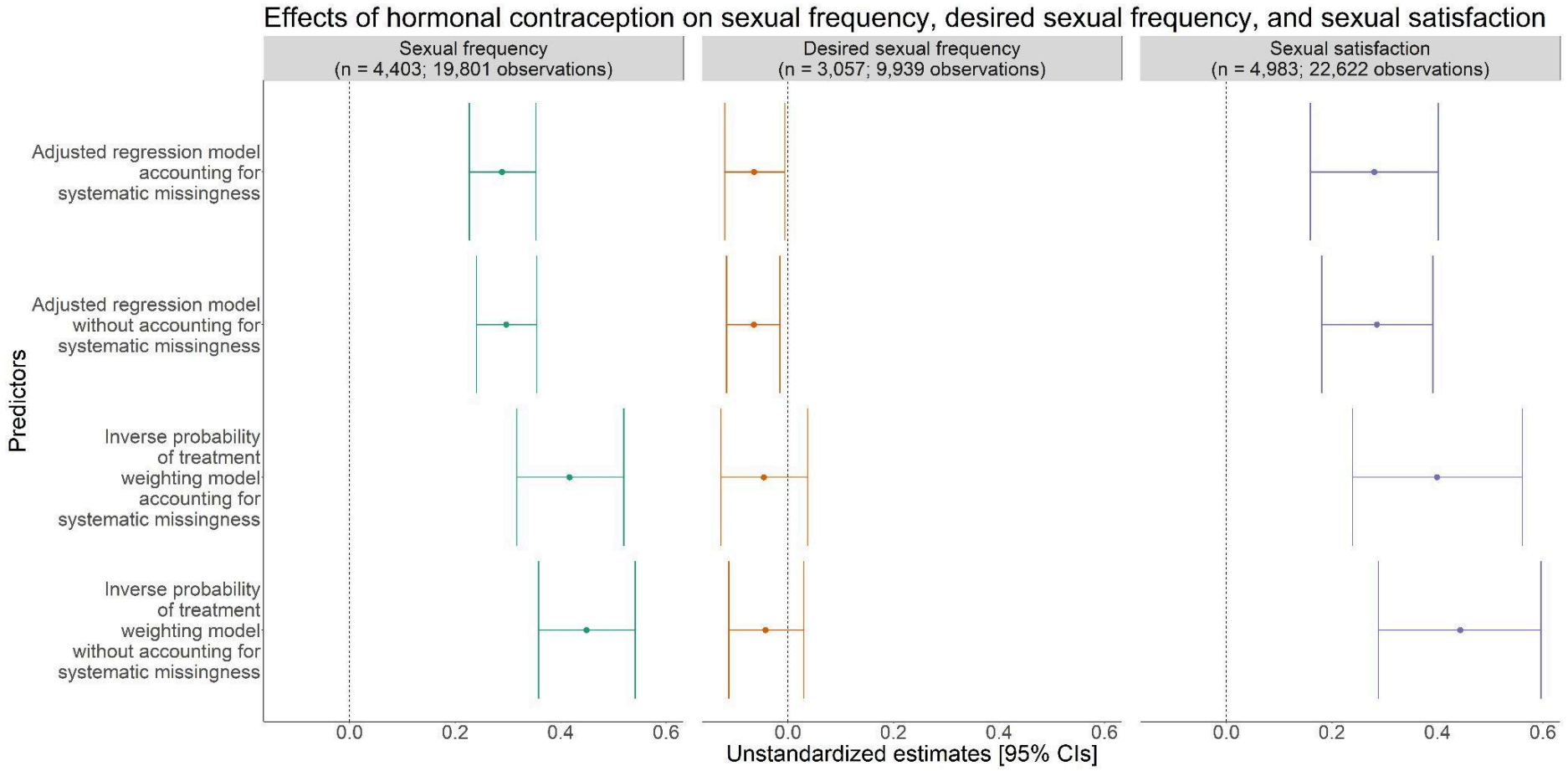
Compared to the estimate based on the adjusted regression analysis (*b* = 0.30 [0.24,~~;~~ 0.35]), the estimate in the IPTW approach for the effects of hormonal contraceptive use on sexual frequency (*b* = 0.42 [0.32,~~;~~ 0.52]) was slightly larger (difference in regression coefficients = 0.12 [0.003,~~;~~ 0.24]). The estimate in the IPTW approach for the effect of hormonal contraceptive use on desired sexual frequency was non-significant (*b* = -0.05 [-0.13,~~;~~ 0.04]), compared to the significant negative effect estimated based on the adjusted regression analysis (*b* = -0.06 [-0.12,~~;~~ -0.02]; however, the two estimates did not differ significantly from each other as the difference in regression coefficients was 0.02 [-0.08,~~;~~ 0.11]). The estimate in the IPTW approach for the effects of hormonal contraceptive use on sexual satisfaction (*b* = 0.40 [0.23,~~;~~ 0.56]) did not differ from the estimate based on the adjusted regression analysis (*b* = 0.29 [0.18,~~;~~ 0.39]; difference in regression coefficients = 0.11 [-0.08,~~;~~ 0.31]).

As we found systematic differences between these two approaches for sexual frequency as an outcome, the question arises whether these are due to the different model classes (regression analyses versus IPTW models) or due to the fact that only the IPTW models accounted for systematic missingness. We thus additionally implemented~~we additionally report the results of~~ (1) adjusted regression analyse~~i~~s accounting for systematic missingness and (2) IPTW models~~approach~~ without accounting for systematic missingness. As can be seen in Figure 5, the differences arise from the different models rather than from

whether or not systematic missingness is accounted for.~~, to illustrate that these systematic differences are not due to differences in accounting for systematic missingness (see Figure 5).~~

Figure 5

Estimated effects of hormonal contraceptive use on sexual frequency, desired sexual frequency, and sexual satisfaction for different modeling approaches



Robustness of Results

Figure 6 contrastssummarizes the estimated effects of hormonal contraceptive use from ouron sexual frequency, desired sexual frequency, and sexual satisfaction for the main analysis, the originally registered main analysis, and all remaining five robustness analseis (which implement different exclusion criteria, see Table 3) for the adjusted regression analysis and the IPTW approach. All analyses supported a positive effect of hormonal contraceptive use on sexual frequency and sexual satisfaction.

For sexual frequency as an outcome in the adjusted regression analysis, results based on the sample for the main analysis (b = 0.30 [0.24,; 0.35]) indicated smaller positive effects than the results based on all other analysis samples. Similarly, the estimate for the effect of hormonal contraceptive use on sexual frequency in the IPTW approach based on the main analysis (b = 0.42 [0.32,; 0.52]) was smaller than the results based on all other analysis samples. As explained above, ourFor the main analseis deviate from the originally registered main analyses in that, all observationswaves in which women indicated that they had never been sexually active, were excluded. In the originally registered main analysis (and in all remaining robustness analyses), for these observations sexual frequency was coded as 0 (*never*) in these waves. Considering the other two outcomes, desired sexual frequency and sexual satisfaction, the estimated effects of hormonal contraceptive use did not differ significantly between the various sets of exclusion criteria. The estimates of the effect of hormonal contraceptive use on desired sexual frequency and sexual satisfaction did not differ significantly between analysis samples within approaches.

Figure 6

Estimated effects of hormonal contraceptive use based on all robustness analyses

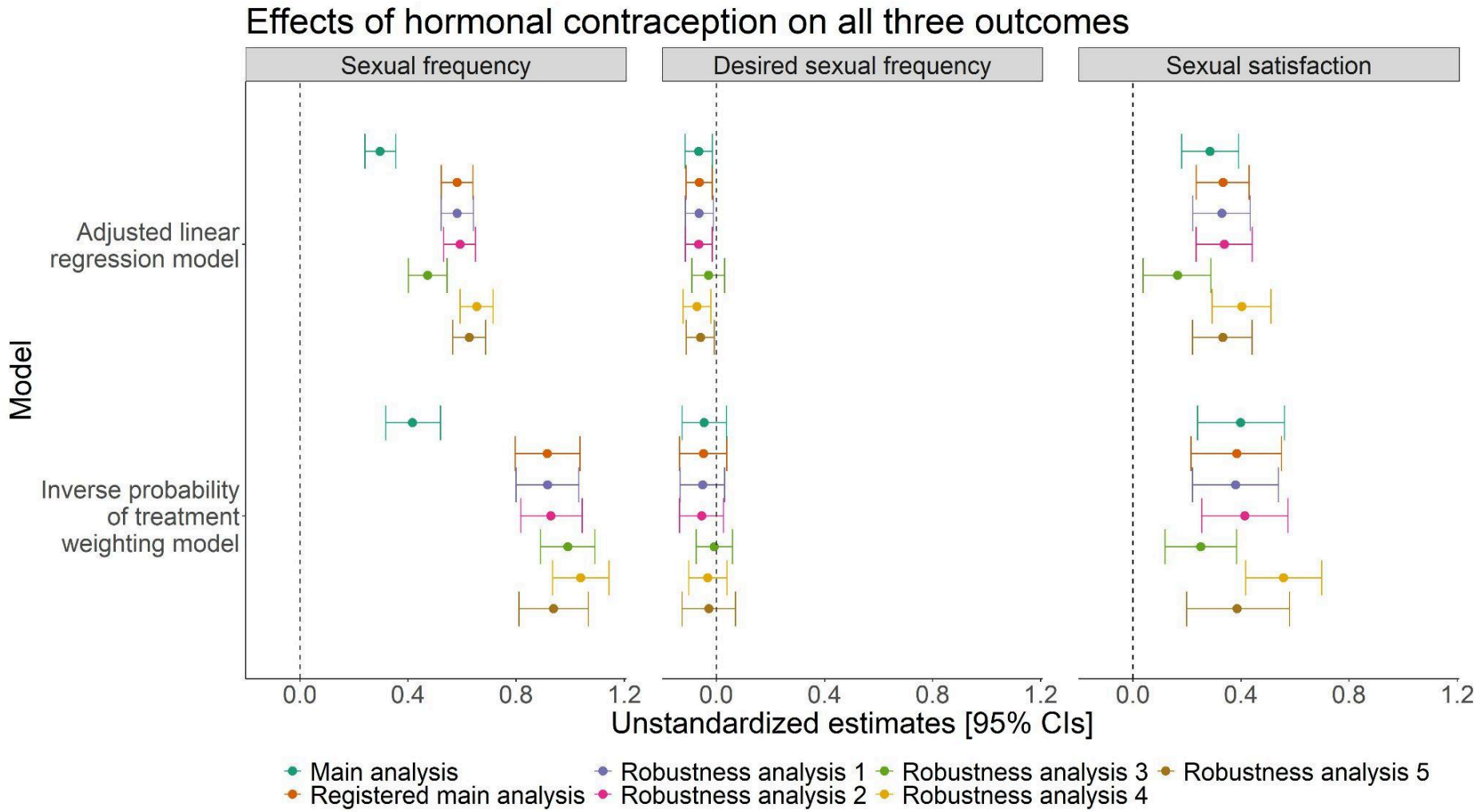


Figure 6 (continued)

Estimated effects of hormonal contraceptive use based on all three robustness analyses.

Note. Additional exclusion criteria compared to the originally registered main analysis in the main analysis (registered as robustness analysis 6): women never sexually active (up to *n* = 5,041~~194~~ and observations~~waves~~ = 23,130~~816~~); robustness analysis 1: women sterilized (up to *n* = 5,586 and observations~~waves~~ = 25,201); robustness analysis 2: partner sterilized (up to *n* = 5,618 and observations~~waves~~ = 25,159); robustness analysis 3: no contraceptive method (up to *n* = 5,114 and observations~~waves~~ = 20,870); robustness analysis 4: intrauterine device as a contraceptive method (up to *n* = 5,358 and observations~~waves~~ = 23,231); and robustness analysis 5: hormonal methods other than the oral contraceptive pill (up to *n* = 5,601 and observations~~waves~~ = 25,023). For more details see section Exclusion Process and Participants.

Sensitivity Analysis

To estimate how sensitive our results areeffects' sensitivity to unobserved confounders, E-values (VanderWeele & Ding, 2017) for the effect of hormonal contraception on sexual frequency, desired sexual frequency, and sexual satisfaction were estimated based on the results of the adjusted regression analysis including observed confounding variables as predictors. The E-Value indicates how strong an association of the unmeasured confounders with both variables (hormonal contraception and the sexuality outcome) *conditional* on all observed covariates would need to be to explain away the effect of hormonal contraception on the sexuality outcome; weaker confounding would not suffice to explain away the effect (wording according to VanderWeele & Ding, 2017). The E-value is interpreted as a risk ratio for the associations between the unmeasured confounders and hormonal contraception as well as the sexuality outcome. The E-Values were 1.59 (outcome: sexual frequency), 1.22 (desired sexual frequency), and 1.57 (sexual satisfaction). The E-value for the effect of hormonal contraception on sexual frequency was 1.59. The E-value for the effect of hormonal contraception on sexual frequency was 1.22. The E-value for the effect of hormonal contraception on sexual satisfaction was 1.57.

Heterogeneity in Treatment Effects of Hormonal Contraceptive Use on Sexuality

We were also interested in treatment heterogeneity: do the effects of hormonal contraceptive use on sexuality vary between women? Based on the adjusted regression analysis, the standard deviation of the treatment effect (i.e., the standard deviation of the random slope of hormonal contraceptive use on the sexuality outcome) was relatively high compared to the overall standard deviations of the outcomes (see Table 4) for sexual frequency (34% of the overall standard deviation) and sexual satisfaction (29% of the overall standard deviation), but not for desired sexual frequency (20% of the overall standard deviation; see the second column of Table 6).

Taking a closer look at the model-implied point estimates of the individual treatment effects, for all three outcomes, some women experienced negative effects and some women

experienced positive effects of hormonal contraceptive use. However, the point estimates of individual treatment effects were estimated very imprecisely, so that only ~~a~~ few of them were significantly different from zero when considered in isolation. Putting these pieces together, while we can say that heterogeneity was high, the number of observations per woman was not sufficient to estimate anyone's treatment effect with certainty (see Table 6). Figure 7 displays selected individual treatment effects (as displaying all of them would not be legible). The figure includes all women~~,~~ who contributed at least 11 observations~~waves~~ to the model (7 observations~~waves~~ for desired sexual frequency) and reported using non-hormonal and hormonal contraception at least once. The large confidence intervals ~~— even for the women who were included in at least 11 (or 7) waves —~~ illustrate the uncertainty in the estimates of the individual treatment effects. To provide an additional quantification of this uncertainty~~To quantify this uncertainty in familiar terms~~, we computed the reliability of the estimated random slopes. To get an estimate of the unweighted reliability, we subtracted the fraction of the uncertainty of the random slope (reflected by the variance across MCMC draws) by the average between-person variance (reflected by the variance of the random slope) from 1 (see Arslan, 2024, for a detailed description of this procedure and the code that we adjusted for our analysis). The reliability of the random slopes for all outcomes was very low (≤ .17, see last column in Table 6), indicating a high uncertainty in the estimates of the individual treatment effects.

Table 6

Heterogeneity in and reliability of treatment effects of hormonal contraceptive use on sexuality

Outcome	ATE [95% CI]	Standard deviation of the random slope	Number of women with nNegative point estimates of ITEs	Number of women with pPositive point estimates of ITEs	Number of women with sSignificant negative ITEs	Number of women with sSignificant positive ITEs	Reliability of the ITE estimates
Sexual frequency	0.30 [0.24,; 0.35]	0.51 [0.44,; 0.58]	302 (7%)	4,101 (93%)	0 (0%)	31 (1%)	0.15
Desired sexual frequency	-0.06 [-0.12,; -0.02]	0.16 [0.04,; 0.28]	2,784 (91%)	273 (9%)	0 (0%)	0 (0%)	0.09
Sexual satisfaction	0.29 [0.18,; 0.39]	0.77 [0.60,; 0.93]	786 (16%)	4,197 (84%)	0 (0%)	45 (1%)	0.17

Note. ATE = average treatment effect (unstandardized regression coefficient *b*); ITE = individual treatment effect. All numbers are based on the adjusted regression analysis. Total sample sizes for the three analyses: sexual frequency: *n* = 4,403 women and 19,801 observationswaves; desired sexual frequency: *n* = 3,057 women and 9,939 observationswaves; and sexual satisfaction: *n* = 4,983 women and 22,622 observationswaves.

Figure 7

*Distribution of selected individual treatment effects of hormonal contraception on a) sexual frequency (*n = *137); b) desired sexual frequency (*n = *181); and c) sexual satisfaction (*n = *193)*

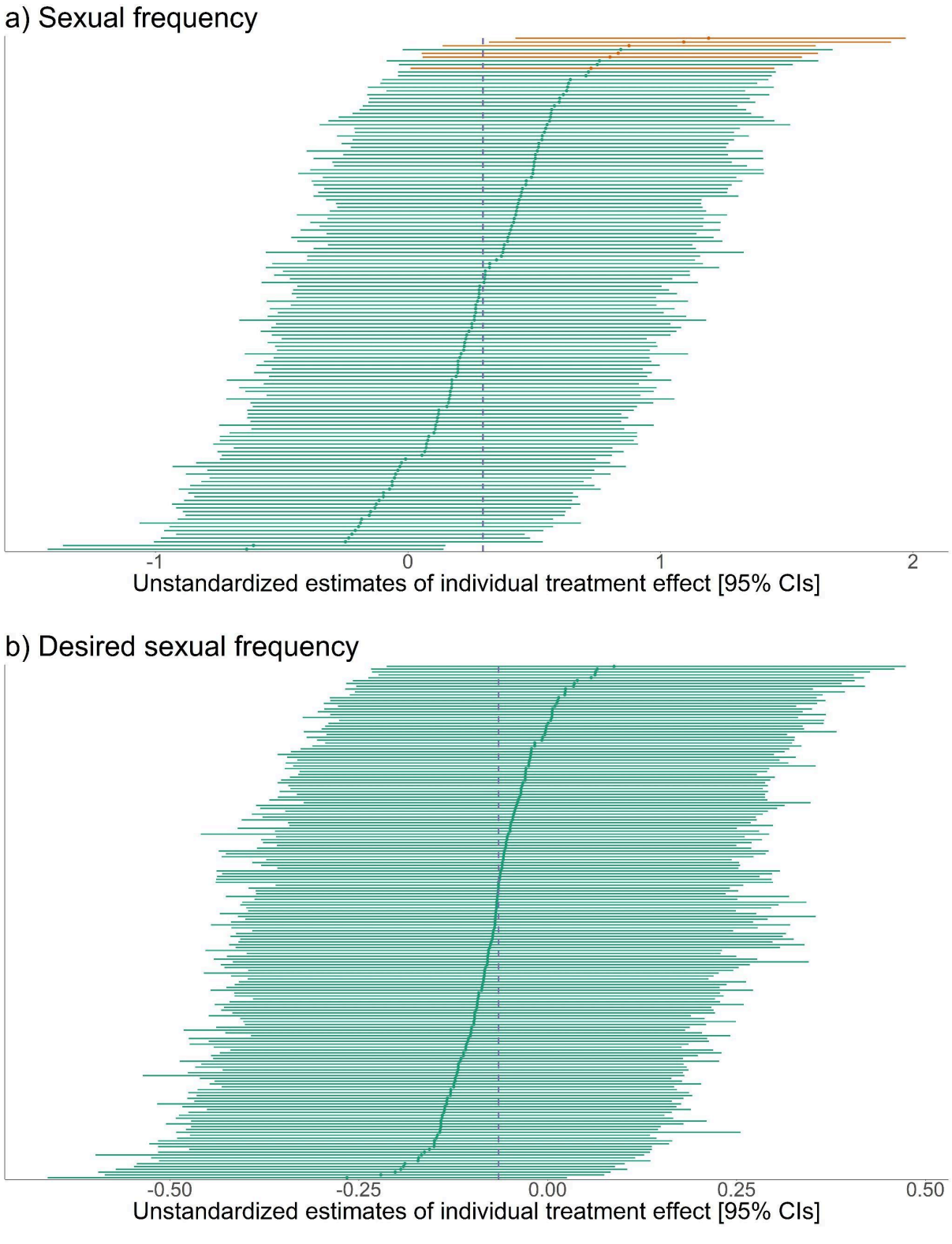
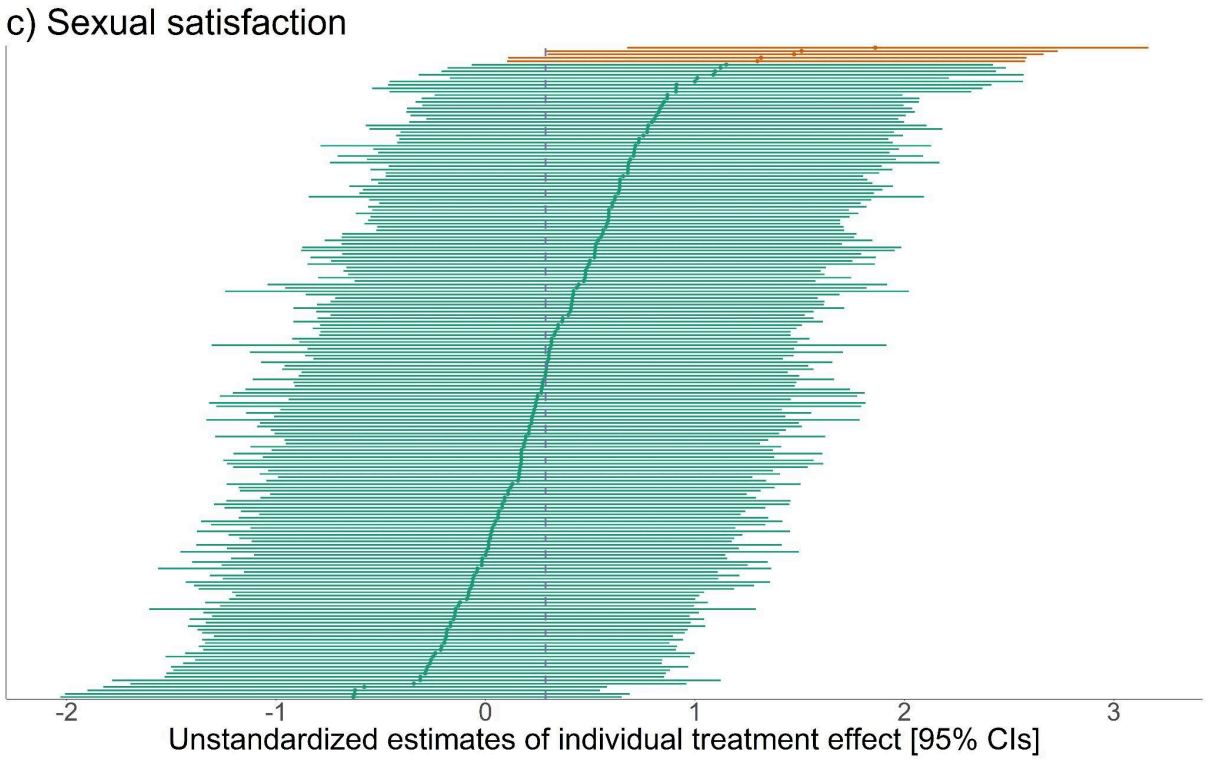


Figure 7 (continued)

*Distribution of selected individual treatment effects of hormonal contraception on a) sexual frequency (*n = *184); b) desired sexual frequency (*n = *197); and c) sexual satisfaction (*n = *244)*



Note. The figure includes all women, who contributed at least 11 observations~~waves~~ to the model (7 observations~~waves~~ for desired sexual frequency) and reported using non-hormonal and hormonal contraception at least once. The green individual treatment effects do not differ significantly from zero, the orange individual treatment effects differ significantly from zero. The dotted purple line indicates the average treatment effect.

Predictors of Individual Treatment Effects

Individual treatment effect estimates for sexual frequency and sexual satisfaction correlated[14] positively with each other (r = .36 [.33,~~;~~ .39], n = 4,346). Individual treatment effect estimates for desired sexual frequency correlated negatively with individual treatment effect estimates for sexual frequency (r = -.24 [-.27,~~;~~ -.20], n = 3,048) and sexual satisfaction (r = -.30 [-.35,~~;~~ -.25], n = 3,034).

To explore the correlation between individual treatment effects and age as well as the correlations between individual treatment effects and Big Five personality traits, we used a "correlate, then average across draws" approach.[15] That is, we correlated individual treatment effect estimates from the adjusted regression analysis with age and the Big Five personality traits at draw level. ~~These correlation analyses were weighted by the inverse of the standard error of the individual treatment effect estimates to propagate uncertainties in their estimation.~~ Finally, these correlations were averaged across draws. They are summarized in the first columns of Table 7.

Using this approach, we found two small correlations for which the 95% credible interval excluded zero. These were between the interindividual treatment effects on sexual satisfaction and a) neuroticism (r = .07 [.04, .11], n = 4,570), meaning that women who scored higher on neuroticism experienced more positive effects of hormonal contraceptive use on sexual satisfaction and b) conscientiousness (r = -.04 [-.07, -.01], n = 4,571), meaning that women who scored higher on conscientiousness experienced less positive effects of hormonal contraceptive use on sexual satisfaction. While some of them were significant, most of them were rather small in magnitude and no common pattern across sexuality outcomes emerged. ~~We found the strongest correlations between interindividual treatment effects on sexual satisfaction and a) neuroticism (r = .17 [.15; .20], n = 4,570),~~

[14] Correlations between the individual treatment effect estimates from the adjusted regression analysis were calculated by weighting the correlation by the inverse of the square root of the multiplication of each standard error of the individual treatment effect estimates to account for uncertainties in their estimation. Confidence intervals are estimated using bootstrapping (10,000 bootstraps).

[15] The relationships based on the originally planned "average across draws, then correlate" approach are reported in Table S6 in the supplement.

~~meaning that women who scored higher on neuroticism experienced more positive effects of hormonal contraceptive use on sexual satisfaction; b) conscientiousness (*r* = -.10 [-.13; -.07], *n* = 4,571), meaning that women who scored higher on conscientiousness experienced less positive effects of hormonal contraceptive use on sexual satisfaction; and c) agreeableness (*r* = -.07 [-.10; -.04], *n* = 4,570), meaning that women who scored higher on agreeableness experienced less positive effects of hormonal contraceptive use on sexual satisfaction.~~ Neither neuroticism~~,~~ nor conscientiousness~~, nor agreeableness~~ showed significant correlations with the individual treatment effects for the other outcomes~~ respectively~~.

Contraceptive Decisions and Individual Treatment Effects

We additionally wanted to investigate whether women guide their contraceptive choices based on their experiences. In fact, our analysis of predictors of hormonal contraceptive use already provided evidence pointing into that direction: women with higher sexual frequency and higher sexual satisfaction were more likely to stick with the method they were previously using. However, our registered analysis went further and asked whether the individual treatment effects (i.e., not just whether women are more or less sexually satisfied, but whether hormonal contraceptives have a negative or positive effect on their sexual satisfaction) correlated with the amount of time that women used hormonal contraceptives. Again, we used the "correlate, then average across draws" approach.[16] That is, we correlated the individual treatment effects~~To answer the question of whether women guide their choice of contraceptive method by deciding against hormonal contraceptives after experiencing adverse effects, we again used the individual treatment effect estimates from the adjusted regression analysis, this time correlating them~~ with the proportion of years of hormonal contraceptive use at draw level. ~~This correlation analysis was again weighted by the inverse of the standard error of the individual treatment effect estimates.~~ Finally, these

[16] The relationships based on the originally planned "average across draws, then correlate" approach are reported in Table S6 in the supplement.

correlations were averaged across draws. None of the “correlate, then average across draws” relationships ~~correlations~~ were significant (see the last column of Table 7).

Table 7

"Correlate, then average across draws" relationships~~Correlations~~ between individual treatment effects and age, Big Five personality traits, as well as years spent on hormonal contraception

	Age	Extraversion	Agreeable-ness	Conscien-tiousness	Neuroticism	Openness	Years spent on HC
ITE Sexual frequency	*r* = -.01 [-.04; .01] ~~***r* = -.04**~~ ~~**[-.07; -.01]**~~	*r* = .0002 [-.03, .03] ~~*r* = -.002~~ ~~[-.03; .03]~~	*r* = -.002 [-.03, .03] ~~*r* = .002~~ ~~[-.03; .03]~~	*r* = -.001 [-.03, .03] ~~*r* = -.002~~ ~~[-.03; .03]~~	*r* = .01 [-.02, .04] ~~*r* = .02~~ ~~[-.01; .05]~~	*r* = .01 [-.02, .04] ~~*r* = .03~~ ~~[-.01; .06]~~	*r* = .01 [-.02, .03] ~~*r* = .02~~ ~~[-.01; .04]~~
	n = 4,403	*n* = 4,037	*n* = 4,035	*n* = 4,037	*n* = 4,036	*n* = 4,036	*n* = 4,403
ITE Desired sexual frequency	*r* = .003 [-.03; .04] ~~*r* = .01~~ ~~[-.03; .04]~~	*r* = -.03 [-.07, .01] ~~***r* = -.09**~~ ~~**[-.12; -.05]**~~	*r* = .01 [-.02, .04] ~~*r* = .02~~ ~~[-.01; .06]~~	*r* = -.002 [-.04, .03] ~~*r* = -.01~~ ~~[-.04; .03]~~	*r* = -.01 [-.04, .03] ~~*r* = -.02~~ ~~[-.06; .02]~~	*r* = -.02 [-.06, .02] ~~***r* = -.05~~ ~~**[-.08; -.01]**~~	*r* = -.002 [-.04, .03] ~~*r* = -.01~~ ~~[-.04; .02]~~
	n = 3,057	*n* = 2,817	*n* = 2,817	*n* = 2,817	*n* = 2,817	*n* = 2,816	*n* = 3,057
ITE Sexual satisfaction	*r* = .01 [-.02; .03] ~~*r* = .02~~ ~~[-.01; .05]~~	*r* = -.02 [-.05, .01] ~~***r* = -.04~~ ~~**[-.07; -.01]**~~	*r* = -.03 [-.06, .002] ~~***r* = -.07~~ ~~**[-.10; -.04]**~~	***r* = -.04** **[-.07, -.01]** ~~***r* = -.10~~ ~~**[-.13; -.07]**~~	***r* = .07** **[.04, .11]** ~~***r* = .17~~ ~~**[.15; .20]**~~	*r* = -.01 [-.03, .02] ~~*r* = -.01~~ ~~[-.04; .02]~~	*r* = .003 [-.02, .03] ~~*r* = .01~~ ~~[-.02; .03]~~
	n = 4,983	*n* = 4,571	*n* = 4,571	*n* = 4,571	*n* = 4,570	*n* = 4,570	*n* = 4,983

Note. HC = hormonal contraception; ITE = individual treatment effect. ~~Correlations were weighted by the inverse of the standard error of the individual treatment effect estimates to propagate uncertainties in their estimation.~~ Square brackets indicate 95% credibility intervals. ~~Square~~

~~brackets indicate 95% confidence interval based on bootstrapping with 10,000 bootstraps.~~ Bold numbers indicate that the 95% credible interval excluded zero.~~significant correlations ($p < .05$).~~

Discussion

Based on a large longitudinal study, we aimed to answer the questions whether hormonal contraceptive use influences women's sexuality as well as whether and to which extent the effects of hormonal contraceptives on sexuality vary between users. Women's hHormonal contraceptive use was highly stable over time inacross the waves of our study. Besides the stability, the current relationship situation, number of children and fertility plans, sexual frequency in the previous wave and sexual satisfaction inat the previous wave, as well as potentially income and education predicted the use of hormonal contraception. Controlling for a number of potential confounding variables, we found strong and robust support for positive average treatment effects of hormonal contraceptive use on sexual frequency and sexual satisfaction across two different analytic approaches and acrossin six robustness analyses. In contrast, We found no robust evidence for a negative effect of hormonal contraceptive use on desired sexual frequency was not robust across both analytic approaches and across all robustness analyses. The potential negative effect of hormonal contraceptive use found in the adjusted regression analysis is ambiguous because of the way desired sexual frequency was measured. The effect could indicate a satisfied sexual desire for hormonal contraceptive users and/or an unfulfilled desire for more sexual intercourse for non-hormonal contraceptive users. BeyondBesides these average treatment effects, we found relatively high heterogeneity for individual treatment effects of hormonal contraception on sexual frequency and sexual satisfaction, (but not onfor desired sexual frequency). However, tThe high uncertainty around individual treatment effects suggests that the number of women and especiallythe number of observationswaves per woman wasmight have been too small to estimate individual treatment effects reliably. In addition, the one year interval between waves might have been too large to estimate individual treatment effects: women might have switched back and forth between methods in between measurement occasionsas switches between methods could have occurred between the measurement occasions. Thesse estimated individual treatment effects were not consistently associated with other interindividual differences (age and Big Five personality) or contraceptive

decisions across the longitudinal study~~(one methodological explanation for this might be the high uncertainty around individual treatment effects)~~.

Descriptive Patterns in Hormonal Contraceptive Use and Predictors of Hormonal Contraceptive Use

Our results suggest that across 5.6 years (mean number of current and previous waves), about one half of the women exclusively reported~~constantly used~~ non-hormonal or hormonal methods, and the other half of the women reported a switch between non-hormonal and hormonal methods. About 60%~~three fourths~~ of these switches were from hormonal to non-hormonal methods and 30%~~the last fourth was~~ from non-hormonal to hormonal methods. Only~~While most women reported no or only one switch during their participation in the PAIRFAM panel dataset,~~ a few women reported several switches. Overall, the percentage of women using hormonal methods declined across the course of PAIRFAM, from 2008 to 2021. These results are in line with a representative survey in Germany comparing contraceptive use between the years 2011 and 2018 showing that the use of oral contraceptives (i.e., the most popular form of hormonal contraception) declined by 6 percentage points while the use of condoms (i.e., the most popular form of non-hormonal contraception) increased by 9 percentage points (Bundeszentrale für gesundheitliche Aufklärung, 2018). The most commonly mentioned concerns about hormonal contraception in Western countries are effects on menstruation, physical side effects, mental health effects and effects on sexuality as well as concerns about future fertility or a general wish to use more natural methods (Le Guen et al., 2021). The discontinuation rate for oral contraceptives was around 28% in an analysis of 23 countries from 1990 to 2008 (Ali & Cleland, 2010). The trend away from hormonal contraception is not generalizable to all countries, though. According to the United Nations, Department of Economic and Social Affairs, Population Division (2022), oral contraceptives are the most common method in 36 countries (e.g., Algeria, Morocco), while condoms are the most common method in 27 countries (including Germany).

There was relatively strong stability in the broad choice of contraceptive method, as indicated by the large predictive contributions of average frequency of hormonal contraceptives and use of hormonal contraceptives in the previous wave.~~The strongest predictor of hormonal contraceptive use was the average frequency of using hormonal contraceptives and above this effect the use of hormonal contraceptives at the previous wave. This is evidence for a relatively strong stability in the broad choice of contraceptive methods.~~ In addition, women were more likely to switch their contraceptive method if sexual frequency and sexual satisfaction were low. ~~This is indicated by the significant main effect and interaction effects of hormonal contraceptive use at the previous wave and sexual frequency or sexual satisfaction, respectively, at the previous wave.~~

Beyond these method- and sexuality-related predictors, the current relationship situation predicted the choice of contraceptive methods. Women in a relationship, especially women who started a new relationship in the last year, were more likely to use hormonal contraceptive methods. Beyond this initial increase in hormonal contraceptive use after the start of a relationship, relationship duration was a negative predictor of hormonal contraceptive use. In addition, ending a relationship negatively predicted the use of hormonal contraception. Besides the current relationship situation, women who had one child (and potentially two children, even though this result was not robust across all analyses) were more likely to use hormonal contraception compared to women who had no children. In line with these results, we found some (albeit non-robust) evidence that women who said that they had completed their fertility plans were more likely to use hormonal contraceptive methods.

We also found some non-robust evidence for negative effects of income and education on hormonal contraceptive use. These results potentially indicate that women with a higher income and more education were less likely to use hormonal contraceptives. As these effects are rather small and did not emerge across~~based on all the performed~~ models, they need to be interpreted cautiously. It is quite possible that the effects of income and education on hormonal contraceptive use differ depending on the details of the health care

system and how accessible it renders such contraceptives (e.g., is a prescription required? Does health insurance cover the cost of the necessary appointments and of the contraceptives?). ~~In addition, these findings – especially concerning the potential negative effect of income – might not be generalizable to women in nations lacking mandatory health insurance, such as the United States, or in countries where contraceptives are typically not covered by health insurance, such as Canada.~~

Overall, the choice of contraceptive method appears to be relatively stable and changes in contraceptive methods are mainly explained by low sexual frequency and sexual satisfaction as well as the current relationship and family situation.

Average Treatment Effects on Sexuality

We found strong support for positive average treatment effects of hormonal contraception on sexual frequency and sexual satisfaction across all analytical approaches and all robustness analyses. When women use hormonal contraceptives, they have a higher sexual frequency and a higher sexual satisfaction then when they use non-hormonal contraceptives, and these effects are robust to the inclusion of the potential confounding effects discussed in the previous section ~~before~~. These results are in line with results from an earlier project based on cross-sectional data that aimed to disentangle, to some extent, causal effects of hormonal contraceptives from selection effects by controlling for potential confounders and estimating the sensitivity to unobserved confounders (Botzet et al., 2021). Applying these methodologies to longitudinal data in the current project advances our understanding of the causal nature of the effect by addressing reverse causality and attrition besides confounding as potential alternative explanations. Based on the performed sensitivity analysis, unobserved confounding would need to be relatively strong to fully account for the estimated effects of hormonal contraceptive use on sexual frequency, desired sexual frequency, and sexual satisfaction.

Our results are in line with existing correlational evidence finding positive effects of hormonal contraception on sexuality concerning sexual functioning (e.g., Oranratanaphan &

Taneepanichskul, 2006), libido (McCoy & Matyas, 1996), sexual satisfaction (e.g., Caruso et al., 2005), and, most strongly, sexual frequency (e.g., Caruso et al., 2005; McCoy & Matyas, 1996). Most importantly, our results are also in line with two other studies finding support for positive effects of hormonal contraceptive use on sexuality in longitudinal datasets (Blumenstock & Barber, 2022; Ott et al., 2008). Taking the available evidence together supports the hypothesis that the use of hormonal contraception positively affects women's sexuality; potentially through several mechanisms, including, for example, overcoming the fear of unwanted pregnancy during sexual activity (Blumenstock & Barber, 2022) and the resolution of painful or troublesome gynecologic disorders (Both et al., 2019).

In contrast~~On the contrary~~, our results are not consistent~~in line~~ with existing experimental evidence ~~literature – mainly from experimental studies –~~ that hormonal contraceptive use negatively affects~~impacts~~ sexual functioning (e.g., Læssøe et al., 2014) and libido (e.g., Lee et al., 2017; Lundin et al., 2018; Zethraeus et al., 2016), as well as sexual activity, arousal, pleasure, orgasm, and lubrication (Smith et al., 2014). Therefore, based on the correlational data, we found no support for the hypothesis that the intervention in~~into~~ the endocrine system through the use of ~~by~~hormonal contraceptives ~~use~~ (Stomati et al., 1998) negatively affects women's sexuality (Both et al., 2019).

The evidence concerning negative average treatment effects on desired sexual frequency (our available measure for libido) is neither constant across analytical approaches nor robustness analyses. It should therefore be interpreted very cautiously, especially because the sample size of available observations~~waves~~ was considerably smaller (desired sexual frequency was only measured from wave 7 to wave 14)~~ and the outcome was measured on a scale that takes actual sexual frequency into account.~~ Even if we take the potential negative effect of hormonal contraceptive use found in the adjusted regression analysis at face value, its interpretation would be ambiguous because of the way desired sexual frequency was measured ("If it was only up to you, would you like to have less or more sexual intercourse compared to the last three months?"). Because the sexual desire item was phrased relative to sexual behavior, we have to interpret this result in light of the

positive effect of HC on sexual frequency. So, potentially, desire is constant, but because frequency is lower among non-HC users, our desire outcome showed a negative effect. Overall, the evidence of an average treatment effect of hormonal contraceptive use on desired sexual frequency seems not very robust and very inconsistent across studies (e.g., Botzet et al., 2021; Lee et al., 2017; Lundin et al., 2018; McCoy & Matyas, 1996; Zethraeus et al., 2016). If an average treatment effect of hormonal contraceptive use on libido exists, it is probably relatively small and large sample sizes are needed to detect it.

We found evidence for strong positive effects of hormonal contraceptive use on sexuality when women are not randomly assigned to one contraceptive method but make free (and potentially to some degree informed) decisions about their contraceptive method. Taking the evidence for positive effects of hormonal contraceptive use and a high heterogeneity together, this makes it very likely that women in our sample, who experience adverse effects of hormonal contraceptive use, are more likely to discontinue using hormonal contraceptives. Therefore, the estimated effect of hormonal contraceptive use on sexuality in the current study, does not recover the average treatment effect across all women, but is closer to the *average treatment effect on the treated*.

Differences in Results Based on Modelling Approach

To our surprise, for the outcome of sexual frequency, we found significant differences between the estimated average treatment effects depending on whether we used~~based on the~~ adjusted regression analysis or~~and~~ the IPTW approach~~ for sexual frequency as an outcome~~. Estimated average treatment effects were slightly larger for the IPTW approach compared to the adjusted regression analysis ~~and this difference could not be explained by accounting for systematic missingness~~ (see Figure 5). As registered in stage 1 of this registered report, adjusted regression analysis was treated as the main analysis and the IPTW approach was treated as a robustness analysis for identifying the average treatment effect. B~~As b~~oth analytic approaches supported positive average treatment effects of

hormonal contraceptive use on sexual frequency, but the question that still remains is how strong this effect actually is.

As outlined in the introduction, both approaches try to estimate the causal effect of a treatment on an outcome and they come with different advantages and disadvantages.[17] Our adjusted regression analysis relies on the ~~untested~~ key assumption that the relationships between the covariates and the sexuality outcome are modeled appropriately (Gutman & Rubin, 2017), while the IPTW approach relies on the ~~untested~~ key assumption that the relationships between the covariates and hormonal contraceptive use are modeled appropriately. For both assumptions, there are reasons to doubt that these assumptions hold true. Considering the regression analyses~~On the one hand~~, relationship duration could affect sexual frequency in a non-linear manner in the form of a sudden drop in sexual frequency after a certain relationship duration, biasing the results ~~of the adjusted regression analysis~~. Considering the IPTW approach~~On the one hand~~, education could affect the choice of contraceptive methods in a non-linear manner in the form of a sudden increase in hormonal contraceptive use after a certain degree of educational attainment, biasing the results ~~of the IPTW approach~~. We cannot discern which modeling approach is more biased here, and thus we cannot clearly favor one over the other.~~rule out any of these biases, making it impossible to decide which approach is the appropriate one (one alternative for future projects might be to apply doubly robust standardization combining the IPTW approach with G-computation, see Robins et al., 2007). The estimated average treatment effects based on the adjusted regression analysis might in addition be biased by extrapolation, i.e., in the regression treatment effects from women who would never consider hormonal contraception in reality are extrapolated (King & Zeng, 2006).~~ Maybe more importantly, ~~In addition,~~ *both* approaches still rely on the key assumption~~s~~ of ~~(1)~~ no unobserved confounders, even if our sensitivity analyses suggest that such confounders would need to be fairly influential to completely explain away the estimated effects.~~(but see the results on the sensitivity of the estimated~~

[17] We are ignoring the first concern outlined in the introduction, ~~focusing on~~ cherry-picking of covariates, as ~~this concern applies to both approaches and should be ruled out as an alternative explanation because~~ we registered our covariates in stage 1 of this registered report.

~~effects to unobserved confounders); (2) positivity; and (3) a correct specification of the underlying models (Thoemmes & Ong, 2016).~~.

As we cannot decide which of the two approaches correctly retrieves the underlying causal effect of hormonal contraceptive use, we would like to use our results as an example to urge researchers to consider different analytical approaches when estimating causal effects and to ~~carefully~~ lay out all of the tested and untested assumptions as well as the specification of the underlying model.

Individual Treatment Effects

Besides the average treatment effects, we were interested in the heterogeneity of effects when estimating individual treatment effects. For both sexual frequency and sexual satisfaction we found support for ~~a~~ relatively strong heterogeneity in individual treatment effects, ranging from negative to positive effects in line with Grahams (2019) summary of anecdotal evidence. This finding questions the idea that hormonal contraceptive use affects sexuality only through a uniform and universal biological pathway.

Where could such heterogeneity come from?~~Possible alternative explanations exist.~~ From a biological perspective, our findings are consistent with~~support~~ the idea that women have varying sensitivity to hormones (Kiesner, 2017) resulting in heterogeneous responses to hormonal contraceptive use. The heterogeneous pattern could also be explained by non-universal psychological mechanisms~~sociological explanations~~. For example, t~~T~~here might be interindividual differences in the expectations and worries about the effectiveness of contraception and the likelihood of unplanned pregnancy (Both et al., 2019). This might result in differential effects of hormonal contraceptive use on sexuality.

At the same time,~~Nevertheless,~~ we found no common pattern across sexuality outcomes for relationships between individual treatment effects and potential predictors of individual treatment effects (including age and Big Five personality). Unfortunately, the high uncertainty around individual treatment effects as indicated by the low reliability makes it impossible to rule out any of the predictors. Future studies interested in the effects of

hormonal contraceptive use on sexuality, should further investigate the heterogeneity in individual treatment effects and explore their relationship with other markers of interindividual differences.

What could these markers of interindividual differences be? First, self-reported side effects (i.e., side effects mentioned by women) might be positively related to experienced side effects (i.e., individual treatment effects). This could provide a first test of whether interindividual differences in self-report questionnaires reflect individual treatment effects at all. Second, instead of a general measure of neuroticism, fear of unwanted pregnancy might be positively related to individual treatment effects, with women with a higher fear of unwanted pregnancy showing positive effects of hormonal contraceptives on sexual frequency and sexual satisfaction. Third, general attitudes toward the use of hormonal contraceptives might also be related to individual treatment effects, with women who are very skeptical about the use of hormonal contraceptives (but who still use them, e.g., for other health reasons) showing negative effects of hormonal contraceptives on sexual frequency and sexual satisfaction. Finally, we would hypothesize that hormone sensitivity (Kiesner, 2017) would be negatively related to individual treatment effects, with women with higher sensitivity experiencing more negative effects of hormonal contraceptives on sexual frequency and sexual satisfaction.

We also did not find significant correlations between the estimated individual treatment effects and the number of years of hormonal contraceptive use over the course of PAIRFAM, which we would interpret as preliminary evidence against potential assortment based on experiences with contraceptive methods. Again, the interpretability of this analysis is limited by the low reliability of the individual treatment effects. Furthermore, the analysis would have at best provided indirect evidence for assortment. Another analysis we conducted can be used to answer a slightly different but related question, namely whether generally low sexual frequency or low sexual satisfaction (which may or may not be attributable to hormonal contraception) predict switches in contraceptive methods. This analysis provides another type of indirect evidence for assortment: In the models predicting

hormonal contraceptive use in the current wave, women were more likely to switch their contraceptive method if they experienced lower levels of sexual frequency and sexual satisfaction. However, another recent study using a similar analysis found no support for women's experiences of lower levels of sexual frequency and sexual satisfaction predicting switches in contraceptive methods (Draxler et al., 2024). We believe that a better empirical test of the underlying idea would require more extensive longitudinal data. In such data, one could use all but the last observation per individual to estimate the individual treatment effect, and then test whether this effect predicts the contraceptive method reported in the last observation of the individual.

Limitations

Our current study was a registered report applying two different analytical approaches and six robustness analyses based on a large panel dataset combining all the relevant information about contraceptive use, covariates, and sexuality. By examining change over time within individuals, we reduced the number of assumptions necessary for causal identification and held many of the potential confounding factors that vary between individuals constant. In addition, the assumed underlying causal network was transparently outlined before performing analyses and the assumptions were questioned in the discussion. Therefore, our study aimed to provide an estimation of the effect of hormonal contraceptive use on sexuality in a naturalistic setting where women decide which contraceptive method to use.

Nevertheless, our study comes with a few limitations. First, the two modeling approaches we used to estimate the average treatment effect differed substantially in their estimations of the average treatment effect questioning the robustness of our results and limiting their generalizability. In future research focusing on identifying causal effects based on correlational data, combining different analytical approaches could be beneficial to determine a range of possible effects (and might additionally inform us about systematic differences between these analytical approaches). Second, our outcome variables were only

measured with one item each and often deviated from items in existing literature (especially in the case of desired sexual frequency as a measure of libido). Future research could try to implement more nuanced measures (e.g., solitary and dyadic libido; masturbation frequency) or measure sexuality in a more extensive way (e.g., sexual functioning, or using a diary method). And third, the time interval between observations (one year)~~waves~~ was potentially too long to capture all of the switches in contraceptive methods that happened between waves. A~~and a~~s we had no indicator of the time between switch and reporting of outcomes, we are unable to conclude whether the observed effects appear only shortly after changing contraceptive methods or after some time has passed. Diary studies (e.g., Blumenstock & Barber, 2022; Ott et al., 2008) investigating very short time frames found support for positive effects of hormonal contraceptive use on sexuality as well;~~. Therefore, we need~~ more research focusing on an intermediate time frame (e.g., one month, one menstrual cycle) would be helpful to understand the effects of hormonal contraceptive use on sexuality further.

Conclusion

The results of our current study support the conclusion that, on average, hormonal contraceptive use has beneficial effects on sexual frequency and sexual satisfaction. In addition, it provides evidence that the heterogeneity in women's sexual responses to hormonal contraceptive use is ~~very~~ high. Because women differ so much, there is promise for a tailored approach to contraception. This would mean that reliable estimates of individual treatment effects help women make informed decisions about which contraceptives have the desired safety and side effect profile for them. To obtain such reliable estimates, we need more granular data. Our study was based on a panel dataset and can be likened to a series of photographs — much can be missed in the moments we did not capture. Next, we should aim for movies, daily data on contraceptive use and sexuality over long time spans.~~This further supports the idea that contraceptive choice is an individual decision that needs to be made with information about available options (including~~

~~the option to switch contraceptive methods if women experience negative side effects). Our study, based on a panel dataset that functions like a series of photographs, raises the call for an even more granular, potentially daily longitudinal dataset (analogous to a movie) on contraceptive behavior and sexuality.~~

Acknowledgements

We would like to thank Tita Gonzalez Avilés, who executed our blind code on the PAIRFAM dataset before we received in-principal acceptance of our programmatic registered report stage 1. Gonzalez Avilés is an independent researcher and is not otherwise involved in this manuscript. We would also like to thank Zoe Rau for diligently reviewing the supportive HTML documents.

Author Contributions: CRediT Statement

Conceptualization: Laura J. Botzet (LJB), Julia M. Rohrer (JMR), Lars Penke (LP) and Ruben C. Arslan (RCA.); **Data curation:** LJB and RCA.; **Formal analysis:** LJB and RCA.; **Funding acquisition:** LP; **Investigation:** LJB and RCA.; **Methodology:** LJB, JMR, and RCA.; **Project administration:** LJB; **Software:** LJB and RCA.; **Supervision:** LP and RCA.; **Validation:** LJB; **Visualization:** LJB; **Writing – original draft:** LJB; **Writing - review & editing:** LJB, JMR, LP and RCA.

References

Ali, M. M., & Cleland, J. (2010). Contraceptive switching after method-related discontinuation: Levels and differentials. *Studies in Family Planning*, *41*(2), 129–133. https://doi.org/10.1111/J.1728-4465.2010.00234.X

Arslan, R.C. (2024, February 26). The reliability of multilevel parameters in Bayesian regressions. https://rubenarslan.github.io/posts/2024-02-15-the-reliability-of-multilevel-parameters-in-bayesian-regressions/

Arslan, R. C., Schilling, K. M., Gerlach, T. M., & Penke, L. (2021). Using 26,000 diary entries to show ovulatory changes in sexual desire and behavior. *Journal of Personality and Social Psychology*, *121*(2), 410–431. https://doi.org/10.1037/pspp0000208

Bafumi, J., & Gelman, A. (2007). Fitting multilevel models when predictors and group effects correlate. *Available at SSRN*. http://dx.doi.org/10.2139/ssrn.1010095

Boozalis, M. A., Tutlam, N. T., Robbins, C. C., & Peipert, J. F. (2016). Sexual desire and hormonal contraception. *Obstetrics and Gynecology*, *127*(3), 563–572. https://doi.org/10.1097/AOG.0000000000001286

Both, S., Lew-Starowicz, M., Luria, M., Sartorius, G., Maseroli, E., Tripodi, F., Lowenstein, L., Nappi, R. E., Corona, G., Reisman, Y. & Vignozzi, L. (2019). Hormonal contraception and female sexuality: position statements from the European Society of Sexual Medicine (ESSM). *The Journal of Sexual Medicine*, *16*(11), 1681–1695. https://doi.org/10.1016/j.jsxm.2019.08.005

Botzet, L. J., Gerlach, T. M., Driebe, J. C., Penke, L., & Arslan, R. C. (2021). Hormonal contraception and sexuality: Causal effects, unobserved selection, or reverse causality? *Collabra: Psychology*, *7*(1). https://doi.org/10.1525/collabra.29039

Botzet, L. J., Rohrer, J. M., Penke, L., & Arslan, R. C. (2023). Hormonal contraceptive use and women's sexuality and well-being: Estimating treatment effects and their heterogeneity based on longitudinal data [Stage 1 registered report]. *Peer Community In Registered Reports*. https://doi.org/10.17605/OSF.IO/KJ3H2

Brüderl, J., Drobnič, S., Hank, K., Neyer, F. J., Walper, S., Alt, P., Borschel, E., Bozoyan, C., Garrett, M., Geissler, S., Gonzalez Avilés, T., Gröpler, N., Hajek, K., Herzig, M., Huyer-May, B., Lenke, R., Lorenz, R., Lutz, K., Minkus, L.,... Wetzel, M. (2021). *The German Family Panel (pairfam)*. GESIS Data Archive, Cologne. ZA5678 Data file Version 12.0.0, https://doi.org/10.4232/pairfam.5678.12.0.0

Bundeszentrale für gesundheitliche Aufklärung (2018). *Verhütungsverhalten Erwachsener – Ergebnisse der Repräsentativbefragung [Contraceptive behavior of adults: Results of the representative study]*. ISBN 978-3-946692-64-5 https://shop.bzga.de/pdf/13317300.pdf

Burrows, L. J., Basha, M., & Goldstein, A. T. (2012). The effects of hormonal contraceptives on female sexuality: A review. *The Journal of Sexual Medicine*, *9*(9), 2213–2223. https://doi.org/10.1111/j.1743-6109.2012.02848.x

Blumenstock, S. M., & Barber, J. S. (2022). Hormonal contraception use and sexual frequency across young women's intimate relationships. *The Journal of Sex Research*, 1–14. https://doi.org/10.1080/00224499.2022.2059649

Caruso, S., Agnello, C., Intelisano, G., Farina, M., Di Mari, L., Sparacino, L., & Cianci, A. (2005). Prospective study on sexual behavior of women using 30 µg ethinylestradiol and 3 mg drospirenone oral contraceptive. *Contraception*, *72*(1), 19–23. https://doi.org/10.1016/j.contraception.2005.02.002

Davis, A. R., & Castaño, P. M. (2004). Oral contraceptives and libido in women. *Annual Review of Sex Research*, *15*(1), 297–320. https://doi.org/10.1080/10532528.2004.10559822

Draxler, C. N., Arslan, R. C., Penke, L., & Botzet, L. J. (2024). Contraceptive satisfaction but not sexual satisfaction or sexual frequency predicts contraceptive switches. *PsyArXiv*. https://doi.org/10.31234/osf.io/aczdm

Eng, P. M., Seeger, J. D., Loughlin, J., Clifford, C. R., Mentor Ba, S., & Walker, A. M. (2008). Supplementary data collection with case‐cohort analysis to address potential confounding in a cohort study of thromboembolism in oral contraceptive initiators matched on claims‐based propensity scores. *Pharmacoepidemiology and Drug Safety*, *17*(3), 297–305. https://doi.org/10.1002/pds.1554

Frye, C. A. (2006). An overview of oral contraceptives: mechanism of action and clinical use. *Neurology*, *66*, 29–36. https://doi.org/10.1212/WNL.66.66_suppl_3.S29

Fuentes, A., Lüdtke, O., & Robitzsch, A. (2021). Causal inference with multilevel data: A comparison of different propensity score weighting approaches. *Multivariate Behavioral Research*, 1–24. https://doi.org/10.1080/00273171.2021.1925521

Gaspard, U. J., Romus, M. A., Gillain, D., Duvivier, J., Demey-Ponsart, E., & Franchimont, P. (1983). Plasma hormone levels in women receiving new oral contraceptives containing ethinyl estradiol plus levonorgestrel or desogestrel. *Contraception*, *27*(6), 577–590. https://doi.org/10.1016/0010-7824(83)90023-9

Graham, C. A. (2019). The pill and women's sexuality. *BMJ*, *364*. https://doi.org/10.1136/bmj.l335

Graham, C. A., Ramos, R., Bancroft, J., Maglaya, C., & Farley, T. M. (1995). The effects of steroidal contraceptives on the well-being and sexuality of women: a double-blind, placebo controlled, two-centre study of combined and progestogen-only methods. *Contraception*, *52*(6), 363–369. https://doi.org/10.1016/0010-7824(95)00226-X

Gutman, R., & Rubin, D. B. (2017). Estimation of causal effects of binary treatments in unconfounded studies with one continuous covariate. *Statistical Methods in Medical Research*, *26*(3), 1199–1215. https://doi.org/10.1177/096228021557072

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Hamaker, E. L., & Muthén, B. (2020). The fixed versus random effects debate and how it relates to centering in multilevel modeling. *Psychological Methods*, *25*(3), 365–379. https://doi.org/10.1037/met0000239

Hill, S. E., & Mengelkoch, S. (2022). Moving Beyond the Mean: Promising Research Pathways to Support a Precision Medicine Approach to Hormonal Contraception. *Frontiers in Neuroendocrinology*, *68*. https://doi.org/10.1016/j.yfrne.2022.101042

Huinink, J.; Brüderl, J.; Nauck, B.; Walper, S.; Castiglioni, L.; Feldhaus, M (2011). Panel Analysis of Intimate Relationships and Family Dynamics (pairfam): Conceptual framework and design. *Zeitschrift für Familienforschung - Journal of Family Research*, *23*, 77–101. Retrieved from http://www.pairfam.de/fileadmin/user_upload/redakteur/publis/Dokumentation/Manuals/Concept%20Paper%20by%20Huinink%20et%20al_en%2C%20ZfF%202011.pdf

Kiesner, J. (2017). The menstrual cycle-response and developmental affective-risk model: A multilevel and integrative model of influence. *Psychological Review*, *124* (*2*), 215–244. https://doi.org/10.1037/rev0000058

King, G., & Zeng, L. (2006). The dangers of extreme counterfactuals. *Political Analysis*, *14*(2), 131–159. https://doi.org/10.1093/pan/mpj004

Lee, J. J. M., Low, L. L., & Ang, S. B. (2017). Oral contraception and female sexual dysfunction in reproductive women. *Sexual Medicine Reviews*, *5*(1), 31–44. https://doi.org/10.1016/j.sxmr.2016.06.001

Le Guen, M., Schantz, C., Régnier-Loilier, A., & de La Rochebrochard, E. (2021). Reasons for rejecting hormonal contraception in Western countries: A systematic review. *Social Science & Medicine*, *284*, 114247. https://doi.org/10.1016/j.socscimed.2021.114247

Lundberg, I., Johnson, R., & Stewart, B. M. (2021). What is your estimand? Defining the target quantity connects statistical evidence to theory. *American Sociological Review*, *86*(3), 532–565. https://doi.org/10.1177/00031224211004187

~~Lundin, C., Danielsson, K. G., Bixo, M., Moby, L., Bengtsdotter, H., Jawad, I., Marrions, L., Brynhildsen, J., Malmborg, A., Lindh, I., & Poromaa, I. S. (2017). Combined oral contraceptive use is associated with both improvement and worsening of mood in the different phases of the treatment cycle—a double-blind, placebo-controlled randomized trial. *Psychoneuroendocrinology*, *76*, 135–143. https://doi.org/10.1016/j.psyneuen.2016.11.033~~

Lundin, C., Malmborg, A., Slezak, J., Danielsson, K. G., Bixo, M., Bengtsdotter, H., Marions, L., Lindh, I., Theodorsson, E., Hammar, M., & Sundström-Poromaa, I. (2018). Sexual

function and combined oral contraceptives: a randomised, placebo-controlled trial. *Endocrine Connections*, *7*(11), 1208--1216. https://doi.org/10.1530/EC-18-0384

Læssøe, N. C., Wåhlin, S., Pedersen, A. T., Kristensen, E., & Giraldi, A. G. (2014). Combined hormonal contraception and women's sexual function: a cross-sectional pilot study in a cohort of Danish women. *Obstetrics and Gynecology: An International Journal*, *2014*, 1–13. https://doi.org/10.5171/2014.616630

Ly, A., Boehm, U., Heathcote, A., Turner, B. M., Forstmann, B., Marsman, M., & Matzke, D. (2017). A flexible and efficient hierarchical Bayesian approach to the exploration of individual differences in cognitive-model-based neuroscience. *Computational Models of Brain and Behavior*, 467–479. https://doi.org/10.1002/9781119159193.ch34

Malmborg, A., Persson, E., Brynhildsen, J., & Hammar, M. (2016). Hormonal contraception and sexual desire: A questionnaire-based study of young Swedish women. *The European Journal of Contraception & Reproductive Health Care*, *21*(2), 158–167. https://doi.org/10.3109/13625187.2015.1079609

McCoy, N. L., & Matyas, J. R. (1996). Oral contraceptives and sexuality in university women. *Archives of Sexual Behavior*, *25*(1). https://doi.org/10.1007/BF02437907

Mishell, D. R., Thorneycroft, I. H., Nakamura, R. M., Nagata, Y., & Stone, S. C. (1972). Serum estradiol in women ingesting combination oral contraceptive steroids. *American Journal of Obstetrics and Gynecology*, *114*(7), 923–928. https://doi.org/10.1016/0002-9378(72)90098-1

Oranratanaphan, S., & Taneepanichskul, S. (2006). A double blind randomized control trial, comparing effect of drospirenone and gestodene to sexual desire and libido. *Journal of the Medical Association of Thailand*, *89*(4), 17–22. Retrieved from

https://pdfs.semanticscholar.org/f426/e6cac8a255632df21482b25e09a7a6350876.pdf

Ott, M. A., Shew, M. L., Ofner, S., Tu, W., & Fortenberry, J. D. (2008). The influence of hormonal contraception on mood and sexual interest among adolescents. *Archives of Sexual Behavior*, *37*(4), 605–613. https://doi.org/10.1007/s10508-007-9302-0

Pastor, Z., Holla, K., & Chmel, R. (2013). The influence of combined oral contraceptives on female sexual desire: a systematic review. *The European Journal of Contraception & Reproductive Health Care*, *18*(1), 27–43. https://doi.org/10.3109/13625187.2012.728643

Riggs, M., Klebanoff, M., Nansel, T., Zhang, J., Schwebke, J., & Andrews, W. (2007). Longitudinal association between hormonal contraceptives and bacterial vaginosis in women of reproductive age. *Sexually Transmitted Diseases*, *34*(12), 954–959. https://doi.org/10.1097/OLQ.0b013e31811ed0e4

~~Robins, J., Sued, M., Lei-Gomez, Q., & Rotnitzky, A. (2007). Comment: Performance of double-robust estimators when" inverse probability" weights are highly variable. *Statistical Science*, *22*(4), 544–559. https://www.jstor.org/stable/27645860~~

Rohrer, J. M., & Lucas, R. E. (2020). Causal effects of well-being on health: It's complicated. *Preprint on PsyArXiv*. https://doi.org/10.31234/osf.io/wgbe4

Rohrer, J. M., & Murayama, K. (2021). These are not the effects you are looking for: Causality and the within-/between-person distinction in longitudinal data analysis. *Preprint on PsyArXiv*. https://doi.org/10.31234/osf.io/tg4vj

Rosenbaum, P. R., & Rubin, D. B. (1983). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal*

Statistical Society: Series B (Methodological), *45*(2), 212–218. https://doi.org/10.1111/j.2517-6161.1983.tb01242.x

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*(3), 169–188. https://doi.org/10.1023/A:1020363010465

Sabatini, R., & Cagiano, R. (2006). Comparison profiles of cycle control, side effects and sexual satisfaction of three hormonal contraceptives. *Contraception*, *74*(3), 220–223. https://doi.org/10.1016/j.contraception.2006.03.022

Schaffir, J. (2006). Hormonal contraception and sexual desire: a critical review. *Journal of Sex & Marital Therapy*, *32*(4), 305–314. https://doi.org/10.1080/00926230600666311

Schleifenbaum, L., Driebe, J. C., Gerlach, T. M., Penke, L., & Arslan, R. C. (2021). Women feel more attractive before ovulation: evidence from a large-scale online diary study. *Evolutionary Human Sciences*, *3*. https://doi.org/10.1017/ehs.2021.44

Skovlund, C. W., Mørch, L. S., Kessing, L. V., & Lidegaard, Ø. (2016). Association of hormonal contraception with depression. *JAMA Psychiatry*, *73*(11), 1154–1162. https://doi.org/10.1001/jamapsychiatry.2016.2387

Smith, N. K., Jozkowski, K. N., & Sanders, S. A. (2014). Hormonal contraception and female pain, orgasm and sexual pleasure. *The Journal of Sexual Medicine*, *11*(2), 462–470. https://doi.org/10.1111/jsm.12409

Stan Development Team (2024). *Stan Reference Manual* (2.34) [Computer software]. https://mc-stan.org

Stomati, M., Genazzani, A. D., Petraglia, F., & Genazzani, A. R. (1998). Contraception as prevention and therapy: sex steroids and the brain. *The European Journal of*

Contraception & Reproductive Health Care, *3*(1), 21–28. https://doi.org/10.3109/13625189809167481

Strufaldi, R., Pompei, L. M., Steiner, M. L., Cunha, E. P., Ferreira, J. A., Peixoto, S., & Fernandes, C. E. (2010). Effects of two combined hormonal contraceptives with the same composition and different doses on female sexual function and plasma androgen levels. *Contraception*, *82*(2), 147–154. https://doi.org/10.1016/j.contraception.2010.02.016

Taggart, T. C., Eaton, N. R., Keyes, K. M., Hammett, J. F., & Ulloa, E. C. (2018). Oral contraceptive use is associated with greater mood stability and higher relationship satisfaction. *Neurology, Psychiatry and Brain Research*, *30*, 154–1162. https://doi.org/10.1016/j.npbr.2018.10.004

Thoemmes, F., & Ong, A. D. (2016). A primer on inverse probability of treatment weighting and marginal structural models. *Emerging Adulthood*, *4*(1), 40–59. https://doi.org/10.1177/2167696815621645

United Nations, Department of Economic and Social Affairs, Population Division (2022). *World Family Planning 2022. Meeting the changing needs for family planning: Contraceptive use by age and method (UN DESA/POP/2022/TR/NO. 4.)*. https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/files/documents/2023/Feb/undesa_pd_2022_world-family-planning.pdf

VanderWeele, T. J., & Ding, P. (2017). Sensitivity analysis in observational research: introducing the E-value. *Annals of Internal Medicine*, *167*(4), 268–274. https://doi.org/10.7326/M16-2607

VanderWeele, T. J., Mathur, M. B., & Chen, Y. (2020). Outcome-wide longitudinal designs for causal inference: a new template for empirical studies. *Statistical Science*, *35*(3), 437–466. https://doi.org/10.1214/19-STS728

Wallwiener, C. W., Wallwiener, L. M., Seeger, H., Schönfisch, B., Mueck, A. O., Bitzer, J., Zipfel, S., Brucker, S. Y., Taran, F.-A., & Wallwiener, M. (2015). Are hormonal components of oral contraceptives associated with impaired female sexual function? A questionnaire-based online survey of medical students in Germany, Austria, and Switzerland. *Archives of Gynecology and Obstetrics*, *292*(4), 883–890. https://doi.org/10.1007/s00404-015-3726-x

Wallwiener, M., Wallwiener, L. M., Seeger, H., Mueck, A. O., Zipfel, S., Bitzer, J., & Wallwiener, C. W. (2010). Effects of sex hormones in oral contraceptives on the female sexual function score: a study in German female medical students. *Contraception*, *82*(2), 155–159. https://doi.org/10.1016/j.contraception.2009.12.022

Wang, Q., Würtz, P., Auro, K., Morin-Papunen, L., Kangas, A. J., Soininen, P., Tiainen, M., Tynkkynen, T., Joensuu, A., Havulinna, A. S., Aalto, K., Salmi, M., Blankenberg, S., Zeller, T., Viikari, J., Kähönen, M., Lehtimäki, T., Salomaa, V., Jalkanen, S.,... Ala-Korpela, M. (2016). Effects of hormonal contraception on systemic metabolism: cross-sectional and longitudinal evidence. *International Journal of Epidemiology*, *45*(5), 1445–1457. https://doi.org/10.1093/ije/dyw147

Watkins, E. S. (2012). How the pill became a lifestyle drug: the pharmaceutical industry and birth control in the United States since 1960. *American Journal of Public Health*, *102*(8), 1462–1472. https://doi.org/10.2105/AJPH.2012.300706

Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *65*(1), 95–114. https://doi.org/10.1111/1467-9868.00374

Zethraeus, N., Dreber, A., Ranehill, E., Blomberg, L., Labrie, F., von Schoultz, B., Johannesson, M., & Hirschberg, A. L. (2016). Combined oral contraceptives and sexual function in women - a double-blind, randomized, placebo-controlled trial. *The Journal of Clinical Endocrinology & Metabolism*, *101*(11), 4046–4053. https://doi.org/10.1210/jc.2016-2032

Zethraeus, N., Dreber, A., Ranehill, E., Blomberg, L., Labrie, F., von Schoultz, B., Johannesson, M., & Hirschberg, A. L. (2017). A first-choice combined oral contraceptive influences general well-being in healthy women: a double-blind, randomized, placebo-controlled trial. *Fertility and Sterility*, *107*(5), 1238–1245. https://doi.org/10.1016/j.fertnstert.2017.02.120

Zimmerman, Y., Eijkemans, M. J. C., Coelingh Bennink, H. J. T., Blankenstein, M. A., & Fauser, B. C. J. M. (2014). The effect of combined oral contraception on testosterone levels in healthy women: a systematic review and meta-analysis. *Human Reproduction Update*, *20*(1), 76–105. https://doi.org/10.1093/humupd/dmt038