POLITICS SYMPOSIUM

Forecasting the 2022 French Presidential Election

Designing Prediction Markets to Forecast Multi-Stage Elections: The 2022 French Presidential Election

Joyce E. Berg, University of Iowa, USA
Thomas S. Gruca, University of Iowa, USA
Thomas A. Rietz, University of Iowa, USA

I. Introduction

The Iowa Electronic Market (IEM) has successfully forecast elections around the world (Berg, et al. 2008). Figure 1 shows a history of election-eve accuracy, comparing forecasts to actual vote shares.¹ The absolute prediction errors for non-US elections averages 2.12%.² Nearly all these forecasts involved single-stage elections and had a local trader base.

Figure 1. Accuracy of IEM markets for US Presidential elections, other US elections, and non-US elections

Many jurisdictions hold two-stage elections in which (typically, many) candidates run in the first round. In majority runoff (MR) elections, any candidate who receives an outright majority in the first-round wins. If no candidate does, the two with the most votes participate in a second-stage (runoff) election. Multi-stage elections create special challenges for designing and running election prediction markets.

The 2022 French Presidential Election is an important and highly contested MR election. We show how to design two-stage election prediction markets in general and implement a prediction market for the 2022 French election. We discuss our design and what we learned

This is a “preproof” accepted article for PS: Political Science & Politics. This version may be subject to change during the production process.

DOI: 10.1017/S104909652200052X
from the unsuccessful market launch. We make two contributions to the literature: (1) We develop a market design with a reasonable number of contracts that forecasts probabilities of greatest interest in two-stage elections. This design also creates some interesting inter-contract pricing relationships that we can use to test market efficiency. (2) We provide more evidence on the importance of having an informed trader base for successful markets.

II. Outcomes in Two-Stage Elections

Two-stage elections have a complex outcome space. It depends on how many candidates run, what combination qualifies for the runoff (if any), and who wins given the specific runoff candidate combination. The number of possible outcomes grows rapidly with the number of first-round candidates: candidates can win outright, any two can qualify for a runoff, and each runoff candidate could win.

Typically, several serious candidates contend against many hopefuls. For example, in October 28-20, 2021, polling for the 2022 French Presidential Election, Harris Interactive asked about voting intentions for 18 potential candidates (Lévy, Bartoli and Gautier 2021), with five candidates receiving 10% or more of poll responses for the first-round election. Suppose 18 candidates ran in the first round. While unlikely, each candidate could win in the first round. In addition, each combination of two might appear in a runoff. For every runoff combination, there are two possible outcomes. This leads to:

\[
\frac{\text{Poss. Oultright Winners}}{18} + \frac{\text{Poss. Runoff Combinations}}{2\times16!} \times \frac{\text{Poss. Outcomes for each Runoff}}{2} = 306
\]  

possible outcomes, most of which have near zero probability.

III. IEM Markets and Contract Design

Open to traders world-wide, IEM markets are real-money, internet-based futures markets where self-selected traders determine prices that forecast election outcomes. The IEM is a purely order driven market: traders place bids and asks for contracts and prices are set when other traders accept outstanding bids or asks. There is no scoring system, no market maker, and no clearing mechanism beyond these mutually agreed upon trades. Prices change when price-determining traders’ beliefs change. Thus, the IEM aggregates information in a manner significantly different from polls, poll averages, or statistical models. Here, we use “winner-takes-all” (WTA) contracts to forecast probabilities of outcomes.

Typically, IEM WTA election markets include contracts for major candidates (e.g., Macron and Le Pen) that each pay $1 if the associated candidate wins and $0 otherwise, and a “rest-of-field” contract that pays $1 if any other candidate wins. Traders can trade “bundles” consisting of one of each contract for $1 directly with the exchange. Using all relevant available information, traders form beliefs about outcome probabilities. Then, they can trade with each other, buying contracts when prices are lower than their estimated probabilities and selling contracts when prices are higher. Thus, shifting expectations change supply, demand, and contract prices.

We could offer two-stage election WTA contracts to simply forecast the overall winner. However, this would not allow us to answer many interesting questions such as: Will anyone win outright in Round 1? If no one does, who will make the runoff? How might different combinations of candidates fare in a runoff?
To run a manageable market, the contract set must span the interesting parts of the outcome space. Three markets effectively do so: a market focused on outright winners, a market on possible runoff combinations, and a market on runoff winners. We show how to construct each market so that contracts span the outcome space of interest, how to minimize the number of contracts, and how to add new contracts as new candidates emerge. While these markets are more complex than typical IEM markets, they are only slightly more complex than combined nomination/conditional vote share markets successfully run for the 1996 US Presidential Election (Berg and Rietz 2003).

A “Round 1” (R1) market forecasts first-round outcomes by listing one contract per major candidate. Define the contract \( R1A \) to pay $1 if candidate \( A \) wins outright in the first round and $0 otherwise. Its price should be:

\[
P_{R1A} = p_{A\text{ wins outright}} \cdot \$1 + (1 - p_{A\text{ wins outright}}) \cdot \$0,
\]

where \( p_i \) is the contract’s price, \( p_i \) is the (risk neutral) probability of the outcome, and “wins outright” means that \( A \) receives >50% of the vote in Round 1. To minimize the number of contracts, list: (1) one contract for each major candidate, (2) a \( R1Unnamed \) contract to cover minor candidates that pays $1 if an unnamed candidate wins outright, and (3) a \( R1Runoff \) contract that pays $1 if no candidate wins outright necessitating a runoff. To add a new named candidate, \( C \), split the \( R1Unnamed \) contract into a new \( R1C \) contract and a new \( R1Unnamed \) contract consisting of remaining unnamed candidates. R1 market prices forecast the probabilities of: a runoff (\( P_{R1Runoff} \)), each named candidate winning outright (\( P_{R1A} \), etc.), and an unnamed candidate winning outright (\( P_{R1Unnamed} \)).

A “Runoff” (R) market forecasts who (if any) candidates qualify for a runoff. Define the contract \( R_AxB \) to pay $1 if candidates \( A \) and \( B \) qualify for the runoff and $0 otherwise. Its price should equal the probability that \( A \) and \( B \) are in a runoff. To minimize the number of contracts and stay consistent with the R1 market: (1) list one contract for each major candidate combination where order does not matter (\( P_{A \text{ and } B \text{ in Runoff}} = P_{B \text{ and } A \text{ in Runoff}} \)), (2) list one contract for each major candidate combined with Unnamed (\( R AxU \), etc.), (3) an Unnamed versus Unnamed contract (\( R UxU \)) and (3) a \( R \text{ None} \) contract that pays $1 if any candidate wins outright in Round 1. In theory, \( P_{R1Runoff} = 1 - P_{R \text{ None}} \), a testable relationship for inter-market efficiency. To add a new candidate, \( C \), split (1) \( R AxU \) into \( R AxC \) and a new \( R AxU \) that represents \( A \) with remaining unnamed candidates, (2) similarly split all remaining named candidates with unnamed candidate combinations, and (3) split \( R UxU \) into \( R CxU \) and a new \( R UxU \) reflecting remaining unnamed candidates. Again, listing only major candidates minimizes the number of contracts.

A “Round 2” (R2) market forecasts which candidate wins the runoff for any possible combination. Define the contract \( R2_ADB \) to pay $1 if candidate \( A \) defeats candidate \( B \) in the runoff and $0 otherwise. Its price should equal (1) the probability that \( A \) and \( B \) are in the runoff times (2) the probability that \( A \) defeats \( B \). In theory, \( P_{R2_ADB} < P_{R AxB} \), which becomes another testable relationship for inter-market efficiency. To minimize the number of contracts and stay consistent with the R market list: (1) one contract for each major candidate combination with one candidate winning (the probability the other candidate wins can be inferred), (2) one contract for each major candidate winning against an unnamed candidate in the runoff (the probability the unnamed candidate wins can be inferred), and (3) one contract (\( R2 \text{ Else} \)) for any other outcome (no runoff, runoffs between unnamed candidates, or first named candidates lose to second named candidates). This creates another testable intermarket efficiency relationship:

\[
P_{R2 \text{ Else}} < P_{R UnnUnn} + P_{R \text{ None}}.
\]

To add a new candidate, \( C \), split (1) \( R2 ADU \) into \( R2 ADC \) and a
new $R2_{ADU}$ that represents $A$ defeating remaining unnamed candidates, (2) similarly split all remaining named candidates with unnamed candidate combinations, and (3) split $R2_{Else}$ into $R2_{CDU}$ and a new $R2_{Else}$ reflecting remaining other possibilities.

IV. Numbers of Candidates, Outcomes, and Contracts

The R1 market requires $N+2$ contracts for $N$ major candidates. Each additional named candidate requires one additional contract. The R market requires $\frac{(N+1)!}{2 \times (N-1)!} + 2 = \frac{4 + (N+1) \times N}{2}$ contracts for $N$ major candidates. Each additional major candidate requires $N+1$ additional contracts. The R2 market requires $\frac{(N+1)!}{2 \times (N-1)!} + 1 = \frac{2 + (N+1) \times N}{2}$ contracts for $N$ major candidates. Each additional major candidate requires $N+1$ additional contracts.

Figure 2 shows the relationship between the number of candidates running, the number of possible outcomes, and the number of contracts needed assuming 20% of the total candidates are major candidates with associated contracts. By focusing only on major candidates and using some contract prices to infer other probabilities, we significantly reduce the number of contracts needed to forecast two-stage elections.

Figure 2. The number of possible outright winners, runoff combinations, and runoff winners

Notes: The number of possible outright winners, runoff combinations, and runoff winners for different numbers of candidates running with the total contracts needed in R1, R and R2 prediction markets assuming 20% of the candidates running are major candidates with associated contracts.

Consider the IEM markets with Macron and Le Pen as the two major candidates in the 2022 French Presidential Election (prospectuses appear in the online appendix). These markets require 13 contracts. Table 1 shows the markets and contracts required with hypothetical prices that might exist (1) before Round 1 and (2) after Round 1 if Macron and Le Pen make the runoff. (Hypothetical prices allow us to discuss how to interpret prices later.)
Table 1: Markets and Contracts needed to forecast the 2022 French Presidential Election with Macro and Le Pen as the major candidates with hypothetical prices before Round 1 and after Round 1.

### Panel A: R1 Market

<table>
<thead>
<tr>
<th>Contracts</th>
<th>Pays $1 if</th>
<th>Hypothetical Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Before Round 1</td>
</tr>
<tr>
<td>R1Macron</td>
<td>Macron wins outright</td>
<td>$0.050</td>
</tr>
<tr>
<td>R1LePen</td>
<td>Le Pen wins outright</td>
<td>$0.040</td>
</tr>
<tr>
<td>R1Unnamed</td>
<td>An unnamed candidate wins outright</td>
<td>$0.010</td>
</tr>
<tr>
<td>R1Runoff</td>
<td>No one wins outright</td>
<td>$0.900</td>
</tr>
</tbody>
</table>

### Panel B: R Market

<table>
<thead>
<tr>
<th>Contracts</th>
<th>Pays $1 if runoff candidates are:</th>
<th>Hypothetical Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Before Round 1</td>
</tr>
<tr>
<td>R_MacxLeP</td>
<td>Macron and Le Pen</td>
<td>$0.600</td>
</tr>
<tr>
<td>R_MacxUnn</td>
<td>Macron and an Unnamed candidate</td>
<td>$0.150</td>
</tr>
<tr>
<td>R_LePxUnn</td>
<td>Le Pen and an Unnamed candidate</td>
<td>$0.090</td>
</tr>
<tr>
<td>R_UnnxUnn</td>
<td>Two Unnamed candidates</td>
<td>$0.060</td>
</tr>
<tr>
<td>R_None</td>
<td>No runoff</td>
<td>$0.100</td>
</tr>
</tbody>
</table>

### Panel C: R2 Market

<table>
<thead>
<tr>
<th>Contracts</th>
<th>Pays $1 if, in runoff,</th>
<th>Hypothetical Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Before Round 1</td>
</tr>
<tr>
<td>R2_MacDLeP</td>
<td>Macron defeats Le Pen</td>
<td>$0.400</td>
</tr>
<tr>
<td>R2_MacDUnn</td>
<td>Macron defeats an Unnamed candidate</td>
<td>$0.125</td>
</tr>
<tr>
<td>R2_LePDUnn</td>
<td>Le Pen defeats an Unnamed candidate</td>
<td>$0.075</td>
</tr>
<tr>
<td>R2_Else</td>
<td>All other outcomes</td>
<td>$0.400</td>
</tr>
</tbody>
</table>

V. How contract prices predict multi-stage election outcomes

Figure 3 shows how market prices forecast probabilities for each pathway for Candidate “A” winning or losing the election. R1 market prices forecast probabilities of A (or any other candidate) winning outright or whether there is a runoff. R market prices forecast whether a runoff will occur and, if so, who is in it. R2 market prices forecast who will win the runoff for each combination when combined with R market prices (with losing probabilities inferred from winning probabilities).
Consider hypothetical “before Round 1” prices from table 1. Some of the forecasts are obvious. The probabilities of outright wins are in Panel A, column 3. The probabilities of runoff combinations are in Panel B, column 3. Dividing each Panel B runoff combination by \( P_{R1Runoff} \) gives the combination probabilities conditional on a runoff. Prices here are consistent with inter-contract/market efficiency: 

\[
P_{R1Runoff} = 1 - P_{R_None}, \quad P_{R2_MacxLeP} < P_{R_MacxUnn}, \quad P_{R2_MacDUnn} < P_{R_MacxUnn}, \quad P_{R2_LePDunn} < P_{R_LePxUnn}, \quad \text{and} \quad P_{R2_Else} > P_{R_UnnxUnn} + P_{R_None}. \]

Inconsistencies would create arbitrage opportunities.

To find the overall winning probability, sum the outcomes where a candidate wins outright with those where a candidate defeats another in a runoff. This gives:

\[
P_{Macron} = P_{R1Macron} + P_{R2_MacDLeP} + P_{R2_MacDUnn},
\]

\[
= 0.050 + 0.400 + 0.125 = 0.575
\]

\[
P_{Le Pen} = P_{R1LeP} + (P_{R_MacxLeP} - P_{R2_MacDLeP}) + P_{R2_LePDunn},
\]

\[
= 0.040 + (0.600 - 0.400) + 0.075 = 0.315
\]

and
\[ p_{\text{Unnamed}} = \frac{P_{R1\text{Unnamed}}}{P_{R1\text{Runoff}}} + \frac{P_{R1\text{UnnnxUnn}}}{P_{R1\text{Runoff}}} + \frac{P_{R2\text{MacxUnn}} - P_{R2\text{MacDUnn}}}{P_{R2\text{LePDUnnn}} - P_{R2\text{LePDUnn}}} \].  

(5)

Find conditional runoff winning probabilities by dividing R2 prices by R prices:

\[ p_{\text{Macron}|\text{Macron & LePen}} = \frac{P_{R2\text{MacDLeP}}}{P_{R2\text{MacxLeP}}} = \frac{0.400}{0.600} = 0.6667 \], and

\[ p_{\text{LePen}|\text{Macron & LePen}} = 1 - \frac{P_{R2\text{MacDLeP}}}{P_{R2\text{MacxLeP}}} = 1 - \frac{0.400}{0.600} = 0.3333 \].  

(6)

(7)

Similar calculations give both Macron and Le Pen a 0.8333 probability of defeating an unnamed candidate in a runoff.

VI. The 2022 French Presidential election markets

In August 2021, the IEM opened the markets shown in table 1 without an established French trader base to see if the election could be forecast without such a trader base. As of January 25 (153 days since opening), only 7 of the 13 contracts have ever traded, so there are insufficient prices to form forecasts. In the R1 market, 155 total contracts had traded on 6 separate days. In the R market, 8 contracts traded on one day. In the R2 market, 105 contracts traded on 3 separate days. In comparison, volume averaged 262 contracts traded per day before Dole became the de facto nominee in the similarly complex combined nomination/conditional vote markets run for the 1996 US Presidential Election. Therefore, we speculate that having no French (informed) trader base results in low trading volume rather than market complexity. However, we cannot know for certain. Although trading may increase closer to the election, we cannot currently make forecasts.

VII. Discussion

Using a prediction market to forecast an election successfully requires several conditions. First, the event being predicted must be clearly and unambiguously defined. In two-stage elections, this means accounting for all outcomes in a complex outcome space, even those unlikely to occur, such as an outright first round win. Yet, the markets must be manageable and understandable for both traders and the exchange, so it is important to minimize the number of contracts.

Second, to reveal information, markets must attract sufficiently many informed traders because accurate market prices depend on “The Wisdom of Crowds” (Surowiecki 2005). To date (Jan. 25, 2022), there has been very little trading in the current IEM 2022 French Presidential Election Markets. Currently, we appear not to have a core group of traders interested in, and informed about, the election outcome. Having a local sponsoring organization to recruit interested traders substantially increases market activity. For example, in 2000, the IEM ran two WTA markets associated with Presidential elections in Mexico and Taiwan, each with four contracts. The Mexican market did not have a sponsoring organization. Over its 61 trading days,
daily volume averaged 48 contracts. The Taiwan market did: the Duke University Center for Chinese Electoral Studies. Over its 38 trading days, daily volume averaged 155 contracts.

The 2022 French Presidential Election markets opened more than 7 months before the first ballots would be cast. Traders have purchased 382 contract bundles, but very few individual contracts have traded. Two possible explanations for low volume are (1) market complexity and (2) no local trader base. The combined nomination/conditional vote share markets for the 1996 US Presidential Election were similarly complex. High volumes in that market lean against the first explanation. Consistent with the second explanation, average daily market volume in the 2000 Taiwan Presidential market (with a trader base) more than tripled the volume in the 2000 Mexican Presidential market (without such a base). In the future, we believe our market design can lead to viable forecasts for two-stage elections, but only with a sponsoring organization that recruits a local trader base.

**DATA AVAILABILITY STATEMENT**
Research documentation and data that support the findings of this study are openly available at the Harvard Dataverse at [https://doi.org/10.7910/DVN/CT7LVB](https://doi.org/10.7910/DVN/CT7LVB)

**CONFLICTS OF INTEREST**
The authors declare no ethical issues or conflicts of interest in this research.

**Endnotes**

1 Replication data for this paper appears in Berg, Gruca and Rietz (2022).
2 The relative price accuracy of the IEM is debated elsewhere. Berg, Nelson and Rietz (2008) find that IEM forecasts are closer to eventual election outcomes than unadjusted polls 74% of the time and the relative accuracy improves at longer forecasting horizons. Erikson and Wlezien (2008) argue that adjusted polls fair better. Their poll adjustment method makes them closer than the IEM 55% of the time.
3 See [http://tippie.uiowa.edu/iem/](http://tippie.uiowa.edu/iem/) (accessed 1/7/22) and numerous publications (e.g., Berg, Nelson and Rietz 2008) for additional information.
4 Because traders should account for all relevant information, this should mitigate issues like pre-electoral alliances and other unknown (to the researcher) impact factors.
5 Technically, the IEM forecasts risk neutral probabilities. Absent significant asymmetric hedging demand, these should approximate true probabilities. Berg and Rietz (2019) study WTA market accuracy. For price ranges and probabilities typical in elections, they find WTA markets accurately forecast probabilities.
6 Were we not interested in who might win outright, we could dispense with the R1 market entirely because the $R\text{\_None}$ contract captures the probability of an outright win without distinguishing who will win.
7 Volume averaged 71 per day over an equivalent period after Dole became the de facto nominee and, presumably, the market became simpler.
References


