Stochastic Models for Predicting US Elections

JW for Guesser

March 15, 2021

Contents

1 Models

We have implemented the models described in [1] and [2].

1.1 Taleb Idea

Taleb is modelizing the decision as a continuous, hidden (dual) random process X_t ($dX_t = \sigma^2 X_t dt + \sigma dW_t$ with variance and drift) mapped onto a decision outcome Y_t via a decision outcome.

Intuition is that the real decision variable exists within an other space but is forced to take a binary decision for elections via continuous decision process as shown in Fig.1.

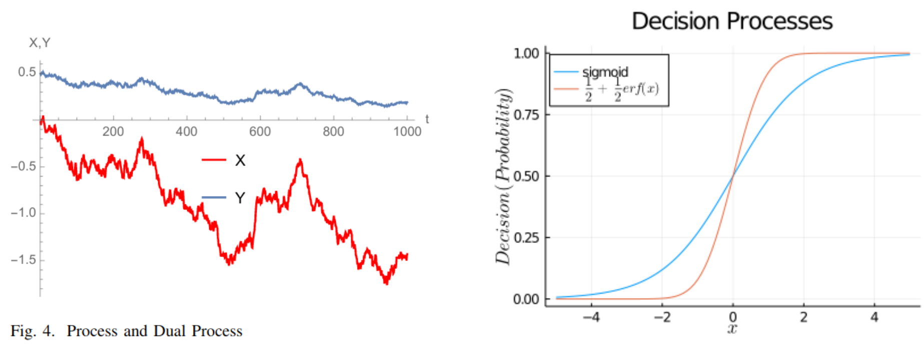

Figure 1: On the left, underlying decision process X_t mapped onto Y_t via a decision function. On the right, decision processes mapping between $X_t \rightarrow Y_t$.

1.1.1 Analytical Model: tal

Mathematical statistical finance knows how to analytically solve such problems: for s the polls annualised volatility (Y_t), the process (X_t) standard deviation reads

$$\sigma(s) = \frac{\sqrt{log(2\pi s^2 e^{2erf^{-1}(2Y_0-1)^2}) + 1}}{\sqrt{2}\sqrt{T-t_0}} \tag{1}$$

and the binarised option value or for a party, the probability to win the election at $T > t_0$ given the current polls fraction Y_0 at t_0 reads

$$B(Y_0, \sigma, T, t_0) = \frac{1}{2} erfc\left(\frac{l - erf^{-1}(2Y_0 - 1)e^{\sigma^2(T-t_0)}}{\sqrt{e^{2\sigma^2(T-t_0)} - 1}}\right), \tag{2}$$

with

- $l = 0.0$ the election threshold
- erf, erf^{-1} standard functions

- s, σ defined above
- T, t_0 election date and evaluation date resp., with $(T - t_0)$ computed as year's fraction

1.1.2 Practical implementation

The above equation is time independent, not depending on $T - t_0$. Furthermore, $l = 0.0$ is unlike what is written in the article but is justified by the results and the interpretation of $X_t = l$ as the quantity whereas a quantitative jumps in Y occurs.

1.2 Fry & Burke

This article is posterior to Taleb's and somewhat simpler. It is based on the following ideas and analogies:

Financial models built on stochastic processes representing market price fluctuations.

- analogy I: market price fluctuations ⇔ votation intentions
- analogy II: % vote intentions ⇔ price of binary options (because later also related to activation percentage)

analogy I is about underlying process modelling, analogy II is about a suitable model

1.2.1 Model 1: fb1

For $P_t = P_{DEM,t}/P_{GOP,t}$, the vote fraction at time t, Binary Call option price corresponds to a probability of DEM winning ¿ 0.5:

$$B_{COP}(P_t, r, \sigma, T, t) = P_t \cdot e^{r - \frac{\sigma^2}{2}(T-t)} e^{-\sigma\sqrt{T-t}\Phi^{-1}(\frac{1}{2}e^(r(T-t)))}, \tag{3}$$

with

- r is the risk free interest rate set at 1.5% (Dowd et al. 2019)
- T, t the "exercise" and "creation" times in years or fraction of
- σ the polls volatility, computed as discussed below
- $\Phi(x)$ is the cumulated distribution fct for a distribution $\mathcal{N}(0., 1.)$

1.2.2 Model 2: fb2

For $P_t = P_{DEM,t}/P_{GOP,t}$, the vote fraction at time t, the Estimated Probability for DEM to win:

$$EP(P_t, r, \sigma, T, t)_{DEM} = e^{-r(T-t)}\Phi\left(\frac{ln(P_t) + r - \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}\right), \tag{4}$$

with

- r is the risk free interest rate set at 1.5% (Dowd et al. 2019)
- T, t the "exercise" and "creation" times in years or fraction of
- σ the polls volatility, computed as discussed below

1.3 Volatility

In practice, both potential implementations below lead to similar results.

1.3.1 Financial evaluation

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=0}^{n}(u_i - \bar{u})^2}, \tag{5}$$

with

- $u_i = \frac{ln(P_i)}{ln(P_{i-1})}$
- $\bar{u}$ the average of u_i's

1.3.2 Fry & Burke

MSE after fit:

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)},$$

with

- y_i's the data points
- $\hat{y}_i$ the data points after fit

2 Model Validation

The above models were tested on synthetic data, controlling the model behaviour in extreme situations (P_t = 1.0, σ = 0.0, ...) and written in two programming languages. We are just reporting this analysis here in Fig.2. Our figure is slightly different than the original results and we need to check this result!

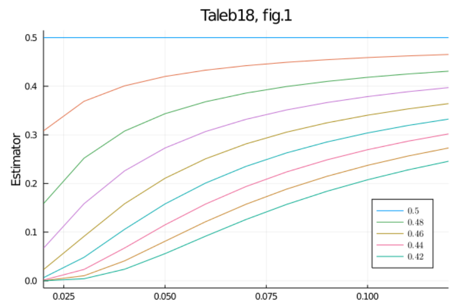

Figure 2: Reproducing [1] fig.1 results with our own code.

3 Methodology and data for models evaluation

3.1 Data

Optimus data, with historical elections polls and results at the national and state level, from 1992 each 4 years to 2016.

3.2 Methodology

Our approach is as follows:

1. run independently models on state and national level elections

2. Benchmark and discuss the models according to 2 typical machine learning scores: **ROC AUC** and **Log Loss** as well as the **De Brier** score.

3. When models are sensitive to the $(T - t_0)$ or number of days to vote, they freeze to the polls number just before the election.
   On the other hand, votes are taking place just after the lates polls, but there is actually a distance caused by biases, behaviours, sub samples, etc...
   As a practical idea to fix it, we are considering 2 additional quantities that we can calibrate on the votation data:

   - ΔT: a synthetic time shift
   - ΔV: a synthetic volatility difference

4 Results

4.1 National Elections

in Fig.3, we show our results for 3 of the last elections. We are running similar evaluations and visualisations for all the data since 1992. This allows to get an intuition of model's behaviours.

polls	fb1	fb2	tal
0.575	0.596	0.603	0.575

Table 1: Accuracy found over all the predictions at the **PRESIDENTIAL** elections.

polls	fb1	fb2	tal
0.728	0.805	0.807	0.750

Table 2: Accuracy found over all the predictions made at the **STATE** elections.

4.2 State Elections

Similarly, in Fig.4, we show similar results for each states.

Unlike the national elections for which we are 8 data sample, we have over 200 local experiments where polls where available. This corresponds to ~ 3000 data points to evaluate our models and benchmark from.

4.3 Scores and Distributions

4.3.1 Prediction Accuracy

Because models are based on polls, it is surprising to observe the accuracy getting improved using statistical models... We summarise this in Table.1 and Table.2.

In our opinion, these improvements are likely to be "artefacts" coming from the behaviour of the models around the threshold value $p = 0.5$. This should be further examined but could be accepted for now[1].

4.3.2 De Brier Scores

De Brier scores are averaged as the quadratic distances between the predictions and outcomes.

We show our results for the states and national levels in Table.4 and Table.3 resp. Furthermore, we show the distributions in Fig.5.

We can conclude that Taleb's model has the best scores for the De Brier score, because this is what our larger data set, from the states elections is telling us.

[1] At the state level, the average is sampled with $\sim$ 5k points, meaning an uncertainty lower than 1%, or 0.01 in the Table.2 and thus significant...

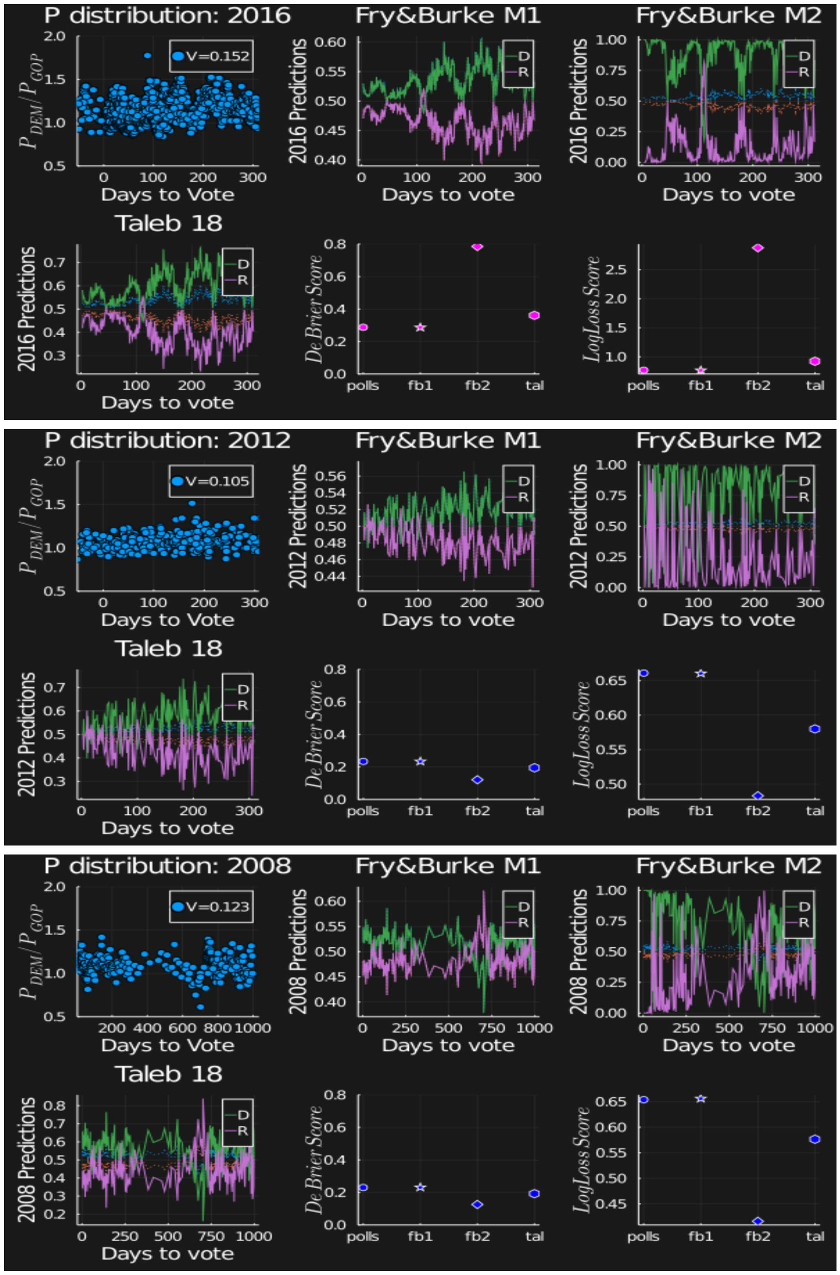

Figure 3: Results summarised for the last **NATIONAL** level elections: We show P_t's distributions and volabilities, then the 3 models: fb1 & fb2 as well as the tal for respectively Fry & Burke 1 and 2 and Taleb's model. The De Brie and Log Loss scores are measured as well.

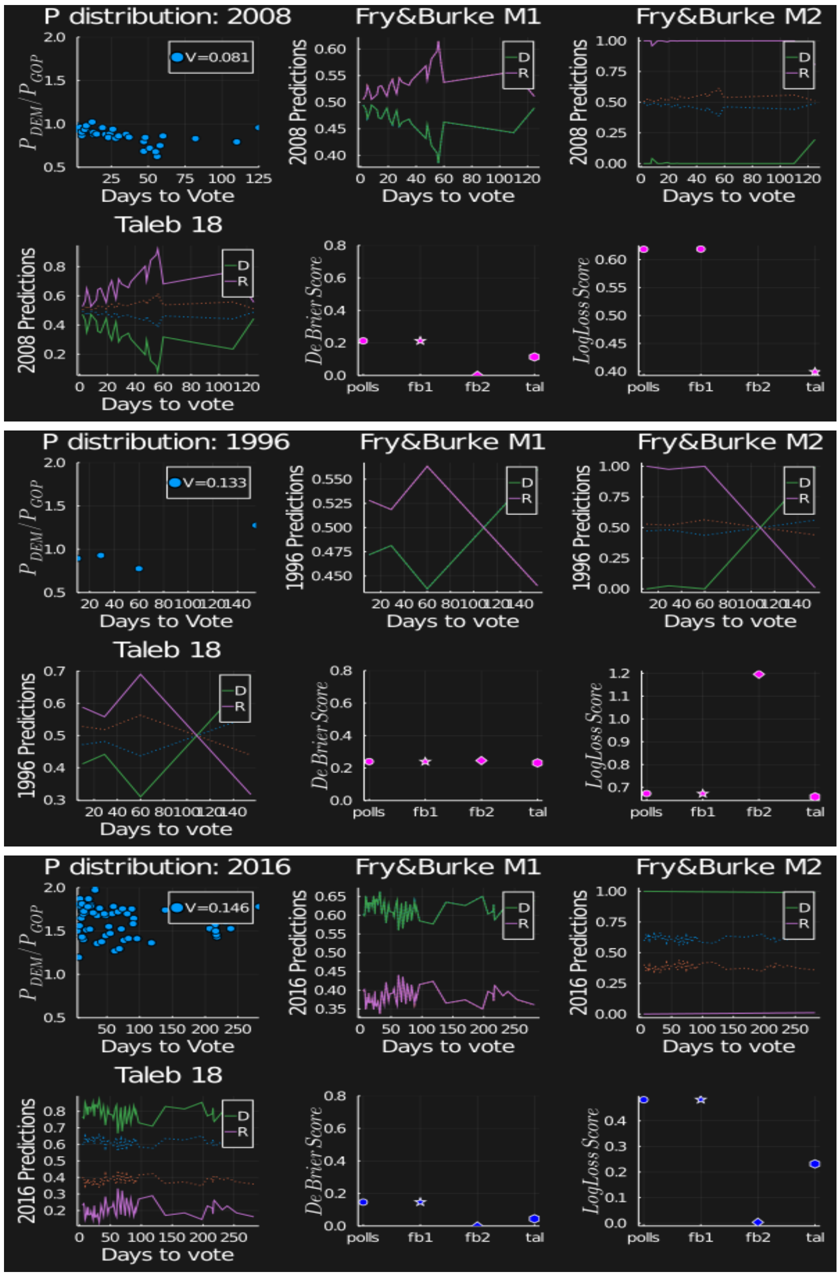

Figure 4: Results summarised for the last **STATE** level elections: for from top to bottom Georgia 2008, Texas 1996 and New York 2016, we show P_t's distributions and volabilities, then the 3 models: fb1 & fb2 as well as the tal for respectively Fry & Burke 1 and 2 and Taleb's model. The De Brie and Log Loss scores are measured as well.

polls	fb1	fb2	tal
0.241	0.241	0.325	0.232

Table 3: De Brier scores found over all the predictions at the **PRESIDENTIAL** elections.

polls	fb1	fb2	tal
0.215	0.214	0.162	0.134

Table 4: De Brier scores found over all the predictions made at the **STATE** elections.

We think that the elections of 2016 impacted the scores for presidential scores. The distribution is interesting because it tells us that fb1 is a more bullish model and Taleb a more stable one.

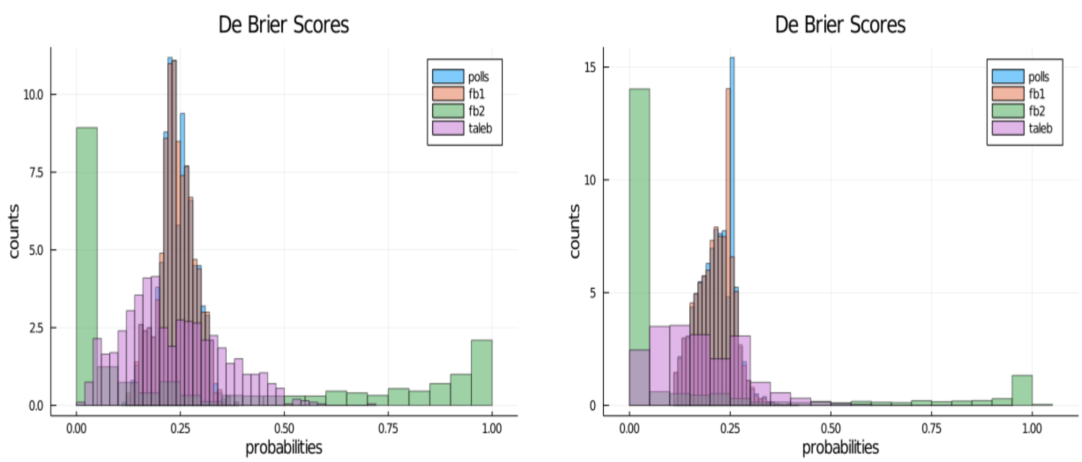

Figure 5: Distributions for De Brier score. On the left for all the presidential elections since 1992 and on the right for the states elections.

4.3.3 Log Loss Scores

Log loss is a standard error evaluation that we don't define here, as the others, the lower the better.

The tables Table.5 and Table.6 summarise our results for the log loss error. Taleb shows the best behaviour.

polls	fb1	fb2	tal
0.675	0.675	1.188	0.654

Table 5: Log Loss scores found over all the predictions at the **PRESIDENTIAL** elections.

polls	fb1	fb2	tal
0.622	0.621	0.978	0.524

Table 6: Log Loss scores found over all the predictions made at the **STATE** elections.

4.3.4 ROC AUC

Finally, we tried to consider ROC AUC, which is an other way to represent model self evaluation. Here the higher the better!

Results are summarise in tables Table.7 and Table.8 and also in Fig.6. Overall, Taleb's model emerges on the top.

5 Optimisation

5.1 Methodology

Using the benchmarks and models presented above, we are introducing 2 synthetic parameters used to approximate the biases within the data as well as the typical "distance" between the last polls and the real votations:

- ΔT: synthetic time shift
- ΔV: synthetic volatility addition

5.2 Parameters exploration

We bump over a range of parameters for ΔT and ΔV. In Fig.7, we found a best combination for the parameters:

- $\Delta T = 800$
- $\Delta V = 0$.

polls	fb1	fb2	tal
0.575	0.596	0.605	0.632

Table 7: ROC AUC scores found over all the predictions at the **PRESIDENTIAL** elections.

polls	fb1	fb2	tal
0.728	0.805	0.876	0.871

Table 8: ROC AUC scores found over all the predictions made at the **STATE** elections.

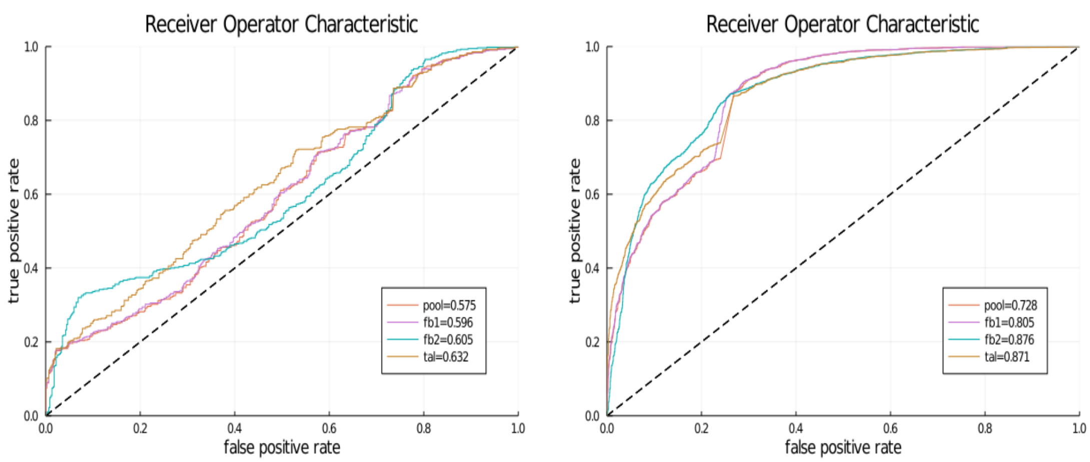

Figure 6: ROC AUC measurements. On the left for all the presidential elections since 1992 and on the right for the states elections.

A few remarks are in order:

1. With these, the model fb2 "beats" tal
2. We found similar results using **PRESIDENTIAL** votes data only

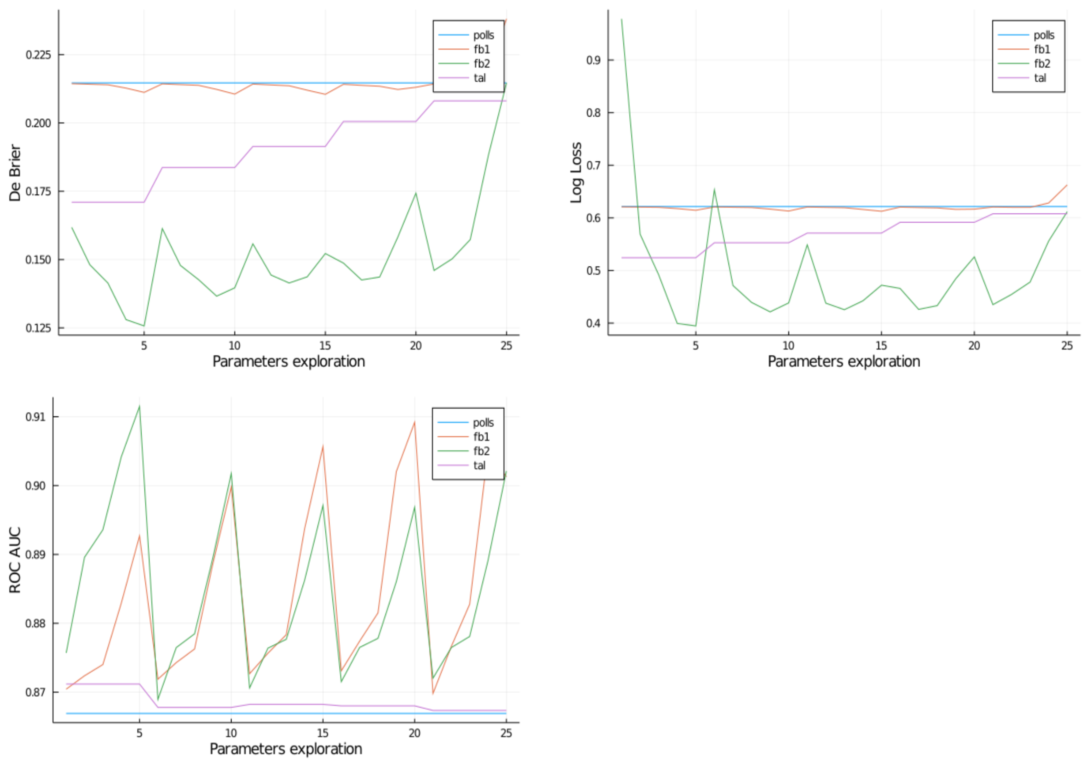

Figure 7: Optimisation: Looking for instance to both De Brier and ROC AUC scores on all the **STATES** level elections, we find an obvious candidate for a best model fb2 with the 5th set of parameters. Above Figs., the lower the better. Below, the higher the better. We conclude that the parameters labelled 5 are the best.

6 Discussion and Conclusion

1. We have presented metrics and a methodology to benchmarks statistical, financial on historical data.
2. We have proven that statistical financial models are a clear, very measurable improvement over raw polls.
3. We show how we improve the model by introducing additional, synthetic volatilities and time shift and how the best parameters can be extracted.
4. We found that the Taleb's model is performing better than the simpler Fry&Burke models
5. We provide with our own optimization of the financial models in using additional parameters. With that we found a very competitive version of the model fb2.

As a result we can propose original models for the election predictions:

- optimised fb2:
- tal

Their Prediction are below shown in Fig.8.

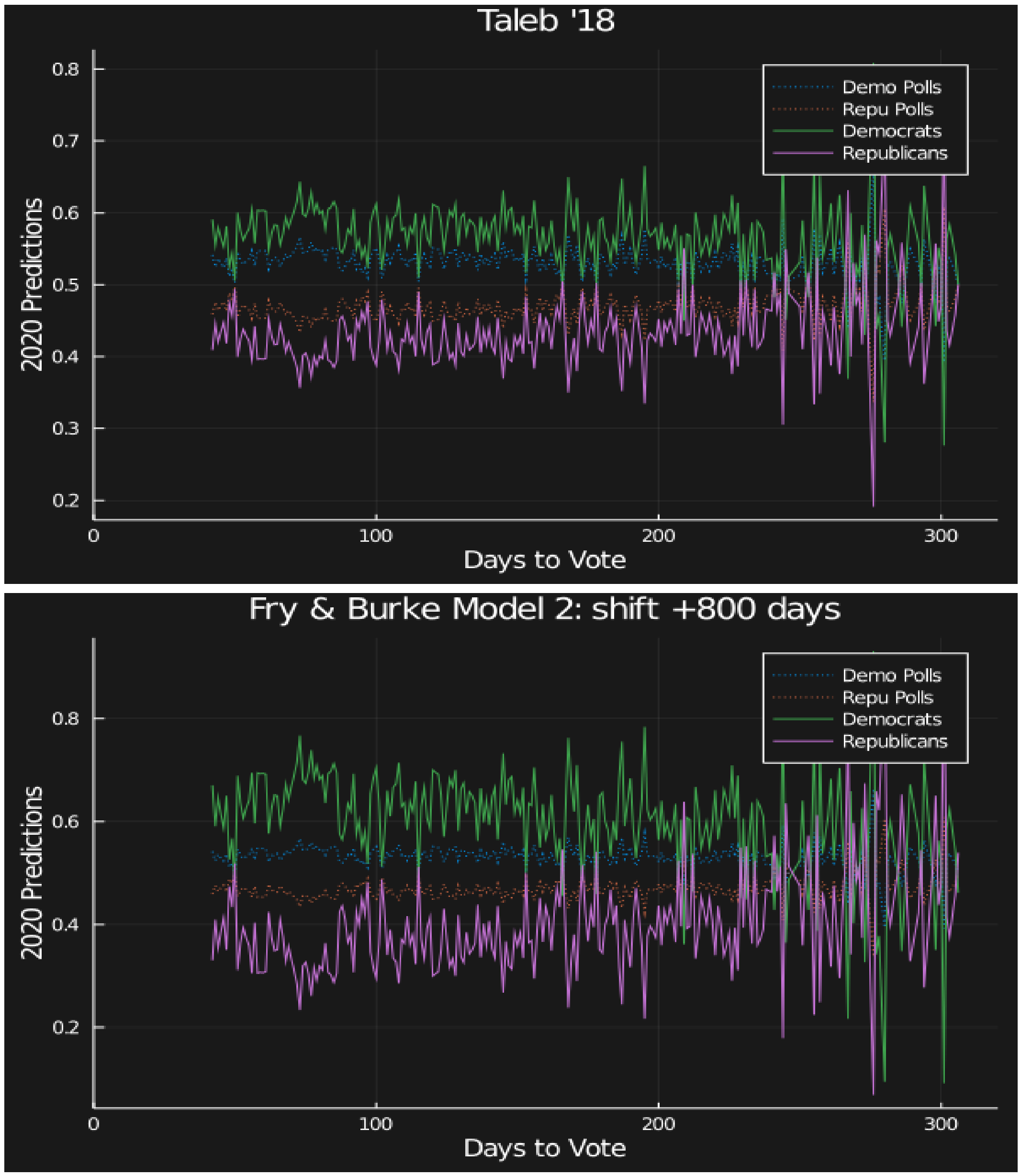

Figure 8: Prediction 2020: Our best Predictors in 2020: the Taleb and our optimised version of the Fry&Burke.

References

[1] N.N.Taleb, *Election Predictions as Martingales: An ArbitrageApproach*, 2018

[2] John Fry & Matt Burke, *An options-pricing approach to election prediction.*, 2020