

A Dynamic Model of Vaccine Compliance: How Fake News Undermined the Danish HPV Vaccine Program

Peter Reinhard Hansen^{a,c} Matthias Schmidtblaicher^{b,d}

^a*University of North Carolina* ^b*European University Institute*

^c*Copenhagen Business School* ^d*QuantCo, Inc.*

April 19, 2019

Abstract

Increased vaccine hesitancy presents challenges to public health and undermines efforts to eradicate diseases such as measles, rubella, and polio. The decline is partly attributed to misconceptions that are shared on social media, such as the debunked association between vaccines and autism. Perhaps, more damaging to vaccine uptake are cases where trusted mainstream media run stories that exaggerate the risks associated with vaccines. It is important to understand the underlying causes of vaccine refusal, because these may be prevented, or countered, in a timely manner by educational campaigns. In this paper, we develop a dynamic model of vaccine compliance that can help pinpoint events that disrupted vaccine compliance. We apply the framework to Danish HPV vaccine data, which experienced a sharp decline in compliance following the broadcast of a controversial TV documentary, and we show that media coverage significantly predicts vaccine uptake.

Keywords: Media, Public Health, Vaccine Uptake, Fake News, Score-Driven Model

JEL Classification: C22, I12, I18, L82.

1 Introduction

Increased opposition to vaccine programs presents an important challenge to public health. Improving the understanding of the causes of vaccine hesitancy is paramount to health authorities. Vaccine scares have, historically, been accompanied by vaccine-critical media stories, prompting the question of how media contributed to vaccine hesitancy, see Offit (2011). More generally, media have been conjectured as drivers of a variety of personal decisions, ranging from asset allocation to migration, see DellaVigna and La Ferrara (2015). Vaccination is an important personal choice and the decision to get vaccinated or not has a clearly defined binary outcome. This makes vaccination data an interesting case study for measuring the influence of media. In this paper we model time variation in vaccine uptake and analyze the extent to which this variation can be explained by related media coverage.

Our paper makes a number of econometric contributions. The first econometric contribution is a time series model for binomial variables with a time-varying success probability. Our model builds on the score-driven framework proposed by Creal et al. (2013), and we are, to the best of our knowledge, the first to develop a score-driven model for binomially distributed time series data. The framework implies an intuitive dynamic structure for the time variation in the binomial coefficient. If the empirical frequency exceeds the expected frequency, the binomial coefficient is subsequently adjusted upwards, and if the empirical frequency falls short of the expected frequency, the binomial coefficient is adjusted downwards. The magnitude of the adjustment is determined empirically. The score-driven model we develop in this paper is not specific to vaccination data and could easily be adapted to other time series that involve aggregated binary outcomes. For instance the modeling of credit rating transitions and default intensities, where time-varying parameter models were proposed using a different approach, see Koopman et al. (2008) and Koopman et al. (2011). A second econometric contribution is the flexible treatment of seasonality. Institutional changes in the way vaccinations were recorded compelled us to model seasonal variation at the monthly frequency, in addition to annual seasonal effects. We introduce a score-driven model of intra-monthly seasonality using an approach that is similar to that in Caivano et al. (2016).

We apply the model to Danish Human Papilloma Virus (HPV)¹ vaccine data. The data

¹HPV denotes a family with more than 100 types of viruses, of which at least 13 types can cause cancer, see WHO (2016). The HPV vaccine is expected to prevent many forms of cancer and the newest HPV vaccine, Gardasil 9, may reduce cervical cancer by as much as 90%. Evidence for the efficacy and safety of the HPV vaccine is strong. A recent Cochrane Systematic Review concluded: “There is high-certainty evidence that HPV vaccines protect against cervical precancer” and “We did not find an increased risk of serious adverse effects”, see

are weekly HPV vaccine initiations by birth year cohort from January 2009 to June 2017. The structure of the data calls for a model of vaccine uptake with two components. The first component characterizes baseline vaccine uptake and is identical for all birth year cohorts. The second component characterizes the variation in vaccine uptake over time, which is key in our analysis. It can be used to monitor vaccine compliance in real time, and it is the time variation in this component that we relate to data on media coverage. The empirical analysis reveals a great deal of time variation in vaccine compliance, including a sudden drop in 2015. Our empirical results show that declines in vaccine compliance began after negative coverage on the HPV vaccine appeared in Danish media and, by augmenting our model to include media coverage, we show that media coverage significantly predicts declines in vaccine compliance. Some newspaper articles were blatantly false and misleading and warrant a *fake news* designation (in the original meaning of these words, which is “false news”). However, most media coverage was merely reporting on alleged side effects or on the declining vaccine uptake that had resulted from this concern. While these news articles individually cannot be said to be false, they may collectively misrepresent the risk of vaccine induced side effects. The largest drop in compliance occurred immediately after a TV documentary was aired on TV2 Denmark on March 26th, 2015. The program was entitled “De Vaccinerede Piger – Syge og Svigtede”, which translates to “The Vaccinated Girls – Sick and Abandoned”. After the program was aired, Danish HPV vaccine compliance fell from over 90% to less than 30%. TV2 Denmark has publicly acknowledged that their documentary contributed to the decline in HPV vaccinations. Our empirical analysis supports this conclusion, because of the sharp decline in HPV vaccine uptake immediately after the documentary aired. The model of HPV vaccine uptake also enables us to quantify the decline in vaccinations, relative to a counterfactual scenario where compliance stayed at the level before the documentary. We estimate that nearly half of the girls born in 2003 postponed HPV vaccination following the TV2 documentary, and many of these girls are still unvaccinated. By the end of our sample period, we estimate that nearly 14,000 of the unvaccinated girls born in 2013, can be attributed to the declining HPV vaccine uptake that followed the TV documentary. For illustration, a 70% reduction in cervical cancer for 14,000 Danish females will, on average, translate into about 100 fewer cases of cervical cancer and 26 fewer deaths. The full consequences of the decline in vaccine uptake may be substantially larger, because several other birth-year cohorts, including girls born in 2004 and 2005, are also behind in vaccine coverage, relative to

Arbyn et al. (2018).

that of older cohorts.

A historical episode that is similar to the one that we investigate in this paper is the decline in DTP vaccinations in the US following the TV program “DPT: Vaccine Roulette”. The program initially aired in 1982 on an NBC affiliate, WRC-TV, and then nation-wide on The Today Show. The program falsely associated the pertussis component of the DTP vaccine with brain damage, and the TV program was followed by extensive media coverage in the US, which speculated that the pertussis vaccine was responsible for epilepsy, intellectual or physical disabilities, and even death, see Offit (2011, chapter 3). The Danish experience with the HPV vaccine program is similar to the onset of the pertussis vaccine scare in the US. As had been the case for the “Vaccine Roulette” program in the US, the Danish documentary was followed by a large number of newspaper articles on the topic.

It has been documented before that media can influence important personal decisions. For instance, Kearney and Levine (2015) showed that the MTV reality show *16 and Pregnant* reduced teen childbearing. Similarly, La Ferrara et al. (2012) found that soap operas that portray small families had a significant impact on fertility. See also DellaVigna and La Ferrara (2015) and references therein. There is also evidence that the revelation of questionable and malicious behavior by health authorities can reduce care-seeking behavior. Specifically, the infamous Tuskegee syphilis experiment reduced the number of physician visits by black men, see Alsan and Wanamaker (2017), colonial era medical malpractice in central Africa is associated with higher levels of distrust in medicine today, see Lowes and Montero (2018), and vaccination uptake fell in Pakistan after the unmasking of CIA’s involvement in a vaccination program that was used in the hunt for Osama bin Laden, see Martinez-Bravo and Stegmann (2018).

There is a nascent public health literature associating HPV vaccine hesitancy with media coverage. Faasse et al. (2017) documented that the number of monthly news articles on the HPV vaccine predicts the number of reported adverse events (AE’s) in New Zealand. In Denmark, we also observe a sharp increase in the number of reported AE’s in 2015. A recent study by Suppli et al. (2018) found that the monthly number of Danish HPV1 vaccinations was uncorrelated with media coverage before July 2013, but negatively correlated with media coverage after July 2013. They study the correlation between media activity and total number of HPV1 vaccinations, and determine a change point in July 2013. A great deal of the variation in the total number of HPV1 vaccinations can be attributed to changes in the the Danish program. A catch-up program made the HPV vaccine freely available to girls born between 1993 and 1995

during the period from October 1, 2008 to December 31, 2010. Another catch-up program made the vaccine freely available to women aged 19-26 in 2012 and 2013. Our analysis is based on cohort specific weekly HPV1 vaccinations for girls aged 12-14 years, which is not influenced by the catch-up programs. We relate vaccine uptake to media coverage by incorporating a measure of media coverage directly in the equation that drives the variation in uptake. In contrast, the analysis in Suppli et al. (2018) is based on a change point analysis of the correlation between the aggregate number of monthly vaccinations and media coverage.

The rest of this paper is organized as follows. In Section 2 we present the core structure of our model in a simplified manner along with some preliminary empirical results for the Danish HPV vaccination data. In Section 3 we present the econometric time-series model with a time-varying vaccine compliance. In Section 4 we incorporate media coverage of the HPV vaccine in the analysis, and show that the intensity of such media coverage predicts the observed time variation in vaccine compliance. We summarize and conclude in Section 5. Mathematical derivations and supplementary empirical results are given in two appendices, and additional empirical results and information about media coverage can be found in a separate Web-Appendix, see Hansen and Schmidtblaicher (2019).

2 Data and Preliminary Analysis

We obtained weekly birth-year cohort specific HPV vaccination data from the Statens Serum Institute (SSI) and birth-year cohort size data from Statistics Denmark. SSI is responsible for the purchase and supply of vaccines to the Danish national vaccination programs, and SSI collects data on vaccination uptake. Two types of HPV vaccines were administered during the sample period from early 2009 to mid-2017, Gardasil and Cervarix. HPV vaccines were initially licensed with 3-dose schedules, but are now administered with just two doses for young adolescents. Our empirical analysis will focus on the number of girls receiving the first dose of the HPV vaccine, which we denote by HPV1. Specifically, we will model the number of girls born in year $c = 1997, \dots, 2005$, who receive HPV1 in week t .

Let X_t denote the aggregate number of HPV1 vaccinated girls by week t , out of a birth-year cohort with N_c girls. The basic idea is that the fraction of vaccinated girls, at time $t = 0, \dots, T$, is approximately given by

$$X_t/N_c \approx \delta \times \Lambda\left(\frac{t}{T}\right),$$

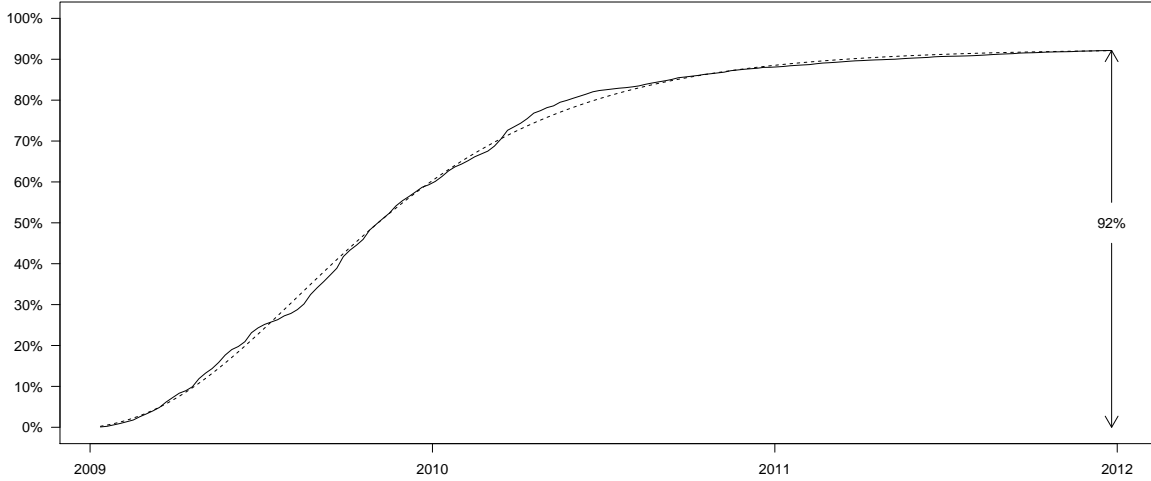


Figure 1: HPV vaccine uptake for birth-year cohort 1997. The solid line displays the cumulative percentage of HPV1 vaccinated girls that were born in 1997, and the dotted line is a simple approximation fitted to the data. The percentage of vaccinated girls by the end of 2012 was 92%.

where $\Lambda(a)$ is an increasing function with $\Lambda(0) = 0$ and $\Lambda(T) = 1$ and where δ is a scalar between zero and one. The parameter δ is key in our analysis, because it can (in a steady state) be interpreted as vaccine compliance/coverage by the end of sample period.

The simple structure, where δ is constant, is illustrated in Figure 1. The solid line shows the percentage of girls born in 1997 that have received the first dose of the HPV vaccine over a three-year period. The dotted line is a curve that is fitted to the data, using the shifted Gompertz distribution to specify $\Lambda(a_t)$, where a_t is the age of cohort 1997 at time t . In conjunction with the estimate of δ (about 92%) the simple model provides an approximation to the vaccine uptake over time for this cohort.

Figure 2 shows vaccine compliance over time for nine cohorts. We observe large discrepancies across birth-year cohorts, with a much lower vaccine uptake for the youngest cohorts. For the four oldest cohorts, 1997-2000, vaccine adoption was high and slightly increasing over time. This was followed by a period of declining compliance starting with cohort 2001. Evidently, it is not possible to accurately describe all cohorts with a common specification such as in Figure 1. But, as we shall see, a modified specification that allows for time variation in the compliance rate describes the data well. The cornerstone of our model is the weekly number of vaccinations for each cohort, which (in a static model) has the expected value $\delta[\Lambda(a_t^c) - \Lambda(a_{t-1}^c)] \times N_c$, where a_t^c is the age of cohort c at time t and N_c is the cohort size. We generalized this model by allowing for time-variation in δ , so that the expected number of vaccinations in week t for cohort c is

given by $\delta_t[\Lambda(a_t^c) - \Lambda(a_{t-1}^c)] \times N_c$.

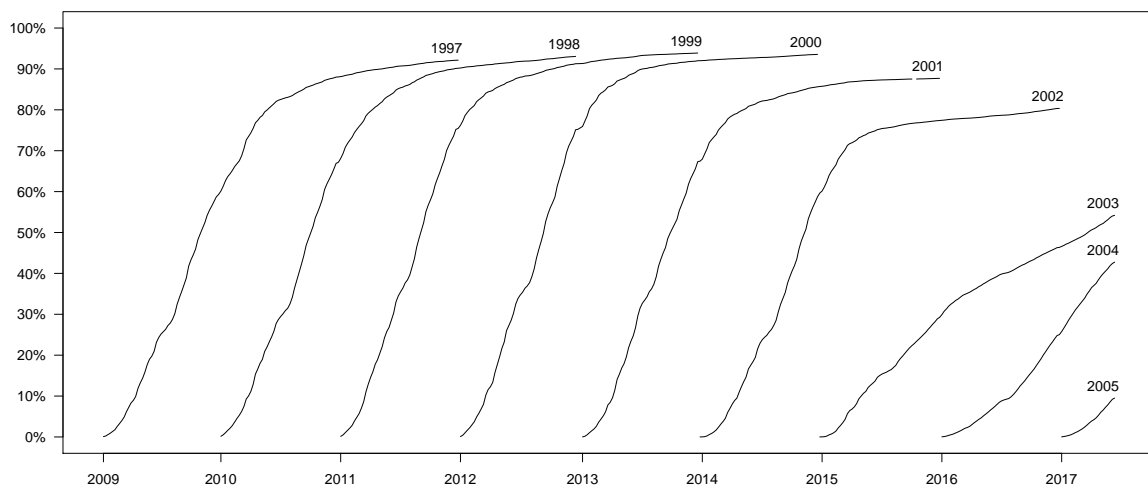


Figure 2: Cumulative HPV vaccine uptake by birth-year cohort. The solid lines display the percentage of HPV1 vaccinated girls for each birth-year cohort over a three-year-period, starting in the year they turn 12 years old (lines labelled by birth-year).

3 Statistical Model

Let $x_{c,t}$ be the number of girls in cohort c that receive HPV1 in week t . The number of vaccinated girls in cohort c at time t is given by $X_{c,t}$, where $X_{c,t} = X_{c,t-1} + x_{c,t}$ with $X_{c,0} = 0$. The number of unvaccinated girls in cohort c that are eligible to receive the vaccine by the end of week t is denoted by $N_{c,t}$, see Appendix C.

The age-variable, a_t^c , for cohort c at time t is, without loss of generality, normalized so that $a = 0$ denotes the beginning of the three-year period and $a = 1$ by the end of the three-year period. To take an example, for those born in year 2000, we have the weekly number of vaccinations for the period primo 2012 to ultimo 2014, such that $a_t^{2000} = 0$ at the beginning of 2012 and $a_t^{2000} = 1$ by the end of 2014. The beginning of the three-year period is motivated by the design of the program, where girls are only offered the vaccine for free once they turn twelve years old, and age twelve is the recommend age to initiate HPV vaccination in Denmark. The number of Danish girls vaccinated before their 12th birthday is therefore negligible.

The basic structure of our model is that the number of vaccinated girls in week t is binomially distributed

$$x_{c,t} | \mathcal{F}_{t-1} \sim \text{bin}(N_{c,t-1}, p_{c,t}), \quad (1)$$

where the dependence across time and cohorts and seasonal effects are embedded in the structure

of $p_{c,t}$. Our model for $p_{c,t}$ is given by

$$p_{c,t}(\theta) = \delta_t(\alpha)\lambda_{c,t}(\beta), \quad \text{with} \quad \lambda_{c,t}(\beta) = \frac{N_c}{N_{c,t-1}}[\Lambda(\beta; a_t^c) - \Lambda(\beta; a_{t-1}^c)], \quad (2)$$

where $\theta = (\alpha', \beta)'$ is the vector of unknown parameters.

The first component, $\delta_t(\alpha) \in (0, 1)$, defines vaccine compliance at time t , whereas the second component, $\lambda_{c,t}(\beta)$, only depends on t though the age of the cohort, a_t^c . So the second term defines the part of $p_{c,t}(\theta)$ which all cohorts have in common, and if vaccine compliance were constant over time then all cohorts would have similar vaccine uptake. This is evidently not the case, as demonstrated in Figure 2.

From the binomial model structure in (1), it follows that the log-likelihood for cohort c in period t is given by

$$\ell_{c,t}(\theta) = \log \binom{N_{c,t-1}}{x_{c,t}} + x_{c,t} \log p_{c,t}(\theta) + (N_{c,t-1} - x_{c,t}) \log(1 - p_{c,t}(\theta)). \quad (3)$$

The maximum likelihood estimators are obtained by maximizing $\ell(\theta) = \sum_{c,t} \ell_{c,t}(\theta)$, with respect to the vector of parameters, $\theta = (\alpha', \beta)'$. To complete the model we need to adopt specifications for $\Lambda(\beta; a)$ and $\delta_t(\alpha)$. For $\Lambda(\beta; a)$ we adopt the cumulative distribution function (cdf) for the shifted and truncated Gompertz distribution, which is given by

$$\Lambda(\beta; a) = \frac{1}{C(\beta)}(1 - e^{-\beta_0 a}) \exp(-\beta_1 e^{-\beta_0 a}), \quad a \in [0, 1],$$

where $C(\beta) = (1 - e^{-\beta_0}) \exp(-\beta_1 e^{-\beta_0})$ is a normalizing constant. Other, more flexible, specifications could be used, and in a preliminary analysis we also experimented with specifications based on the Weibull distribution and the Beta distribution. The shifted Gompertz was adopted because it had the best empirical fit.

Possible time-variation in $\delta_t(\alpha)$, which represents vaccine compliance across time, can be modeled in many ways. In the Web-Appendix we present results for several alternative approaches, such as the case where δ_t is piecewise constant with structural changes and cases where δ_t is a deterministic function of time. The best empirical fit is obtained with the specification for $\delta_t(\alpha)$ we present next.

3.1 A Score-Driven Model for $\delta_t(\alpha)$

We model δ_t using an autoregressive model for the logit transformed variable, $\tilde{\delta}_t = \log\left(\frac{\delta_t}{1-\delta_t}\right)$,

$$\tilde{\delta}_t = \alpha_0 + \alpha_1 \tilde{\delta}_{t-1} + \alpha_2 \tilde{s}_{t-1}, \quad (4)$$

where \tilde{s}_t (to be made precise later) signals the direction in which compliance may have changed as well as the magnitude. The logit transformation allows us to model $\tilde{\delta}_t$ as an unrestricted real-valued parameter, and the inverse transformation, $\delta = e^{\tilde{\delta}}/(1 + e^{\tilde{\delta}})$, will ensure that δ stays within its boundaries between zero and one. Equation (4) defines an observation-driven model based on the generalized autoregressive score framework by Creal et al. (2013). Score-driven models have been very successful in modeling time-varying parameters in econometric models, and are the underlying structure of many empirically established models, such as the GARCH model, by Bollerslev (1986). In the present context, the score-driven model adjusts the value of δ_t directly in response to the number of vaccinations, $x_{c,t}$, deviating from the expected number. The adjustment is defined by the score which is the derivative of the log-likelihood, suitably scaled by the expected curvature of the log-likelihood function.

The vector of unknown parameters in $\delta_t(\alpha)$ is given here by $\alpha = (\alpha_0, \alpha_1, \alpha_2, \delta_0)'$, where δ_0 is the starting value for δ_t . Here α_1 is a measure of the persistence in δ_t and α_2 measures how strongly the model responds to the signal provided by \tilde{s}_t . There is typically a high degree of persistence in score-driven models, so we also consider the restricted variant of the model, where $(\alpha_0, \alpha_1) = (0, 1)$, which corresponds to the case where $\tilde{\delta}_t$ is a very persistent (unit root) process.

The signal, \tilde{s}_t , is key in this model. If, for instance, the number of vaccinated girls exceeds the expected number of vaccinations, it indicates that δ_t has increased in value, and intuitively we would want (α_2 times) \tilde{s}_t to be positive in this situation. The score-driven framework by Creal et al. (2013) employs a natural signal that is deduced from the score of the log-likelihood function, $s_t = \partial \ell_t / \partial \tilde{\delta}$, weighted by a term that is defined by the curvature of the log-likelihood, $h_t = \partial^2 \ell_t / \partial \tilde{\delta}^2$. Specifically,

$$\tilde{s}_t = \frac{\sum_c s_{c,t}}{\sqrt{-\sum_c \mathbb{E}_{t-1} h_{c,t}}} = \frac{1}{\sqrt{\sum_c \frac{\lambda_{c,t}^2 N_{c,t-1}}{p_{c,t}(1-p_{c,t})}}} \sum_c \lambda_{c,t} N_{c,t-1} \left(\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}} \right), \quad (5)$$

where $\hat{p}_{c,t} = x_{c,t}/N_{c,t-1}$. The expression (5) is derived in Appendix A.1, but the interpretation is intuitive. In a week where more individuals are vaccinated than expected, i.e. $\hat{p}_{c,t} > p_{c,t}$ for

all c , then $\tilde{s}_t > 0$, and this would be an indication that δ_t may have increased in value, and visa versa in the event $\tilde{s}_t < 0$. So we should expect the estimate of α_2 to be positive, which is indeed the case in our empirical analysis.

Our empirical analysis revealed that there was a need to account for seasonal variation in the weekly vaccination rate. The most obvious seasonal effect is associated with the summer vacation and winter holidays, where the number of vaccinations is distinctly below that of neighboring weeks. This seasonal effect is pronounced and can be seen in Figures 1 and 2. The second seasonal effect is specific to the way in which the weekly vaccination data were collected over time. During the first part of the sample period, the reported number of vaccinations is higher towards the end of the month. After conferring with a medical professional, we discovered that the way in which vaccines are recorded has changed during the sample period. Before November 15, 2015, vaccines were recorded when physicians billed for the vaccines, which resulted in an over-recording of vaccines towards the end of the month. The introduction of an electronic vaccine registry (Det Danske Vaccinationsregister) resolved this issue, starting November 15, 2015. Our modeling of seasonal effects is detailed in Appendix B.

Table 1: Estimates for Score-Driven Model

	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\ell(\hat{\theta})$
Unrestricted	0.000 (0.087)	0.992 (0.022)	0.054 (0.017)	7.250 (0.038)	3.030 (0.088)	-20122
$\alpha_0 = 0, \alpha_1 = 1$	0	1	0.066 (0.004)	7.240 (0.018)	3.010 (0.003)	-20318

Note: Standard errors in parentheses.

The key parameters of the estimated score-driven model are presented in Table 1, and the resulting time series for vaccine compliance is presented in Figure 3, where we plot δ_{t+1} against time t (because δ_{t+1} is observable at time t). In Table 1 we also report estimates of the restricted model where we impose the restrictions $\alpha_0 = 0$ and $\alpha_1 = 1$, which resembles a local-level model for $\tilde{\delta}_t$. This specification is strongly rejected by the data. Estimates of all parameters in the model are presented in Table B.1 in the Appendix.

Figure 3 shows that vaccine compliance is estimated to be a bit lower during the first year in our sample (year 2009). Then compliance increases and stays above 95% for a three-year period. The first noticeable decline in $\hat{\delta}_t$ is seen in 2013 and vaccine compliance is relatively low and volatile until the fall of 2014, after which compliance recovers to about 95% again. The

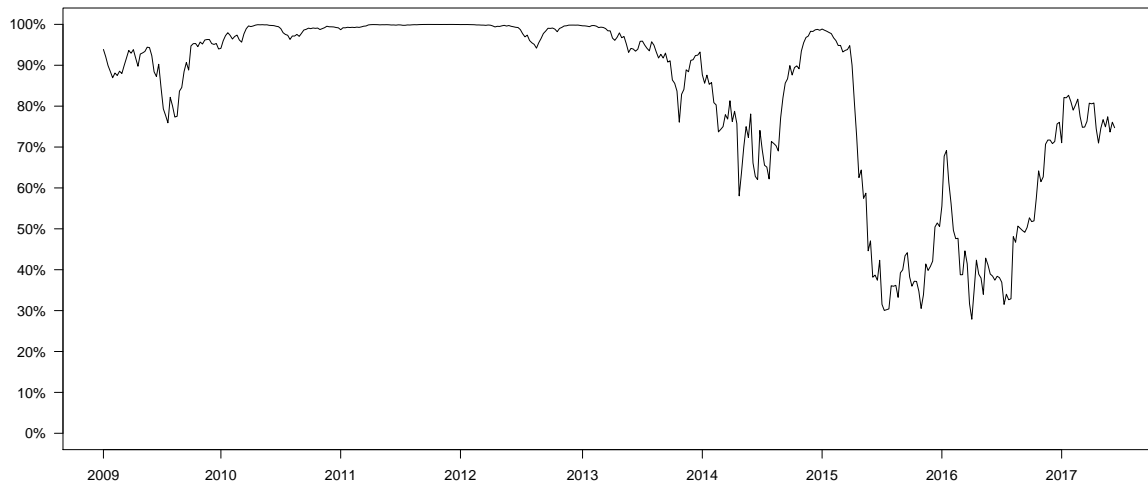


Figure 3: The figure shows vaccine compliance, δ_t , as estimated by the score-driven model. Aside from the first year after the introduction of the HPV vaccine, the first noticeable decline in vaccine uptake is observed in 2013, followed by a rebound in 2014. The largest decline is observed in the second quarter of 2015 where δ_t falls abruptly from about 95% to just over 30%.

most drastic shift in compliance is observed in the second quarter of 2015 where $\hat{\delta}_t$ abruptly falls to just over 30%. Compliance stays low for an extended period, aside from a brief spike in early 2016. Only in late 2016 does compliance begin to recover. Towards the end of the sample period, June 2017, it hovers at about 75%. In Section 4 we analyze the variation in $\hat{\delta}$ in greater detail by relating it to media data.

4 The Influence of Media

In this section we integrate the Danish media coverage of the HPV vaccine in the analysis. We obtained data on media coverage on the HPV vaccine from Infomedia, which is a searchable media database with comprehensive coverage of the Danish media, targeting negative stories. Specifically, we obtained the weekly number of media articles containing the keywords “HPV” and “bivirkning” (English: “side effect”) along with a list of excluding keywords that served to preclude positive and irrelevant news stories, detailed in the Web-Appendix.

Figure 4 displays the weekly number of HPV1 vaccinations for girls born in 1997 or later (upper panel) and the weekly media count, denoted m_t (lower panel), along with the four-week moving averages for both variables. The weekly vaccination data in the upper part of Figure 4 has clear-cut seasonal effects, with relatively few number of vaccinations during the summer and end-of-year vacations. The vaccination data also display the pronounced end-of-month effects that existed prior to the introduction of the electronic vaccine registry on November 15, 2015.

The lower part of Figure 4 shows that most of the media coverage occurred in 2015, starting in March 2015.

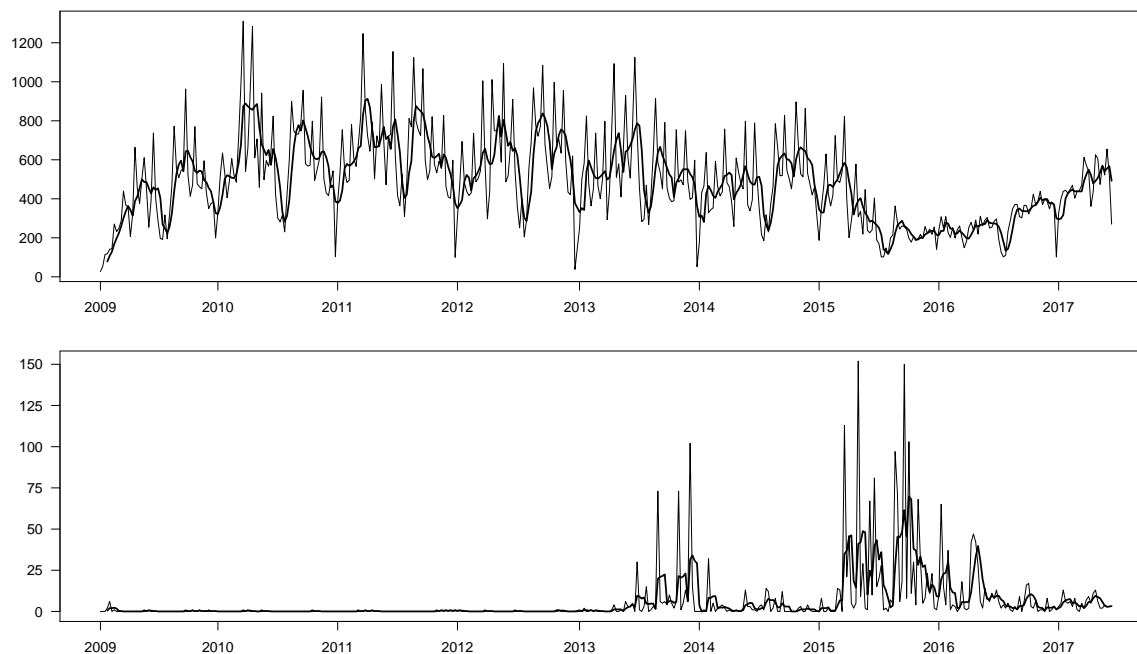


Figure 4: The upper panels show the weekly number of HPV1 vaccination of girls born in 1997 or later, $\sum_c x_{c,t}$. The lower panel is the number of media stories, m_t , on HPV vaccine related suspected side effects. The solid lines are 4-week moving averages.

Most of the media articles that define m_t are merely reporting on alleged side effects or reporting on the declining vaccine uptake as it occurred. Only a few articles can appropriately be labeled as “fake news”, such as an article in Metroxpress on June 11, 2015, where the headline read: “Doctors: One in 500 get seriously ill from the HPV vaccine”. Another example is the article in Information on May 30, 2016 that (contrary to scientific consensus) presented the view that “there is no scientific evidence that the HPV vaccine prevents cervical cancer”. Most articles were merely reporting on factual information, such as the declining vaccine uptake or on the number of alleged side effects. Such as the article in Berlingske on August 31, 2015: “Now there are fewer being HPV vaccinated than MMR vaccinated”, in which it was speculated that stories about side effects might have caused the decline in HPV vaccinations. Another example is an article in Ekstrabladet on September 24, 2015: “More than 1,500 girls supposedly have HPV side effects”. This article reported that the total number of suspected HPV vaccine side effects had risen to 1,586. An extensive list of articles on this subject is presented in the Web-Appendix.

Below we incorporate the media activity variable into the model, before turning our attention

to a TV documentary that is referenced in many of the media stories after March 2015.

4.1 The Effect of Media Coverage on Vaccine Compliance

In this section, we include the media variable, m_t , in the model to study if some of the observed variation in vaccine compliance can be explained by media coverage on the topic. We achieve this by augmenting the score-driven model to include $\tilde{m}_t = \log(1 + m_t)$ as an explanatory variable in the dynamic equation for vaccine compliance, $\tilde{\delta}_{t+1} = \alpha_0 + \alpha_1\tilde{\delta}_t + \alpha_2\tilde{s}_t + \alpha_3\tilde{m}_t$. The influence of media activity on vaccine compliance is measured by the parameter α_3 . If media activity were to reduce vaccine uptake, then we should expect $\alpha_3 < 0$, whereas if media activity had no effect on compliance then we should expect $\alpha_3 = 0$.

Estimating the augmented model by maximum likelihood yields:

$$\tilde{\delta}_{t+1} = \underset{(0.011)}{0.182} + \underset{(0.007)}{0.948}\tilde{\delta}_t + \underset{(0.006)}{0.050}\tilde{s}_t - \underset{(0.018)}{0.080}\tilde{m}_t, \quad (6)$$

where standard errors are given in parentheses below the estimates. The parameter of interest is estimated to be negative, $\hat{\alpha}_3 = -0.080$, and is significant. The significance of including the media variable in the model is also affirmed by the value of the log-likelihood function that increases from -20122 to -19954 . Additional estimates of this specification are presented in the first row of Table 2. The Table also reports the estimates for two alternative specifications, where the media variable is defined as, respectively, $\tilde{m}_t = m_t$ and $\tilde{m}_t = \sqrt{m_t}$. The complete set of parameter estimates of all specification are presented in Table B.2 in the Appendix.

Table 2: Estimates for Score-Driven Model with Media Variable

	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\ell(\hat{\theta})$
$\tilde{m}_t = \log(1 + m_t)$	0.182 (0.011)	0.948 (0.007)	0.050 (0.006)	-0.080 (0.018)	7.24 (0.019)	3.000 (0.003)	-19954
$\tilde{m}_t = m_t$	0.087 (0.003)	0.967 (0.014)	0.052 (0.012)	-0.004 (0.000)	7.24 (0.010)	3.000 (0.021)	-20033
$\tilde{m}_t = \sqrt{m_t}$	0.160 (0.007)	0.952 (0.004)	0.052 (0.012)	-0.045 (0.013)	7.24 (0.015)	3.000 (0.005)	-19971

Note: Parameter estimates and standard errors in parentheses, for the model $\tilde{\delta}_{t+1} = \alpha_0 + \alpha_1\tilde{\delta}_t + \alpha_2\tilde{s}_t + \alpha_3\tilde{m}_t$ for different specifications of \tilde{m}_t .

The qualitative results are the same for all specifications. The estimated coefficient for media activity, $\hat{\alpha}_3$, is negative and significant in all cases. We focus on the results for the specification with $\tilde{m}_t = \log(1 + m_t)$, because it has the largest value of the log-likelihood function.

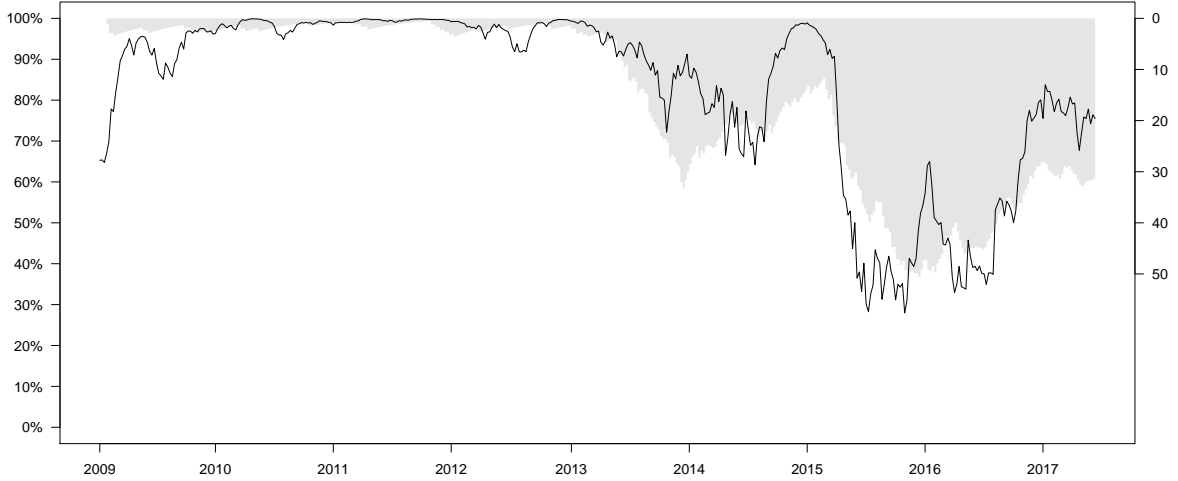


Figure 5: Vaccine compliance in the model augmented with media coverage. The solid line (left axis) is the estimated vaccine compliance in the augmented score-driven model. The shaded area (right axis) is the exponentially weighted moving average of the media variable, m_t .

A detailed review of Danish media coverage of the HPV vaccine, detailed in the Web-Appendix, reveals that coverage was overwhelmingly positive until April 2013. The first article in mainstream media that associated the HPV vaccine with serious side effects was published in Politiken, a leading Danish newspaper, on April 17th, 2013, which coincides with the first episode with relatively low values of δ_t . The article in Politiken had a story about a girl who had a number of symptoms, including frequent headaches, dizziness, and tiredness, which her parents said were caused by the HPV vaccine. The article was followed by a series of articles in the same newspaper that discussed the possibility of a link between the HPV vaccine and serious adverse events. An article in Politiken on May 3, 2013 raised doubt about the efficacy of the vaccine, and accused the Danish Health Authority for being “absolutely misleading” in their information about the HPV vaccine and cervical cancer.

In the model, (6), the media variable has a direct impact on $\tilde{\delta}_{t+1}$, and also a large indirect impact on $\tilde{\delta}_{t+2}$, $\tilde{\delta}_{t+3}$, etc. because the so-called *impulse response* function is given by $d\tilde{\delta}_{t+1}/d\tilde{m}_{t-j} = \alpha_3\alpha_1^j$, $j = 0, 1, \dots$, and α_1 is estimated to be close to 1. The aggregate media impact on $\tilde{\delta}_{t+1}$ is given by $\alpha_3 \sum_{j \geq 0} \alpha_1^j \tilde{m}_{t-j} = \alpha_3 M_t$ where $M_t = \sum_{j \geq 0} \alpha_1^j \tilde{m}_{t-j}$, which is an exponentially weighted average of past media activity. In Figure 5 we present the estimated time series of vaccine compliance, $\hat{\delta}_t$, and the aggregated media activity variable, M_t , based on the estimated model (6). The variation in $\hat{\delta}_t$ is similar to that in the model without the media variable, because $\hat{\delta}_t$ is primarily driven by the weekly number of vaccinations. The largest decline in vaccine compliance occurred in March 2015 which coincides with a large increase in the

number of media articles that related the HPV vaccine to possible side effects. Many of these articles are related to the TV documentary that we focus on in the next Section.

4.2 Vaccine Uptake following the TV2 Documentary on March 26, 2015

On March 26, 2015, TV2 Denmark aired a documentary on the HPV vaccine. The documentary, entitled “The Vaccinated Girls – Sick and Abandoned”, presented personal and emotionally charged stories of girls who claimed that their sicknesses were caused by the HPV vaccine. The documentary was viewed by nearly 10% of the population aged 12 years and above (466,000 viewers). In a newspaper article, the TV2 editor responsible for the documentary acknowledged that the documentary contributed to the large decline in HPV vaccine uptake in 2015, see Sjöberg (2018). This view is supported by the raw vaccination data and our empirical analysis. While our results cannot establish causality, the coincidence between the TV documentary and the drop in uptake represents a smoking gun in the absence of any good alternative explanation. The results based on the augmented model, (6), show that the number of media stories significantly predicts the decline in vaccine uptake. Many of these media stories in 2015 were referencing the TV2 documentary. In fact, a media search on infomedia.dk reveals that the TV2 documentary was referenced 255 times in Danish media during 2015, of which 170 occurred before the end of April. At least 41 of the references to the documentary were in stories produced by TV2. So the spike in media activity, m_t , observed after March 2015 can be ascribed to the TV2 documentary.

The present framework enables us to compare actual vaccine uptake with counterfactual scenarios, such as the scenario where vaccine compliance had remained constant at the level it had before the TV2 documentary aired. This is illustrated in Figure 6, where the solid line shows the difference between expected number of vaccinated girls under full compliance ($\delta = 1$) and actual number of vaccinated girls (i.e. $\Lambda(a_t^c) \times N_c$ minus $X_{c,t}$) for girls born in 2003. The vertical line in early 2015 is the time that the TV2 documentary was aired, and we observed a distinct break in the number of unvaccinated girls. The dotted line presents the corresponding number of unvaccinated girls in the hypothetical scenario where δ_t remained constant at the value it had by the end of March 2015 (95%). Under the hypothetical scenario, we would have expected about 2,000 unvaccinated girls (born in 2003) by the end of our sample period. The actual number turned out to be substantially larger, and the discrepancy between the solid line and the dotted line shows how many fewer girls were vaccinated. At the peak there were

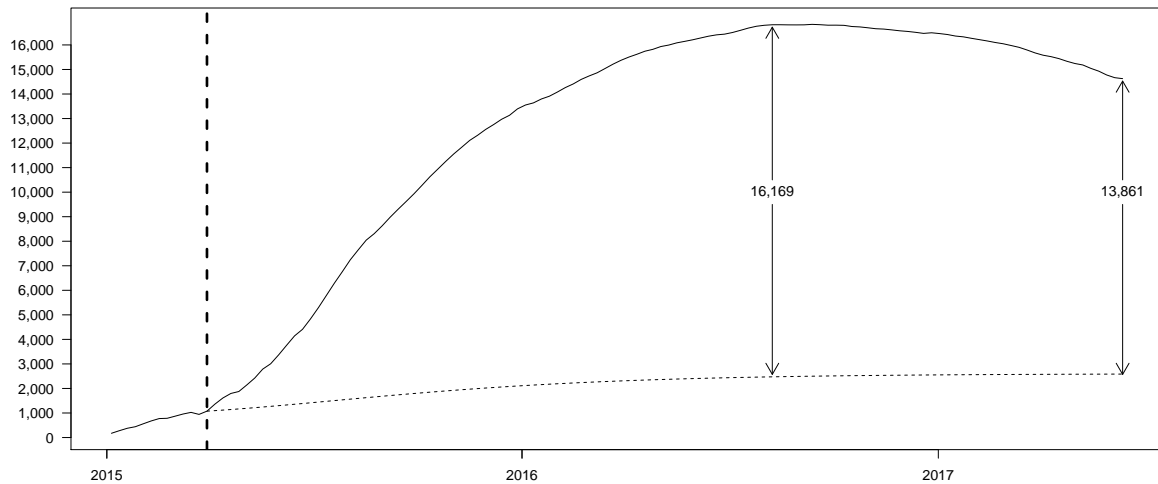


Figure 6: Missing vaccinations for girls born in 2003. The solid line displays the number of unvaccinated girls, born in 2003, relative to a hypothetical baseline with 100% compliance. The dotted line is the corresponding number of missing vaccinations had compliance been stable at 95%. The difference between the two lines is the additional number of unvaccinated girls resulting from the sharp decline in vaccine uptake starting in late March 2015.

more than 16,000 additional unvaccinated girls in cohort 2003 alone. Some of these girls were vaccinated later, so that the excess number of unvaccinated girls, born in 2003, is over 13,800 by the end of our sample period (June 2017).

The implications of 13,800 unvaccinated girls can be estimated as follows. Annually, in Denmark, there are about 375 new cases of cervical cancer and about 100 deaths caused by cervical cancer. Moreover, about 6,000 cone biopsies are made annually as a preventive measure to cervical cancer. These figures may be compared with a cohort size of 36,737 women (the average cohort size for women currently between 35 and 49 years old that are the cohorts with the highest incidence of cervical cancer). The HPV vaccines that were offered in the Danish vaccine program during our sample period protect against the HPV virus types 16 and 18, that are responsible for about 70% of the cases of cervical cancer. So, if we assume that the vaccine is 100% effective against HPV types 16 and 18, then 13,800 unvaccinated girls will translate into about 99 preventable cases of cervical cancer and 26 preventable deaths. Moreover, if we make the (quite conservative) assumption that HPV 16 and 18 are the cause of 50% of abnormal cell changes that are removed by a cone biopsy, then 13,800 unvaccinated girls translate into 1,127 preventable cone biopsies.

4.3 Some Relevant International Comparisons

We have established that Danish media coverage on the HPV vaccine and suspected side effects is a significant predictor of declining vaccine uptake in Denmark. That Danish media coverage played a key role in decreasing HPV vaccine uptake is further bolstered by the fact that the HPV vaccine programs in neighboring countries, Sweden and Norway, did not see much variation in uptake. A recent study by Amdisen et al. (2018) compared the Danish HPV vaccine uptake for birth-year cohorts 1998-2000 with that of cohorts 2001-2003. Interestingly they found that the decline in vaccine uptake was significantly smaller for immigrants, which is consistent with Danish media coverage being a key determinant for the decline in HPV vaccine uptake.

It could be hypothesized that Danish media were merely reporting on an unusually high number of suspected side effects. This hypothesis is, however, not consistent with the data. We have obtained the monthly number of adverse events (AE) reports for Denmark and annual figures from Norway. It is important to note that an AE can be self-reported and does not establish that the reported symptoms were caused by the vaccine. As of 2018, only three cases were deemed sufficiently plausible to justify compensation. Before 2015, the number of AE reports per 10,000 doses was similar for Denmark (between 3 and 17) and Norway (between 8 and 17). In 2015, this statistic rose to 153 in Denmark, while nothing out of the ordinary was observed in Norway, see Table 3. From the monthly data, shown in the Web-Appendix, it is also clear that the high number AE reports in 2015 pertains to the period after the TV2 documentary aired, not the period before. There were a total of 822 AE reports in Denmark in 2015, of which 11 and 15 were reported in January and February, respectively. In the following month, when the TV2 documentary first aired, this figure jumped to 53, then to 70 in April, before it peaked at 150 in May of 2015. The timing suggests that increased levels of media coverage precede increased reporting of AE, similar to the case of New Zealand, see Faasse et al. (2017). It therefore supports the view that the high number of AE reports in 2015 is due to a hesitancy-induced false association between vaccinations and causally unrelated symptoms, rather than a localized occurrence of theretofore unrecognized actual side effects. In the Web-Appendix, we also investigate Google Search activity related to vaccine side effects in Denmark and Norway.

There is evidence that the TV2 documentary influenced vaccine uptake beyond Danish borders. An Irish documentary, "Cervical Cancer Vaccine - Is it safe?", aired on December 14, 2015 on the channel TV3 in Ireland. The Irish documentary included many segments from the

Table 3: Media Coverage and Number of Adverse Events Reports

Year	Media Count	Adverse Events (AE)	AE per 10k doses	AE per 10k Norway
2010	4	67	13	17
2011	8	43	3	12
2012	5	95	3	11
2013	415	512	10	8
2014	140	192	17	11
2015	1329	822	153	12
2016	489	307	109	10

Note: Danish annual figures on: HPV vaccine related newspaper articles; Adverse events (AE) reported for the HPV vaccine; and AE per 10,000 doses of HPV vaccine given. Date retrieved from Lægemiddelstyrelsen (2016) and Sundhedsstyrelsen (2017). For comparison, we include the number of AE per 10,000 doses of HPV vaccine in Norway, based on an estimated 76,000 doses per year.

Danish documentary with English subtitles, and HPV vaccine compliance fell from 89.7% for the cohort vaccinated before the documentary aired to 55.8% for the cohort vaccinated after the documentary. In addition, TV2 Denmark made their documentary available for free on YouTube, where subtitles in various languages were added in due course. The decision to make the documentary freely available on YouTube was uncharacteristic for TV2 Denmark, because the station typically keeps its contents behind a paywall.

It is interesting to compare the Danish media coverage with a similar episode in the US. In 2013 the TV network, CBS, aired a controversial TV program on the HPV vaccine in 2013, to which US media responded very differently from that in Denmark. On December 4, 2013, the CBS TV talk show, *Katie*, covered what the host referred to as the “HPV vaccine controversy”. Similar to the TV2 documentary, the program presented emotionally charged and personal stories of individuals who claimed having been injured by the HPV vaccine, and only scantily mentioned the scientific consensus on the matter. The reaction from the US media was prompt with scolding criticism. On the day of its airing, several media outlets criticized the talk show in articles with castigating headlines such as: “Katie Couric Hands Her Show Over to Anti-Vaccination Alarmists” (Slate), “Katie Couric puts the anti-vaccination movement into the mainstream” (LA Times), and “Is Katie Couric the Next Jenny McCarthy?” (TIME), where the TIME article concludes with: “Couric’s misdeeds are all the worse given that she’s taken much more seriously than [...]”. Shortly after, on December 10, 2013, the TV host apologized and conceded that her program had been “too anti-vaccine and anti-science”. The instant media criticism in the US and Katie Couric’s prompt apology may have averted a subsequent decline in HPV vaccine uptake in the US, similar to the one observed in Denmark in 2015. In contrast,

it was only after the TV2 documentary had been nominated for a prestigious journalistic prize, nine month later, that Danish media expressed the first criticism of the TV2 documentary.

5 Summary and Discussion

Vaccination is arguably one of the most important public health achievements in history. However, its success critically relies on the personal decisions to comply with recommended vaccine schedules. Vaccine refusal is increasingly becoming a problem in many countries, as exemplified by ongoing outbreaks of measles in the US and several European countries. Understanding the mechanisms that determine vaccine hesitancy is therefore of great importance to health authorities, including the influence that media have in the personal decision to comply with, or deviate from, recommended vaccine schedules.

In this paper we have developed a dynamic model for vaccine compliance that is driven by discrepancies between expected and actual vaccination rate. The model could serve as a tool for health authorities to monitor vaccine hesitancy in real time. We applied the model to Danish HPV vaccination data, which has experienced a great deal of variation in vaccine uptake since the vaccine was introduced in the Danish childhood vaccination program in 2009. The econometric model we have proposed in this paper is the first score-driven model for binomially distributed variables. The model could be adapted to other time series involving aggregated binary outcome variables and is therefore not specific to vaccine data.

Our empirical analysis supports the view that Danish media played an important role in the collapse in the Danish HPV vaccination program. Our results show that the first decline in HPV vaccine uptake coincides with the period where Danish media began running stories that associated the HPV vaccine with serious side effects, and the largest decline in vaccine uptake occurred immediately after a vaccine-critical TV documentary. The evidence that media coverage influenced vaccine uptake is strengthened further by the empirical results from the augmented specification that includes a media coverage variable. For this specification we find that the media variable is a significant predictor of HPV vaccine uptake. The larger the number of media stories that associate the HPV vaccine with side effects, the lower is the expected HPV1 vaccine uptake in subsequent weeks. The empirical evidence points to the TV2 documentary as the main culprit for the sharp decline in vaccine uptake in 2015. Primarily because the sharp decline in HPV vaccinations occurred immediately after the documentary aired, and because there is no good alternative explanation for the decline to have occurred

at that particular point in time. This is also consistent with the empirical results from the augmented specification, because a very large number (255) of media stories in 2015 referenced the TV2 documentary, including 41 stories produced by TV2 Denmark. In February 2018, TV2 Denmark acknowledged that their documentary contributed to the demise of the HPV vaccination program, but emphasized that this had not been the intention of the documentary.

The model framework introduced in this paper makes it possible to quantify the effect that the decline in vaccine compliance has had on the vaccination program. We estimate that more than 16,000 girls born in 2003 delayed HPV vaccination as a result of the sharp decline in vaccine uptake after March 2015. Many of these girls have since been vaccinated. On May 10th, 2017, the Danish health authorities and the Danish Cancer Society started a campaign, “Stop HPV”, in an attempt to rehabilitate the HPV vaccine program. The latest data, as of March 2019, shows that the campaign has partially succeeded in increasing vaccine uptake. For example, 37% of cohort 2005 received HPV1 during 2017 (the year they turned twelve), and 46% of cohort 2006 received HPV1 during 2018. These figures compare favorably to 30% and 25%, which were the corresponding percentages for cohorts 2003 and 2004, respectively. However, they fall well short of the vaccine uptake before 2013. For comparison, about 75% of the girls born in 1999 and 2000 were HPV1 vaccinated during the year they turned 12 years old. The campaign has also had a positive effect on cohort 2003, of which 80% are HPV1 vaccinated as of March 2019. However, this figure is also well short of the coverage for older cohorts, which is 94% for girls born in 1998, 1999, or 2000.

Our empirical analysis of Danish HPV vaccination data is based on anonymized data, which limits a deeper analysis of the individuals that declined HPV vaccination. It would be interesting to investigate whether media effects vary with socio-economic characteristics. Following the infamous (and retracted) study that led the public to suspect a relationship between the MMR vaccine and autism, vaccine compliance fell throughout the UK (and elsewhere). Interestingly, Anderberg et al. (2011) found that the decline in MMR vaccine compliance was more pronounced in areas with a higher fraction of educated individuals and a higher average income. It would also be interesting to investigate how prior knowledge about factual information on vaccines affects sensitivity with respect to media stories. In an online experiment with French voters, Rodriguez et al. (2018) found that respondents who had accurate beliefs about immigration statistics were not misled by fake news, while those with incorrect priors often were. Lastly, social media may also have played an important role, albeit there is mixed evidence on the influence of social

media. For instance, Allcott and Gentzkow (2017) observed that survey respondents were as likely to believe in fake stories that had circulation on social media, as they were likely to believe in fake stories that had not circulated on social media. In the case of Danish HPV vaccination, social media likely played some role in the decline in HPV vaccine uptake. Social media were a catalyst for much of the media coverage. Links to the TV2 documentary (on YouTube) and articles that related the HPV vaccine to serious side effects were frequently shared on social media. We leave an investigation into the role of social media and HPV vaccinations for future research.

Acknowledgements

We thank Statens Serum Institute for providing the Danish vaccination data, Gridt Vig Find for collecting data on Danish media websites, and Anna Lunde for explaining how vaccinations are reported by physicians and how this protocol has changed over time. We also thank Patrick Conway, Luca Flabbi, and seminar participants at the University of Konstanz, PUC in Rio de Janeiro, the European University Institute, Michigan State University, Paris Nanterre University, and conference participants at the 28th EC2 conference in Amsterdam, the 2018 Kochi International Seminar, and the 2019 INET Cambridge Score Conference for valuable comments. Potential conflicts of interest: None.

References

- Allcott, H. and Gentzkow, M. (2017), ‘Social media and fake news in the 2016 election’, *Journal of Economic Perspectives* pp. 211–236.
- Alsan, M. and Wanamaker, M. (2017), ‘Tuskegee and the health of black men’, *The Quarterly Journal of Economics* **133**(1), 407–455.
- Amdisen, L., Kristensen, M. L., Rytter, D., Mølbak, K. and Valentiner-Branth, P. (2018), ‘Identification of determinants associated with uptake of the first dose of the human papillomavirus vaccine in Denmark’, *Vaccine* **36**, 5747–5753.
- Anderberg, D., Chevalier, A. and Wadsworth, J. (2011), ‘Anatomy of a health scare: Education, income and the MMR controversy in the UK’, *Journal of Health Economics* **30**, 515–530.
- Arbyn, M., Xu, L., Simoens, C. and Martin-Hirsch, P. P. L. (2018), ‘Prophylactic vaccination against human papillomaviruses to prevent cervical cancer and its precursors’, *Cochrane Database of Systematic Reviews* **5**.

- Bollerslev, T. (1986), ‘Generalized autoregressive heteroskedasticity’, *Journal of Econometrics* **31**, 307–327.
- Caivano, M., Harvey, A. and Luati, A. (2016), ‘Robust time series models with trend and seasonal components’, *SERIEs* **7**, 99–120.
- Creal, D. D., Koopman, S. J. and Lucas, A. (2013), ‘Generalized autoregressive score models with applications’, *Journal of Applied Econometrics* **28**, 777–795.
- DellaVigna, S. and La Ferrara, E. (2015), Economic and social impacts of the media, in S. P. Anderson, J. Waldfogel and D. Strömberg, eds, ‘Handbook of Media Economics’, Vol. 1, Elsevier, pp. 723–768.
- Faasse, K., Porsius, J. T., Faasse, J. and Martin, L. R. M. (2017), ‘Bad news: The influence of news coverage and Google searches on Gardasil adverse event reporting’, *Vaccine* **35**, 6872–6878.
- Hansen, P. R. and Schmidtblaicher, M. (2019), ‘Web-appendix for ‘a dynamic model of vaccine compliance: How fake news undermined the Danish HPV vaccine program’’, <https://sites.google.com/site/peterreinhardhansen/>.
- Kearney, M. S. and Levine, P. B. (2015), ‘Media influences on social outcomes: The impact of MTV’s 16 and pregnant on teen childbearing’, *American Economic Review* **105**, 3597–3632.
- Koopman, S. J., Lucas, A. and Monteiro, A. (2008), ‘The multi-state latent factor intensity model for credit rating transitions’, *Journal of Econometrics* **142**, 399–424.
- Koopman, S. J., Lucas, A. and Schwaab, B. (2011), ‘Modeling frailty-correlated defaults using many macroeconomic covariates’, *Journal of Econometrics* **162**, 312–325.
- La Ferrara, E., Chong, A. and Duryea, S. (2012), ‘Soap operas and fertility: Evidence from brazil’, *American Economic Journal: Applied Economics* **4**, 1–31.
- Lægemiddelstyrelsen (2016), ‘Nyt om birkninger, marts 2016’, *Retrieved from: https://laegemiddelstyrelsen.dk*.
- Lowes, S. and Montero, E. (2018), ‘The legacy of colonial medicine in central africa’. Working Paper.
- Martinez-Bravo, M. and Stegmann, A. (2018), ‘In vaccines we trust? the effects of the cia’s vaccine ruse on immunization in pakistan’. Working Paper.
- Offit, P. A. (2011), *Deadly Choices: How the Anti-Vaccine Movement Threatens Us All*, Basic Books, New York.
- Rodriguez, O. B., Guriev, S. M., Emeric, H. and Zhuravskaya, E. (2018), ‘Facts, alternative facts, and fact checking in times of post-truth politics’. Available at SSRN: <https://ssrn.com/abstract=3004631> or <http://dx.doi.org/10.2139/ssrn.3004631>.
- Sjöberg, S. S. (2018), ‘Tv 2 om hpv-dokumentar: Vi er medansvarlige for at forældre undlod at vaccinere deres børn’, *Retrieved from: https://www.fyens.dk*.

Sundhedsstyrelsen (2017), ‘Børnevaccinationsprogrammet. årsrapport 2016’, *Retrieved from: <https://www.sst.dk>*.

Suppli, C. H., Hansen, N. D., Rasmussen, M., Valentiner-Branth, P., Krause, T. G. and Mølbak, K. (2018), ‘Decline in HPV-vaccination uptake in Denmark – the association between HPV-related media coverage and HPV-vaccination’, *BMC Public Health* **18**.

WHO (2016), ‘Human papillomavirus (hpv) and cervical cancer’, *World Health Organization Fact Sheet*.

A Appendix: Derivations of Various Results

A.1 Score for Vaccine Compliance

In this section we establish the result in (5). We seek the first and second derivatives of the log-likelihood function with respect to $\tilde{\delta}$. First observe that

$$\frac{\partial p_{c,t}}{\partial \tilde{\delta}_t} = \frac{\partial}{\partial \tilde{\delta}_t} \frac{e^{\tilde{\delta}_t}}{1+e^{\tilde{\delta}_t}} \lambda_{c,t} \eta_t = \frac{e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^2} \lambda_{c,t} \eta_t,$$

where $\lambda_{c,t} = \frac{N_c}{N_{c,t-1}} [\Lambda(\beta; a_t^c) - \Lambda(\beta; a_{t-1}^c)]$. Thus from (3) we have that

$$s_{c,t} = \frac{\partial \ell_{c,t}}{\partial \tilde{\delta}_t} = \frac{\partial p_{c,t}}{\partial \tilde{\delta}_t} \left[\frac{x_{c,t}}{p_{c,t}} - \frac{N_{c,t-1} - x_{c,t}}{1-p_{c,t}} \right] = \frac{e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^2} \lambda_{c,t} \eta_t N_{c,t-1} \left[\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}} \right],$$

where $\hat{p}_{c,t} = x_{c,t}/N_{c,t-1}$. The score for $\tilde{\delta}_t$ is therefore given by

$$s_t = \sum_c s_{c,t} = \frac{e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^2} \sum_c \lambda_{c,t} \eta_t N_{c,t-1} \left[\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}} \right].$$

Next, for the second derivative, we have

$$h_{c,t} = \frac{\partial^2 \ell_{c,t}}{\partial \tilde{\delta}_t^2} = \frac{\partial}{\partial \tilde{\delta}_t} \frac{e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^2} \lambda_{c,t} \eta_t N_{c,t-1} \left[\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}} \right].$$

Now

$$\frac{\partial}{\partial \tilde{\delta}_t} \frac{e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^2} = \frac{e^{\tilde{\delta}_t} (1+e^{\tilde{\delta}_t})^2 - e^{\tilde{\delta}_t} e^{\tilde{\delta}_t} 2(1+e^{\tilde{\delta}_t})}{(1+e^{\tilde{\delta}_t})^4} = \frac{-e^{3\tilde{\delta}_t} + e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^4} = \frac{e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^2} \frac{1-e^{\tilde{\delta}_t}}{1+e^{\tilde{\delta}_t}},$$

and $\frac{\partial}{\partial \tilde{\delta}_t} \left[\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}} \right] = -\frac{e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^2} \lambda_{c,t} \eta_t \left[\frac{\hat{p}_{c,t}}{p_{c,t}^2} + \frac{1-\hat{p}_{c,t}}{(1-p_{c,t})^2} \right]$. Using that $\mathbb{E}_{t-1}(\hat{p}_{c,t}) = p_{c,t}$, we have

$$-\mathbb{E}_{t-1} h_{c,t} = \frac{e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^2} \lambda_{c,t} \eta_t N_{c,t-1} \frac{e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^2} \lambda_{c,t} \eta_t \frac{1}{p_{c,t}(1-p_{c,t})} = \left[\frac{e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^2} \right]^2 \frac{\lambda_{c,t}^2 \eta_t^2 N_{c,t-1}}{p_{c,t}(1-p_{c,t})},$$

hence

$$\left(-\sum_c \mathbb{E}_{t-1} h_{c,t}\right)^{-1/2} = \left(\left[\frac{e^{\tilde{\delta}_t}}{(1+e^{\tilde{\delta}_t})^2}\right]^2 \sum_c \frac{\lambda_{c,t}^2 \eta_t^2 N_{c,t-1}}{p_{c,t}(1-p_{c,t})}\right)^{-1/2} = \frac{(1+e^{\tilde{\delta}_t})^2}{e^{\tilde{\delta}_t}} / \sqrt{\sum_c \frac{\lambda_{c,t}^2 \eta_t^2 N_{c,t-1}}{p_{c,t}(1-p_{c,t})}}.$$

Thus $\tilde{s}_t = (-\sum_c \mathbb{E}_{t-1} h_{c,t})^{-1/2} \sum_c s_{c,t}$ equals

$$\frac{\sum_c \lambda_{c,t} \eta_t N_{c,t-1} \left(\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}}\right)}{\sqrt{\sum_c \frac{\lambda_{c,t}^2 \eta_t^2 N_{c,t-1}}{p_{c,t}(1-p_{c,t})}}} = \frac{\sum_c N_c \tilde{\lambda}_{c,t} \left(\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}}\right)}{\sqrt{\sum_c \frac{N_c^2}{N_{c,t-1}} \frac{\tilde{\lambda}_{c,t}^2}{p_{c,t}(1-p_{c,t})}}} \simeq \frac{\sum_c \tilde{\lambda}_{c,t} \left(\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}}\right)}{\sqrt{\sum_c \frac{1}{N_{c,t-1}} \frac{\tilde{\lambda}_{c,t}^2}{p_{c,t}(1-p_{c,t})}}},$$

where $\tilde{\lambda}_{c,t} = \Lambda(\beta; a_t^c) - \Lambda(\beta; a_{t-1}^c)$.

B Model of Seasonal Effects

In this appendix we detail the part of the model that accounts for the seasonal effect in the weekly vaccination data, and present the corresponding empirical results.

B.1 Score for Seasonal Component

We model seasonal effects by enhancing the model for $p_{c,t}$ with a third component. The binomial parameter is now decomposed as

$$p_{c,t}(\theta) = \delta_t(\alpha) \lambda_{c,t}(\beta) \eta_t(\gamma), \quad (\text{B.1})$$

so that the parameter vector is given by $\theta = (\alpha', \beta', \gamma')'$.

To account for time variation in seasonality, we model $\tilde{\eta} = \log(\eta)$ by a separate score-driven model

$$\tilde{\eta}_t = g_{0,t} \sin(2\pi(z_t^m + g_{1,t}))(1 - z_t^a) + \gamma_2 z_t^a,$$

where $g_{i,t} = g_{i,t-1} + \gamma_i \tilde{s}_{g,i,t-1}$ for $i = 0, 1$. Analogous to (5), $\tilde{s}_{g,i,t-1}$ is the scaled score with respect to the seasonality parameters $g_{0,t}$ and $g_{1,t}$. Specifically,

$$\tilde{s}_{g,0,t} = \tilde{s}_t \text{sign}(\sin(2\pi(z_t^m + g_{1,t}))(1 - z_t^a)), \quad (\text{B.2})$$

$$\tilde{s}_{g,1,t} = \tilde{s}_t \text{sign}(g_{0,t} \cos(2\pi(z_t^m + g_{1,t}))(1 - z_t^a)). \quad (\text{B.3})$$

The expressions for (B.2) and (B.3) are derived below. In the estimation of the model, we treat the initial values for $(g_{0,0}, g_{1,0}) = (\gamma_3, \gamma_4)$ as free parameters, with domains $\gamma_3 \geq 0$ and $\gamma_4 \in [0, 1]$, respectively.

The seasonal variables, z_t^m and z_t^a , are defined as follows. First, z_t^m represents the location of week t in the month, as defined by the date of the Monday of that week, divided by the number of days in the months. For example, a week with a Monday on January 12th translates into $z_t^m = 12/31$. Second, z_t^a is a binary variable that takes the value one during the summer vacation period (week numbers 28 to 31) as well as the two weeks around Christmas/New Year (week numbers 52 and 1).

Next, we derive (B.2) and (B.3) for the case $z_t^a = 0$. We seek the first and second derivatives of the log-likelihood function with respect to $g_{0,t}$ and $g_{1,t}$. First note that $\frac{\partial p_{c,t}}{\partial g_{0,t}} = p_{c,t} \sin(2\pi(z_t^m + g_{1,t}))$ and $\frac{\partial p_{c,t}}{\partial g_{1,t}} = p_{c,t} g_{0,t} 2\pi \cos(2\pi(z_t^m + g_{1,t}))$. From (3) we have that

$$s_{g,1,c,t} = \frac{\partial \ell_{c,t}}{\partial g_{0,t}} = \frac{\partial p_{c,t}}{\partial g_{0,t}} \left[\frac{x_{c,t}}{p_{c,t}} - \frac{N_{c,t-1} - x_{c,t}}{1-p_{c,t}} \right] = p_{c,t} N_{c,t-1} \left[\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}} \right] \sin(2\pi(z_t^m + g_{1,t}))$$

and $s_{g,2,c,t} = \frac{\partial \ell_{c,t}}{\partial g_{1,t}} = p_{c,t} N_{c,t-1} \left[\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}} \right] g_{0,t} 2\pi \cos(2\pi(z_t^m + g_{1,t}))$. For the second derivatives, note that $\frac{\partial}{\partial g_{0,t}} \left[\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}} \right] = -p_{c,t} \left[\frac{\hat{p}_{c,t}}{p_{c,t}^2} + \frac{1-\hat{p}_{c,t}}{(1-p_{c,t})^2} \right] \sin(2\pi(z_t^m + g_{1,t}))$ and

$$\frac{\partial}{\partial g_{1,t}} \left[\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}} \right] = -p_{c,t} \left[\frac{\hat{p}_{c,t}}{p_{c,t}^2} + \frac{1-\hat{p}_{c,t}}{(1-p_{c,t})^2} \right] g_{0,t} 2\pi \cos(2\pi(z_t^m + g_{1,t})).$$

Since $\mathbb{E}_{t-1}(\hat{p}_{c,t}) = p_{c,t}$, we therefore have that

$$-\mathbb{E}_{t-1} \frac{\partial^2 \ell_{c,t}}{\partial g_{0,t}^2} = N_{c,t-1} \frac{p_{c,t}^2}{p_{c,t}(1-p_{c,t})} \sin^2(2\pi(z_t^m + g_{1,t}))$$

and

$$-\mathbb{E}_{t-1} \frac{\partial^2 \ell_{c,t}}{\partial g_{1,t}^2} = N_{c,t-1} \frac{p_{c,t}^2}{p_{c,t}(1-p_{c,t})} g_{0,t}^2 4\pi^2 \cos^2(2\pi(z_t^m + g_{1,t})).$$

Hence

$$\left(-\sum_c \mathbb{E}_{t-1} \frac{\partial^2 \ell_{c,t}}{\partial g_{0,t}^2} \right)^{-1/2} = \left(\sin^2(2\pi(z_t^m + g_{1,t})) \sum_c \frac{p_{c,t}^2 N_{c,t-1}}{p_{c,t}(1-p_{c,t})} \right)^{-1/2}$$

and

$$\left(-\sum_c \mathbb{E}_{t-1} \frac{\partial^2 \ell_{c,t}}{\partial g_{1,t}^2} \right)^{-1/2} = \left(g_{0,t}^2 \cos^2(2\pi(z_t^m + g_{1,t})) \sum_c \frac{p_{c,t}^2 N_{c,t-1}}{p_{c,t}(1-p_{c,t})} \right)^{-1/2}.$$

Finally,

$$\begin{aligned}\tilde{s}_{g,0,t} &= \left(-\sum_c \mathbb{E}_{t-1} \frac{\partial^2 \ell_{c,t}}{\partial g_{0,t}^2} \right)^{-1/2} \sum_c s_{g,1,c,t} = \frac{\sum_c \lambda_{c,t} N_{c,t-1} \left[\frac{\hat{p}_{c,t}}{p_{c,t}} - \frac{1-\hat{p}_{c,t}}{1-p_{c,t}} \right] \sin(2\pi(z_t^m + g_{1,t}))}{\sqrt{\sin^2(2\pi(z_t^m + g_{1,t})) \sum_c \frac{\lambda_{c,t}^2 N_{c,t-1}}{p_{c,t}(1-p_{c,t})}}} \\ &= \tilde{s}_t \text{sign}(\sin(2\pi(z_t^m + g_{1,t}))),\end{aligned}$$

and, similarly, $\tilde{s}_{g,1,t} = \tilde{s}_t \text{sign}(g_{0,t} \cos(2\pi(z_t^m + g_{1,t})))$. In the case where $z_t^a = 1$, we define $\tilde{s}_{g,0,t} = \tilde{s}_{g,1,t} = 0$.

B.2 Empirical Results related to Seasonal Effects

Table B.1: Estimates for Score-Driven Model

$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\delta}_0$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\gamma}_4$	$\ell(\hat{\theta})$
0.000 (0.087)	0.992 (0.022)	0.054 (0.017)	0.939 (0.005)	7.25 (0.038)	3.03 (0.088)	0.001 (0.007)	0.002 (0.022)	-0.757 (0.087)	0.246 (0.042)	0.384 (0.081)	-20122
0	1	0.066 (0.004)	0.883 (0.000)	7.24 (0.018)	3.01 (0.003)	0.001 (0)	0.002 (0.000)	-0.679 (0.008)	0.285 (0.011)	0.465 (0.010)	-20318

Note: Standard errors in parentheses.

Table B.2: Estimates for Score-Driven Model with Media Variable

	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\delta}_0$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\gamma}_0$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\gamma}_4$	$\ell(\hat{\theta})$
(1)	0.182 (0.011)	0.948 (0.007)	0.050 (0.006)	-0.080 (0.018)	0.668 (0.004)	7.24 (0.019)	3.000 (0.003)	0.001 (0.000)	0.002 (0.000)	-0.755 (0.006)	0.304 (0.010)	0.488 (0.008)	-19951
(2)	0.087 (0.003)	0.967 (0.014)	0.052 (0.012)	-0.004 (0.000)	0.766 (0.003)	7.24 (0.010)	3.000 (0.021)	0.001 (0.000)	0.002 (0.000)	-0.756 (0.018)	0.287 (0.013)	0.491 (0.004)	-20033
(3)	0.160 (0.007)	0.952 (0.004)	0.052 (0.012)	-0.045 (0.013)	0.685 (0.003)	7.24 (0.015)	3.000 (0.005)	0.001 (0.004)	0.002 (0.000)	-0.753 (0.005)	0.284 (0.008)	0.489 (0.006)	-19971

Note: Standard errors in parentheses.

C The Number of Girls Eligible for HPV1 Vaccination

In Denmark the HPV vaccine is given for free after the twelfth birthday. So the number of girls eligible for their first vaccine dose (HPV1) increases gradually over the year in which the cohort turns twelve. We adapt the definition of $N_{c,t}$ accordingly. Specifically we let the number of unvaccinated girls be $N_{c,t} = \lfloor g(a_t^c) N_c \rfloor - X_{c,t}$ where $g(a_t^c)$ is a positive and increasing function that fulfills $g(0) = 0$ and $g(a) = 1$ for $a \geq \frac{1}{3}$ and where $\lfloor \cdot \rfloor$ is the integer part operator, and N_c is the birth-year cohort size.

We do not have information about the birthday of individuals in our anonymized data, but we can approximate the distribution of birthdays over the year for each cohort. Monthly birth statistics show that the birth distribution over the calendar year is almost identical across

cohorts. So we estimate $g(a)$ using daily births for 2007. These data were obtained from Statistics Denmark (Danmarks Statistik). The estimation of $g(a)$ is fully detailed in the Web-Appendix.

C.1 Robustness

It is worth mentioning that this time-of-birth correction has only a negligible effect on the estimation results. In fact, we also estimated the model using $g(a) = 1$ for $a \in [0, 1]$ (which assumes all girls are eligible for the vaccine when $a = 0$), as well as with $g(a) = \min(3a, 1)$, which assumes an even distribution of births over the year. The alternative specifications for g did not change the empirical results in any substantive way.

As an additional robustness check, we have also estimated the model where $x_{c,t}$ is assumed to be conditionally Poisson distributed (rather than conditionally Binomially distributed). These results were also very similar to the ones reported here with the Binomial specification. Finally, we also experimented with the specification for $\Lambda(\beta; a)$, and obtained very similar results with a truncated Weibull cumulative distribution function.