Vaccine Hesitancy and Differential Susceptibility to Media Coverage: A Critical Documentary Led to Substantial Reductions in Human Papillomavirus Vaccine Uptake in Denmark*

Maria Knoth Humlum, Ph.D. (corresponding author)

mhumlum@econ.au.dk

Department of Economics and Business Economics, Aarhus University

Niels Skipper, Ph.D.

nskipper@econ.au.dk

Department of Economics and Business Economics, Aarhus University

Peter Rønø Thingholm, Ph.D.

pthingholm@econ.au.dk

Department of Economics and Business Economics, Aarhus University

Word count: 4,155

* *Financial support for this study was provided by the Independent Research Fund Denmark (DFF | Social Sciences), grant no. 8106-00013B, and the Carlsberg Foundation, grant no. CF17-0408. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.*

ABSTRACT

Objectives: To investigate whether negative media coverage of the human papillomavirus (HPV) vaccine led to a decrease in the uptake of the first dose of the HPV vaccine (HPV1) in Denmark, and importantly, whether some groups of individuals were more susceptible to negative media coverage.

Methods: We measured HPV vaccine uptake of 12-year old girls born in 2001 to 2004 using Danish administrative data. A quasi-experimental design was employed to assess whether a documentary that was critical of the HPV vaccine and aired in March 2015 affected HPV uptake.

Results: The documentary led to a quick and substantial decrease in the monthly propensity to vaccinate, which dropped three percentage points---or about 50%---in response to the documentary. Responses differed substantially across subgroups and girls from families with high socioeconomic status (SES) were more susceptible to the negative media coverage.

Conclusions: Susceptibility to negative media coverage varied substantially across subgroups, highlighting the need for policy makers to appropriately target and differentiate initiatives to improve vaccine compliance rates.

Keywords: Media coverage; vaccine uptake; HPV; vaccine hesitancy.

Introduction

Vaccine hesitancy was placed on the World Health Organization (WHO) list of top ten threats to global health in 2019. Despite wide availability of vaccines in large parts of the world, many individuals choose to go unvaccinated or to have their children go unvaccinated. Even in developed countries, parental concerns about vaccine safety remain, and the vaccination rates are often below the target levels set by national health organizations.[1] Vaccine hesitancy and specifically concerns about vaccine safety and a lack of knowledge about vaccines has been identified as key barriers to HPV vaccination.[1-2]

The HPV vaccine was officially incorporated in the Danish Childhood Vaccination Program for girls born 1996 or later. Initiation of HPV vaccination was recommended for girls at age 12. All vaccines in the Danish Childhood Vaccination Program are administered by local general practitioners. The vaccines and the administration hereof are free-of-charge. In Denmark, the compliance rate with the Childhood Vaccination Program has traditionally been high, but compliance with respect to the human papillomavirus (HPV) vaccine has recently deteriorated. In 2017, about 90% of girls in the 1996-2001 birth cohorts had received the first HPV shot---compared to about 60% for the 2004 birth cohort.[3] It is widely believed that this reduced compliance is related to a substantial amount of negative publicity received by the HPV vaccine in the Danish national media. Earlier evidence has also established an association between the media coverage of the HPV vaccine and HPV uptake[4] or argued for an association[5]. Associations have also been found between media coverage and vaccine uptake in other contexts, most notably the measles, mumps and rubella (MMR) vaccine.[6-8]

On March 26, 2015, a documentary with the title “The vaccinated girls” aired on national television in Denmark. The documentary was very critical of the HPV vaccine and the national health authorities. It included segments with girls who suffered from unexplainable pain and fatigue. Some girls had symptoms resembling postural orthostatic tachycardia syndrome (POTS) which were hypothesized to be caused by the vaccine. These postulated side effects of the HPV vaccine were unsubstantiated by clinical evidence at the time, and later studies have not found a link between the HPV vaccine and these potential side effects.[9-11]

While earlier studies have documented an association between media coverage and vaccine uptake, causal evidence is rare. To establish causality is challenging in these types of settings where typically only observational data are available. One study used aggregate data to establish a structural break in the time series of number of HPV vaccinations administered at the time of the documentary.[12] We improved on earlier studies by using Danish administrative individual-level data covering entire cohorts of young girls to establish a causal relationship between HPV uptake in Denmark and media coverage (in this case the airing of a documentary that was critical of the HPV vaccine and led to substantial media coverage and an intense public debate).

First, we took advantage of vaccine eligibility rules and quasi-experimental variation in the timing of the documentary to estimate effects of the documentary on the uptake of the HPV1 vaccine shot. We were able to credibly identify and isolate the effect of one major media event and its repercussions on the individual vaccine uptake of 12-year-old girls.

Second, in order to determine whether some groups of individuals were more affected by the documentary than others, we estimated the response to the documentary of different subgroups of individuals. It has previously been shown that high-SES parents respond more strongly to media coverage in the context of vaccines.[6] From the perspective of policy makers, it is important to know whether adverse media coverage can impact vaccine uptake, but also whether specific groups of individuals are affected by media coverage since the latter can help policy makers appropriately target initiatives to increase vaccination rates.

Methods

We investigated the effect of the documentary on HPV1 uptake using Danish administrative registers hosted by Statistics Denmark. The administrative registers cover the entire Danish population and provide detailed information about all Danish residents including sociodemographic characteristics and physician contacts. All Danish residents have a unique personal identifier which enabled us to link the different registers and follow individuals over time.

For the purpose of this analysis, we had access to data at the individual level on the month of vaccine uptake, exact date of birth and a range of demographic variables. In addition, we were able to link individuals with their parents and siblings. This rich data set allowed us to estimate the effect of the documentary and to investigate whether there were differential responses across groups of girls with different sociodemographic characteristics at the same time.

EMPIRICAL STRATEGY One major issue in observational studies of the relationship between media coverage and vaccine uptake is the existence of potential confounders that may lead to a spurious relationship between media coverage and vaccine uptake. While some confounders can be controlled for, other confounders may be unobserved. Especially, there may be (unobserved) variables that move in the same direction as media coverage and vaccine uptake in the analysis of aggregate data, making it challenging to draw causal inference.

We used a Regression Discontinuity Design (RDD)---a quasi-experimental study design which offers researchers the possibility to draw causal inference in cases where treatment is a discontinuous function of another variable (the forcing variable).[13] At a specific threshold of the forcing variable, the probability of treatment jumps discontinuously. A discontinuous jump or drop in the probability of the outcome variable at the threshold captures the treatment effect (at least locally). As a study design based on quasi-randomized variation in treatment, RDD has substantial benefits compared to standard multivariate regression analysis and is increasingly used in many fields of research.[14-16]

We exploited the fact that a documentary that was critical of the HPV vaccine aired at the end of March 2015. The Danish childhood vaccination program recommends the HPV vaccine for girls turning 12. Exposure to the documentary was jointly determined by the timing of the documentary and date of birth which determined eligibility. Vaccination decisions made before April 2015 were unaffected by the documentary (unexposed). Vaccination decisions in April 2015 or later were potentially affected by the documentary (exposed).

We applied the RDD to a monthly panel data set. The first month of observation was the first month of eligibility - the month of the girl's 12th birthday. Each girl was observed a maximum of 12 times in the panel corresponding to the year following the girl's 12th birthday. In a given month a girl was unexposed or exposed depending on whether or not the month of observation was before or after the airing of the documentary.

We investigated whether or not the airing of the documentary induced a discontinuity in vaccination uptake (HPV1) where the forcing variable was the distance from the time of the airing of the documentary to the month of observation. April 2015 was the first post-documentary month and the forcing variable took the value zero for this month. To estimate the direct effect of the documentary as well as any indirect effects occurring due to increased media attention, public debate etc., we did not include April 2015 in the analysis. This type of analysis is sometimes referred to as a "donut" RDD.[17]

Absent other factors that changed at exactly the same point in time, the regression discontinuity estimate was interpreted as the effect of the documentary on HPV1 vaccine uptake. Also, a crucial assumption in the RDD was that girls who were eligible in the months prior to the documentary and girls who were eligible in the months following the documentary were similar in terms of both observed and unobserved confounders. Since eligibility was uniquely determined by date of birth, a violation of the assumption would occur only if girls could themselves determine their date of eligibility.

For the purpose of establishing the effect of the documentary on the vaccination rate in the first year of eligibility, we compared a cross-section of a cohort of girls that were unexposed to the documentary with a cross-section of a cohort of girls that were fully exposed to the documentary. In the context of this retrospective cohort study, we applied machine learning methods to estimate treatment effects for different subgroups of individuals. Specifically, we used nonparametric causal forests to estimate conditional average treatment effects (CATEs).[18-19] A brief and informal description of this methodology can be found in Appendix B.

DEFINITION OF KEY VARIABLES We observed the timing of the HPV1 vaccine shots in the administrative data at the monthly level. Based on this, we defined the dependent variable in our analysis as an indicator variable taking the value 1 if the girl had received the HPV1 shot in the month of observation or before and 0 otherwise. This outcome variable was used in the RDD analysis with panel data. To measure uptake in the retrospective cohort study, we defined an indicator variable equal to 1 if a girl received HPV1 within 12 months of her 12th birthday and 0 otherwise.

The key independent variable in the analysis was an indicator for whether or not the girl-month observation was exposed to the documentary or not. In the RDD analysis, the indicator for exposure took the value 1 if the month of observation was in April 2015 or later and 0 if the month of observation was before April 2015. Similarly, in the retrospective cohort study, the treatment indicator took the value 1 if an individual was exposed to the documentary within the first 12 months of eligibility for the vaccine and 0 otherwise.

The control variables included child and family characteristics as well as maternal characteristics. We constructed indicator variables for birth order (first, second, or third or higher), whether or not a girl had an older sister, whether or not a girl had an older sister who received the HPV vaccine, whether or not one of the girl's parents was a health professional, maternal age (<35, 35-40, 40-45, >45), maternal marital status, maternal country of origin (native, immigrant, descendant), maternal highest completed education (at most high school, vocational, bachelor, master), maternal income quartile and whether or not the mother had any labor market income.

STATISTICAL ANALYSIS We constructed a monthly panel data set of girls born in 2001 to 2004 who resided in Denmark in the year of their 12th birthday.

For the purpose of the RDD individual panel data analysis, we focused on girls that became eligible in a 3-year window (+/- 1.5 year) around the time of the documentary, i.e., in the period from October 2013 to October 2016. To estimate the effect of the documentary on uptake of HPV1 using an RDD, we estimated a linear probability model and included the exposure indicator, the distance from the documentary (the forcing variable) and an interaction of the two as regressors. The interaction allowed for the slopes to differ on either side of April 2015. The coefficient on the exposure indicator captured the effect of the documentary on HPV1 uptake. In this type of RDD analysis where the running variable is time, standard inference is not sufficient.[20] Standard errors were clustered at the individual level to account for the likely correlation in the error terms due to the panel data setup.

For the purpose of estimating the effect of the documentary on the yearly vaccination uptake, we only used the cross-sectional information available and excluded girls that were both exposed and unexposed during the 12 months after becoming eligible. This enabled the comparison of girls that were either exposed or unexposed during their entire first year of eligibility for the vaccine. We modeled the relationship between exposure to the documentary (treatment) and uptake of HPV1 during the first year of eligibility as a linear probability model.

Both models were estimated using Ordinary Least Squares. The control variables included child and family characteristics (birth order, older sister (0/1), older sister received HPV1 (0/1), either parent was a health professional (0/1), maternal age, maternal marital status, maternal country of origin, maternal education level, maternal income).

To assess the robustness of our results, we performed a number of auxiliary analyses. First, we modeled the dichotomous outcome using a logistic regression model to ensure that the results were not driven by the functional form of the chosen model. Second, we made an alternative specification of the linear probability model that did not include any child and family characteristics as control variables. For the RDD analysis, the inclusion or exclusion of control variables should not affect the results much. Third, we conducted the RDD analysis with two alternative bandwidths; that is we restricted the sample to vaccination rates between April 2014 and April 2016 or January 2015 and July 2015, respectively, to check that the choice of bandwidth was not driving the results. Fourth, we conducted the RDD analysis including April 2015. Fifth, a difference-in differences specification was estimated using the window around April 2013 as a comparison group. This was done to alleviate any concern of residual confounding from time-varying factors (e.g., season-of-birth or calendar month effects) that can be present in RDD in time analyses.[20] Finally, as a last sensitivity analysis, we graphed the fraction of children aged 0-5 who were up-to-date with their vaccination in the Childhood Vaccination Program over time around the airing of the documentary. This was done as an informal investigation to see if the documentary had any spillovers to other vaccination programs or led to vaccine hesitancy in general (see Appendix Exhibits A2-A4 and A7 for more details on the robustness checks).

SAS version 9.4 was used for data management and Stata SE version 15.1 along with R version 3.6 was used for the statistical analysis.

Results

For the RDD analysis, the monthly panel consisted of 1,195,440 person-month observations representing 129,569 girls who were living in Denmark and turned 12 from November 2012 to October 2016. For the cohort study, we constructed a cross-section of a subset of these girls (N=65,053) that were either entirely exposed (turned 12 from May 2015 to April 2016) or entirely unexposed (turned 12 from May 2013 to April 2014) during their first year of eligibility for the HPV vaccine. Exhibit 1 provides a summary of the characteristics of the girls in the two samples split by their exposure to the documentary. Girls that were not exposed to the documentary had much higher HPV initiation rates (70-80%) than girls that were exposed to the documentary (30-45%). Girls in the RDD sample and the cohort sample exhibited fairly similar characteristics. For the unexposed girls in the RDD sample, 37 percent were first-born, 15 percent had an older sister who received the HPV vaccine and about 2 percent had a parent who was a health professional. 39 percent of the mothers were between 40 and 45 years old, 84 percent were native Danes, and about 40 percent of the mothers had completed either a bachelor's or master's degree, while 81 percent of the mothers had an income. We performed balance tests based on the covariates for the retrospective cohort study (see Exhibit 1) and we concluded that the exposed and unexposed girls were fairly similar in terms of observed characteristics.

EFFECTS OF DOCUMENTARY ON HPV UPTAKE To illustrate the effect of the documentary on vaccination uptake, we plotted the average monthly vaccination uptake (HPV1) by distance from the time of the documentary along with fitted probabilities from a linear probability model (Exhibit

2). Prior to April 2015, the monthly vaccination rate was relatively stable at a level of about 6 percentage points (slope=0.04 pp; 95% CI=[-0.06 ; 0.14]). In April 2015---after the airing of the documentary and while the ensuing public debate was soaring---the monthly vaccination rate was 4.8 percentage points. After April 2015, the monthly vaccination rate stabilized at a new level of about 3 percentage points (with a slightly decreasing trend, slope=-0.04 pp; 95% CI=[-0.08 ; 0.00]). Graphically, the estimated treatment effect corresponded to the vertical distance between the fitted regression lines at the discontinuity point. We found an almost immediate drop of about 3 percentage points---or 50%---in the monthly vaccination uptake from March 2015 to May 2015.

The aim of the study was to estimate both the direct effect of the documentary as well as any indirect effects occurring due to increased media attention, public debate etc. Therefore, we excluded 31,796 person-period observations in April 2015 from the analysis. Formal RDD regression confirmed the graphical result and the estimated drop in the monthly propensity to vaccinate in response to the documentary was 3.5 percentage points, 95% CI=[-3.4,-3.6]. We included the following independent variables in the RDD regression: the exposure indicator, the forcing variable and an interaction of the two, see Appendix Exhibit 1 for full regression results. The estimated drop was statistically significant at the 1 percent level. The results were not sensitive to the inclusion of covariates.

Turning to the cohort sample, the estimated drop in the yearly propensity to vaccinate was 44.7 percentage points, 95% CI=[-46.4,-43.2], see Appendix Exhibit A1 for full regression results. This constituted a drop of 58.2% compared to the pre-documentary level of 76.8 percentage points. Although the estimated drop in the yearly propensity to vaccinate was based on a slightly different sample and estimation method than the estimated drop in the monthly propensity to vaccinate, the two approaches yielded fairly similar results, since a monthly drop of 3.5 percentage points corresponded to a yearly drop of 41.9 percentage points, 95% CI=[-40.3,-43.5].

The results from the sensitivity analysis using a difference-in-differences design showed a drop in the monthly propensity to vaccinate of 3.2 percentage points, 95% CI=[-3.0;-3.4] (See Appendix Exhibit A2 for a graphical representation and Appendix Exhibit A3 for the estimation results).

The graph showing the fraction of children aged 0-5 who were up-to-date with the Childhood Vaccination program is shown in Appendix Exhibit A4. There is no visible evidence that the take-up of other vaccines declined following April 2015.

SUBGROUP RESPONSES TO THE DOCUMENTARY Based on the estimated drop in the yearly propensity to vaccinate, we identified three groups of individuals according to their susceptibility to the documentary (Exhibit 3). Group 1 (48% of sample) consisted of individuals who were estimated to be highly responsive to the documentary in the sense that their estimated Conditional Average Treatment Effects (CATEs) were numerically large. The mean exposure effect in group 1 was -.51. Correspondingly, groups 2 (34% of sample) and 3 (18% of sample) consisted of individuals who were less responsive to the documentary with mean exposure effects of -.41 and -.27, respectively (Appendix Exhibit A5).

We compared the characteristics of individuals in groups 1 (high susceptibility) and 3 (low susceptibility) to establish the characteristics of individuals who were more or less susceptible to the influence of the documentary (Exhibit 4). We found evidence of substantial heterogeneity in the effects of the documentary. For example, first-born girls were 13 percentage points more likely to have high versus low susceptibility. Other groups of girls that were found to have high (versus low) susceptibility are second-born girls, girls with older mothers (30 and older), girls with mothers with vocational or higher education, and girls with higher-income mothers. On the other hand, we found that for immigrant girls the probability of having high susceptibility is about 40 percentage points *lower* than the probability of having low susceptibility. For second-generation immigrants the corresponding number was about 10 percentage points. The only other characteristics (out of the characteristics that we included in the analysis) that we found to be associated with low susceptibility to the documentary was whether or not the girl had an older sister and if either parent was a health professional. If a girl had an older sister who did not receive the vaccine, the probability that the girl had high (versus low) susceptibility to the documentary was about 30 percentage points lower, and having a health professional as a parent was associated with a 5 percentage point reduction in the probability of having high compared to low susceptibility. In comparison, having an older sister who already received the HPV vaccine was associated with 7 percentage point increase in the probability of having high compared to low susceptibility. The full set of estimated coefficients are reported in Appendix Exhibit A6.

Discussion

We documented a sharp decrease in the monthly HPV1 vaccination rate of about 3 percentage points corresponding to an almost 50% reduction in the monthly vaccination rate around the time of the airing of a vaccine-skeptical documentary on Danish national television. We found evidence of substantial heterogeneity in the individual responses to the negative information shock that the documentary constituted.

The groups of individuals that we found to be highly susceptible to the documentary were to a large extent groups that have been identified as having high uptake of vaccine in earlier studies, for example girls of lower birth order and girls from advantaged families in terms of education and income.[5] At the same time girls with immigrant background---who had relatively low compliance rates at baseline---were less susceptible to the negative information shock provided by the documentary. This is in line with earlier studies that have found that parents with high socioeconomic status are more responsive to media coverage.[6,21] Girls with an older sister who received the HPV vaccine were more susceptible to the negative information shock. While this may seem counterintuitive, having an older sister who received the vaccine is positively correlated with SES, i.e., other factors that also predict higher susceptibility.

This study highlights the vulnerability of even high-coverage vaccine programs to vaccine-skeptical media coverage. While our results do not suggest spillovers of the media coverage of the HPV vaccine to vaccines for younger children, spillovers to the concurrent MMR booster at age 12 have previously been documented[22]. The results underpin the need for national health authorities to not only design information campaigns when compliance rates are low, but also to think about preemptive information strategies to handle these types of negative information shocks. Earlier studies suggest that information campaigns are relatively ineffective in promoting vaccine compliance[23] and that existing knowledge about effective informational campaigns is scarce[24].

While a range of efforts---including widespread informational campaigns---have been made by the Danish health authorities, HPV compliance rates are still below 80% for the 2003-2004 birth cohorts and have thus not returned to their pre-documentary levels.[25] In the spring of 2015, reminders to parents were introduced in the Childhood Vaccination Program. While these reminders have proven to be effective in increasing compliance rates in general[26], even the combination of reminders and informational campaigns have not been sufficient in terms of recuperating HPV compliance for the affected cohorts.[27]

In order to appropriately target and design any initiatives including informational campaigns, policy makers need to identify at-risk groups, i.e., groups that are particularly susceptible to a negative information shock. Importantly, this study documents that the at-risk group in terms of susceptibility to media coverage and the at-risk group in terms of baseline compliance are inherently different. Consequently, appropriate strategies in terms of improvement of baseline compliance are not necessarily appropriate strategies to improve vaccine compliance in the wake of negative media coverage.

LIMITATIONS The study was subject to the following limitations. First, uptake of HPV1 was measured based on administrative data on services delivered by physicians that are covered by public health insurance. To the extent that some individuals may have gotten the vaccine from other sources or physicians failed to report a vaccination, we will underestimate the vaccination rates. However, it has previously been assessed that the coverage of the register is high.[28]

Second, we attributed the large decline in the HPV1 vaccination rate to the airing of the documentary and the ensuing public debate. While it is possible that another event affecting HPV1 uptake took place at exactly the same time, we rendered this unlikely given the almost immediate and very sharp decline in HPV1 uptake between March 2015 and May 2015. The documentary has also been highlighted as a central factor in previous studies.[4-5]

Third, in the retrospective cohort study, it was a concern that unobserved factors affecting HPV vaccine uptake may differ across the cohorts. However, our analysis controlled for a range of control variables and the two cohorts were tightly spaced. Further, the differences-in-differences analysis showed almost identical results to those of the RDD analysis. Thus, omitted variable bias was assessed to be negligible.

Conclusion

Vaccine hesitancy has been identified as a key barrier to vaccination, and ensuring public trust in vaccines is therefore central to obtain desired vaccine compliance rates. We found that a vaccine-skeptical documentary and the ensuing public debate led to substantial decreases in HPV vaccine uptake in Denmark. The response to the documentary in terms of uptake of HPV vaccine uptake was highly heterogeneous across subgroups. These results underscore the need for policymakers to appropriately target and differentiate initiatives to improve vaccine compliance rates.

Acknowledgements

This study was funded by the Independent Research Fund Denmark (DFF | Social Sciences), grant no. 8106-00013B. Support for research infrastructure for this study was provided by the Carlsberg Foundation, grant no. CF17-0408. The study was approved by the Danish Data Protection Agency. This approval constitutes the necessary legal requirements, and informed consent from participants is not required. The authors would like to thank Julie Cullen, Dorte Rytter and Agne Suziedelyte for comments on an earlier version of the paper.

Data availability statement

The data used in this study are administrative data hosted by Statistics Denmark. Access to personal data is restricted. The authors can facilitate data access with permission of Statistics Denmark upon reasonable request. Data access can be obtained from Aarhus only.

Conflict of interests

The authors declare that there is no conflict of interest.

REFERENCES

1. Omer SB, Salmon DA, Orenstein WA, deHart MP, Halsey N. Vaccine Refusal, Mandatory Immunization, and the Risks of Vaccine-Preventable Diseases. *The New England Journal of Medicine.* 2009; 360(19): 1981-1988.

2. Holman DM, Benard V, Roland KB, Watson M, Liddon N, Stokley S. Barriers to Human Papillomavirus Vaccination Among US Adolescents: A Systematic Review of the Literature. *JAMA Pediatr.* 2014;168(1):76–82.
3. Sundhedsstyrelsen. Børnevaccinationsprogrammet. Årsrapport 2017. *Danish Health Authority*. 2018.
4. Suppli CH, Hansen ND, Rasmussen M, Valentiner-Branth P, Krause TG, Mølbak K. Decline in HPV-vaccination uptake in Denmark – the association between HPV-related media coverage and HPV-vaccination. *BMC Public Health*. 2018; 18:1360: 1-8.
5. Amdisen L, Kristensen ML, Rytter D, Mølbak K, Valentiner-Branth P. Identification of determinants associated with uptake of the first dose of the human papillomavirus vaccine in Denmark. *Vaccine*. 2018; 36: 5747-5753.
6. Anderberg D, Chevalier A, Wadsworth J. Anatomy of a health scare: Education, income and the MMR controversy. *Journal of Health Economics*. 2011; 30: 515-530.
7. Chang LV. Information, education, and health behaviors. Evidence from the MMR vaccine autism controversy. *Health Economics*. 2018; 27(7): 1043-1062.
8. Hansen ND, Mølbak K, Cox IJ, Lioma C. Relationship Between Media Coverage and Measles-Mumps-Rubella (MMR) Vaccination Uptake in Denmark: Retrospective Study. *JMIR Public Health Surveill*. 2019; 5(1):e9544.
9. Feiring B, Laake I, Bakken IJ, Greve-Isdahl M, Wyller VB, Håberg SE, Magnus P, Trogstad L. HPV vaccination and risk of chronic fatigue syndrome/myalgic encephalomyelitis: A nationwide register-based study from Norway. *Vaccine*. 2017;35:4203-4212.
10. Hviid A, Svanström H, Scheller NM, Grönlund O, Pasternak B, Arnheim-Dahlström, L. Human papillomavirus vaccination of adult women and risk of autoimmune and neurological diseases. *J Intern Med.* 2018;283:154-165.

11. Arbyn M, Xu L, Simoens C, Martin-Hirsch PPL. Prophylactic vaccination against human papillomaviruses to prevent cervical cancer and its precursors. Cochrane Database of Systematic Reviews. 2018; 5 (CD009069).

12. Hansen PR, Schmidtblaicher M. A Dynamic Model of Vaccine Compliance: How Fake News Undermined the Danish HPV Vaccine Program. *Journal of Business & Economic Statistics*. Forthcoming DOI: 10.1080/07350015.2019.1623045.

13. Lee DS, Lemieux T. Regression Discontinuity Designs in Economics. *Journal of Economic Literature*. 2010; 48(2): 281-355.

14. Smith LM, Strumpf EC, Kaufman JS, Lofters A, Schwandt M, Lévesque LE. The Early Benefits of Human Papillomavirus Vaccination on Cervical Dysplasia and Anogenital Warts. *Pediatrics*. 2015; 135:e1131.

15. Dague L. The effect of Medicaid premiums on enrollment: A regression discontinuity approach. *Journal of Health Economics*. 2014;37:1-12.

16. Humlum MK, Vejlin RM. The responses of youth of a cash transfer conditional on schooling: A quasi-experimental study. *Journal of Applied Econometrics*. 2013;28:628-649.

17. Barreca A, Guldi M, Lindo J, Waddell, G. Saving babies? Revisiting the effect of very low birth weight classification. *The Quarterly Journal of Economics.* 2011;126(4):2117-2123.

18. Athey S, Tibshirani J, Wager S. Generalized Random Forests. *Annals of Statistics.* 2019;47(2):1148-1178.

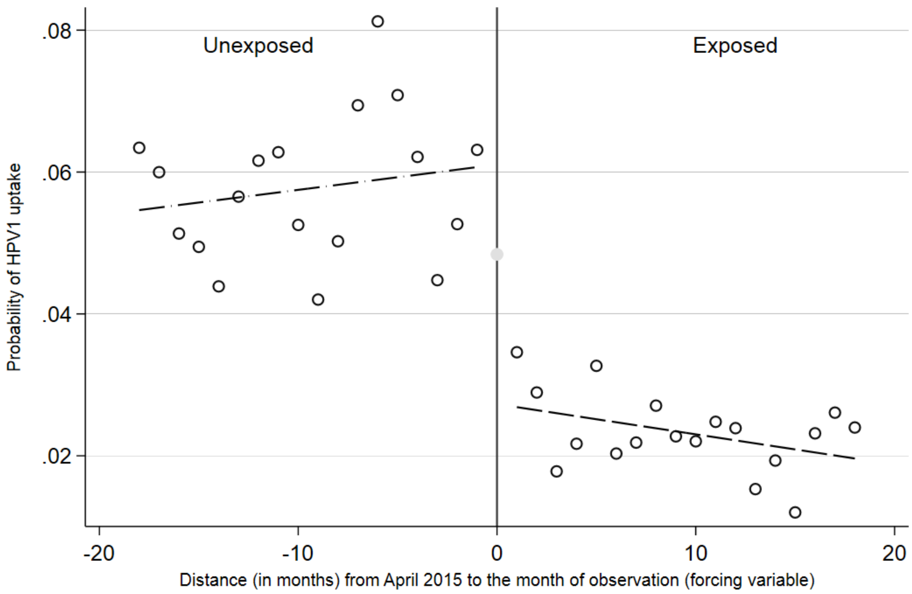
19. Athey S, Wager S. Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*. 2019;113(523):1228-1242.

20. Hausman C, Rapson DS. Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics*. 2018; 10:533-552.

21. Lübker CL, Lynge E. Stronger responders—uptake and decline of HPV-vaccination in Denmark, *European Journal of Public Health*. 2019: 29(3): 500–505.
22. Gørtz M, Brewer NT, Hansen P R, Ejrnæs M. The contagious nature of a vaccine scare: How the introduction of HPV initiation lifted and eroded MMR vaccination in Denmark. *Vaccine*. 2020;38(28):4432-4439.
23. Nyhan B, Reifler J, Richey S, Freed GL. Effective messages in vaccine promotion: a randomized trial. *Pediatrics*. 2014;133(4):e835-42.
24. Sadaf A, Richards JL, Glanz J, Salmon DA, Omer SB. A systematic review of interventions for reducing parental vaccine refusal and vaccine hesitancy. *Vaccine*. 2013;31(40):4293-304.
25. Sundhedsstyrelsen. Børnevaccinationsprogrammet. Årsrapport 2018. *Danish Health Authority*. 2019.
26. Suppli CH, Rasmussen M, Valentiner-Branth P, Mølbak K, Krause TG. Written reminders increase vaccine coverage in Danish children – evaluation of a nationwide intervention using the Danish Vaccination Register, 2014-2015. *Euro Surveillance.* 2017;22(17):1-8.
27. Hansen PR, Schmidtblaicher M, Brewer NT. Resilience of HPV vaccine uptake in Denmark: Decline and recovery. *Vaccine*. 2020(7): 1842-1842. Oktawia P, Wójcik JS, Mølbak K, Valentiner-Branth P. Validation of the 5-year tetanus, diphtheria, pertussis and polio booster vaccination in the Danish childhood vaccination database. *Vaccine*. 2013;31(6):955-959.
28. Oktawia P, Wójcik JS, Mølbak K, Valentiner-Branth P. Validation of the 5-year tetanus, diphtheria, pertussis and polio booster vaccination in the Danish childhood vaccination database. *Vaccine*. 2013;31(6):955-959.

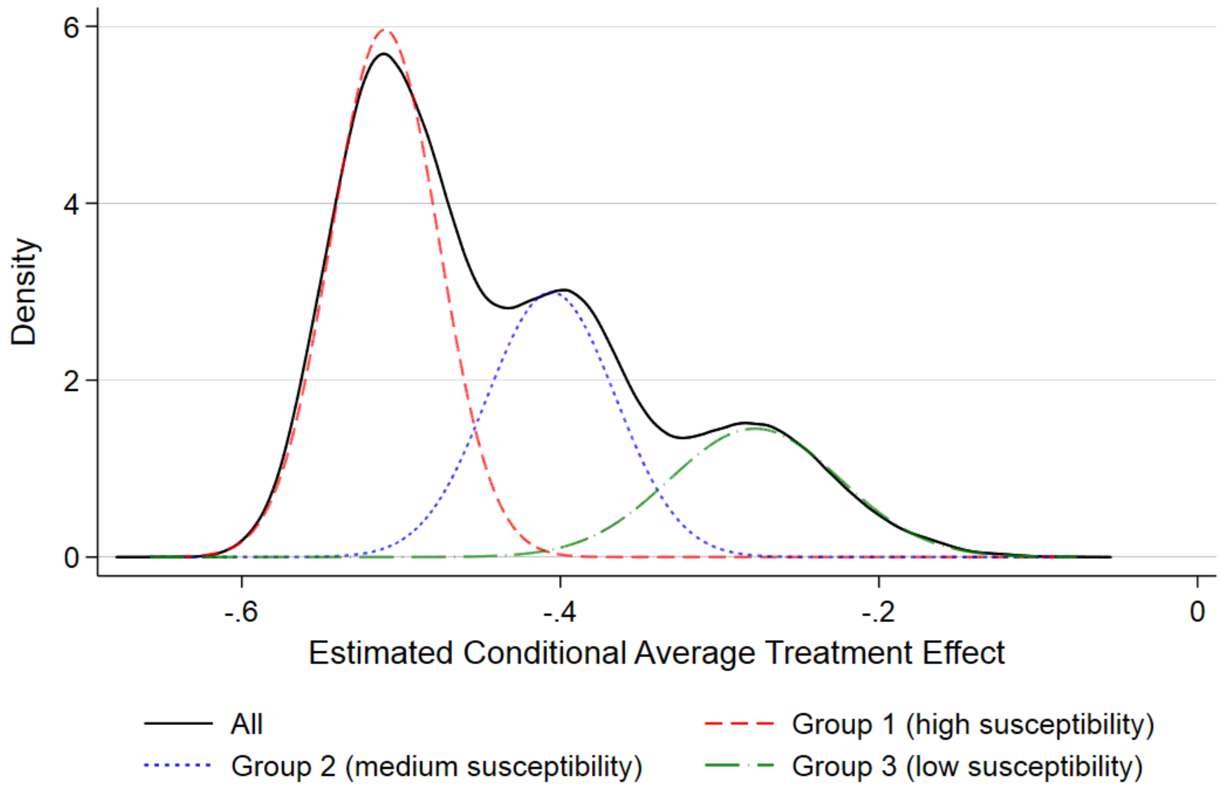
FIGURES

Exhibit 2: Monthly HPV1 uptake by exposure to documentary



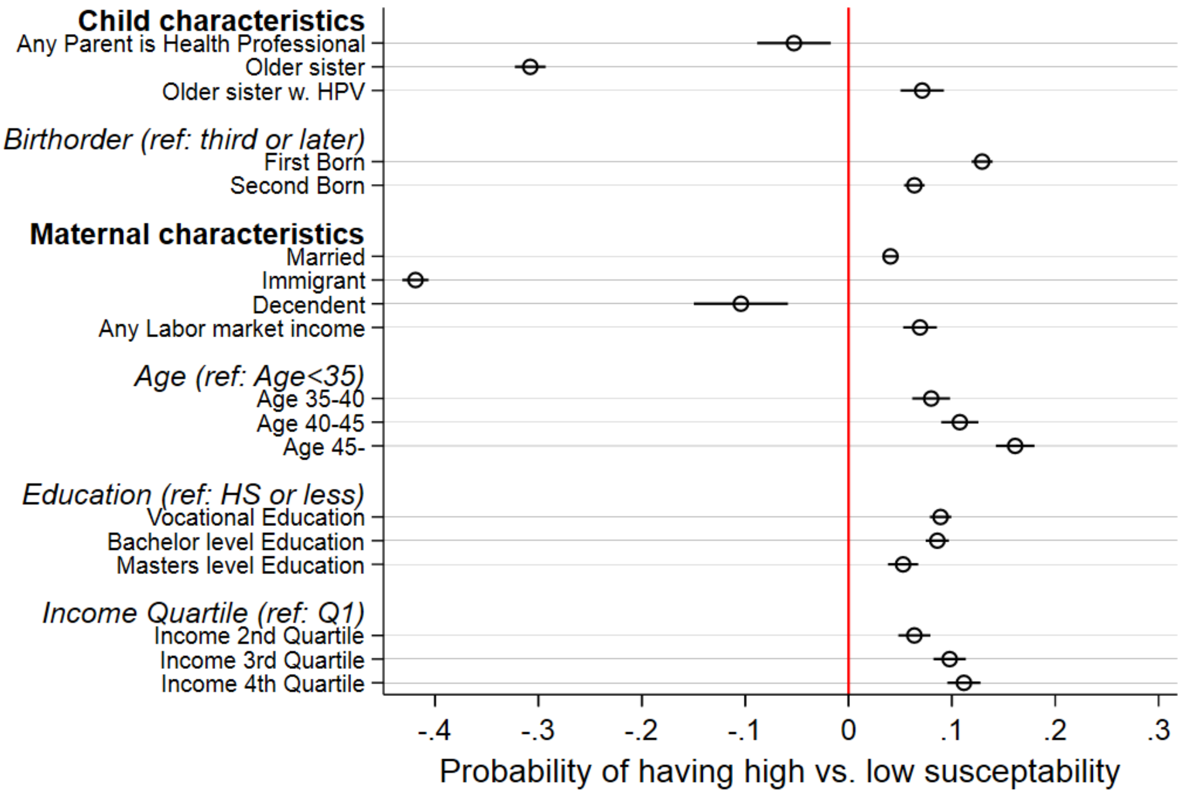
SOURCE Authors' analysis of Danish administrative data from October 2013 to October 2016. **NOTES** Association between monthly vaccination uptake (HPV1) and the distance (in months) from April 2015 to the month of observation. The vertical line marked April 2015. Dots marked the probability of getting the vaccine in a given month. Dots were overlaid with fitted regression lines from the linear probability model with the following independent variables: the treatment indicator, the forcing variable and the interaction of the two.

Exhibit 3: Distribution of Estimated Conditional Average Treatment Effects of Documentary



SOURCE Authors' analysis of Danish administrative data on girls who turned 12 in the periods May 2013 to April 2014 and May 2015 to April 2016. **NOTES** Due to missing information on either mother or father, the number of observations included in the heterogeneity sample is 62,920. Conditional Average Treatment Effects (CATEs) were estimated using nonparametric causal forests.

Exhibit 4: Characteristics of individuals with high and low susceptibility to the documentary



***SOURCE** Authors' analysis of Danish administrative data on girls who turned 12 in the periods May 2013 to April 2014 and May 2015 to April 2016. **NOTES** Difference in probability of having high vs. low susceptibility for various subgroups. The plotted coefficients were from an OLS regression of the estimated CATEs on characteristics (Appendix Exhibit A3). All variables were categorical and reference categories were listed in the figure. Number of observations was 32,481 and 11,584 in the high and low susceptibility groups, respectively. Conditional Average Treatment Effects (CATEs) were estimated using the cohort sample (62,920 individuals). Due to missing information on either mother or father, the number of observations included in the heterogeneity sample is 62,920.*

TABLES

Exhibit 1: Baseline Characteristics of 12-year-old Girls in Denmark

Variables	RDD Sample				Cohort Sample			
	Unexposed (N=579,213)	Exposed (N=616,227)	Diff.	95% CI	Unexposed (N= 32299)	Exposed (N=32754)	Diff.	95% CI
Outcome Variables								
Initiation of HPV at age 13	0.75	0.41	-0.34	(-0.33 ; -0.35)	0.77	0.32	-0.45	(-0.44 ; -0.45)
Control variables								
Child and family characteristics								
Birth order								
First	0.37	0.38	0.02	(0.02 ; 0.00)	0.36	0.38	0.02	(0.03 ; 0.01)
Second	0.36	0.35	-0.01	(0.00 ; -0.01)	0.36	0.35	-0.01	(0.00 ; -0.02)
Third or more	0.28	0.27	-0.01	(0.00 ; -0,01)	0.28	0.27	-0.01	(0.00 ; -0,01)
Older sister	0.21	0.21	-0.00	(-0.00 ; 0.00)	0.21	0.21	-0.00	(-0.00 ; 0.00)
Older sister w. HPV vaccine	0.15	0.11	-0.04	(-0.04 ; -0.04)	0.16	0.11	-0.05	(-0.05 ; -0.04)
Parent is health professional	0.02	0.02	0.00	(0.00 ; 0.00)	0.02	0.02	0.00	(0.00 ; 0.00)
Mothers characteristics								
Age								
-35	0.06	0.05	-0.01	(0.00 ; -0.01)	0.06	0.05	-0.01	(-0.01 ; -0.01)
35-40	0.22	0.21	-0.01	(-0.01 ; -0.02)	0.23	0.21	-0.02	(-0.01 ; -0.03)
40-45	0.39	0.40	0.01	(0.01 ; 0.00)	0.39	0.39	0.00	(0.01 ; 0.00)
45-	0.33	0.35	0.02	(0.02 ; 0.01)	0.32	0.35	0.03	(0.03 ; 0.02)
Married	0.67	0.66	-0.01	(0.00 ; -0.02)	0.67	0.65	-0.02	(-0.02 ; -0.01)
Ethnicity								
Native	0.84	0.84	0.00	(0.00 ; -0.01)	0.85	0.84	-0.01	(-0.01 ; -0.02)
Immigrant	0.13	0.12	-0.01	(0.00 ; -0.01)	0.13	0.12	0.00	(0.00 ; -0.01)
Descendant	0.01	0.01	0.00	(0.00 ; 0.00)	0.01	0.01	0.00	(0.00 ; 0.00)
Highest completed education								
Vocational	0.35	0.32	-0.01	(-0.01 ; -0.02)	0.35	0.33	-0.02	(-0.03 ; -0.02)
Bachelor's	0.30	0.32	0.01	(0.02 ; 0.01)	0.30	0.32	0.02	(0.02 ; 0.01)
Master's	0.10	0.12	0.01	(0.02 ; 0.01)	0.10	0.12	0.02	(0.02 ; 0.01)
Income quartile								
1st	0.25		-0.01	(-0.01 ; -0.02)	0.24	0.24	-0.01	(0.00 ; -0.01)
2nd	0.22	0.20	-0.01	(-0.01 ; -0.02)	0.22	0.20	-0.02	(-0.01 ; -0.02)
3rd	0.25	0.24	-0.02	(-0.01 ; -0.02)	0.23	0.21	-0.03	(-0.02 ; -0.03)
4th	0.26	0.30	0.04	(0.05 ; 0.03)	0.20	0.25	0.05	(0.04 ; 0.06)
Any labor market income	0.81	0.82	0.01	(0.02 ; 0.01)	0.81	0.82	0.01	(0.00 ; 0.02)

SOURCE Autmhors' analysis of Danish administrative data from 2012-2017. **NOTES** The study population comprised girls who turned 12 in November 2012 to October 2016 and who were living in Denmark in the year they turned 12. Initiation of HPV vaccine series was conditional on initiation within 12 months of becoming eligible. All characteristics were measured in the year the child turns 12. The RDD sample included each person-month observation in calculating the means. The RDD sample represented 129,569 girls that were observed for 9.2 months on average. 62,937 (66,632) girls were unexposed (exposed) at the time of the first observation.

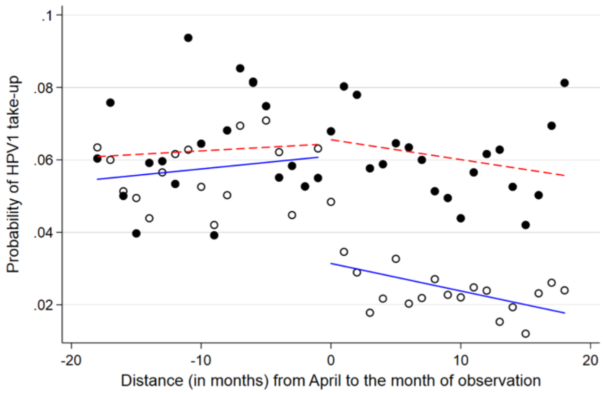
APPENDIX A: EXHIBITS

Appendix Exhibit A1: Impacts of documentary on monthly and yearly vaccination uptake (HPV1)

	Exposure effect		
	Monthly HPV1 uptake	Yearly HPV1 uptake	
	RDD sample	Cohort study	
Exposure Effect (95% CI)	-0.035 (-0.036 ; -0.034)	-0.447 (-0.464; -0.432)	-0.434 (-0.449 ; -0.419)
Observations	1,163,644	65,053	65,053
R-squared	0.04	0.14	0.25
Maternal characteristiccs	No	No	Yes
Child and famliy characteristics	No	No	Yes

SOURCE Authors' analysis of Danish administrative data on girls who turned 12 in the periods November 2012 to April 2016. **NOTES** The exhibit reports the estimated exposure effects on monthly, column (1) and yearly vaccination uptake, columns (2) and (3). In column (1) standard errors are clustered at the individual level. In columns (2) and (3) standard errors are clustered at the maternal municipality level.

Appendix Exhibit A2: Difference-in-differences



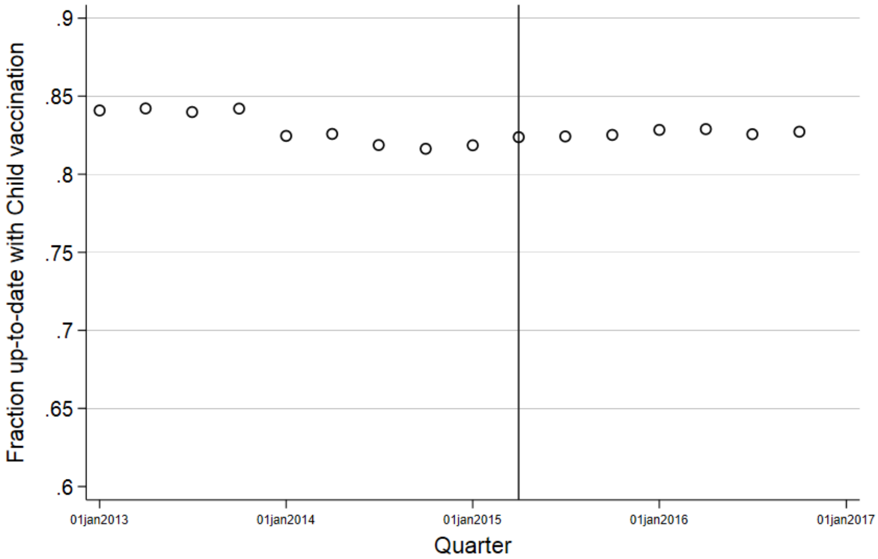
SOURCE Authors' analysis of Danish administrative data on girls who turned 12 in the period May 2011 to April 2016. Individuals in the "Exposed" (hollow dots, solid lines) group turned 12 in the period from May 1st 2013 to October 1st 2016, and individuals in the "Unexposed" group (solid dots, dashed lines) turned 12 from May 1st 2011 to April 30th 2014. The exposed group is observed +/-18 months around April 2015 whereas the individuals in the unexposed group is observed +/-18 months around April 2013. **NOTES** The graph depicts the pre- and post-trends for the exposed and unexposed group.

Appendix Exhibit A3: Difference-in-Difference estimates

	(1) Unexposed		(2) Exposed		(3) Dif-in-Dif	
	Coef	95% CI	Coef	95% CI	Coef	95% CI
Post	0.0004	(-0.001 ; 0.002)	-0.0343	(-0.036 ; -0.033)		
TreatxPost					-0.0315	(-0.032 ; -0.031)
Observations	1,181,953		1,163,644		2,248,965	
R-squared	0.001		0.010		0.006	

SOURCE Authors' analysis of Danish administrative data on girls who turned 12 in the period May 2011 to April 2016. Individuals in the "Exposed" group turn 12 in the period from May 1st 2013 to October 1st 2016, and individuals in the "Unexposed" group turned 12 from May 1st 2011 to April 30th 2014. The exposed group is observed +/-18 months around April 2015 whereas the individuals in the unexposed group is observed +/-18 months around April 2013
NOTES This table shows the difference-in-differences estimation results. Column 1 estimates the difference between the pre- and post-period for the unexposed group. Column 2 shows the pre and post difference for the exposed group. Column 3 shows the difference-in-differences estimate. All columns include demographic controls. Standard errors are clustered at the individual level.

Appendix Exhibit A4: Fraction of children aged 0-5 who are up-to-date in terms of their vaccinations in the Childhood Vaccination Program over time.



SOURCE Authors' analysis of Danish administrative data on children age 0-5. **NOTES** The graph shows the fraction of children that are up-to-date with the Childhood Vaccination program on a quarterly level.

Appendix Exhibit A5: Susceptibility groups identified using mixture-model on estimated CATEs

Group	(1) High Responsive	(2) Middle Responsive	(3) Low Responsive
Mean CATE (SD)	-0.51 (0.03)	-0.41 (0.04)	-0.27 (0.04)
Number of individuals	32,481	18,855	11,584
Fraction relative to total	48%	34%	18%

SOURCE Authors' analysis of Danish administrative data on girls who turned 12 in the periods May 2013 to April 2014 and May 2015 to April 2016 **NOTES** This table describes the three latent groups identified in the conditional treatment effect distribution. Due to missing information on either mother or father, the number of observations included in the heterogeneity sample is 62,920.

Appendix Exhibit A6: Impact of child, family and maternal characteristics on the probability of having either high, medium, or low susceptibility to the documentary

		(1)		(2)		(3)	
		1 vs. 3		1 vs. 2		2 vs. 3	
		Coef.	95% Confidence	Coef.	95% Confidence	Coef.	95% Confidence
Child characteristics							
Health professional parent		-0.053	(-0.097 ; -0.009)	-0.126	(-0.167 ; -0.085)	0.091	(0.046 ; 0.135)
Any older sister		-0.308	(-0.326 ; -0.29)	-0.334	(-0.352 ; -0.316)	-0.059	(-0.077 ; -0.041)
Older sister with HPV Vaccine		0.071	(0.046 ; 0.097)	-0.251	(-0.27 ; -0.232)	0.284	(0.266 ; 0.302)
Birth order	First born	0.129	(0.118 ; 0.14)	0.273	(0.262 ; 0.284)	-0.113	(-0.13 ; -0.096)
	Second born	0.064	(0.053 ; 0.075)	0.101	(0.091 ; 0.111)	0.0081	(-0.004 ; 0.02)
Maternal Characteristics							
Married		0.041	(0.032 ; 0.049)	-0.01	(-0.017 ; -0.003)	0.038	(0.027 ; 0.05)
Ethnicity	Immigrant	-0.419	(-0.433 ; -0.405)	-0.38	(-0.395 ; -0.365)	-0.14	(-0.155 ; -0.125)
	Descendant	-0.104	(-0.155 ; -0.053)	-0.205	(-0.251 ; -0.159)	0.098	(0.043 ; 0.153)
Any labor income		0.069	(0.051 ; 0.087)	0.058	(0.04 ; 0.076)	0.065	(0.043 ; 0.087)
Age	30-35	0.08	(0.06 ; 0.1)	0.179	(0.159 ; 0.199)	-0.124	(-0.149 ; -0.099)
	35-40	0.108	(0.088 ; 0.128)	0.222	(0.203 ; 0.241)	-0.14	(-0.165 ; -0.115)
	40+	0.161	(0.141 ; 0.181)	0.294	(0.274 ; 0.314)	-0.127	(-0.153 ; -0.101)
Education	Vocational	0.089	(0.078 ; 0.1)	0.091	(0.081 ; 0.102)	0.011	(-0.004 ; 0.026)
	Bachelor	0.086	(0.074 ; 0.098)	0.047	(0.035 ; 0.058)	0.054	(0.038 ; 0.071)
	Masters	0.053	(0.037 ; 0.068)	0.007	(-0.007 ; 0.022)	0.049	(0.026 ; 0.072)
Income Quartile	Q2	0.064	(0.047 ; 0.081)	0.045	(0.029 ; 0.061)	0.036	(0.013 ; 0.058)
	Q3	0.098	(0.081 ; 0.115)	0.086	(0.069 ; 0.102)	0.031	(0.008 ; 0.055)
	Q4	0.112	(0.094 ; 0.13)	0.094	(0.077 ; 0.11)	0.043	(0.019 ; 0.066)
Observations		44,065		51,336		30,439	
R-squared		0.293		0.382		0.123	

SOURCE Authors' analysis of Danish administrative data on girls who turned 12 in the periods May 2013 to April 2014 and May 2015 to April 2016. **NOTES** OLS estimates of probability of being (1) high susceptible vs low, (2) high vs. medium and (3) medium vs low. 95% confidence intervals in parentheses. Standard errors are clustered at the maternal municipality level.

Appendix Exhibit A7: Robustness checks

	(1)	(2)	(3)	(4)	(5)
	RDD Logistic	RDD Controls	RDD 12 month BW	RDD 3 month BW	RDD - no don
	Coef (95% CI)	Coef (95% CI)	Coef (95% CI)	Coef (95% CI)	Coef (95% CI
Exposed	-0.033 (-0.031; -0.034)	-0.035 (-0.037; -0.033)	-0.035 (-0.037; -0.033)	-0.029 (-0.031; -0.027)	-0.028 (-0.030; -0
Observations	1,163,644	1,163,644	772,344	190,799	1,195,440
Yearly Effects	-0.4	-0.42	-0.41	-0.34	-0.34

SOURCE Authors' analysis of Danish administrative data on girls who turned 12 in the period November 2012 to April 2016. **NOTES** This table shows alternative specifications of our main specification. Column 1 estimates the RDD with a logistic regression and reports the marginal effect. Column 2 includes controls to our baseline regression discontinuity design. Columns 3 and 4 reduce the bandwidth in the RDD to 12 and 3 months, respectively. Column 5 reports standard RDD estimates including April 2015. The row titled "Yearly Effects" scales the estimate in columns 1-5 by 12 months to allow for comparison with the results from the cohort study. 95% confidence intervals are reported in parentheses. Standard errors are clustered at the individual level.

APPENDIX B: ESTIMATION OF CONDITIONAL AVERAGE TREATMENT EFFECTS

Causal forests rely on modern non-parametric methods to estimate individual conditional treatment effects. It is a method that is well suited for our setting, as we have a relatively large number of covariates, and no a priori theoretical ranking of their importance to the outcome. The idea behind a causal forest is relatively straight forward (for details, see Athey & Wager, 2019). Fundamentally, a regression tree recursively splits the set of covariates into L leaves. The outcome prediction for any observation is then the mean outcome in the leaf that the observation would be assigned to based on its observable characteristics. Like the regression tree, the causal tree also splits the covariates into leaves, but among other restrictions additionally impose that at least a fixed fraction, α, of the observations in each leaf must be either treated or non-treated. Hence one can estimate a treatment effect by comparing the outcomes for treated and non-treated in the same leaf. Given a procedure for estimating a causal tree, a causal forest is an ensemble of B such trees. Each of these B trees produces a leave-specific treatment effect, and by averaging across trees a distribution of individual conditional treatment effects can be estimated. Each causal tree is grown based on a random subset of "training" observations, and the averaging across trees reduces variance in the treatment effect estimate (Breiman, 2001). Athey and Wager (2019) presents conditions under which asymptotic confidence intervals can be constructed for the individual treatment effects.

We implement the causal forest with honest trees (Athey & Imbens, 2016), grow 2,000 trees and impose that at a minimum 5% of individuals in any leaf must be either treated or non-treated (α=0.05).

References

1. Athey S, Wager S. Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*. 2019;113(523):1228-1242.
2. Athey, S Imbens G. Recursive partitioning for heterogenous causal effects. *Proceedings of the National Academy of Sciences.* 2016 113 (27) 7353-7360
3. Breiman L. Random Forests. *Machine Learning*. 2001, 45, 5-32