



Managers should not stop checking alternative data once a data source confirms their preconceptions or indicates that a strategy is working great.



# Metrics Gone Wrong: What Managers Can Learn from the 2016 US Presidential Election

Raoul Kübler and Koen Pauwels

## KEYWORDS

**Metrics, Dashboards, Decision-Making, Polls, Probabilistic Models, User-Generated Data**

## THE AUTHORS

### **Raoul Kübler**

Junior Professor of Marketing, Marketing Center  
Münster, Germany  
[raoul.kuebler@uni-muenster.de](mailto:raoul.kuebler@uni-muenster.de)

### **Koen Pauwels**

Distinguished Professor of Marketing,  
Northeastern University  
Boston, MA, USA  
[k.pauwels@northeastern.edu](mailto:k.pauwels@northeastern.edu)

**The age of data – boon or bane?** ✕ In the last decade, we witnessed an explosion of data availability. Humankind creates more data each day than we did in the last 20,000 years altogether. Despite all this data, it's not its size but what you do with it that matters. Marketers often start from the available data to brainstorm potential uses, instead of asking the right questions, and then dig in deep: how do we find the adequate answer, which type of data do we need, where do we get this data, and how do we access, process, and combine this data with our existing insights? This leads to important issues such as "Which source is reliable?" or "Which data is richer in information?" Don't stop once a data source confirms your preconceptions or indicates your strategy is working great – check alternative data sources to ensure your conclusions are valid. Our analysis of the 2016 US presidential election illustrates what could happen if you don't.

### **Let the user speak – the power of alternative data sources**

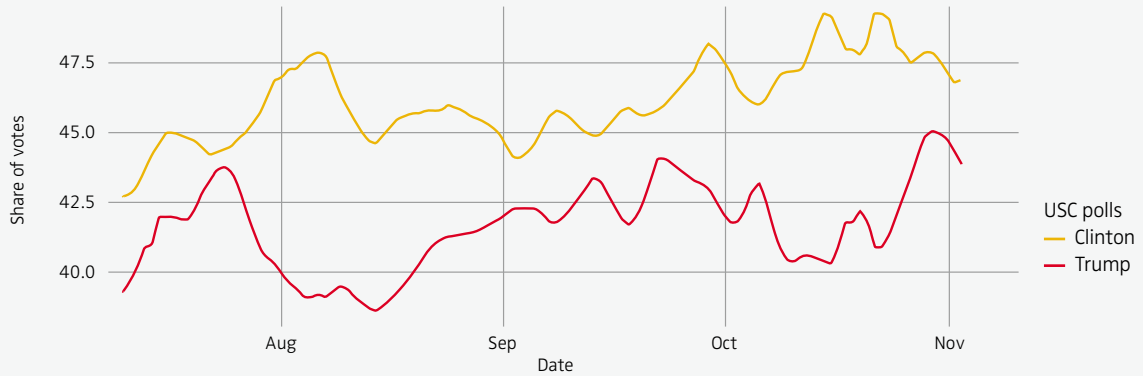
✕ Could any campaign manager have known better? Given that campaign managers – like marketing managers – often base their decision-making on a limited set of performance metrics, finding the right polls or marketing metrics becomes essential. How about looking at other data sources and variables that – like polls – show voter engagement and preferences? In the political arena, donations, media coverage, social media followership, engagement, and sentiment may similarly indicate how well a candidate is doing. In addition, most of these variables are available for free and can easily be "harvested" (Figure 3).

## BOX 1

## The 2016 US presidential election: The false comfort of being in the lead

Remember the morning of November 9, 2016, waking up to the “surprise” that Donald Trump would become the 45th US president? The vast majority of national and state polls showed his opponent, Hillary Clinton, in a comfortable lead throughout the whole race. Campaign managers could choose from almost 100 different polls, of which 80% predicted a comfortable win for Hillary Clinton. Figure 1 shows the averaged polls across all publicly available polling.

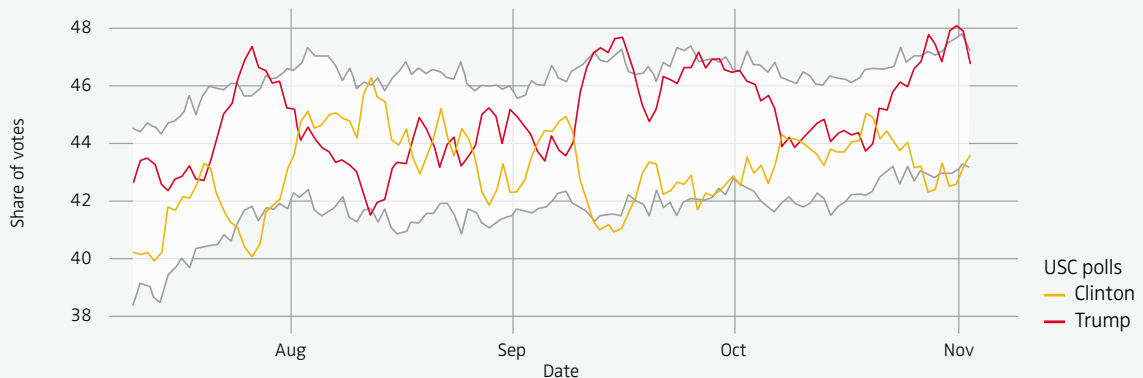
**FIGURE 1 > Summary of traditional polls**



Traditional polls for Trump and Clinton by Polltracker

In contrast, Figure 2 depicts the probabilistic polls by the University of Southern California (USC), showing a tight race – the overlapping confidence intervals between the grey lines – with Trump ahead for most of the campaign period, including the final weeks. A key difference between the USC poll and most of the other polls was that participants were not only asked to indicate their favorite candidate but also to indicate how likely they were to vote. Weighting these two factors makes the information richer and paints campaign managers a more accurate picture of the impact of important events. For instance, and clearly visible in the chart, many Clinton-leaning voters indicated a lower likelihood to vote after Clinton’s “Basket of Deplorables” leaked video (September 12th) and FBI director Comey’s letter to Congress about the FBI investigation into Clinton’s emails (October 28th). But wait, didn’t Clinton win the popular vote in the end? Yes, and Figure 2 shows a resurgence of Clinton in the last days, bringing us back into the area of insignificant difference between the candidates’ probabilistic poll numbers. This information is not commonly displayed by most polling institutes, but it is the key to combating over-optimism.

**FIGURE 2 > Probabilistic polls**



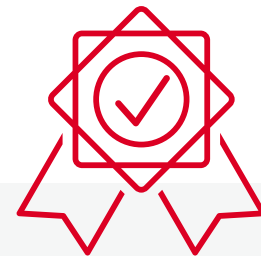
USC probability weighted polls – 95% confidence interval in white

»  
*Trusting a single “silver bullet” metric does not just lead to surprises; it can also mislead your decision-making.*  
 «

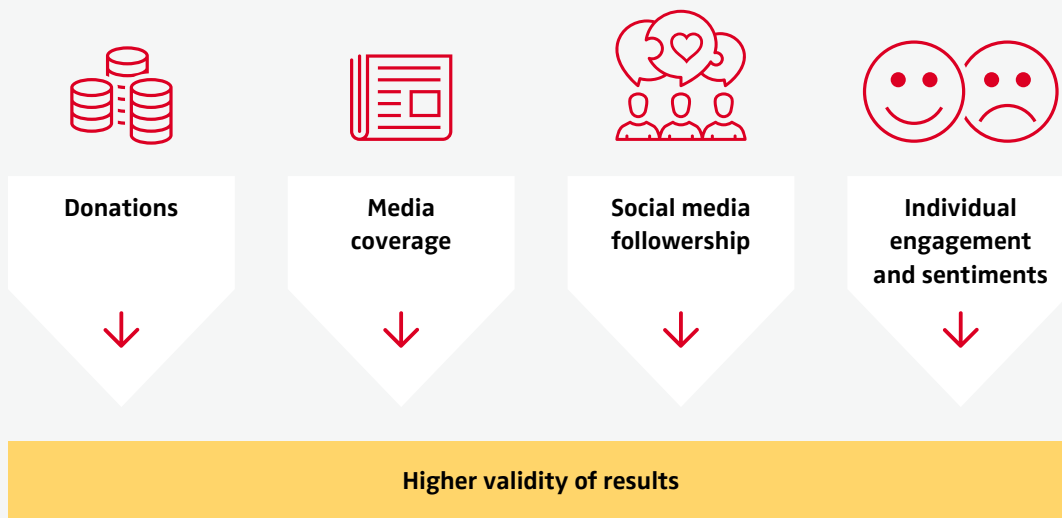
While the majority of traditional polls painted an overly optimistic picture for Hillary Clinton, the majority of user-generated data clearly indicated that the predicted landslide win of the democratic candidate was in jeopardy. Donations still may have comforted the Democratic campaign managers, as Clinton received much more in donations than her rival throughout the whole campaign. The amount of news media coverage in the 4 months prior to the election, however, showed a different picture. To obtain this information, we conducted a text mining and topical analysis of tweets by 56 major US news outlets prior to the election. This data clearly shows that Trump dominated the media and received much free publicity from the center and left-of-center media. Also, the left and left-leaning media focused more on the inter-party rivalry between Clinton and Sanders, while the right and right-leaning media focused more on highlighting Trump’s strengths and the weaknesses of his democratic opponent.

Social media deliver an even clearer picture. Throughout the whole campaign, Trump showed substantially more followership and higher growth than Clinton, again indicating that Trump enjoyed more momentum than indicated by the majority of the traditional polls (Figure 4).

A topical analysis of what users posted or commented on the two candidates’ social media pages reveals that the majority of social engagements were in favor of Donald Trump. Not only was user sentiment better for Trump, but the amount of disinformation spread trying to hurt Hillary Clinton was equally high on both of the candidates’ pages.



**FIGURE 3** > Alternative variables and data sources for ensuring higher validity of metrics





*Combining different data that is linked to a similar outcome helps to predict the actual outcome and to drive it with appropriate action.*



### **Blinded by the wrong numbers: A jeopardy for sound decision-making** ✕

Validating the bigger picture with alternative data sources is not limited to politics. The latest marketing research shows that online consumer behavior metrics can enrich and sometimes replace traditional funnel metrics. Trusting a single “silver bullet” metric does not just lead to surprises; it can also mislead your decision-making. Econometric models can help disentangle a complex web of dynamic interactions and show immediate and lagged effects of marketing or political events. Our model for the 2016 election clearly showed the impact of external events, user-generated content, campaign actions, and media coverage for both candidates, explaining the poll gap between the two candidates at different points in time.

### **Don't be fooled by data – Lessons for marketers** ✕

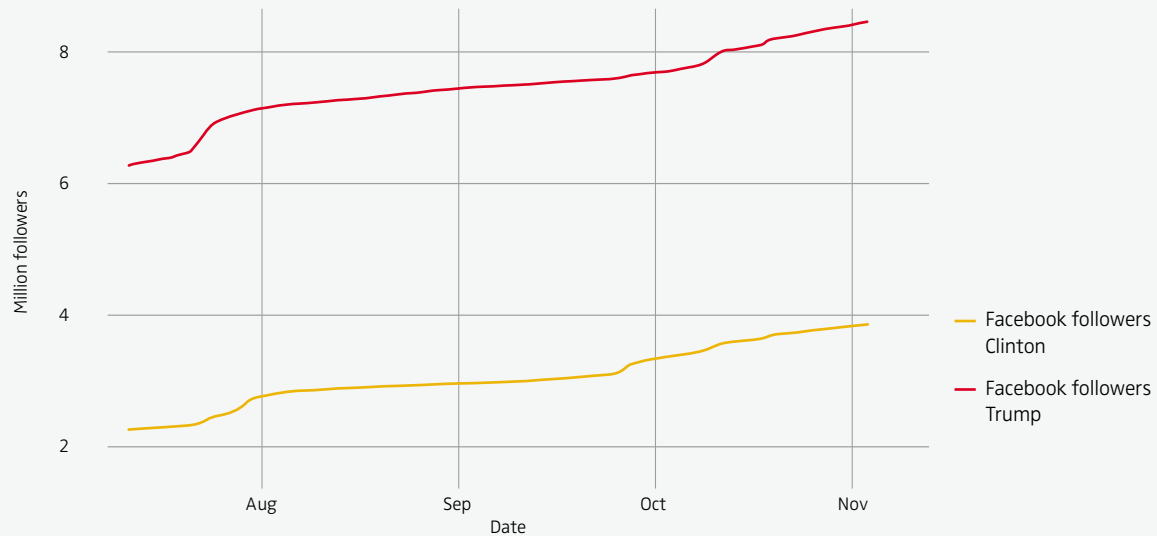
> **Assess numbers and forecasts critically** ✕ A key lesson from the 2016 elections for marketers is to stay as critical of numbers and forecasts as possible. Therefore, we recommend a healthy dose of skepticism when assessing insights presented to you. One way to do so in the avalanche of data is to verify existing insights and to check validity against alternative data sources. Combining different data that is linked to a similar outcome – in our scenario data on voter preference – helps to predict the actual outcome, to explain it and to drive it with appropriate action. As pointed out in our example, such data may be gained from users directly online – e.g., reviews, social media comments and posts, online forums, and other sources, such as statistical databases.

### > **Rely on marketing theory to evaluate suspicious or contradicting “evidence”** ✕

Another key challenge arises if the gathered data does not give a unanimous result. In this case, management's key responsibility is to use its expert judgment. A first step in the right direction is to check for face validity. Does the algebraic sign of the estimated effect ring true to you? As humans, we have the uncanny ability to integrate many different signals, from anecdotes and feelings to current data and the interpretation of past events. While an econometric analysis is typically better at pinpointing the magnitude and the duration of an effect, managers can easily tell whether it should be positive or negative. In many cases, simple marketing theory can be tremendously helpful. For example, if you find that your intended marketing performance variable increases when you increase prices, you may be skeptical about having the right measure at hand. Similarly, simple correlation analyses may help you to understand how variables work together and behave



**FIGURE 4** > Number of Facebook followers of Hillary Clinton and Donald Trump



together. Again, a first face validity control may be helpful to screen out suspicious effects that go against marketing theory or your own experience.

> **Use dashboards based on econometric modeling** ✕

Finally, marketers are well advised to develop company- or brand-specific dashboards, which should be based on econometric models. Relying on established procedures and the help of econometric methods, such as vector autoregressive models, not only may help managers to identify and track key performance variables but may also be helpful to understand which data sources bring meaningful information to a decision-maker's table: As suggested by Pauwels in his 2014 book, "It's not the size of your data, but what you do with it."

Using such approaches to continuously monitoring your company's data environment and controlling the reliability and validity of available data for decision-making will finally enable you not to be blinded or overwhelmed by the richness of data available to you. In other words: To avoid being lured into the dark side of decision-making, shed some light on your data, and think critically about its utility. Then your marketing will be great again – seriously great! ✕



FURTHER READING

**Kübler, R. V.; Colicev, A.; & Pauwels, K. H. (2020):** "Social Media's Impact on the Consumer Mindset: When to Use Which Sentiment Extraction Tool", *Journal of Interactive Marketing*, Vol. 50, 136-155.

**Kübler, R.; Pauwels, K.; & Manke, K. (2020):** "How Social Media drove the 2016 US Presidential Election: a longitudinal topic and platform analysis". [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3661846](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3661846)

**Pauwels, K. (2014):** *It's Not the Size of the Data – it's how You Use it: Smarter Marketing with Analytics and Dashboards*. Amacom.