eAppendix1: "Increased incidence rate of trauma- and stressor-related disorders in Denmark following the Breivik attacks in Norway"

SUPPLEMENTAL METHODS

For this study we used data from the Danish Psychiatric Central Research Register (DPCRR),[1] which is a well-established source of data for research purposes.[2] The DPCRR contains diagnoses on all individuals assessed/treated for mental disorders at psychiatric services (inpatient-, outpatient-, as well as emergency room settings) in Denmark (including Greenland and The Faroe Islands). For these contacts, a primary diagnosis is assigned by the treating psychiatrist in accordance with the criteria from the 10th edition of the International Classification of Disease (ICD-10).[3] In this study, we based our analyses on acute contacts with psychiatric services by excluding all contacts that were labeled as "non-acute" or "unknown status", as well as all outpatient contacts. This was done in order to avoid potential misclassification of contacts, which were planned prior to the Breivik attacks, but carried out and registered in the DPCRR after the attacks. The study was approved by the Danish Data Protection Agency, the State Serum Institute and Statistics Denmark.

Based on the results of our analysis of the 9/11 attacks,[4] the outcome of our study was the incidence of diagnoses of "Reaction to severe stress, and adjustment disorders" (ICD-10 code: F43), which we will refer to as "trauma- and stressor-related disorders".

In accordance with our recent study of the effects of the 9/11 attacks,[4] we generated a time series counting the daily number of reaction to severe stress, and adjustment disorders diagnoses registered in the DPCRR from January 1, 1995 to December 31, 2012. We then aggregated these data to the weekly level in order to reduce noise. After aggregation each observation reflected the average daily number of diagnoses in a given week (we refer to this as the "incidence rate") adding up to a total of 940 weekly observations. We then carried out a time series intervention analysis using the week of the Breivik attacks – which occurred on Friday July 22, 2011 – as the intervention period.[5] The incidence data was aggregated such that each week began on a Saturday, thereby ensuring that the first post-intervention week would begin immediately after the attacks. In the analyses reported, we used the full pre-intervention from 1995 until the day of the attack, but in a number of robustness checks we reran the model with various shorter pre-intervention periods (see page 25 of this document).

In order to test whether the Breivik attacks were associated with increased incidence rates of mental disorders, we followed the intervention analysis approach given in Cryer and Chan[6] and originally introduced by Box and Tiao.[7] The central idea of the Box and Tiao approach to intervention analysis is to view the time series being studied as consisting of two parts. One being the natural or unperturbed process, which represents how the time series would have behaved in the absence of an intervention, and the other being the perturbed process, which represents the actual time series after the intervention (in this case the Breivik attacks) has acted upon it.

If the intervention timing is exogenous to the time series – that is, if the incident rates of trauma- and stressor-related disorders in Denmark were not substantially affected by factors that determined the timing of the Breivik attacks, and if the incidents rates themselves did not affect that timing – the pre-intervention part of the time series can be used to obtain an unbiased estimate of the unperturbed series. This counterfactual series can then be compared to the actual post-intervention time series to estimate the effect of the intervention.

Formally, the Box and Tiao approach models the time series as a combination of an autoregressive integrated moving average model process (ARIMA) and a transfer function of a binary variable that estimates how the intervention acts upon this process.

The generic (non-seasonal) intervention model can be written as follows:*

$$y_t = \mu + \frac{\theta(B)}{\phi(B)}\varepsilon_t + \frac{\omega(B)}{\delta(B)}B^k x_t$$

Where:

y is the time series in question or a difference of it. In our case the first difference of the weekly incidence rate of trauma- and stressor-related disorders in Denmark between 1995 and 2012.

t indexes time.

μ is a constant.

B is the backshift operator, defined such that: $B^k y_t = y_{t-k}$.

$\phi(B)$ is the autoregressive parameter, which consists of a polynomial of order p in the backshift operator such that: $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 \ldots \phi_p B^p$.

$\theta(B)$ is the moving average parameter, which consists of a polynomial of order q in the backshift operator such that: $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 \ldots - \theta_q B^q$.

ε is a white noise disturbance term (i.e. $\varepsilon \sim i.i.d.(0, \sigma^2)$).

x is the intervention variable.

$\omega(B)$ is the numerator polynomial of the transfer function for the intervention variable, which consists of a polynomial in the backshift operator of order f such that: $\omega(B) = \omega_0 + \omega_1 B + \omega_2 B^2 \ldots + \omega_f B^f$.

$\delta(B)$ is the denominator polynomial of the transfer function for the intervention variable, which consists of a polynomial in the backshift operator of order g such that: $\delta(B) = 1 - \delta_1 B - \delta_2 B^2 \ldots - \delta_g B^g$.

This expression can be written more succinctly as:

$$y_t = \mu + v(B)x_t + N_t$$

Where $v(B)$ is the transfer function and N_t is the ARIMA noise process. The latter term can be thought of as the unperturbed process; the path of the time series had there been no intervention. It serves as the counterfactual for the analysis and given the exogeneity of x, the deviation from this process, $v(B)$, yields the causal effect of the intervention.

To estimate $v(B)$ for the effect of the Breivik attacks on the incidence rate of trauma- and stressor-related disorders, we went through the following steps:

1) Identified and estimated a tentative ARIMA model for N_t (including any seasonal components) using all weeks prior to the intervention week as input.

In this step, we went through the standard Box-Jenkins procedure to identify tentative values of p,d, and q and (see e.g. Box and Jenkins[8] or Cryer and Chan[6]). We did this both manually and through the algorithm developed by Hyndman and Khandakar.[9] We then inspected the time series for seasonality visually and with both the Canova-Hansen and the Osborn-Chui-Smith-Birchenhall tests.[10]

2) Specified intervention time and transfer function.

In this next step, we specified a shape for the transfer function $v(B)$ and added the remainder of the time series in order to estimate its parameters. After inspecting the time series around the intervention, we hypothesized that the effect

* The exposition extends directly to the seasonal case.

would be parabolic, i.e. increasing gradually after the attacks before returning to pre-intervention levels As a first cut, we modelled the effect as a gradual, but permanent level shift by specifying the transfer function as an autoregressive moving average model (ARMA) of order (1,0) with a step input. The step input variable was equal to 1 all post-intervention weeks and zero in all pre-intervention weeks, including the week of the attacks. To account for a delay in the onset of the effect, we sat $k = 1$ in the main specifications, lagging the step one week. We also applied a parabolic function, since it could model the seemingly temporary increase while maintaining a gradual climb. To do this, we replaced the transfer function with two simple functions of intervention time, defined as:

$$linear = \begin{cases} 0 & if\ t \leq T^* \\ t - T^* & if\ t > T^* \end{cases}$$
$$quadratic = \begin{cases} 0 & if\ t \leq T^* \\ (t - T^*)^2 & if\ t > T^* \end{cases}$$

Where T^* denotes the intervention week. Again we sat $k = 1$ in the main specifications and thus lagged these function one period.

The two function of time elapsed after the intervention were added to the ARIMA model for the unperturbed process as regular exogenous regressors, and together they implied zero deviation from the unperturbed process before the intervention and a parabolic deviation after the intervention. The parabolic model provided a somewhat better fit than the model with the step function and also better captured the decline of the effect at the end of the time series. We therefore focused on the results from the parabolic model in the paper.[*]

After estimation we inspected the residuals from the models through the same Box-Jenkins methodology as in 1) to check whether the ARIMA specification for N_t needed to be modified. This was not the case for any of the models.

3) Tested alternative functions

After estimating the parameters of the two type of functions of intervention time, we tried expanding these with further terms to determine if a better fit for the effect existed. We evaluated specifications according to the AIC values for the model as a whole and the significance of the parameter estimates. The transfer function was expanded with both AR and MA terms, but none improved on the initial (1,0) specification. The parabolic function was similarly expanded by iteratively increasing the degree of the polynomial. The lowest AIC was achieved with a six-degree polynomial, which accurately captured the second hump of the effect (see figures below and in the manuscript).

4) Visually inspected fit of model with and without transfer function

We then used the chosen specifications for $v(B)$ and N_t along with pre-intervention data to predict the post-intervention values of y. In order to visualize whether and how modelling the intervention had improved our ability to predict the incidence rates, we computed both a set of one-step and dynamic predictions from a model with $v(B)$ and N_t together and from one with just N_t on its own. We subsequently graphed these along with the observed y's.

5) Calculated intervention effect

Using the estimates from the final model of step 3), we calculated the estimated effect of the intervention upon the values of y.

6) Carried out placebo tests

In accordance with our recent study of the effects of the 9/11 attacks,[4] we supplemented the Box and Tiao procedure with a series of "placebo" tests. These tests involved estimating the exact same intervention models as in the main analysis, but doing so with the intervention time moved to a different date than that of the Breivik attacks. The logic behind these tests was simply that if the effects we had estimated for the attacks were truly caused by that event, then we should not be able to find similar effects for other dates. If we were to find seemingly strong effects for a large

[*] It is important to note that neither of the two models are well suited to provide forecasts out of the sample, since the step function implies a rather implausible permanent shift, and the parabolic function, at least without further modifications, simply keeps decreasing as time goes to infinity. Their purpose here is solely to model the in-sample increase.

number of dates without any actual interventions, it would cast doubt upon the validity of the effect found for the Breivik attacks and indicate that our result could be due to mere random fluctuations in the data.

We did this by first assigning an intervention to the time series 800 weeks before the true intervention and then using that period as the intervention date to run the exact same models as for the true intervention, while also limiting the analysis to having the same post-intervention period length of 76 weeks. After noting the coefficients of both the step function and the quadratic function estimated with this placebo intervention and this data window, we repeated the process with a new placebo intervention assigned 50 weeks later into the time series. This continued until we reached the last 50-week increment where this was possible, namely the period of the true intervention. This approach resulted in 16 placebo tests for every run, which meant that for each one we could expect between zero and one significant coefficients at the 0.05 level purely due to chance (1/16 = 0.0625).

For both the step and the quadratic function, we applied this procedure in two ways. One in which we always kept the same ARIMA specification that we had chosen for the true intervention, and one in which we chose the ARIMA specification specifically for each placebo intervention using the Hyndman-Khandakar algorithm.[9] The reason for the second type was that since the pre-intervention period was different for each placebo intervention it was possible that a different model of the error process, N_t, was required.

SUPPLEMENTAL RESULTS

Main analysis:

eTable 1 presents the results of the analysis of the true intervention. The model with the step function is given in column (1) and the one with the quadratic function is given in (2).

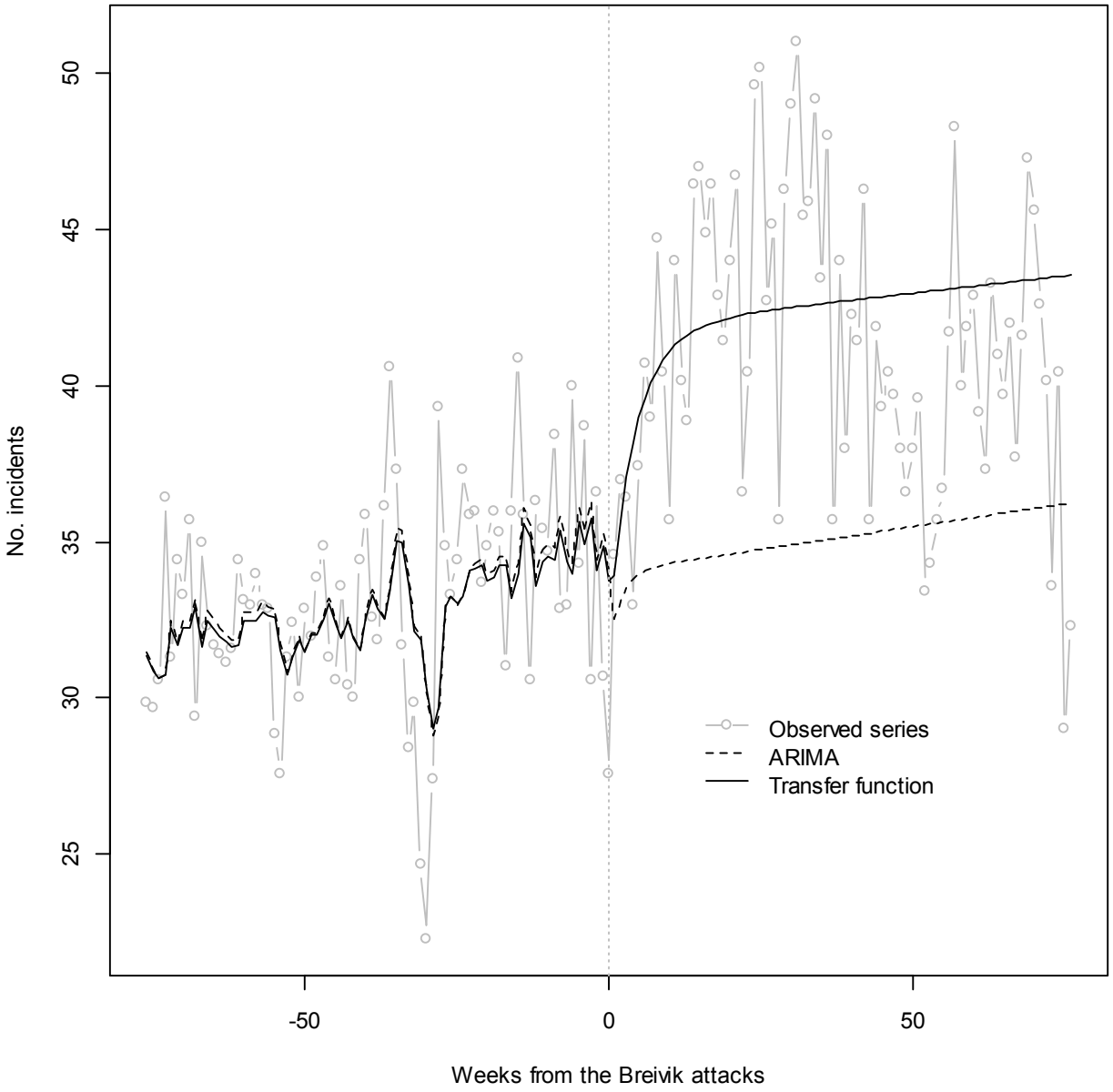
eTable 1: Trauma- and stressor-related disorders following the Breivik attacks

	Dependent variables in levels	
	ARIMA(1,1,2)	ARIMA(1,1,2)
	(1)	(2)
Step (AR1)	0.810***	
	(0.080)	
Step (MA0)	1.484**	
	(0.575)	
Linear trend		0.460***
		(0.072)
Quadratic trend		-0.006***
		(0.001)
Drift	Yes	Yes
Obs.	939	939
Log-likelihood	-2303.52	-2299.94
AIC	4619.03	4613.88

Note: *p<0.05; **p<0.01; ***p<0.001
ARMA- and drift-coefficients omitted

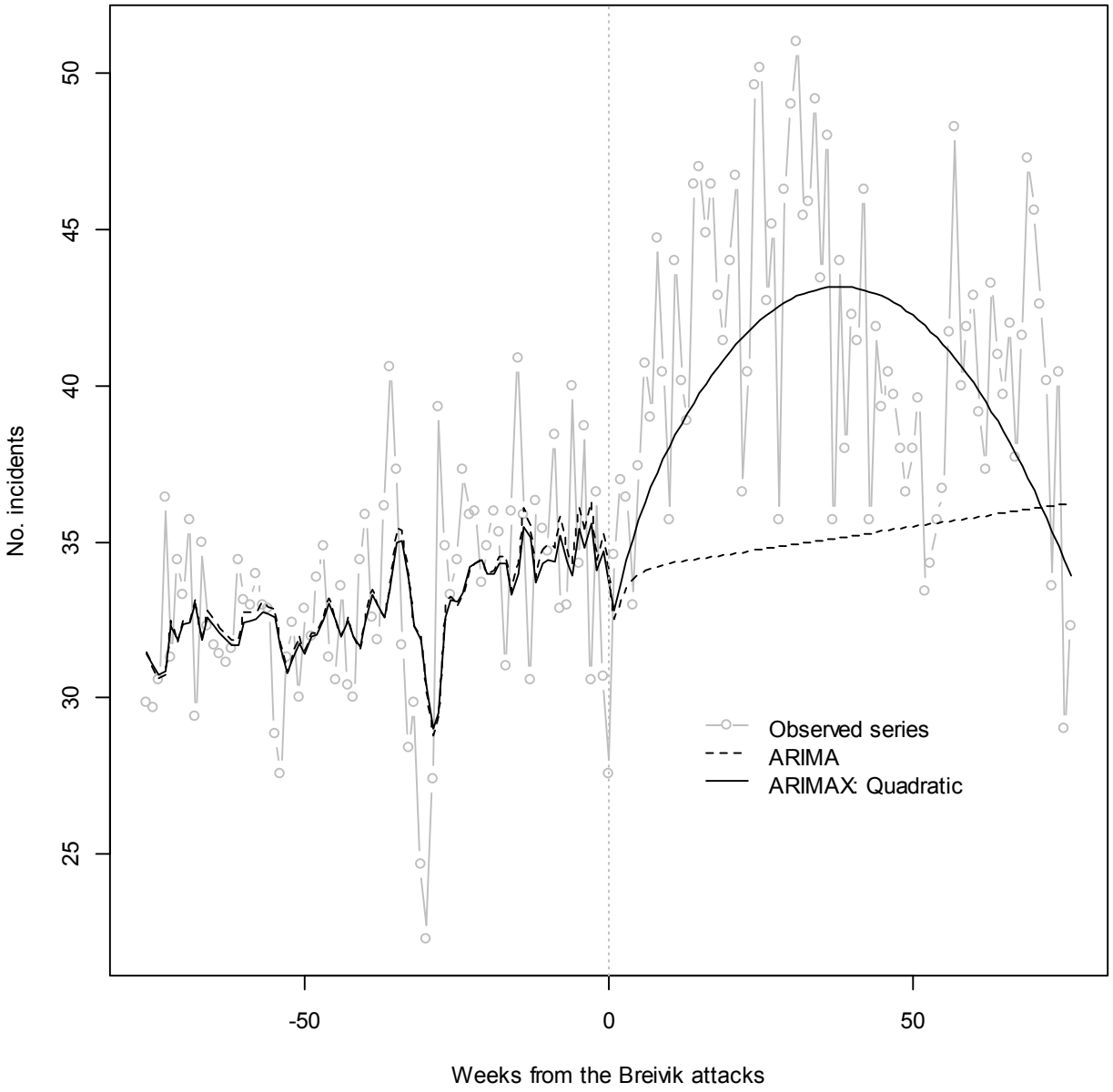
The table shows the strong effect of the Breivik attacks on the incidence rate of trauma- and stressor-related disorders in Denmark. For both the step and the quadratic operationalization the steps outlined above led to a ARIMA(1,1,2) specification with drift, and both models estimate a large and gradual increase in the incidence rate beginning right after the attacks. The step function estimates an additional 3918 cases over the 76 weeks following the attacks (a 22 % increase compared to the previous 76 weeks), while the quadratic function only estimates 2736 additional cases (a 16 % increase). The quadratic function fits slightly better, however, and provides a more realistic prediction of the time series – i.e. it does not predict a permanent level increase. We plot one-step and dynamic predictions from both types below:

eFigure 1: One-step and dynamic predictions of the incidence rate of trauma- and stressor-related disorders (step function).



Using a time window of about one and a half years before and after the Breivik attacks, the figure plots the observed number of trauma- and stressor-related disorders (grey line with open circles) along with predictions from both the transfer function (solid) and a pure ARIMA model (dashed line). The ARIMA model is of order (1,1,2) and includes a drift term. Up until the first week after the attacks (week = 0) the predictions are one-step, meaning that the models use data from all relevant previous periods in predicting the value of the series in a given period. After week zero the predictions become dynamic and are thus based solely on pre-intervention data and the parameters of the function of intervention time.

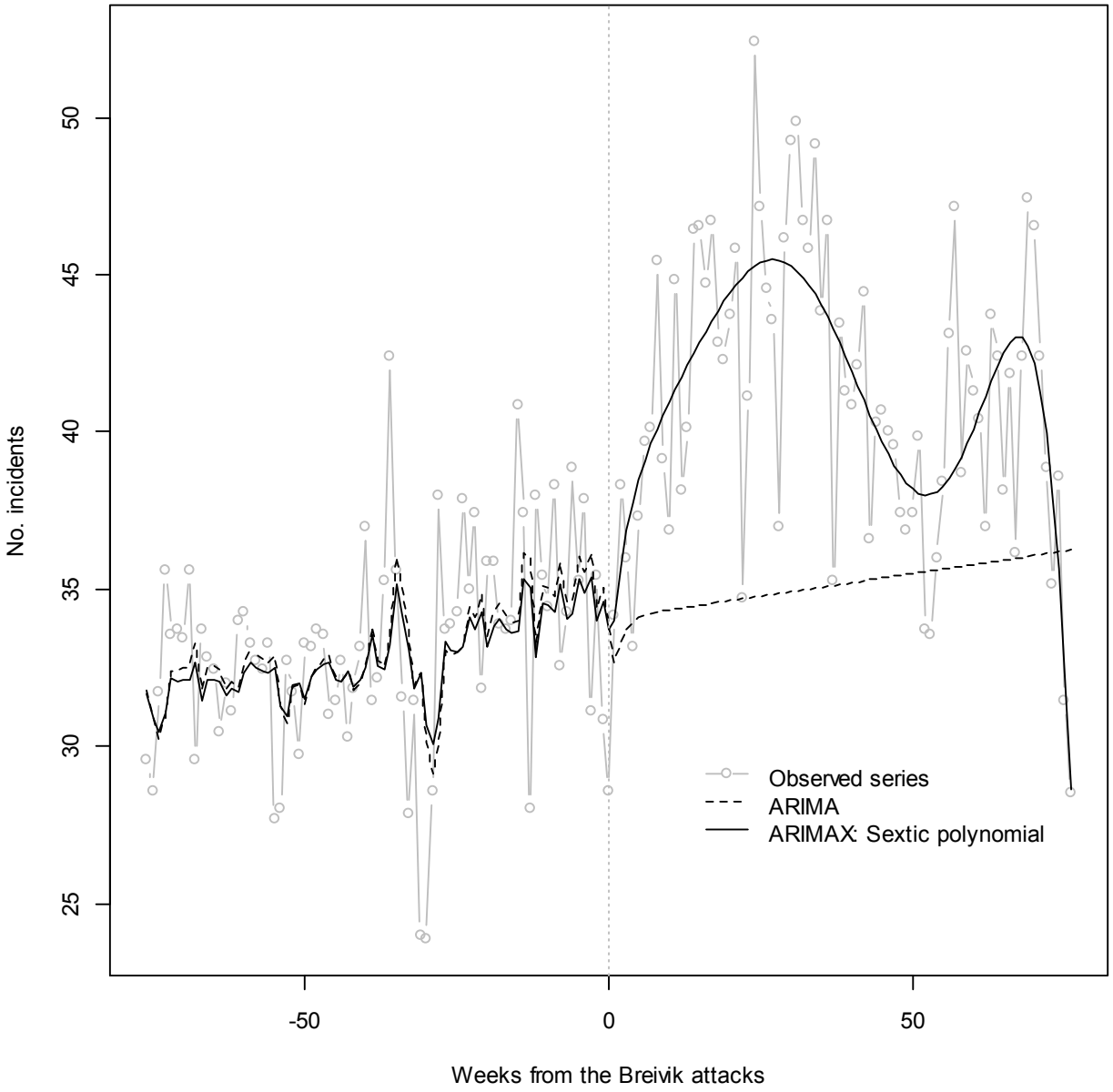
eFigure 2: One-step and dynamic predictions of the incidence rate of trauma- and stressor-related disorders (quadratic function).



The plot applies the same procedure as in eFigure 1, but instead of the transfer function with a step input, it adds the quadratic function described above to the ARIMA specification as an exogenous regressor, making it an ARIMAX model.

The plots show that the step function captured the immediate, relatively steep increase quite well, whereas the quadratic function did better when it comes to the series returning to its pre-intervention level. It is also evident that both did much better than the pure ARIMA model, which severely underpredicted most of the post-intervention observations.

eFigure 3: One-step and dynamic predictions of the incidence rate of trauma- and stressor-related disorders (sextic polynomial).

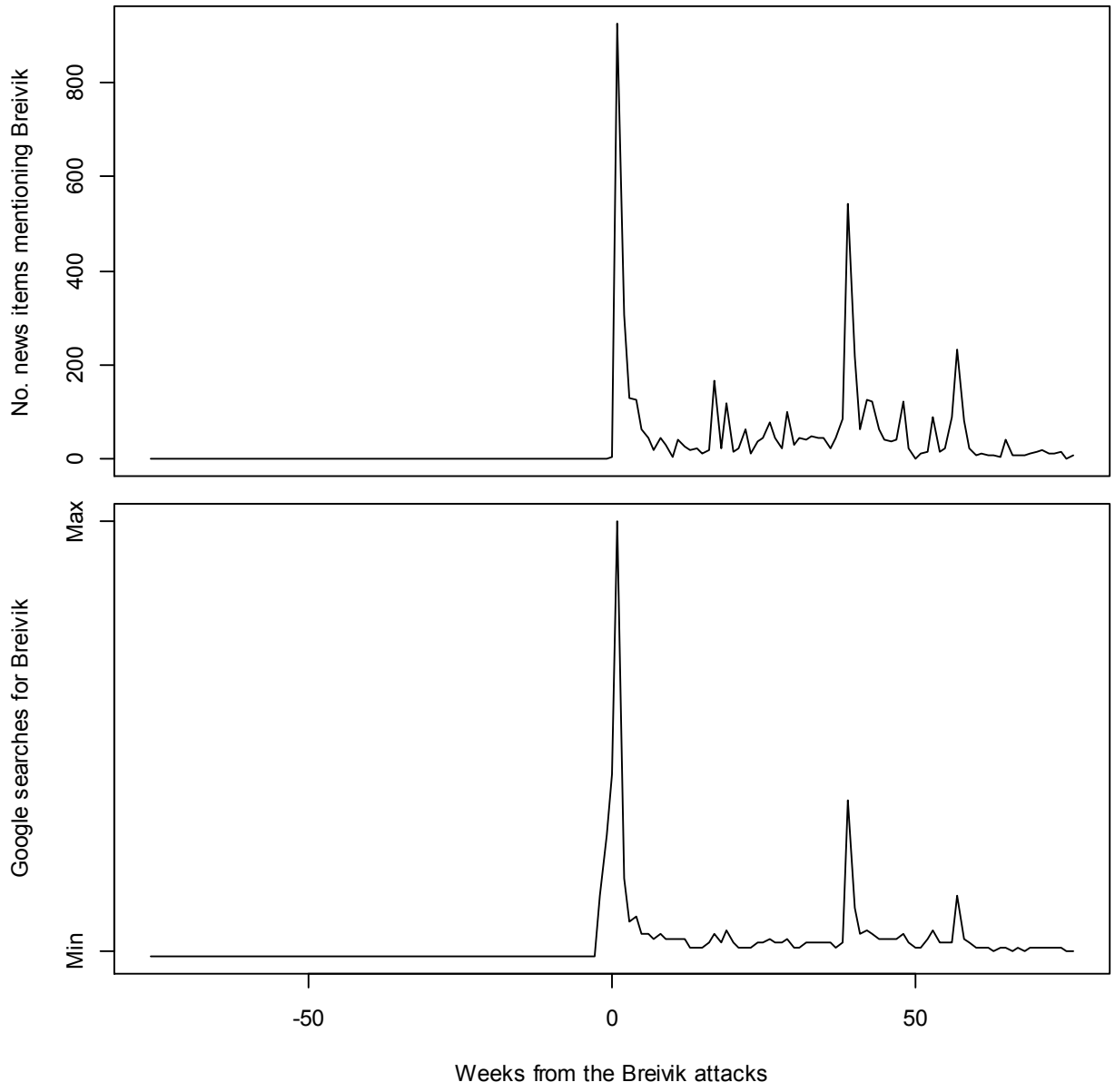


The plot applies the same procedure as in eFigure 2, but adds four terms to the polynomial in the ARIMAX specification, making it a sextic (six-degree) polynomial.

Unsurprisingly, the sextic polynomial responded more closely to movements of the post-intervention time series than the quadratic function. It rose more steeply immediately after the intervention and – importantly – captured the second surge of the series, which occurred at the time of the verdict in the Breivik trial. Although we would certainly not want to forecast out of sample with this highly flexible model either, it clearly characterizes the in-sample, biphasic post-intervention development of the incidence rate quite well.

Trends in the coverage of Breivik:

eFigure 4: News items and Google searches for ”breivik”.



The top panel plots the average daily number of news items in the Danish media (newspapers, tv, webmedia, radio) that according to the Danish media database “Infomedia” mentioned the word “Breivik”. The bottom panel plots the relative frequency of Danish Google searches for the term “Breivik”.

The plots in eFigure 4 reveal that the number of news items mentioned Breivik is very strongly correlated with the public interest in the term, thus supporting the use of the frequency of news mentions as a measure of public exposure to the attacks and their perpetrator. Correlating the two raw time series yields an R^2 of 0,73.

Formal test of the association between media attention and incidence rate:

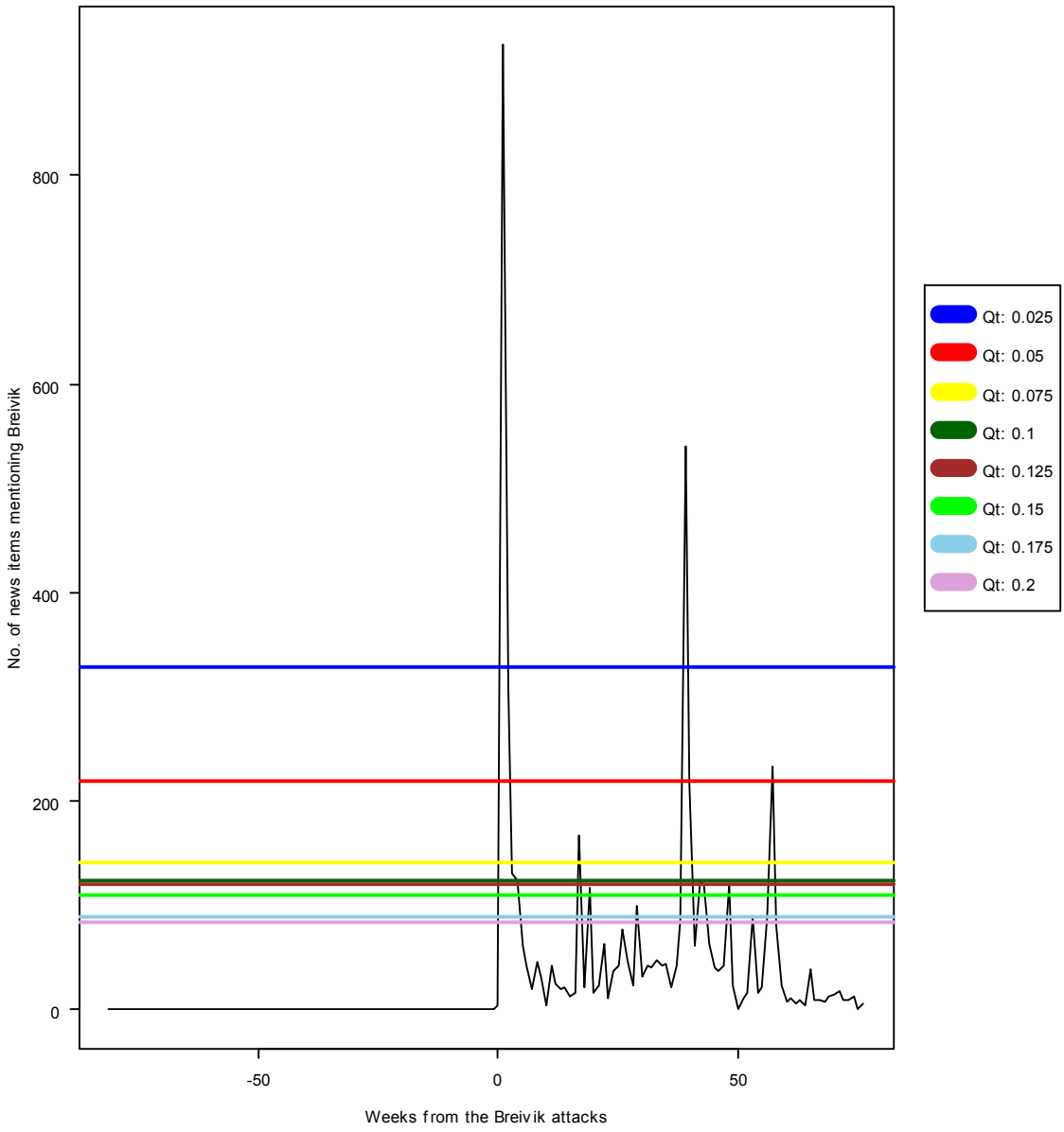
We tested the association between media attention and the incidence of trauma- and stressor-related disorders through an intervention model that used the timing of major spikes in the search interest or news coverage as interventions instead of the date of the actual attacks.

The logic behind these models was simply that if the effects of the attacks on mental health did indeed work through media attention, we should expect the incidence rates to increase following major increases in attention – irrespective of the attacks themselves.

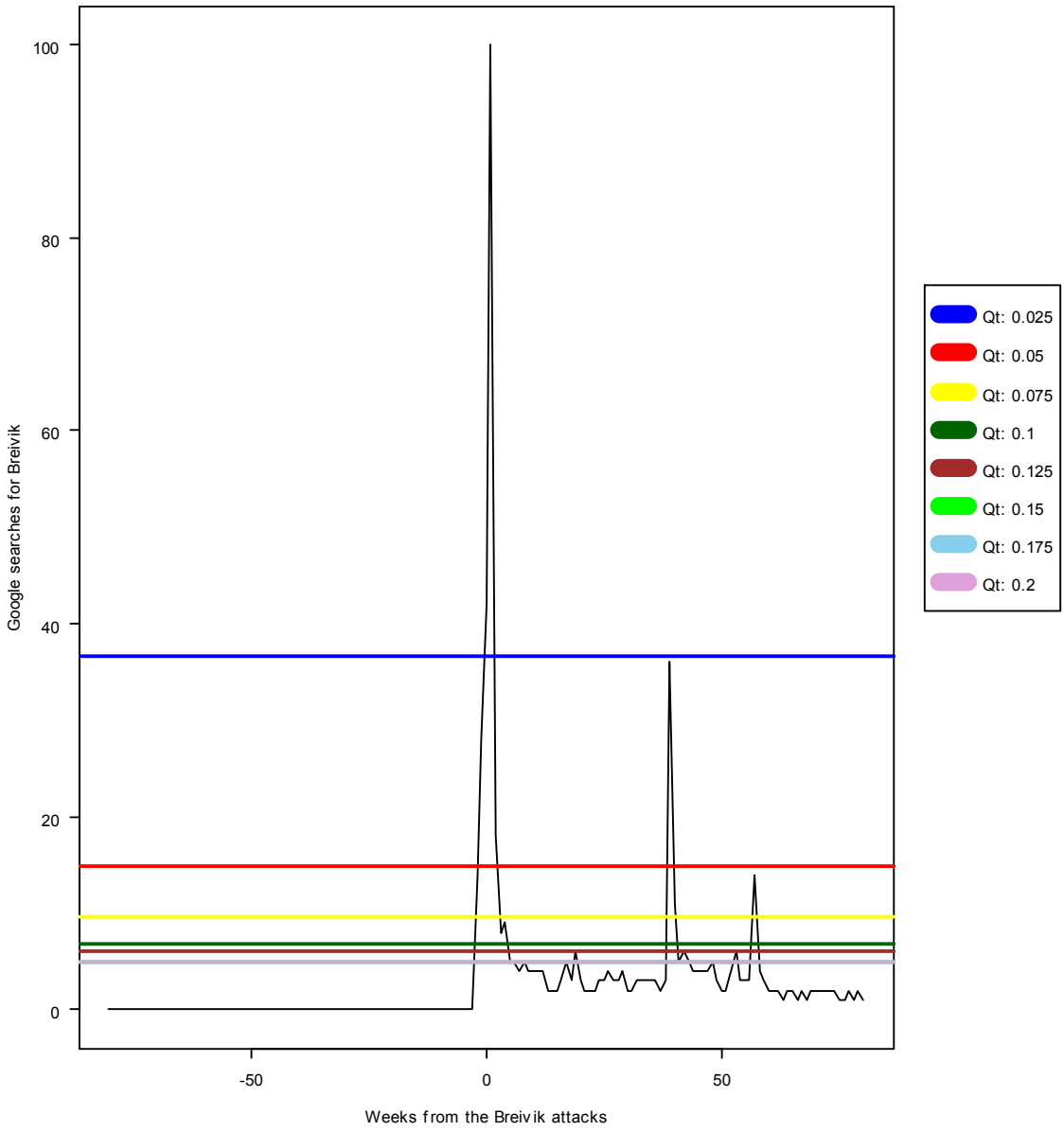
The models took the same form as the models using the true intervention week, but replaced this with the weeks showing spikes in attention. For each of these "spike weeks" we constructed a linear and a quadratic term analogous to the ones constructed above for the true intervention. We assigned a spike in weeks where the number of news items or Google searches exceeded a certain quantile of their post-intervention frequency distribution. As the choice of this quantile was somewhat arbitrary, we show results for a number of different thresholds below. Finally, we limited the models to only having a maximum of one spike every five weeks, such that periods with several weeks of high attention would just be assigned a spike in the week with the most attention. We maintained the noise specification as ARIMA(1,1,2) with a drift term.

eFigures 5 and 6 show which spikes are included in the different models by plotting the quantiles on top of the time series for news items and Google searches. eFigures 7 and 8 display one-step and dynamic predictions from the models along with the number of spikes included in each model.

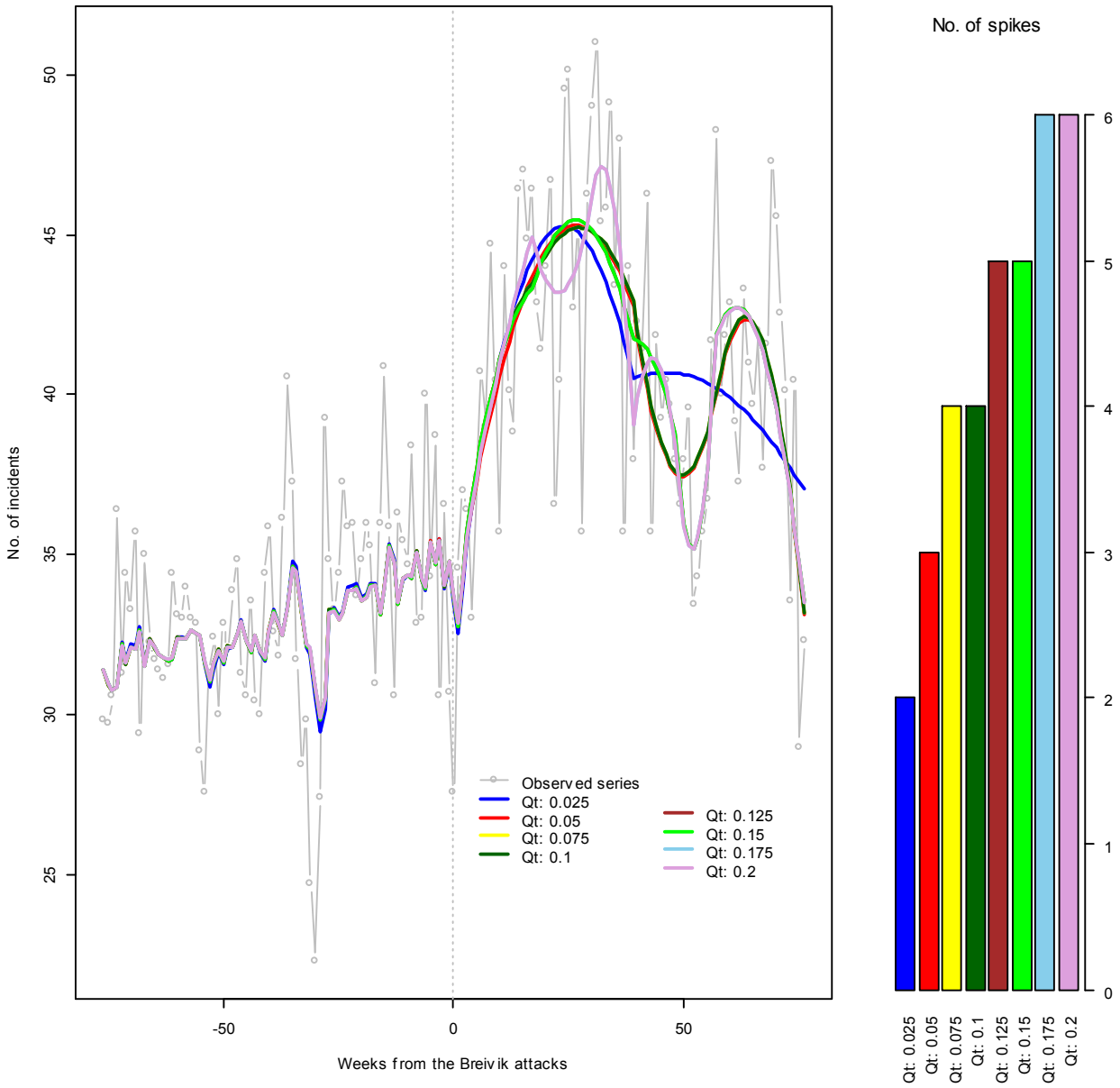
eFigure 5: Spikes in news items mentioning Breivik



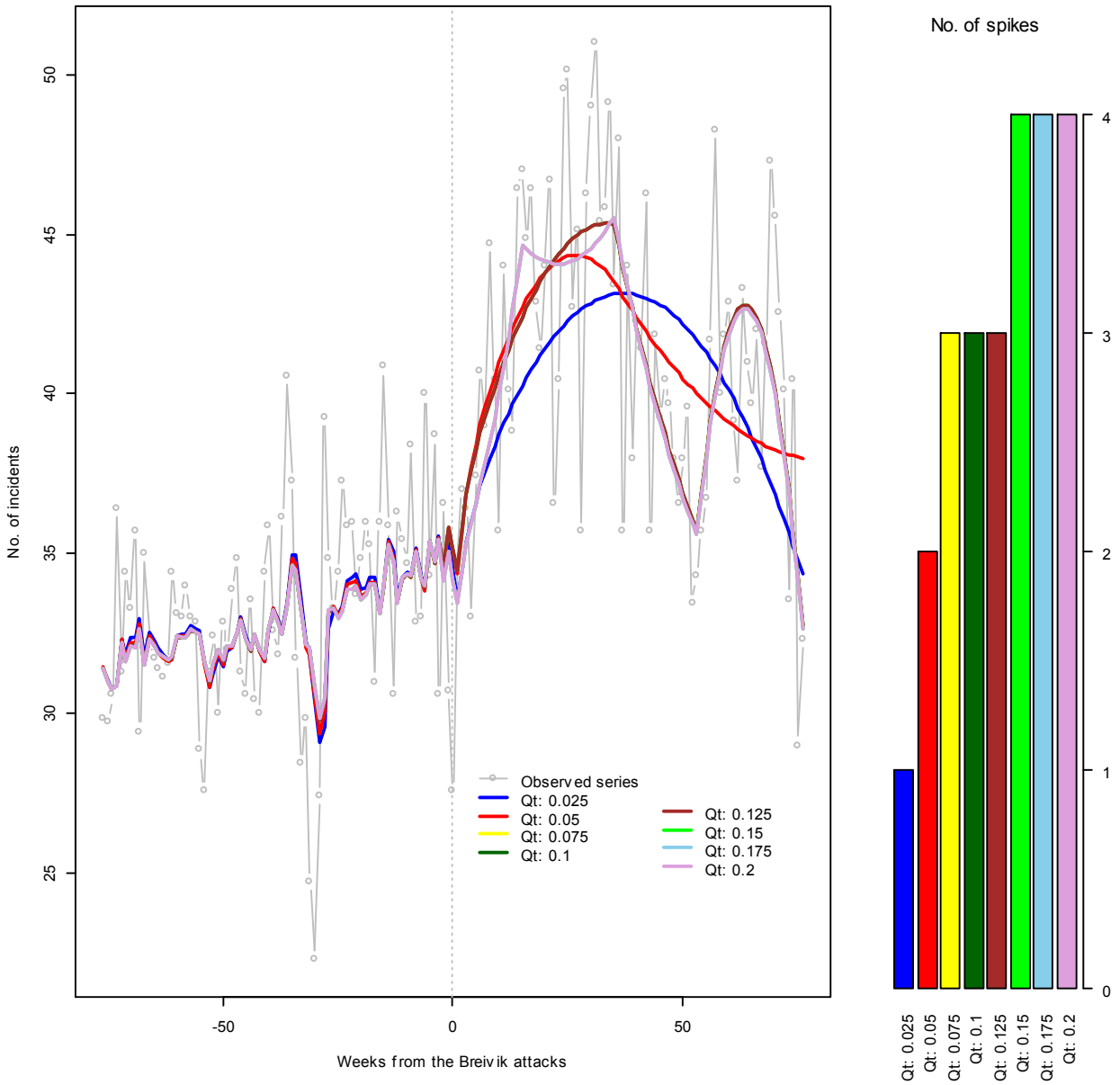
eFigure 6: Spikes in Google searches for Breivik



eFigure 7: One-step and dynamic predictions using spikes in news items



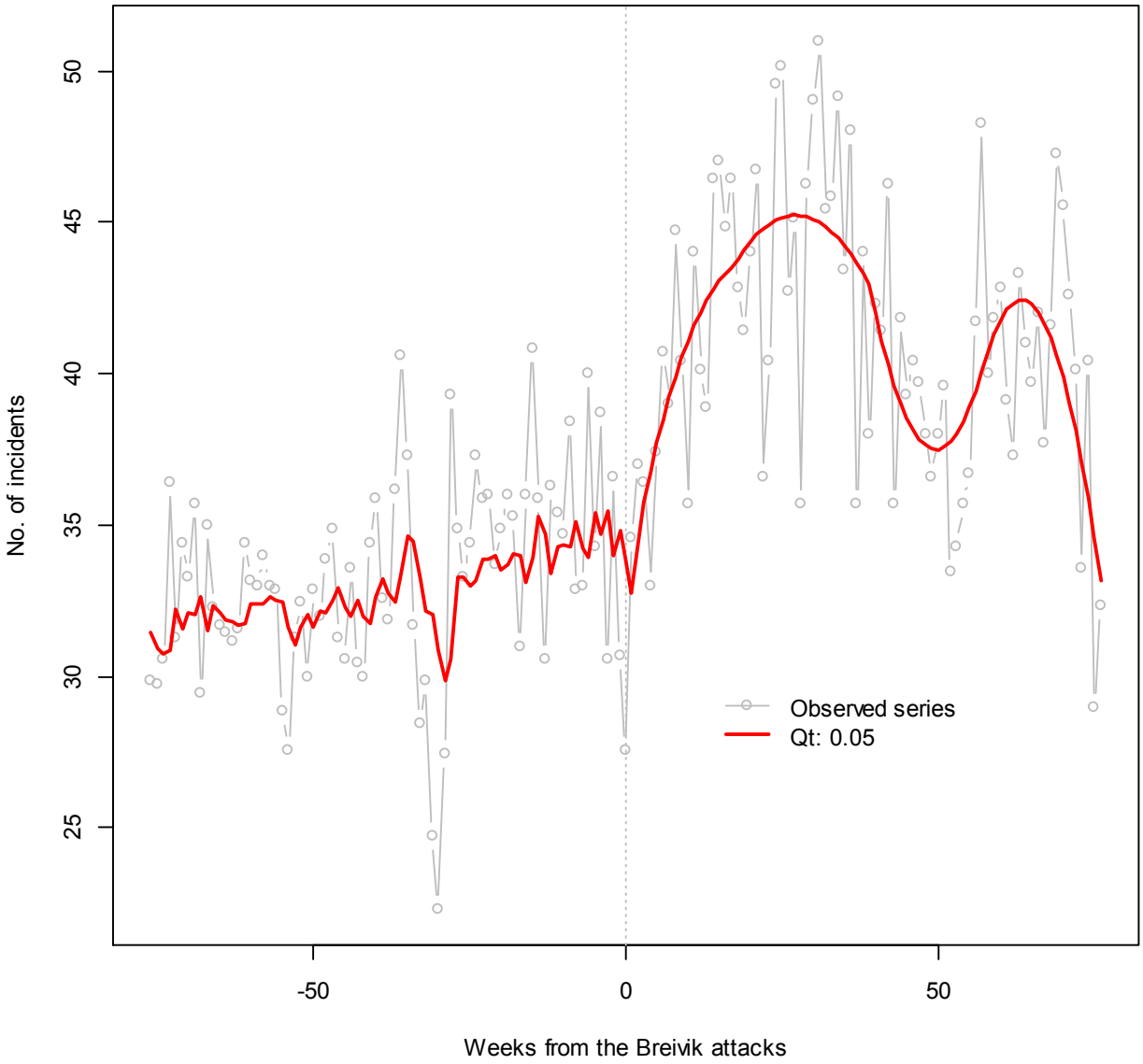
eFigure 8: One-step and dynamic predictions using spikes in Google searches



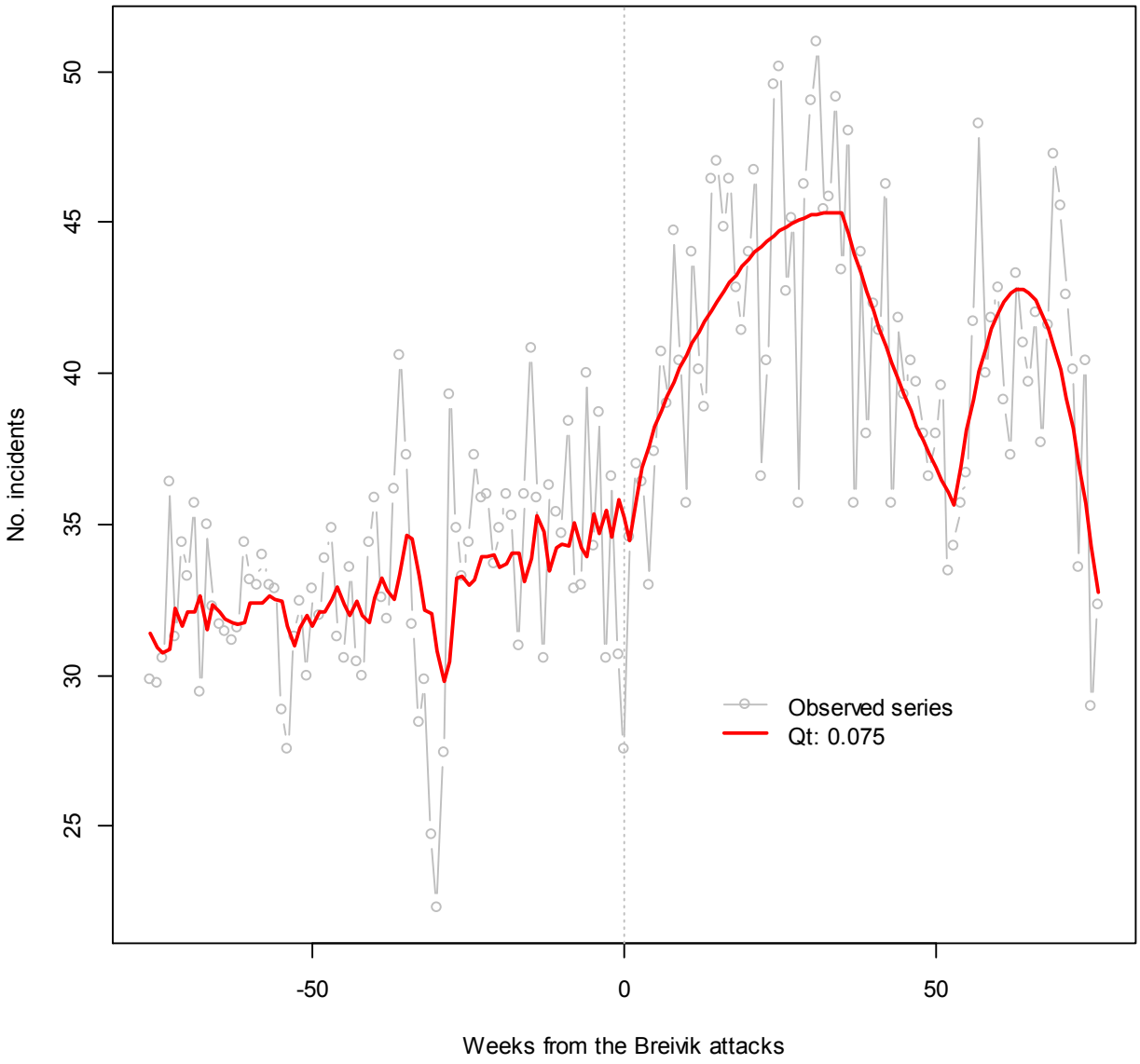
As is evident from the both the models for news items and those for Google searches, once we assigned three or more spikes (using the weeks in the top 5 % and 7.5 %, respectively), the models clearly picked up the biphasic trend that was estimated using the sextic polynomial. Most of the predictions from the models with three or more spikes lay right or almost on top of each other.

For clarity, we plot predictions from just the three-spike models in eFigures 9 and 10. The predictions in eFigure 9 are also plotted in Figure 3 in the paper.

eFigure 9: One-step and dynamic predictions using three spikes in news items



eFigure 10: One-step and dynamic predictions using three spikes in Google searches



The root mean squared error (RMSE) of the models with just these three spikes were both on par with the RMSE from the model using the sextic polynomial and were substantially lower than the RMSE for the parabolic model (the RMSE was 3.64 for the news model, 3.58 for the searches model, 3.56 for the sextic polynomial-model, and 4.66 for the parabolic model).

This means that we could accurately model the development in the incidence rate after the Breivik attacks, by completely ignoring the actual attacks and only using the media attention surrounding them. This was particularly interesting for the second surge in the incidence rate, which appears to stem from the media coverage of the sentencing of Breivik.

It should be noted that unlike the models using the actual attacks, the models using spikes in media attention should not be given any causal interpretation. The central identifying assumption of intervention analysis is the as-if-randomness of the timing of the intervention relative to the determinants of the time series under study. This is quite plausibly true for the Breivik attacks, since the decision of the perpetrator to attack on the 22nd of July 2011 is unlikely to affected by any of the other determinants of the incidence rates of mental disorders in Denmark, nor are these incidence rates

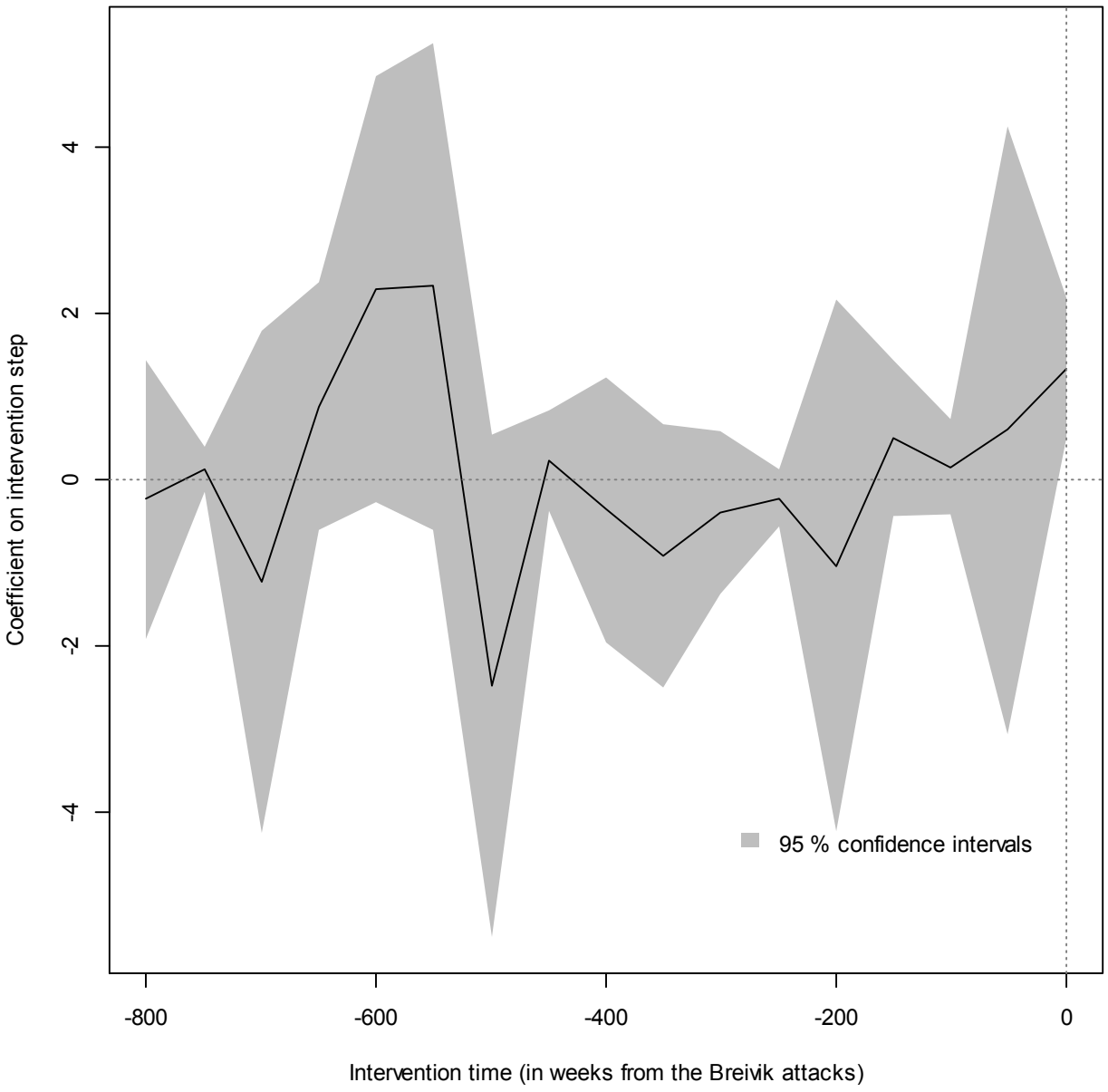
themselves likely to have affected his decision. However, the timing of the surges in media attention given to the attacks and the subsequent trial was certainly a function of the timing of the attacks, and, given that the attacks affected the incidence rates in Denmark, the as-if-randomness of assignment does not hold for media attention. In classical regression terms, the Breivik attacks is an omitted variable in the spike-models above.

These models should instead be seen as a way of getting some traction on the way in which the Breivik attacks could have affected the mental health of people in a different country: Through the media.

Results of placebo tests:

Here we present the results of the four different combinations of placebo tests described above.

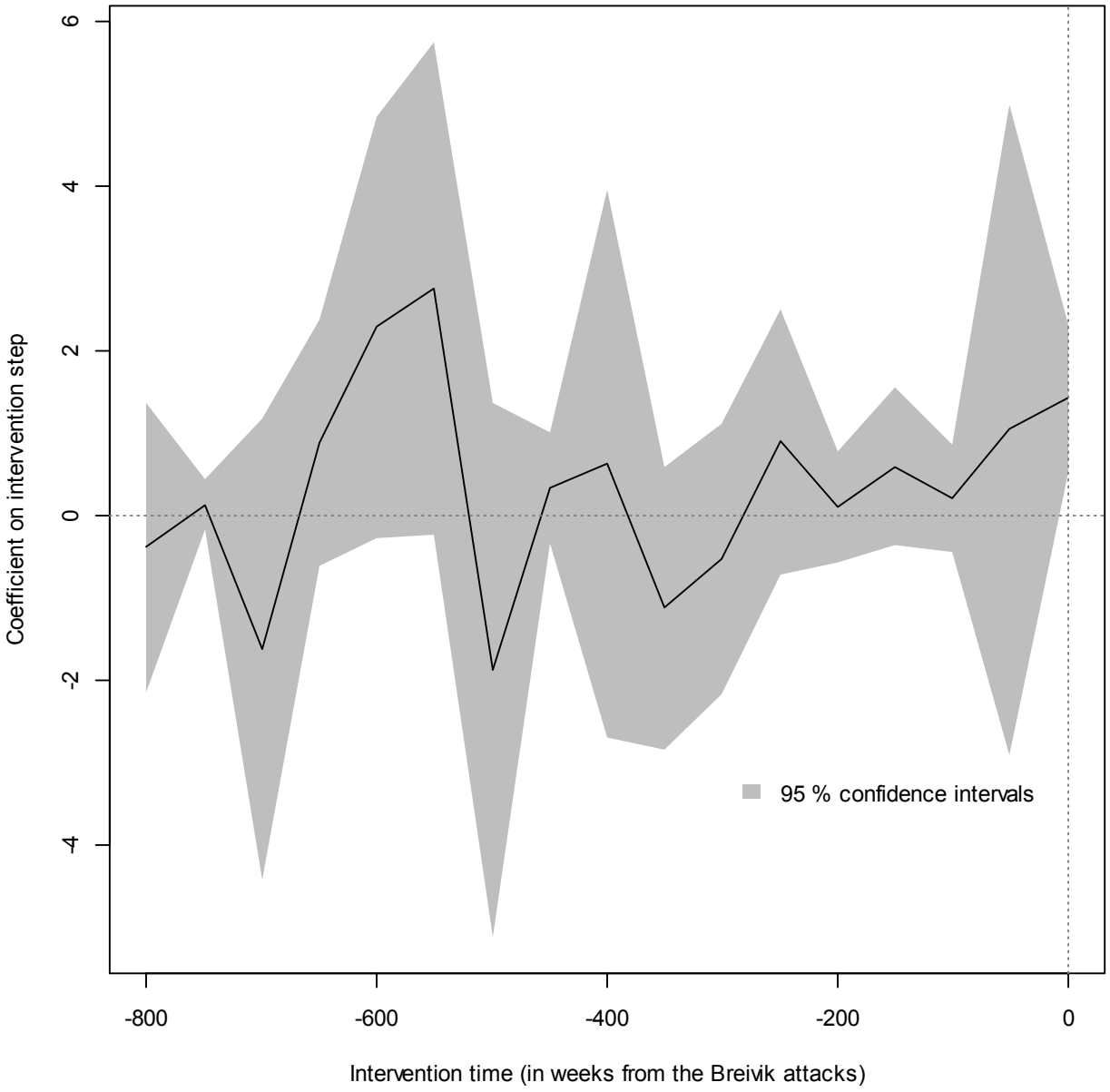
eFigure 11: Placebo interventions, step function, ARIMA(1,1,2) with drift



This test used the same ARIMA(1,1,2) specification as the main analysis to estimate the effect of placebo interventions at different points in the time series, each 50 weeks apart. The plot shows the coefficient on the moving average term of the step function along with its 95 % confidence interval.

The plot demonstrates that none of the 16 placebo interventions showed any sign of a non-zero change. Only the true intervention (at week zero) came out significantly different from zero.

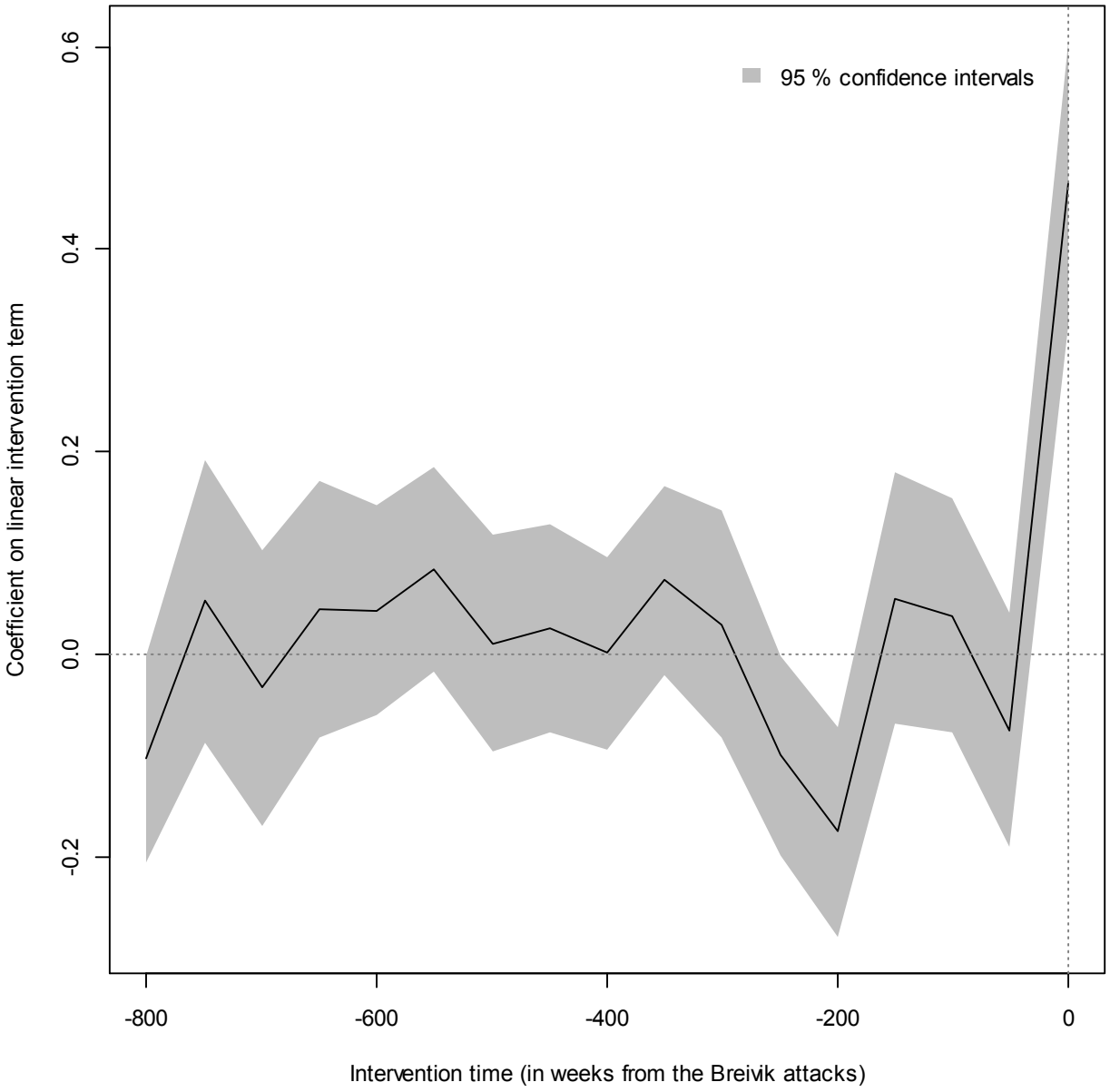
eFigure 12: Placebo interventions, step function, variable ARIMA spec.



To estimate the effect of the placebo interventions this test used a variable ARIMA specification determined for each intervention by the Hyndman-Khandakar algorithm. These specifications were allowed to include a drift parameter. Again, the plot shows the coefficient on the moving average term of the step function along with its 95 % confidence interval.

Allowing the ARIMA specification to adjust to fit each placebo test did not make much of difference. The true intervention still remained the only significant one.

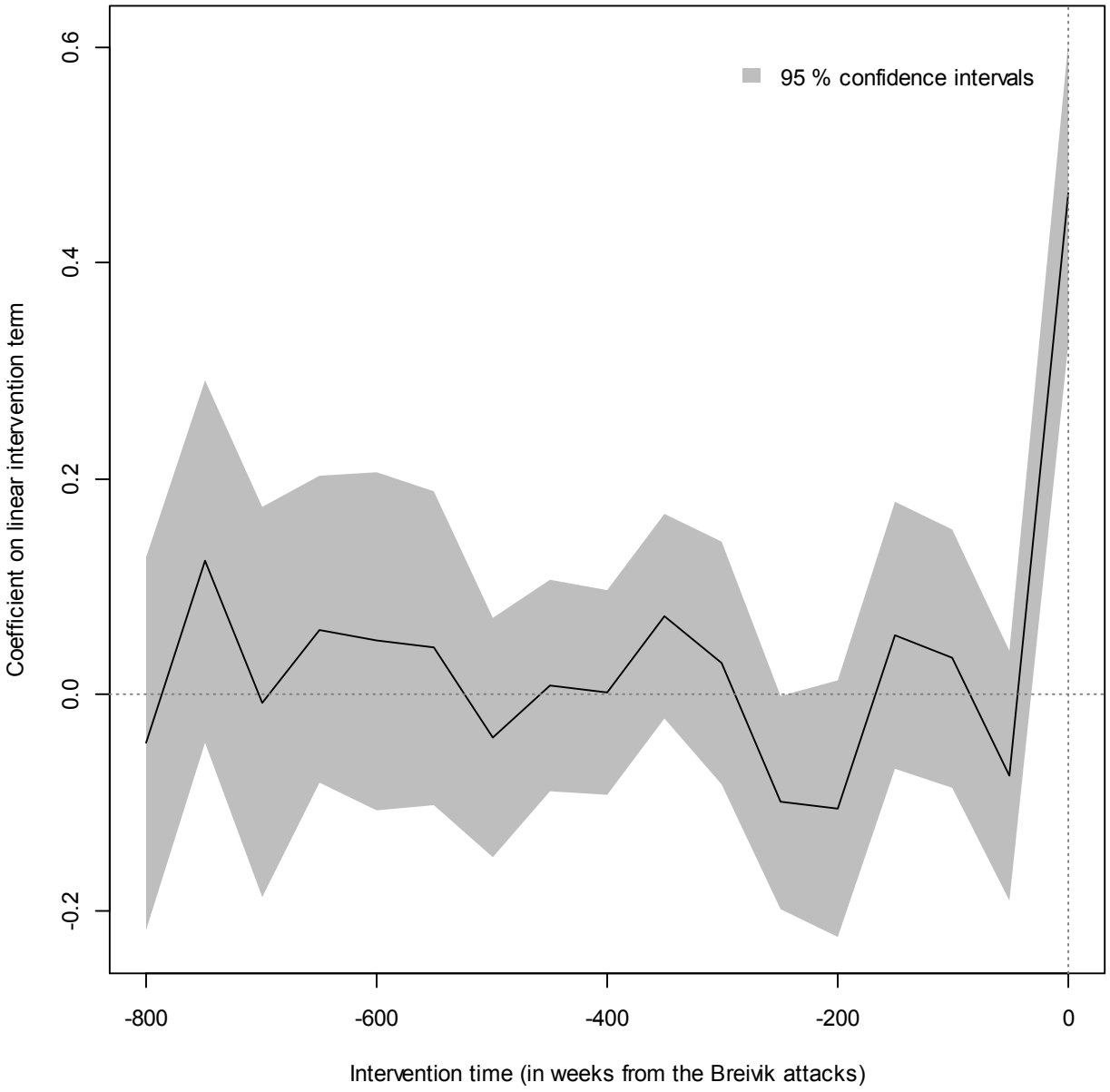
eFigure 13: Placebo interventions, quadratic function, ARIMA(1,1,2) with drift



This test used the same ARIMA(1,1,2) specification as the main analysis to estimate the effect of placebo interventions at different points in the time series, each 50 weeks apart. The plot shows the coefficient on the linear term of the quadratic function along with its 95 % confidence interval.

For the quadratic function, we plotted the coefficient on the linear term since this term estimated the post-intervention change, whereas the quadratic term estimated the rate of reversal to the outset. Here, one out of the 16 placebo interventions came out significant at the 0.05-level, which, as we note above, was within the number of significant coefficients we could expect due to pure randomness. Furthermore, this significant change was negative and of a much smaller magnitude than that of the true intervention.

eFigure 14: Placebo interventions, quadratic function, variable ARIMA spec.



To estimate the effect of the placebo interventions, this test again used a variable ARIMA specification determined for each intervention by the Hyndman-Khandakar algorithm. These specifications were allowed to include a drift parameter. Once more, the plot shows the coefficient on the moving average term of the step function along with its 95 % confidence interval

Again the results were very similar with the variable ARIMA specification. One small placebo intervention came out significant at the 0.05 level, and the true intervention remained quite large and highly significant.

In summary, the placebo tests provided strong evidence that the observed increase in the incidence rate of trauma- and stressor-related disorders following the Breivik attacks were not a product of chance, since similar increases could not be detected at other, unrelated points in the time series.

Inspection of other mental disorders following the Breivik attacks:

To check if the Breivik attacks were followed by an increase in trauma- and stressor-related disorder specifically, rather than an increase in psychiatric disorders generally, we carried out a similar intervention analysis on a combined time series of incidence rates of all other mental disorders in Denmark.

To determine the ARIMA-specification we went through the steps detailed above. The model with the step function is given in column (1) and the one with the quadratic function is given in (2).

eTable 2: All other diagnoses following the Breivik attacks

	Dependent variables in levels			
	ARIMA(3,1,2)	ARIMA(3,1,2)	ARIMA(3,1,2) last two obs. excl.	ARIMA(3,1,2) last two obs. excl.
	(1)	(2)	(3)	(4)
Step (AR1)	0.849***		0.849***	
	(0.162)		(0.157)	
Step (MA0)	1.665		1.612	
	(1.808)		(1.685)	
Linear trend		0.970*		0.629
		(0.387)		(0.373)
Quadratic trend		-0.016***		-0.009*
		(0.005)		(0.005)
Drift	Yes	Yes	Yes	Yes
Obs.	939	939	937	937
Log-likelihood	-3762.04	-3756.79	-3739.01	-3738.16
AIC	7540.08	7531.58	7494.02	7494.32

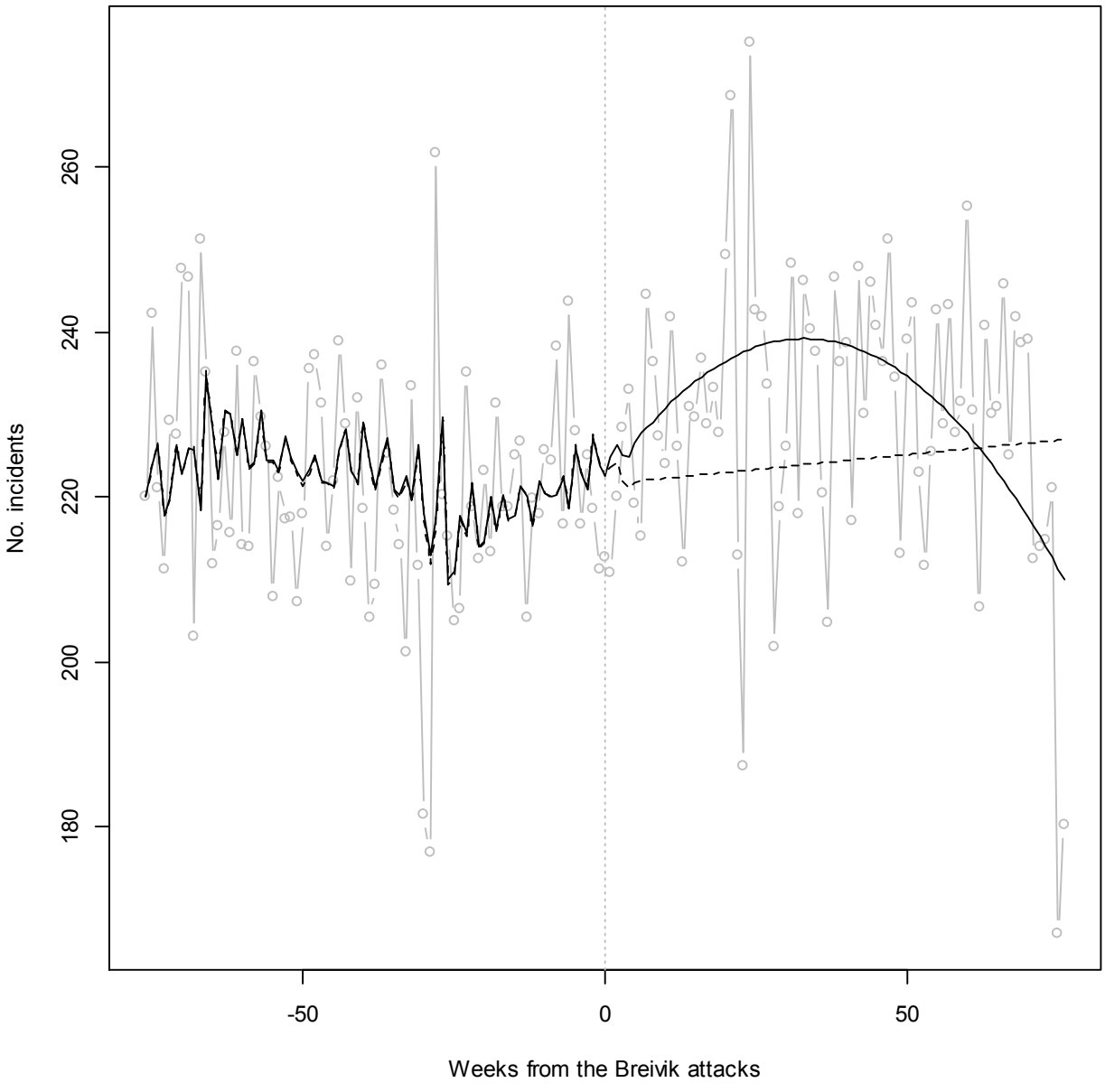
Note: *p<0.05; **p<0.01; ***p<0.001

ARMA- and drift-coefficients omitted

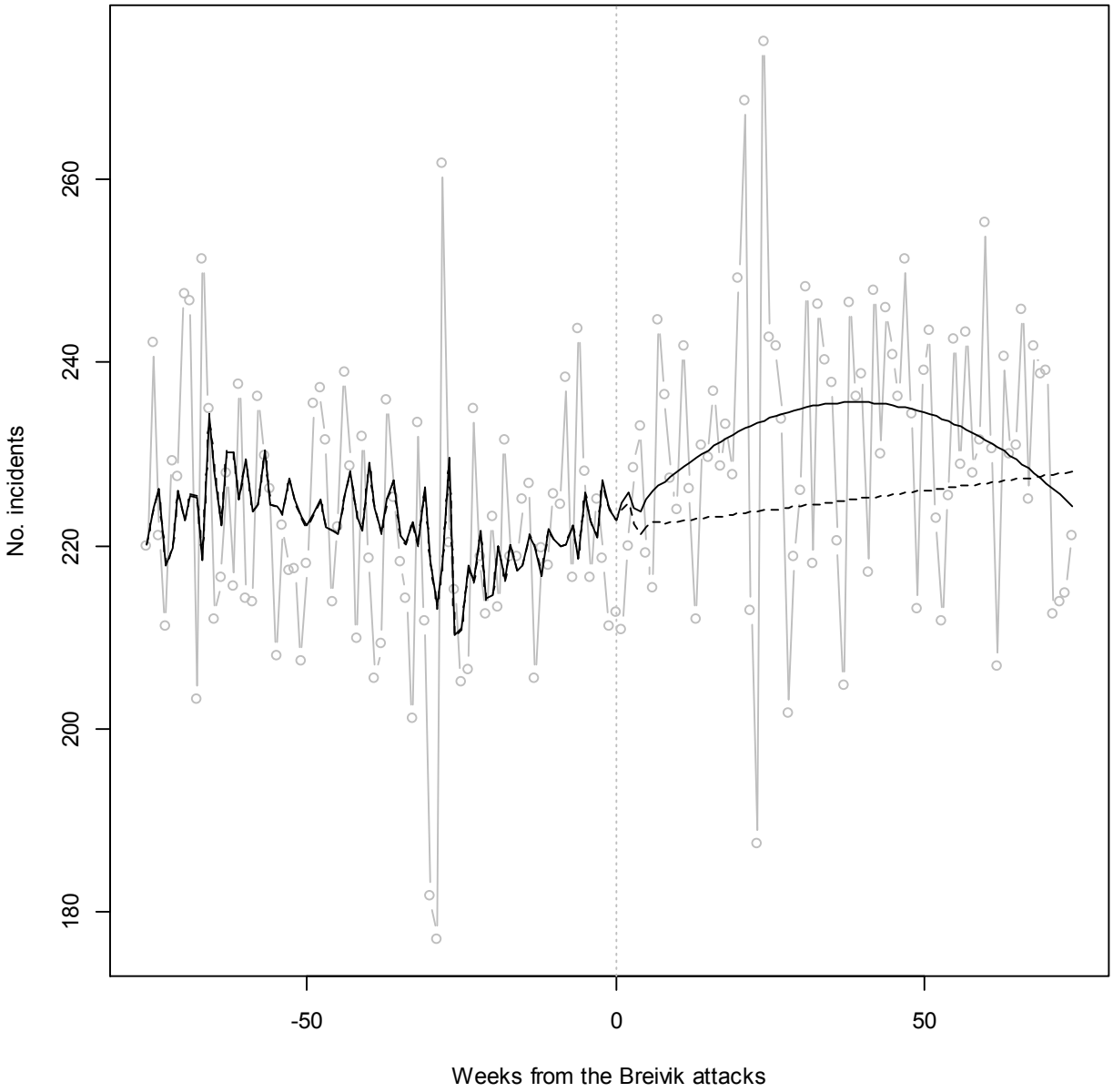
The step function showed no increase in the incidence rate of other diagnoses. The quadratic function estimated a non-zero parabolic increase, but since the incidence rate for other diagnoses was about eight times as high as the incidence rate for trauma- and stressor-related disorders, a coefficient of about twice the size indicated a comparatively small increase. Further, it was almost exclusively driven by the last two weeks, when the incidence rate of other diagnoses

dropped sharply. Columns 3 and 4 estimated the same models without these last two observations, and could no longer distinguish the increase from zero (note that this was not the case for trauma- and stressor-related disorders). eFigures 15 and 16 plot one-step and dynamic predictions from columns 2 and 4.

eFigure 15: One-step and dynamic predictions of other diagnoses



eFigure 16: One-step and dynamic predictions of other diagnoses (excl. last two obs.)



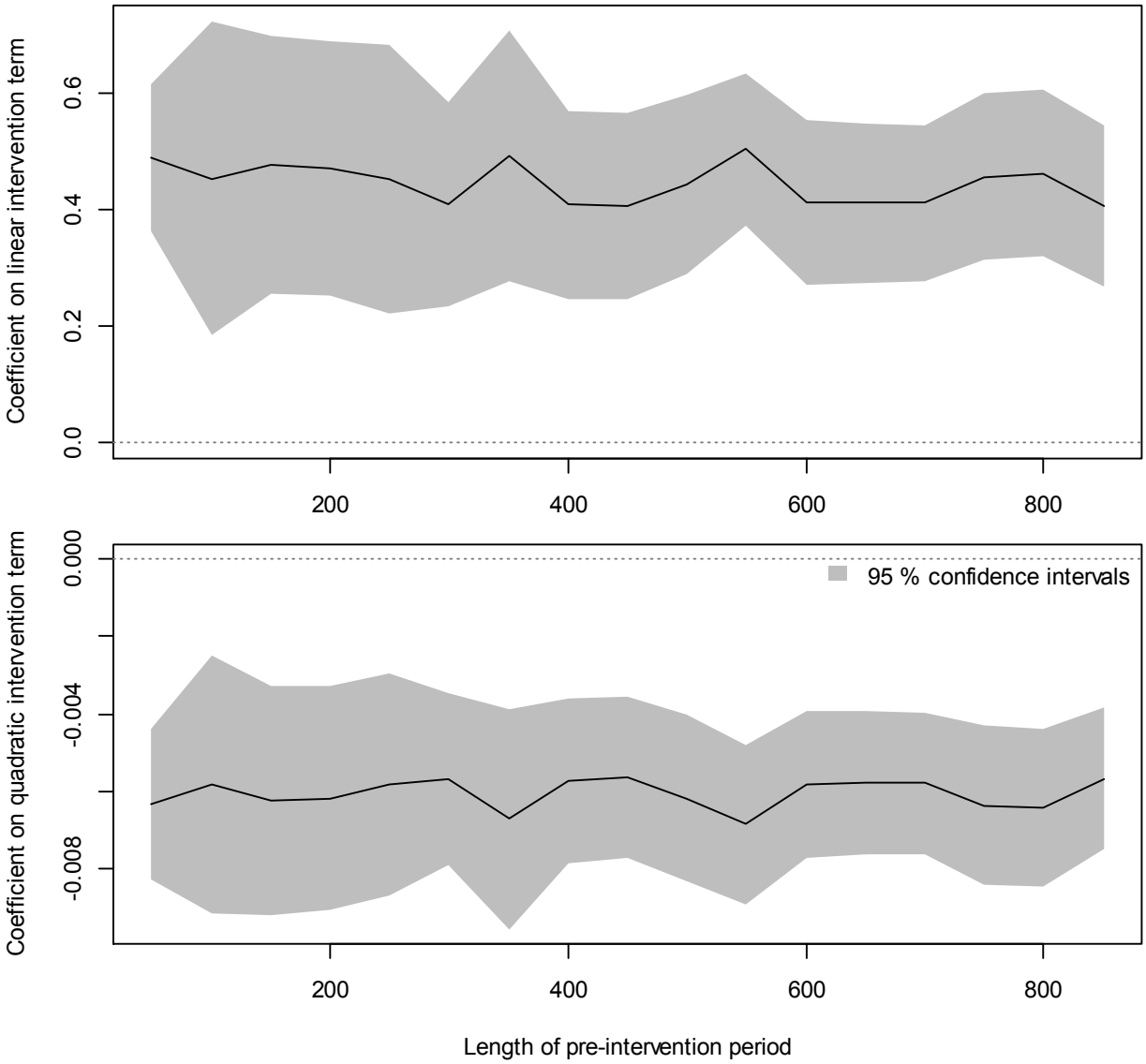
In sum, there appeared to be some evidence of a general parabolic increase in other diagnoses following the Breivik attacks, but nowhere near as large or as robust as the increase estimated for trauma- and stressor-related disorders.

Estimating main models with limited pre-intervention time series:

To check it the results were an artefact of the long pre-intervention time series being used, we ran a series of intervention models with limits on the length of the pre-intervention period.

Below, we plot the coefficients of the parabolic model while progressively shortening the pre-intervention part of the time series; down to a minimum of 50 weeks.

eFigure 17: Quadratic model with different lengths of pre-intervention time series



Regardless of the length of the pre-intervention time series, both the linear and quadratic terms remain both at the same size and at the same level of significance.

REFERENCES

1. Mors O, Perto GP, Mortensen PB. The Danish Psychiatric Central Research Register. *Scand J Public Health*. 2011;39(7 Suppl):54-57.

2. Munk-Jorgensen P, Ostergaard SD. Register-based studies of mental disorders. *Scand J Public Health*. 2011;39(7):170-174.

3. World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders. Diagnostic criteria for research. WHO. Geneva. . 1993.

4. Hansen BT, Ostergaard SD, Sonderskov KM, Dinesen PT. Increased incidence rate of trauma and stressor related disorders in Denmark following the 9/11 attacks. *Am J Epidemiol*. 2016. 184(7):494-500.

5. Metcalfe R, Powdhavee N, Dolan P. Destruction and distress: Using a quasi-experiment to show the effects of the September 11 attacks on mental well-being in the United Kingdom. *Econ J*. 2011;121:81-103.

6. Cryer J, Chan K. Time Series Analysis: With Applications in R. *Springer, New York*. 2010.

7. Box GEP, Tiao GC. Intervention analysis with applications to economic and environmental problems. *J Amer Statist Assoc*. 1975.

8. Box GEP, Jenkins GM. Time series analysis, forecasting and control. *Holden-Day, San Francisco*. 1970.

9. Hyndman RJ, Khandakar Y. Automatic time series forecasting: the forecast package for R. *J Stat Softw*. 2008;27:1-22.

10. Canova F, Hansen BE. Are seasonal patterns constant over time? A test for seasonal stability. *J Bus Econ Stat*. 1995;12:237-252.