Title: **Pharmaceutical advertising biases media reports on drug safety**

Short title: Pharmaceutical advertising biases media reports

Authors: **Florens Focke[1], Alexandra Niessen-Ruenzi[1], Stefan Ruenzi[1]**

Author Affiliation: [1]Department of Finance, University of Mannheim, 68161 Mannheim, Germany

Corresponding author: Stefan Ruenzi, Department of Finance, University of Mannheim, 68161 Mannheim, Germany, phone: (+49) 621 181 1646, e-mail: ruenzi@bwl.uni-mannheim.de

Classification: Social Sciences - Economic Sciences

Keywords: Media bias, advertising, drug safety

Abstract: We examine whether advertising by pharmaceutical companies biases newspaper articles on drug safety. The media plays an important role in the dissemination of health information to the general public. However, the dependency of the media on advertising results in severe conflicts of interest: Media outlets have to decide between pleasing advertising clients by reporting positively about them and informing their readers objectively about their client's products. We show that commercial pressure arising from pharmaceutical advertising biases newspaper reports on drugs. Newspapers are less likely to mention side effects of drugs produced by their advertising clients or to report on US Food and Drug Administration alerts regarding these drugs. Finally, newspapers generally write less negatively about drugs of their advertising clients. Our findings have important public health implications: Given the broad reach of mass media, accurate media reports on drugs are an important element of informing the public about drug safety.

Significance statement: The media plays an important role in informing the general public on scientific advances. However, media reports on health issues have been criticized for being inaccurate, superficial, or inappropriately optimistic. In this manuscript, we provide the first empirical evidence that the commercial pressure arising from advertising biases media reports on pharmaceutical drugs. We show that newspapers that have received advertising from a drug's manufacturer are significantly less likely to report on potential harms of the drug. Our results highlight that the dependency of many media outlets on advertising revenues can contribute to the poor quality of health reports. Thus, they also provide a new argument against direct-to-consumer-advertising.

Text:

The mass media are an important source from which the public obtains information on drugs and medical procedures *(1-4)*. Media-based health information affects health decision-making and medical advice-seeking behavior *(5)*. Furthermore, recent studies point out the poor quality of many articles about health issues, particularly those covering drugs *(6, 7)*. Articles are frequently described as inaccurate, superficial, or inappropriately optimistic *(8-10)*.

At the same time, the media is highly dependent on advertising revenue and the pharmaceutical industry is a major advertiser.[*] From 1999 to 2012, the industry spent more than $120 billion on advertising in the US alone, making it the second largest advertiser after the automobile industry. Given that media reports have a strong impact on health behaviors, it is important to understand whether pharmaceutical advertising biases these reports. In the specific case of the danger of smoking, for example, early studies found evidence for a relationship between reports in magazines and advertising by tobacco companies *(11, 12)*. In contrast, whether advertising might also contribute to the perceived poor quality of media reports on pharmaceutical drugs has not previously been investigated.

We establish a causal link between advertising pressure from the pharmaceutical industry and a deterioration of the quality of newspaper reports on drug safety. Following the recently advocated approach of big data analysis in public health *(13)*, we combine two large databases for our analysis. First, we use the Kantar Stradegy database to obtain the daily dollar amount that each pharmaceutical company listed in the Drugs@FDA database spends on advertising in US newspapers (see table S1). Second, we collect all articles on drugs published in each of these newspapers from LexisNexis. This yields a large sample of 81,656 articles on pharmaceutical drugs between 1999 and 2012.We then classify each article according to whether it mentions side effects of the drug that is covered or FDA safety alerts regarding this drug. Using computer-linguistic text analysis, we determine the general tone of each article. For each article, we calculate the advertising spending of the drug's manufacturer at the newspaper publishing this article over the past 7, 14 and 30 days (see table S2).

We estimate two fixed effects regression model specifications to assess the causal impact of advertising on the content of articles on drugs of advertising clients. The first model includes drug-month fixed effects, which means that we only compare articles written about the same drug across newspapers at the same time. The second model additionally includes newspaper fixed effects to control for differences in the general writing styles of newspapers. Hence, any effect we find must be driven by within-newspaper changes in the way they report on a particular drug due to changes in short-term advertising. Consequently, these results should provide a lower limit for the true impact of advertising on drug reporting. Using high-dimensional fixed effects regressions helps us to establish causality.

The result of the analysis of the relation between advertising revenues and newspaper reports on side effects is illustrated in fig. 1 (for underlying estimation results using conditional logit models, see table S3). We find that newspapers are less likely to report on a drug's side effects if they obtained more advertising dollars from the drug's producing company over the past 14 days; this is irrespective of whether we use drug-month and newspaper fixed effects (dashed line) or only drug-month fixed effects (solid line). This effect is statistically significant ($p<0.01$) and has a meaningful size. For example, a newspaper that receives an additional $100,000 of advertising in the 14 days prior to the publication of an article on its advertising client's drug is

[*] Organization for Economic Co-operation and Development (OECD), The Evolution of News and the Internet, Report by the Working Party on the Information Economy (2010; http://www.oecd.org/sti/ieconomy/45559596.pdf).

9% to 20% less likely to mention side effects of this drug (based on the unconditional probability of an article mentioning side effects of 10%). The effect increases with higher amounts of advertising dollars.

Figure 2 adds another dimension of biased media reporting by showing that newspapers are also less likely to report on FDA safety alerts regarding their advertising clients' drugs (for underlying regression results, see table S5). This analysis is based on all articles that appeared within three days before and 14 days after the official safety alert. The effects are even stronger than those we find for side effects: Receiving $100,000 for advertising over the past 14 days is associated with a 23% to 33% decrease in the probability that the newspaper covers an FDA safety alert. The baseline probability is 62%, so that the effect again implies a substantial reduction.

In our last set of results, we show how the general tone of an article is related to previous advertising expenditures. We use an automated textual analysis procedure to classify the tone of each drug-related article. We use the bag-of-words approach to classify the tone of each article. In the first step, this approach to textual analysis consists of creating a dictionary containing words that have a negative connotation in drug-related articles. Then, an automated computer program is used to compute the fraction of negative words in each article that appears in the dictionary. This simple and easily replicable method has been used, for example, in psychology, political science or business research *(14-16)*. The word list we used to classify the tone of an article is provided in table S7. To test whether a newspaper reports more favorably on drugs marketed by its advertisers, we run ordinary least squares fixed effects regressions. As is common in textual analysis, we focus on negative tone and use the fraction of negative words that appear in an article according to our word list as the dependent variable. We expect the advertising coefficients to have a negative sign, indicating that more advertising is associated with less negative words.

In fig. 3, we illustrate the coefficients from regressions including drug-month as well as drug-month and newspaper fixed effects (for underlying results, see table S8). The coefficients using only drug-month fixed effects or additionally newspaper fixed effects are both significantly negative ($p<0.01$) and indicate that higher advertising revenues are associated with a decrease in the proportion of negative words used in an article, i.e., newspaper articles on drugs are written less negatively if the producer spends more on advertising.

Our results are robust to variations in the time period over which we measure advertising. They hold for advertising aggregated over seven days to up to ninety days, but tend to get weaker the longer the time period used for aggregation (see tables S3 and S5). Our results are also robust with respect to alternative specifications of the statistical model (i.e., estimating linear probability models instead of conditional logit models; see tables S4 and S6).

Our results indicate that pharmaceutical advertising reduces the extent to which the media reports on potential harms of their advertising clients' drugs. As a result of the biased reporting, the public's perception of drug safety is likely to be distorted and public health can eventually be adversely affected. Thus, our findings also add to the mounting evidence of the adverse effects of the pressure exerted by the pharmaceutical industry on healthcare research and practice *(17)*.

We anticipate that our essay will start new discussions on the regulation of advertising by pharmaceutical companies. For example, direct-to-consumer advertising (DTCA) of prescription drugs, which is only allowed in the U.S. and in New Zealand, has been criticized for downplaying safety information and for causing overmedication *(17-20)*. Our results provide a new argument against DTCA, showing it creates additional distortions beyond its direct impact on health behaviors. Furthermore, our results call for a stricter separation of the marketing and the editorial departments of newspapers, particularly when it comes to critical news issues like drug safety.

While we focus on newspaper articles on drug safety, the conflicts of interest documented in this study may also apply to other media, like internet platforms, television or magazines. The pharmaceutical industry is the largest advertiser in magazines and the second largest in television. Hence, the effect we find for newspapers---where the pharmaceutical industry is only the 11th largest advertiser---might actually be a lower bound of the influence the industry has on these other media channels. Thus, caution is warranted in the interpretation of reports on drug safety in all advertising-financed media outlets.

References and Notes:

1 T. Johnson (1998) Medicine and the media. N. Eng. J. Med. 339:87-92.

2 V. Entwistle (1995) Reporting research in medical journals and newspapers. BMJ 310:920-923.

3 A. Wilson, B. Bonevski, A. Jones, D. Henry (2009) Media reporting of health interventions: Signs of improvement, but major problems persist. PLOS One 4:1-5.

4 A. Cassels, et al. (2003) Drugs in the news: An analysis of Canadian newspaper coverage of new prescription drugs. Can. Med. Assoc. J. 168:1133-1137.

5 S. Dentzer (2009) Communicating medical news – pitfalls of health care journalism. N. Eng. J. Med. 360:1-3.

6 A. Wilson, J. Robertson, P. McElduff, A. Jones, D. Henry (2010) Does it matter who writes medical news stories? PLOS Med. 7:1-5.

7 R. Moynihan, et al. (2000) Coverage by the news media of the benefits and risks of medications, N. Eng. J. Med. 342:1645-1650.

8 M. Shuchman, M. S. Wilkens (1997) Medical scientists and health news reporting: A case of miscommunication. Ann. Intern. Med. 126:976-982.

9 J. Fishman, T. Ten Have, D. Casarett (2010) Cancer and the media: How does the news report on treatment and outcomes? Arch. Intern. Med. 170:515-518.

10 L. Schwartz (2004) The Media Matter : A Call for Straightforward Medical Reporting. Ann. Intern. Med. 140:226–228.

11 K.E. Warner (1985) Cigarette advertising and media coverage of smoking and health. N. Eng. J. Med. 312:384-388.

12 K.E. Warner, L.M. Goldenhar, C.G. McLaughlin (1992) Cigarette Advertising and magazine coverage of the hazards of Smoking - A statistical analysis. N. Eng. J. Med. 326:305-309.

13 M. Khoury, J. Ioannidis (2014) Big data meets public health. Science 346:1054-1055.

14 Y.R. Tausczik, J.W. Pennebaker (2010) The psychological meaning of words: LIWC and computerized text analysis methods. J. Lang. Soc. Psych. 29:24–54.

15 J. Grimmer, B.M. Stewart (2013) Text as data: The promise and pitfalls of automatic content analysis methods for political texts. Polit. Anal. 21:267–297.

16 T. Loughran, B. McDonald (2011) When is a liability not a liability? Textual analysis, dictionaries, and 10Ks. J. Finance 66:35–64.

17 E. Stamatakis, R. Weiler, J. Ioannidis (2013) Undue industry influences that distort healthcare research, strategy, expenditure and practice: A review. Eur. J. Clin. Invest. 43:469–475.

18 M.F. Hollon (2005) Direct-to-consumer advertising - A haphazard approach to health promotion. JAMA 293:2030-2033.

19 E. Berndt (2005) To inform or persuade? Direct-to-consumer advertising of prescription drugs. N. Engl. J. Med. 352:325-328.

20 J. Donohue, B. Cevasco, M. Rosenthal (2007) A decade of direct-to-consumer advertising of prescription drugs. N. Eng. J. Med. 357:673-681.

21 J.W. Pennebaker, M.R. Mehl, K.G. Niederhoffer (2003) Psychological aspects of natural language use: our words, our selves. Annual review of psychology 54:547–77.

22 M. Arellano, B. Honoré (2001) Panel data models: some recent developments. Handbook of econometrics 5:3229-3296.

Acknowledgments: Authors' names are listed alphabetically; the authors contributed equally to this work. The work was supported by grant number RU 1210/9-1 of the German Research Foundation (DFG). The authors declare no competing financial interests. Florens Focke is employed at Merck KGaA, Darmstadt. However, the work in this article has been conducted completely independently of his employment.

Fig. 1: Effect of advertising on side effect mentions

This figure presents the results from conditional logit regressions of a dummy variable equal to one if a newspaper article mentions side effects, and zero otherwise, on past advertising. The lines indicate the change in probability for different levels of advertising in the 14 days prior to the publication of the article. The solid (dashed) line displays results from a conditional logit regression using *drug-month* (*drug-month* and *newspaper*) fixed effects. This effect is statistically significant at the 1% (5%) level using *drug-month* (*drug-month* and *newspaper*) fixed effects.

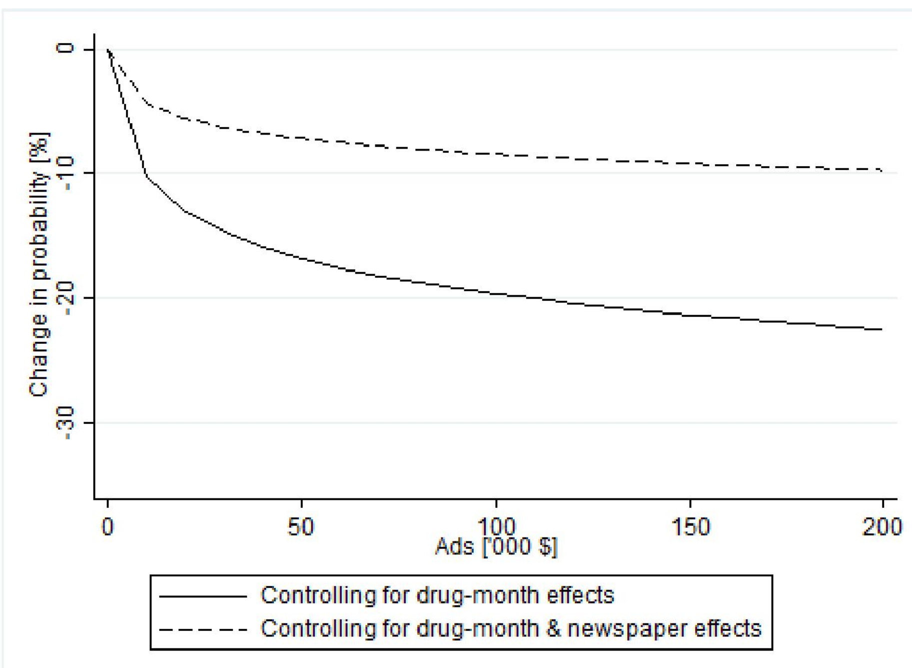


Fig. 2: Effect of advertising on safety alert coverage

This figure presents the results from conditional logit regressions of a dummy variable equal to one if a newspaper article covering an FDA safety alert, and zero otherwise, on past advertising. The lines indicate the change in probability for different levels of advertising in the 14 days prior to publication. The solid (dashed) line displays results from a conditional logit regression using *drug-month* (*drug-month* and *newspaper*) fixed effects. This effect is statistically significant at the 1% level in both specifications.

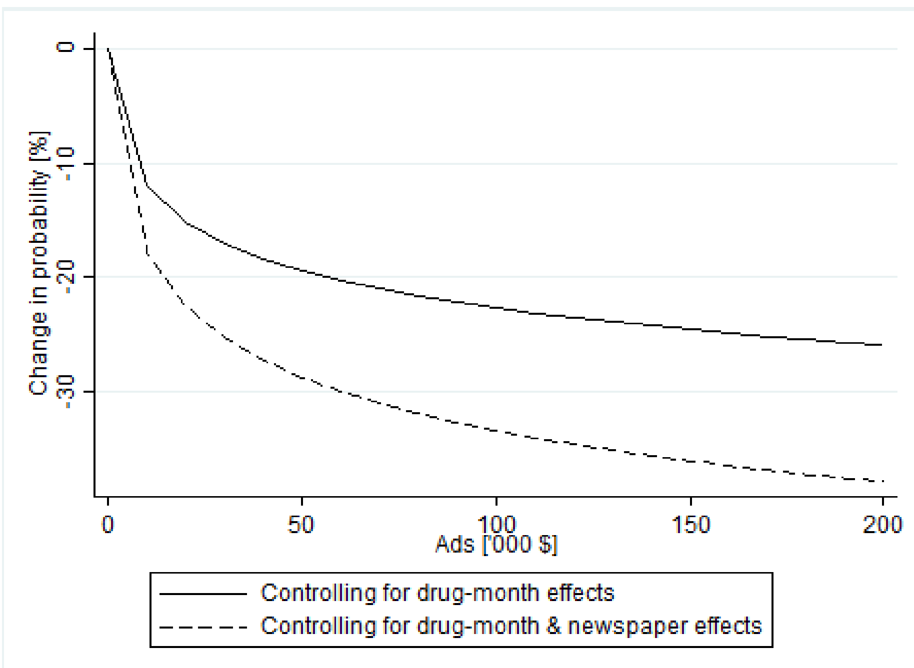
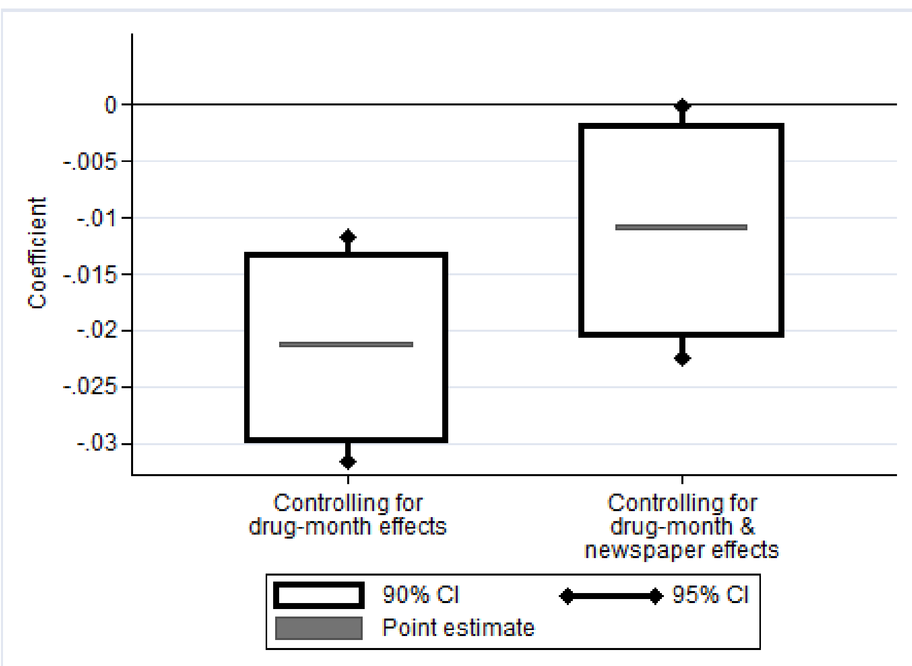


Fig. 3: Effect of advertising on article tone

This figure presents the results from ordinary least squares regressions of the tone of newspaper articles on past advertising. Point estimates and confidence intervals are from regressions including interacted *drug-month* and interacted *drug-month* and *newspaper* fixed effects as indicated in the figure. The figure shows results from a specification with the log of advertising in the 14 days prior to the publication of the article as the independent variable. According to both specifications, the amount of advertising spent by a drug's manufacturer at a newspaper is associated with a reduction in the fraction of negative words about the drug. This effect is statistically significant at the 1% (10%) level using *drug-month* (*drug-month* and *newspaper*) fixed effects.



Supplementary Materials:

Materials and Methods

Tables S1-S8

References (*21-22*)

Supplementary Materials

Focke, Niessen-Ruenzi, Ruenzi: Pharmaceutical advertising biases media reports on drug safety

Materials and Methods:

Data. Our sample is based on all drugs listed in the Drugs@FDA database.[†] We obtain the name of each drug as well as the name of its manufacturer(s). We clean the drug names to exclude name components which would likely not be used in newspaper articles referring to the drug (e.g., "in plastic container" or "preservative free").

We obtain advertising data from the media-monitoring firm Kantar Media from 1999 to 2012. These data contain all advertisements that are published in a core set of 155 important US newspapers. Their proprietary database, "Kantar Media Stradegy", contains the advertisement's publication date, the news outlet in which it is published, and the firm that commissioned the advertisement. In addition, each advertisement is associated with a precise cost estimate using "rate cards" that indicate advertising prices depending on size, product categories, and days of the week or newspaper sections. We link this dataset with the drug manufacturers from the Drugs@FDA database by their names and, where necessary, by their most advertised products. We identify 171 drug manufacturing companies that advertise in US newspapers according to the Kantar database. We use the total advertising expenditures (rather than just expenditures for a specific drug) by these companies in a specific newspaper in our analysis. In doing so, we assume that a newspaper's reports about a specific firm's drugs is driven by the overall advertising budget this firm spends at the specific newspaper.

In the next step, we obtained newspaper articles on each drug from the newspaper archive LexisNexis. Out of the 155 US newspapers that are covered by Kantar Media, 39 newspapers are also covered by LexisNexis. These are listed in table S1. All major national newspapers as well as many of the most important regional newspapers were included. The article search was conducted using the drug name and the SmartIndexing subject field of LexisNexis, which allows us to condition solely on articles covering drug-related topics. SmartIndexing is a technology developed by LexisNexis that classifies articles based on their content.[‡] We require that LexisNexis classify an article as making a major reference to pharmaceutical drugs; we exclude articles that refer to illegal drugs or drug trafficking. We identified 81,656 articles that meet these criteria, covering 1,651 different drugs. In some cases more than one article on a drug was published in the same issue of a newspaper; we aggregated these to arrive at 79,205 observations at the drug-newspaper-day level.

We classify whether an article covers side effects based on the LexisNexis SmartIndexing field "Drug Interactions & Side Effects", setting $I_{SideEffects}$ to 1 if the article mentions drug side effects, and zero otherwise. This dummy variable is the dependent variable in our side effects regressions reported in tables S3 and S4, which also underlie the results in fig. 1. To verify that the

---

[†] US Food and Drug Administration (FDA), Drugs@FDA database (2014; http://www.accessdata.fda.gov/scripts/cder/drugsatfda).

[‡] LexisNexis, Introduction to LexisNexis SmartIndexing Technology (2015; http://www.lexisnexis.com/infopro/resource-centers/product_resource_centers /b/smartindexing/archive/2013/09/11/what-is-lexisnexis-smartindexing-technology.aspx).

LexisNexis SmartIndexing field correctly identifies articles on side effects, we randomly selected 500 articles that were classified as covering side effects and manually checked their accuracy. Out of these, only 7.60% of articles did not cover side effects, verifying that LexisNexis reliably classifies articles. Moreover, of those articles covering side effects, the overwhelming majority (88.31%) discussed them in a critical manner. Therefore, we conclude that articles classified as mentioning side effects generally do so in a negative manner.

For our analysis on safety alerts, we use the FDA MedWatch Safety Alerts for Human Medical Products database.[§] Under this program, the FDA publishes warnings such as newly discovered adverse reactions, side effects, product recalls or market withdrawals. We obtain information on all safety alerts issued by the FDA from 2002 to 2012 and link them with articles on the drugs affected by the alert. We consider all articles appearing within three days before and fourteen days after the alert was issued. We start three days prior to the actual FDA announcement to allow for early leakage of the alert information. There are several examples in our article database where a safety alert was expected for the following days in a newspaper article. We collect articles up to two weeks after the FDA alert to capture continued coverage over a period of time after the actual alert announcement date. We identified 3,284 articles on 306 safety alerts. We manually coded each individual article as mentioning the alert or not, setting $I_{FDAalert}$ to 1 if an article covers the alert, and zero otherwise. This dummy variable is the dependent variable in our safety alert regressions reported in tables S5 and S6, which also underlie the results in fig. 2.

To analyze the tone of each article, we employ a bag-of-words technique which is frequently used in linguistic analysis *(14)*. Prior research on the bag-of-words approach has stressed the importance of a context-specific word list *(21)*. Thus, to compile a word list with negative words specific to the medical domain, three independent coders reviewed 100 randomly drawn articles on drugs and collected all words that had a negative connotation in the context of each article. In the next step, the authors manually checked all words that were mentioned by at least one coder individually. To determine whether a word is truly negatively connoted in a medical context, we used the Keyword-in-Context feature of WordStat. This program displays the word with its context in the articles in our sample. We only retained words in our list if they still appeared negative after considering the context in which they are typically used. For ease of replication of our results, we include the complete word list in table S7. For each article, we then use the textual analysis program LIWC which counts the number of words from the word list appearing in an article and divides it by the total number of words in the article. The resulting variable, *Negativity*, which is the fraction of negative words in an article, is used as the dependent variable in our regressions on the impact of advertising on an article's tone reported in table S8. These regressions also underlie the results in fig. 3.

We combined our data on advertising and newspaper coverage of drugs to create our main dataset. For each day with an article on a drug, our dataset includes both information on the article as well as on past advertising by the manufacturer at a given newspaper. For our main analysis, we aggregate past advertising over the previous two-week period.

Statistical analysis. Data were analyzed using the statistical software package Stata (see www.stata.com). For the analysis of side effect mentions and of general article tone, we used our full sample of articles. For the analysis of FDA safety alerts, we considered only articles that

[§] US Food and Drug Administration (FDA), MedWatch Safety Alerts for Human Medical Products (2014; http://www.fda.gov/Safety/MedWatch/SafetyInformation/ SafetyAlertsforHumanMedicalProducts/).

were published within three days before and fourteen days after the official announcement date of the alert.
The dependent variables in our analysis of side effect mentions and safety alert coverage are binary variables. As our identification strategy relies on the inclusion of a large number of high-dimensional fixed effects, we use a conditional logit model rather than a simple logit model *(22)*. In all our models we include drug-month interacted fixed effects. The regression model is:

$$y_{ijt} = \alpha + \beta_1 Log(14\text{-}day\ Ad\ \$)_{ijt} + u_{im} + \varepsilon_{ijt}, \qquad (1)$$

where i denotes a particular drug, j denotes a particular newspaper, t denotes the day and m denotes the month, respectively. The dependent variable, y_{ijt}, corresponds to either $I_{SideEffects}$ for the analysis of side effect mentions or $I_{FDAalert}$ for the analysis of safety alerts. These two variables are equal to one if an article on drug i in newspaper j on day t mentions side effects or FDA safety alerts, respectively, and zero otherwise. *Log(14-day Ad \$)* is the sum of advertising expenditures of the pharmaceutical company producing drug i in newspaper j over the 14 days before the publication of the article. u_{im} represents drug-month interacted fixed effects.

Including drug-month interacted fixed effects enables us to compare articles written on the same drug at the same point in time. By doing that we control for time-varying characteristics of the drug, such as the general current perception of its efficacy and safety or the severity of a safety alert issued by the FDA. Thus, including fixed effects allows us to better address causality in the relationship between advertising expenditures and newspaper reporting. For example, it could be that new research shows that a specific drug has more serious side effects than previously thought. The pharmaceutical company producing this drug could then decide not to actively market this drug anymore, leading to a reduction in their overall advertising budget. At the same time, newspapers might become aware of the new research results and eventually cover them in articles or generally write more negatively about this drug. Then, without including drug-month fixed effects, we might find a positive coefficient for the impact of advertising on, e.g., side effect mentions, although there is no causal relationship. Including drug-month fixed effects controls for such patterns. Then, any effect we find for an impact of advertising on article content must be driven by across-newspaper differences in advertising and article content.

One remaining concern with only including drug-month fixed effects is that some newspapers might have a general tendency to report more favorably about drugs and that drug companies advertise more with such newspapers for reasons unrelated to their reporting style. Thus, in a more conservative regression specification, we additionally include newspaper fixed effects (besides drug-month fixed effects) to control for differences in the general writing styles of newspapers. In this specification, we thus add a v_j term to equation (1). The newspaper fixed effect, v_j, controls for (time-invariant) general newspaper reporting tendencies regarding drugs. Of course, such reporting tendencies could themselves be caused by advertising. Including newspaper fixed effects is therefore restrictive by only considering the impact of within-newspaper changes in advertising on newspaper reporting. In this more restrictive specification, the coefficient on our independent variable, the log of past advertising dollars, identifies the effect of advertising on media coverage in different newspapers, controlling for time-varying drug characteristics and general newspaper reporting tendencies. Consequently, these results should generally provide a lower limit for the true impact of advertising on drug reporting. Any effect we document is now driven by within-newspaper changes in the way they report on a particular drug due to changes in short-term advertising and allows us to establish causality.

In fig. 1 and 2, we present the percentage changes in the probability of an article's mentioning side effects and safety alerts, respectively. Lagged advertising revenues are aggregated over 14 days before an article is published. These figures draw on the regression results reported in tables S3 and S5, Column 2. To ensure robustness of our results, we also aggregate advertising revenues over 7 days (Column 1) and 30 days (Column 3) before an article is published, respectively. We find a statistically and economically similar effect across all horizons and in both the analysis of side effects and FDA safety alerts. Note that only drug-months in which at least one but not all articles on a drug cover side effects can be included in the estimation of the conditional logit model. Hence, the number of observations in Tables 4 and 6 is lower than the full sample of about 79,000 observations.

In addition to conditional logit models, we also estimate linear probability models as an alternative statistical approach. In contrast to conditional logit models, linear probability models have the advantage that they can be used to calculate predicted probabilities and at the same time accommodate a large number of fixed effects. However, predicted values of a linear probability model are not bound to the unit interval and are thus harder to interpret. Therefore, our main specification relies on a conditional logit model that allows us to directly interpret the regression coefficients. However, to ensure the stability of our results, we also estimate equation (1) as a linear probability model. Results are reported in tables S4 and S6. We use the same fixed effects and advertising variables as for our conditional logit regressions. The effect sizes are slightly larger compared to the conditional logit regressions in our analysis of side effect mentions and somewhat smaller in our analysis of FDA safety alerts. All effects in the linear probability models are still highly statistically significant, confirming our results from the conditional logit regressions. We therefore conclude that our results are robust to using either econometric method.

In our analysis of the impact of advertising on newspaper article tone, we again use the full sample of articles. The regression model follows the same structure as in equation (1). The dependent variable y_{ijt} is *Negativity*, i.e., the fraction of negative words in an article. We estimate this model using ordinary least squares (OLS) regressions with the same fixed effects and advertising variables as in our other tests. table S8 presents the results from this model, which also underlie fig. 3.

Table S1 **Newspapers used in our analysis**	
Atlanta Journal-Constitution	Passaic North Jersey Herald News
Austin American-Statesman	Philadelphia Daily News
Bergen Record	Philadelphia Inquirer
Boulder Daily Camera	Pittsburgh Post-Gazette
Buffalo News	Pittsburgh Tribune-Review
Charleston Gazette	Providence Journal
Contra Costa Times	Richmond Times-Dispatch
Daily Herald Arlington Heights	Salt Lake Tribune
Dayton Daily News	San Bernardino Sun
Denver Post	San Jose Mercury News
Deseret Morning News	St Louis Post-Dispatch
Inland Valley Daily Bulletin	St Paul Pioneer Press
Los Angeles Daily News	Tampa Bay Times
Minneapolis Star Tribune	Tampa Tribune
New York Daily News	Tulsa World
New York Post	USA Today
New York Times	Wall Street Journal
Omaha World Herald	Washington Post
Orange County Register	West Chester Daily Local News
Palm Beach Post	

Table S2 **Summary statistics**						
	Mean	Median	Std.dev.	p95	p5	N
Article # (month)	1.63	1	1.67	4	1	28,310
Word # ('000)	0.72	0.58	0.55	1.67	0.15	79,250
Negativity	2.04	1.76	1.37	4.61	0.31	79,250
$I_{SideEffects}$	0.10	0	0.3	1	0	79,250
$I_{FDAalert}$	0.62	1	0.49	1	0	3,284
7-day Ads ('000 $)	3.88	0	30.34	0	0	77,864
14-day Ads ('000 $)	7.76	0	51.64	13.97	0	77,794
30-day Ads ('000 $)	16.45	0	90.95	82.88	0	77,541

All statistics, except for $I_{FDAalert}$, are based on the full sample of articles. *Article # (month)* is the sum of articles published on a drug in a month across all newspapers. The number of observations of this variable therefore is the number of drug-months in the sample. *Word # ('000)* is the number of words per article in thousands. *Negativity* is the fraction of negative words in an article, where negative words are those listed in Table 2. $I_{SideEffects}$ is an indicator equal to one if an article is classified as covering side effects according to LexisNexis' SmartIndexing technology, and zero otherwise. $I_{FDAalert}$ is only measured in the three days before and fourteen days after the official release of an FDA safety alert. It is an indicator equal to one if an article covers an FDA safety alert, and zero otherwise. *7 (14/30) day Ads ('000 $)* is the sum of advertising expenditures in thousands of the company manufacturing a drug in the 7 (14/30) days before the article is published.

Table S3 **Advertising and side effect mentions (conditional logit model)**			
Dependent variable: $I_{SideEffects}$			
	Drug-month fixed effects		
	(1)	(2)	(3)
Log(7-day Ad $)	-0.1011***		
	(-4.87)		
Log(14-day Ad $)		-0.0864***	
		(-4.93)	
Log(30-day Ad $)			-0.0904***
			(-6.38)
No. observations	27,026	26,986	26,911
Log-likelihood	-8485.63	-8469.42	-8437.40
Effect size for 100k in ads	-22,91%	-19.68%	-20.56%
	Drug-month and newspaper fixed effects		
	(1)	(2)	(3)
Log(7-day Ad $)	-0.0546**		
	(-2.45)		
Log(14-day Ad $)		-0.0452**	
		(-2.32)	
Log(30-day Ad $)			-0.0489***
			(-3.18)
No. observations	27,026	26,986	26,911
Log-likelihood	-8234.03	-8218.75	-8190.58
Effect size for 100k in ads	-10.19%	-8.54%	-9.24%

The table presents results from a conditional logit model with $I_{SideEffects}$ as the dependent variable. The analysis is based on the full sample of articles. $I_{SideEffects}$ is an indicator equal to one if an article is classified as covering side effects according to LexisNexis' SmartIndexing technology, and zero otherwise. *7 (14/30) day Ads* is the sum of advertising expenditures of the company manufacturing a drug in the 7 (14/30) days before the article is published. Effect size is the percentage change in the probability of an article's mentioning side effects associated with advertising spending of 100,000 $ by the manufacturer of the drug in the newspaper publishing the article. *Drug-month* is an interacted fixed effect. Standard errors are clustered by drug. *t*-statistics in parentheses. ** and *** denote significance at the 5% and 1% level, respectively.

Table S4 **Advertising and side effect mentions (linear probability model)**			
Dependent variable: $I_{SideEffects}$			
	Drug-month fixed effects		
	(1)	(2)	(3)
Log(7-day Ad $)	-0.0072***		
	(-5.56)		
Log(14-day Ad $)		-0.0064***	
		(-5.38)	
Log(30-day Ad $)			-0.0065***
			(-7.28)
No. observations	77,864	77,794	77,541
Adj. R^2	0.1979	0.1985	0.1989
Effect size for 100k in ads	-32.38%	-28.85%	-28.79%
	Drug-month and newspaper fixed effects		
	(1)	(2)	(3)
Log(7-day Ad $)	-0.0043***		
	(-3.17)		
Log(14-day Ad $)		-0.0036***	
		(-2.72)	
Log(30-day Ad $)			-0.0037***
			(-3.71)
No. observations	77,864	77,794	77,541
Adj. R^2	0.2055	0.2061	0.2064
Effect size for 100k in ads	-19.60%	-16.41%	-16.52%

The table presents results from a linear probability model with $I_{SideEffects}$ as the dependent variable. The analysis is based on the full sample of articles. $I_{SideEffects}$ is an indicator equal to one if an article is classified as covering side effects according to LexisNexis' SmartIndexing technology, and zero otherwise. *7 (14/30) day Ads* is the sum of advertising expenditures of the company manufacturing a drug in the 7 (14/30) days before the article is published. Effect size is the percentage change in the probability of an article mentioning side effects associated with advertising spending of 100,000 $ by the manufacturer of the drug in the newspaper publishing the article. *Drug-month* is an interacted fixed effect. Standard errors are clustered by drug. *t*-statistics in parentheses. **,*** denote significance at the 5% and 1% level, respectively.

Table S5 **Advertising and FDA Safety Alerts mentions (conditional logit model)**			
Dependent variable: $I_{FDAalert}$			
	Drug-month fixed effects		
	(1)	(2)	(3)
Log(7-day Ad $)	-0.1122***		
	(-2.74)		
Log(14-day Ad $)		-0.0999***	
		(-2.80)	
Log(30-day Ad $)			-0.0746**
			(-2.04)
No. observations	2,407	2,407	2,407
Log-likelihood	-996.34	-996.12	-997.17
Effect size for 100k in ads	-25.32%	-22.65%	-17.05%
	Drug-month and newspaper fixed effects		
	(1)	(2)	(3)
Log(7-day Ad $)	-0.1656***		
	(-3.51)		
Log(14-day Ad $)		-0.1532***	
		(-3.33)	
Log(30-day Ad $)			-0.1226***
			(-2.84)
No. observations	2,407	2,407	2,407
Log-likelihood	-950.17	-949.56	-950.96
Effect size for 100k in ads	-35.91%	-33.33%	-27.18%

The table presents results from a conditional logit model with $I_{FDAalert}$ as the dependent variable. The analysis is based on all articles that appear within three days before and 14 days after the official release of an FDA MedWatch Safety Alert. $I_{FDAalert}$ is an indicator equal to one if an article covers an FDA safety alert, and zero otherwise. *7 (14/30) day Ads* is the sum of advertising expenditures of the company manufacturing a drug in the 7 (14/30) days before the article is published. Effect size is the percentage change in the probability of an article mentioning the FDA alert associated with advertising spending of 100,000 $ by the manufacturer of the drug in the newspaper publishing the article. *Drug-month* is an interacted fixed effect. Standard errors are clustered by drug. *t*-statistics in parentheses. **,*** denote significance at the 5% and 1% level, respectively.

Table S6 **Advertising and FDA Safety Alerts mentions (linear probability model)**			
Dependent variable: $I_{FDAalert}$			
	Drug-month fixed effects		
	(1)	(2)	(3)
Log(7-day Ad $)	-0.0140**		
	(-2.22)		
Log(14-day Ad $)		-0.0130**	
		(-2.56)	
Log(30-day Ad $)			-0.0098*
			(-1.70)
No. observations	3,284	3,284	3,284
Adj. R^2	0.4778	0.4779	0.4775
Effect size for 100k in ads	-10.38%	-9.62%	-7.27%
	Drug-month and newspaper fixed effects		
	(1)	(2)	(3)
Log(7-day Ad $)	-0.0212***		
	(-2.86)		
Log(14-day Ad $)		-0.0202***	
		(-3.24)	
Log(30-day Ad $)			-0.0168**
			(-2.32)
No. observations	3,284	3,284	3,284
Adj. R^2	0.4880	0.4883	0.4878
Effect size for 100k in ads	-15.72%	-14.90%	-12.36%

The table presents results from a linear probability model with $I_{FDAalert}$ as the dependent variable. The analysis is based on all articles that appear within three days before and 14 days after the official release of an FDA MedWatch Safety Alert. $I_{FDAalert}$ is an indicator equal to one if an article covers an FDA safety alert, and zero otherwise. *7 (14/30) day Ads* is the sum of advertising expenditures of the company manufacturing a drug in the 7 (14/30) days before the article is published. Effect size is the percentage change in the probability of an article mentioning the FDA alert associated with advertising spending of 100,000 $ by the manufacturer of the drug in the newspaper publishing the article. *Drug-month* is an interacted fixed effect. Standard errors are clustered by drug. *t*-statistics in parentheses. **,*** denote significance at the 5% and 1% level, respectively.

Table S7 **Word list used in tone analysis (variable *Negativity*)**

ABUSE*	DANGER*	HEADACHE*	OUTCRY	SLEEPLESSNESS
ACCUSE*	DEADLY	HIGH-PRICED	OUT-OF-POCKET	STROKE
ACUTE	DEATH*	HORRIBLE	OVERDIAGNOSED	STRUGGLE*
ADDICTION	DEBILITATING	HOSPITALIZATIONS	OVERDOSE	SUFFER*
ADDICTIVE	DECEPTIVE*	HOSTILITY	OVERLOAD	SUICIDE*
ADVERSE	DELAY	HURT*	OVERPRESCRIBE*	SUING
AFRAID	DENIED	IGNORE*	OVERREACT	SUSPEND*
AGGRESSIVE	DENOUNCE	ILLEGAL	OVERUSE*	SWOLLEN
AGONIZE*	DENY	ILL-PREPARED	PAIN	TAINTED
ALARMING	DESPERATE*	IMPAIR*	PAINFUL	TENSIONS
ALLEGEDLY	DETRIMENTAL	IMPLICATE*	PANIC	TERRIBLE
ANGERED	DEVASTATING	INADEQUATE	PARANOID	THREATEN*
ANGRY	DICEY	INCAPABLE	PENALITES	TOOTHLESS
ANXIETY	DIE*	INFLAMMATION	PLAGUED	TOXIC*
ANXIOUS	DIFFICULT*	INSOMNIA	PLUMMET*	TROUBLE
AWFUL	DILEMMA	INSTABILITY	PLUNGE*	TROUBLING
BALLOONING	DISAGREE*	INTERACTIONS	POISON	ULCERS
BAN	DISAPPOINTED	INTERFERE*	POOR	UNBEARABLE
BANNED	DISAPPOINTING	INTIMIDATE*	PRECAUTION	UNCLEAR
BANS	DISASTER	INTOXICAT*	PROBLEM*	UNCOMFORTABLE
BATTLE	DISPUTE*	INVESTIGATION	PROHIBIT*	UNEXPECTED
BATTLING	DISRUPT*	IRREPARABLE	PROSECUTION	UNEXPLAINED
BURDEN	DISTURBING	IRRESPONSIBLE	PUNISHABLE	UNFAIR
CATASTROPHE	DIZZY	IRREVERSIBLE	PUNISHING	UNFORTUNATELY
CAUSE*	DOPED	IRRITABILITY	QUESTIONED	UNHAPPY
CAUTION	DOWNING	IRRITATE*	RASH	UNPLEASANT
CHARGES	DOWNPLAY	JEOPARDIZE*	REACTIONS	UNPREDICTABLE
CHOPPY	DOWNTURN	JUGGLING	REBOUND	UNREASONABLE
CLAIMS	DROWSY	JURY	RECALL	UNRESPONSIVE
CLASH	DRUG-INDUCED	KILL*	RECKLESS	UNSAFE
COERCIVE	DRUG-RESISTANT	LABEL*	REFUSE	UNWILLINGNESS
COMPLAIN*	ERUPTED	LAGGING	RELAPSE	URGED
COMPLICATION*	ESCALATE	LAMBAST*	REMOVAL	VERDICT
CONCERN*	EXPIRATION	LAMENTED	RESISTANCE	VICTIM*
CONFLICT	EXPIRING	LAWSUIT*	RESISTANT	VIOLATE
CONGESTION	FAIL*	LETHAL	RESTATE	VIOLENT
CONTAMINATED	FALSE*	LIFE-THREATENING	RESTLESS*	VOMIT*
CONTROVERS*	FATAL	LIMITED	RESTRICTING	VULNERABLE
CO-PAID	FATALITIES	LOOMING	RIGGED	WARN*
CO-PAY*	FATIGUE	LOST	RIGGING	WHEEZING
COST*	FEAR*	MEAGER	RISKIER	WITHDRAW*
CRANKY	FECKLESS	MIGRAINE	RISKY	WORRI*
CRIMP	FELONY	MISDIAGNOSED	SAD	WORRY
CRIPPLING	FEVER	MISLEAD	SAFETY	WORSE*
CRISIS	FLAW*	MISPRESCRIBED	SCARE*	WORTHLESS
CRITICI*	FOOLISH	MISREAD	SEIZURE	WREAK*
CRITICS	FRUSTRATE*	MISUSED	SERIOUS	WREAKS
CURBED	GLITCH*	NASTY	SETTLE	WRONG*
CURSE*	HARM*	NAUSEA	SEVERE*	YANK*
CUTBACKS	HARSH	NEGATIVE	SHORTAGE*	
DAMAGE*	HAVOC	NEGLECT*	SKEPTICAL	

* indicates that all inflections to this word stem are included in the word list.

Table S8 **Advertising and article tone (OLS model)**

Dependent variable: *Negativity*

	Drug-month fixed effects		
	(1)	(2)	(3)
Log(7-day Ad $)	-0.0199***		
	(-3.51)		
Log(14-day Ad $)		-0.0216***	
		(-4.27)	
Log(30-day Ad $)			-0.0207***
			(-6.46)
No. observations	77,864	77,794	77,541
Adj. R^2	0.3275	0.3277	0.3274
	Drug-month and newspaper fixed effects		
	(1)	(2)	(3)
Log(7-day Ad $)	-0.0103*		
	(-1.70)		
Log(14-day Ad $)		-0.0112**	
		(-1.96)	
Log(30-day Ad $)			-0.0104***
			(-2.96)
No. observations	77,864	77,794	77,541
Adj. R^2	0.3381	0.3382	0.3378

The table present results from an ordinary least squares regression with *Negativity* as the dependent variable. The analysis is based on the full sample of articles. *Negativity* is the fraction of negative words in an article, where negative words are those listed in Table 2. 7 (14/30) day Ads is the sum of advertising expenditures of the company manufacturing a drug in the 7 (14/30) days before the article is published. *Drug-month* is an interacted fixed effect. Standard errors are clustered by drug. *t*-statistics in parentheses. **,*** denote significance at the 5% and 1% level, respectively.