8-2013

Pundits: The confidence trick: Better confident than right?

Ben O. Smith
University of Nebraska at Omaha, bosmith@unomaha.edu

Jadrian Wooten

What Drives Demand for Pundits?

Ben Smith and Jadrian Wooten

We live in a world full of pundits, professionals hired to make predictions in various media outlets. Some pundits make predictions about finance, some about politics and some about sports. Often times, it appears they are not particularly accurate. In 2008, Jim Cramer famously endorsed Bear Stearns on his television show, Mad Money, one week before their collapse, which was in the early stages of the financial crisis [1]:

> *Bear Stearns is fine... Bear Stearns is not in trouble. Don't be silly... don't move your money.*

Cramer later attempted to justify his remarks by suggesting that being too negative would have resulted in a mass panic. In fact, a popular novice investment strategy is to invest in Cramer-endorsed stocks to see small overnight gains, and then short sell the stocks to capitalize on the surge in prices [2]. Sites like *Pundit Tracker* suggest that the most famous pundits are also among some of the worst performing pundits.

Whenever pundits are extremely accurate, the public often considers them an outlier among pundits, as was the case of Nate Silver's accurate prediction of the 2012 presidential election. In the 2008 presidential election, statistician Nate Silver was able to correctly predict the winner for 49 of the 50 states, as well as the winner for all 35 senatorial races. While often considered mild compared to most pundits, even Nate Silver has the occasional bravado of a confident pundit. After being criticized by an MSNBC pundit about his prediction that

Barack Obama had a 75% change of winning re-election, Silver responded with a charitable wager of $1,000 to benefit the Red Cross, insisting his prediction was accurate [3]. Ironically, Silver's confident wager was published publicly on his Twitter account (@fivethirtyeight):

> *If you think it's a toss–up, let's bet. If Obama wins, you donate $1,000 to the American Red Cross. If Romney wins, I do. Deal?*

If pundits are not held to high stands for accuracy, what is it that makes pundits so popular with the public? This was the original question we set out to answer. We were inspired by a basic idea from psychology: people do not like uncertainty. Uncertain situations make people uncomfortable, and people are willing to give something up to avoid uncertainty. If this holds true for media pundits, the public would allow some inaccuracy for an increased sense of certainty. This forms our hypothesis, when tested statistically, we should see increased popularity from confident pundits, when controlling for all other factors.

The Process

Sadly, stocks do not have a terminal date. Therefore a pundit can always claim we simply did not wait long enough for the stock to go up or down as they predicted. While Jim Cramer may have inspired out investigation, we turn to the world of sports pundits to test our hypothesis. When a pundit makes a prediction about a game, they are either right or wrong and we know exactly when we will record an outcome either way. This makes sports the ideal subject matter for testing the popularity of pundits.

Once the subject matter was resolved, we focused on an ideal datasource. We could watch a whole lot of television and read a lot of newspapers to collect every prediction we see. However, this poses three problems:

1. Sites like *Pundit Tracker* already record pundit predictions on a regular basis
2. Regardless of how much television we watch, we will miss a large number of predictions while we sleep and go to work, which means our dataset will be relatively small

3. This will only examine predictions made by *professional* pundits from a limited number of media outlets

Specifically looking at the last issue, a reasonable person might argue that, in fact, the public does not want confident pundits. Perhaps, the news/entertainment industry simply *thinks* we want confidence and therefore, we do not have a choice. What we need is a datasource where both professionals and amateurs alike make predictions on games: this is one of the reasons we use Twitter.

Twitter has a number of interesting properties that make it ideal for this type of analysis. First, the cost of providing content to one more subscriber for the tweet producer is costless. Pundit wish to serve as many people as possible, which means that the number of followers is entirely up to the individuals who decide to follow each pundit. This allows us to observe the preferences of the people following the pundit and not the preferences of the pundit themselves.

Twitter is also fairly representative of the general population [4]. While the network is more urban and younger than the U.S. population, it matches educational attainment and income rather well. Because Twitter is a real market, and not an artificial environment, it avoids a set of observer bias effects common to surveys and experiments.

Further, both professionals and amateurs alike can make predictions, and Twitter conveniently separates those groups for all of their users. Because famous people commonly have a problem where they might be impersonated on the web, Twitter has created the "verified account." Twitter checks to make sure that the claimed user, is in fact, the user. Anyone that is famous because they are on TV, write for a newspaper or on the radio is going to have this special flag.

Additionally, each user has a biography section where they have the opportunity to describe themselves. We can examine this information for claims of sports expertise[1]. In the

[1]There are many ways to claim expertise, we use a technique called "regular expression" to capture as

final analysis, someone who claims to be a sports expert, and is not verified, we classify as *amateur* pundits. Someone claiming to be a sports expert with a verified account is classified as a *professional* pundit.

The next step is to collect the data from Twitter. Twitter allows users to programmatically "watch" a set of words. After you register a set of words, a tweet is sent in real time to your computer if the words occur in the tweet. Starting one week prior the 2012 baseball playoffs, we started collecting every tweet containing any of the team names, nicknames, or city names. This resulted in over a billion items, most of which were not predictions, but that is okay because only predictions matched our regular expressions[2].

If you think about sentence structure, only a few words in each sentence matter in terms of determining the meaning of the entire sentence. We only need to be able to identify these key elements of the sentence. One method of doing this is to build a large table of regular expressions, a technique where a number of phrases can be generalized. A simplified example of a regular expression would be:

```
\b(Bears)(?:(?!(\b((not)|(won[’]t))\b)).)*\b((destroy)|(annihilate))\b.+\b(Dogs)\b
```

Which would match any phrase that says that the Bears will destroy or annihilate the Dogs. But, unlike a normal search, this structure allows for variations on that theme and will still be picked up by the regular expression (e.g. "the bears will totally destroy the dogs"); however, it specifically disallows variations with the opposite meaning ("the bears will not destroy the dogs"). Making a large table of these expressions results in us picking up all[3] true predictions.

We then determine the strength of the each of these predictions by using the work of Chklovski and Pantel [5], who ranked the strength of words. This allowed us to mark a

many of these as possible. Regular expression is discussed in more detail a little later in the document.

[2] Regular-Expressions.info is a good site to learn regular expression syntax

[3] To the best of our knowledge, we continued to add expressions until our list matched all the forms of predictions we were seeing on Twitter.

prediction like "the bears will destroy the dogs" as confident, while marking a prediction like "the bears will beat the dogs" as not confident.

Using regular expression, along with the strength of phrases, we thinned our billion plus mass of collected tweets down to 1.6 million predictions with information about followers, confidence, accuracy, verified status and many other aspects of the tweeting account. At this point, the problem becomes a simple regression problem.

Using a technique called a Box-Cox parameter transformation estimation, we found that a log model best fits the data. We then use both Ordinary Least Squares (OLS) and iterated Generalized Methods of Moments (iGMM) to estimate the importance of accuracy and confidence while controlling for other observable factors, such as age of the account, engagement in the Twitter community and the number of tweets per year.

The Findings

The public appears to heavily value confidence and places a much smaller, although still positive, emphasis on accuracy. Among professional pundits, predicting every baseball playoff game correctly would only result in $3\frac{1}{2}\%$ increase popularity while being consistently confident would result in an almost 17% increase in popularity. Ideally, we would all appreciate confident pundits with perfect accuracy, but realistically it is difficult to be perfectly accurate. Pundits can, however, control their confidence, which results in many pundits making strong predictions regardless of the probability of their statement being accurate. By focusing on confidence, pundits on Twitter result in higher follower counts, which is a rough gauge of consumer demand.

Imagine two roulette tables side–by–side with two equally competent gamblers placing bets for a red outcome. While each has approximately a 50% chance of being correct[4], people will gravitate to the more animated player. Gamblers trying to confidently predict the red

[4]Because roulette tables have green slots for 0 and 00, the true probability of landing on a red/black slot is actually 48.6%.

landing will usually gather a crowd, while the quieter gambler will play by himself.

Amateur pundits experience a similar story with an increase of 7% for perfect accuracy compared to almost 20% for being consistently confident. The ability for amateur pundits, especially for those operating as aspiring bloggers, to gain followers is pivotal to their success. This result is particularly important when compared to the professional pundits. Because the two results are not particularly dissimilar, we can reject the idea that it is the media who wants confident pundits. Amateurs on Twitter have no intermediary (like a network or newspaper) so they directly serve the public's desires.

Regardless of the punditry status, one thing is certain, the most popular pundits may not be right all the time, but they tend to make their predictions more confidently than the other pundits.

Implications

Pundits are confident because it is what the public demands. At first, people just assume that pundits are confident because the networks are asking them to operate that way, but the results are similar for amateur pundits. These pundits have no intermediary dictating their confidence level, and yet, they attract a larger following when being confident. So while ESPN or CNBC have ultimate discretion over which pundits host their shows, the demand for pundits is actually derived from the viewing audience.

When you watch a pundit on television, remember that their job is to maximize eyeballs, not accuracy. Their employer is in the business of selling advertising, which means the networks will choose pundits that provide the most advertising revenue: which is not necessarily the most accurate pundit – it could just be the most confident.

References

[1] H. Post. (Mar. 2008). Cramer called bear stearns "fine" just a week ago, [Online]. Available: http://www.huffingtonpost.com/2008/03/17/cramer-called-bear-stearn_n_91878.html.

[2] J. Engelberg, C. Sasseville, and J. Williams, "Market madness? the case of mad money," *Management Science*,

[3] M. Sullivan. (Nov. 2012). Under attack, nate silver picks the wrong defense, [Online]. Available: http://publiceditor.blogs.nytimes.com/2012/11/01/under-attack-nate-silver-picks-the-wrong-defense/.

[4] (Aug. 2012). Pew internet: trend data (adults), [Online]. Available: http://pewinternet.org/Static-Pages/Trend-Data-(Adults)/Usage-Over-Time.aspx.

[5] T Chklovski and P Pantel, "Verbocean: Mining the web for fine-grained semantic verb relations," *Proceedings of Conference on Empirical Methods in Natural Language Processing*, vol. 4, pp. 33–40, 2004.

Figures

Table 1: A Sample of Twitter Verified Accounts During the 2012 Major League Baseball Playoffs

Account	Predictions	Confidence	Accuracy
ESPN	27	52%	41%
SportsCenter	32	44%	31%
Baseball Tonight	39	44%	28%
Mike Greenberg	13	53%	15%
Lou Holtz	9	11%	33%
Chris Rose	11	36%	36%

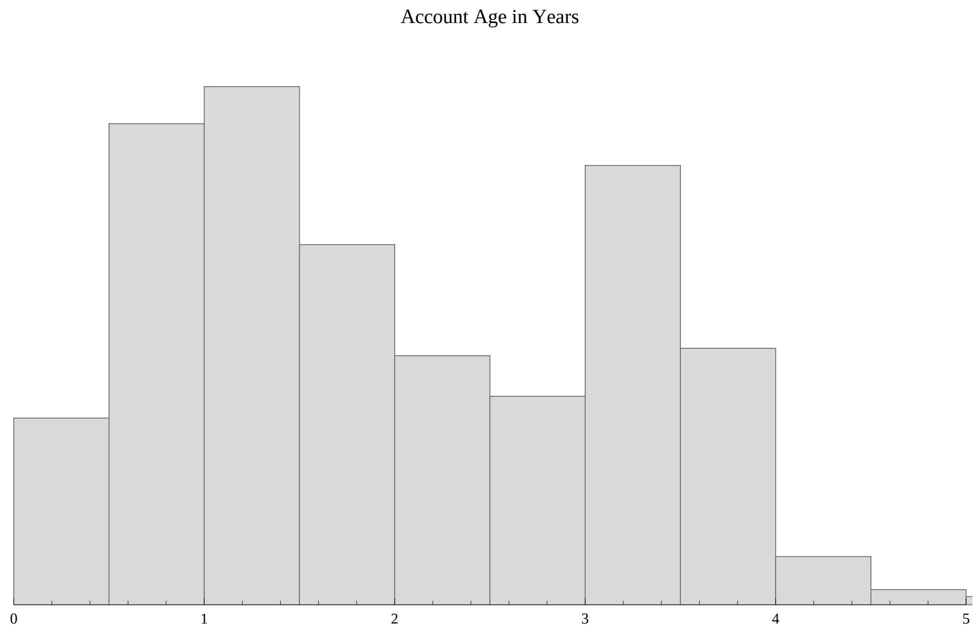

Figure 1: Distribution of account age in years. Outliers (less than 1%) have been removed for readability. While few accounts are still active since Twitter's debut, accounts created approximately three years ago represent the "boom" in Twitter's popularity.

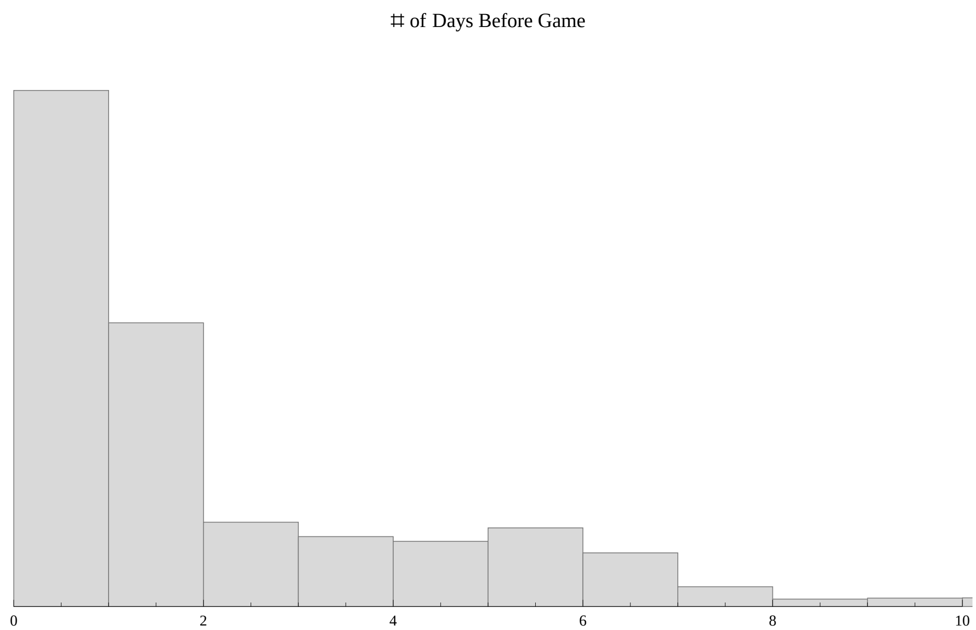

Figure 2: Difference between prediction time and game start time in days. Outliers (less than 1%) have been removed for readability. The vast majority of predictions are made the day before the game, which might represent Twitter users collecting information before making their prediction.