

# Note méthodologique technique – Modèle de simulation d’une élection présidentielle à deux tours

## 1 Structure générale

On modélise une élection présidentielle à deux tours au moyen d’un modèle dynamique latent simulé par Monte Carlo.

Le modèle vise à approximer la distribution des issues électorales sous incertitude, en combinant :

- une structure idéologique persistante,
- une dynamique stochastique de campagne,
- une tendance structurelle lente (drift),
- un mécanisme endogène de vote utile,
- un modèle de report de voix au second tour.

On note :

- $K$  le nombre de candidats,
- $B$  le nombre de blocs idéologiques,
- $T$  la durée de la campagne (en jours),
- $S$  le nombre de simulations Monte Carlo.

Dans tout le modèle, l’utilisateur peut sélectionner pour chaque candidat  $k$  deux paramètres :

- $\mathbf{w}_k$  le profil idéologique du candidat  $k$ ,
- $\delta_k$  le dynamique structurelle du candidat  $k$ .

Le modèle est non bayésien : les paramètres sont calibrés à partir de données historiques (présidentielle 2022, législatives 2024).

## 2 Blocs idéologiques

Chaque candidat  $k$  est associé à un profil idéologique (que l’utilisateur peut choisir) :

$$\mathbf{w}_k \in \mathbb{R}^B.$$

La matrice

$$W_{ideo} \in \mathbb{R}^{K \times B}$$

encode l’appartenance partielle ou totale aux blocs.

### Justification

Cette représentation permet :

- d’autoriser des candidats positionnés entre plusieurs blocs,
- d’éviter une segmentation rigide de l’espace politique,
- de modéliser des transferts continus plutôt que discrets.

La proximité idéologique est mesurée par similarité cosinus :

$$W_{ij} = \frac{\mathbf{w}_i \cdot \mathbf{w}_j}{\|\mathbf{w}_i\| \|\mathbf{w}_j\|}.$$

La matrice  $W$  est utilisée comme noyau de diffusion idéologique dans l’ensemble du modèle.

### 3 Dynamique latente du premier tour

#### 3.1 Paramétrisation logit relative

Les parts de vote  $v_{k,t}$  appartiennent au simplexe :

$$\sum_{k=1}^K v_{k,t} = 1.$$

On passe dans l'espace des logits relatifs afin de supprimer la contrainte d'addition à un :

$$\eta_{k,t} = \log \left( \frac{v_{k,t}}{v_{K,t}} \right).$$

Cette transformation garantit :

- la positivité des parts,
- la contrainte de somme à un,
- une dynamique additive dans l'espace latent.

#### 3.2 Processus stochastique

$$\eta_t = \eta_{t-1} + u_t^{rel}.$$

Il s'agit d'une marche aléatoire multidimensionnelle dans l'espace des logits.

Les innovations brutes sont :

$$\tilde{u}_t = W_{ideo} \mu_t + \varepsilon_t,$$

où

$$\mu_t \sim \mathcal{N}(0, \Sigma_f), \quad \varepsilon_t \sim \mathcal{N}(0, D).$$

**Interprétation :**

- $\nu_t$  représente des chocs macro-idéologiques affectant un bloc entier ;
- $\varepsilon_t$  représente des chocs idiosyncratiques propres à chaque candidat.

Les matrices de variance sont :

$$\Sigma_f = \sigma_f^2 I_B, \quad D = \text{diag}(\sigma_{\varepsilon,1}^2, \dots, \sigma_{\varepsilon,K}^2).$$

**Calibration** Les paramètres sont estimés par maximum de vraisemblance afin de reproduire la volatilité observée lors de la présidentielle de 2022 :

$$\sigma_f = 1.368569 \times 10^{-2}, \quad \sigma_{\varepsilon,k} = 1.910283 \times 10^{-2}.$$

### 4 Drift structurel

Chaque candidat possède une dynamique structurelle  $\delta_k$  pouvant être choisi par l'utilisateur.

Après centrage :

$$\delta^0 = \delta - \bar{\delta} \mathbf{1}.$$

Ce centrage garantit que le drift conserve la masse électorale totale.

La matrice de redistribution est définie par :

$$S_{ij} = \frac{(W_{ij})^{\lambda_{drift}}}{\sum_{m \neq i} (W_{im})^{\lambda_{drift}}},$$

avec

$$\lambda_{drift} = 5,$$

ce qui accentue les transferts locaux.

Le drift net ajouté aux innovations est :

$$u_t = \tilde{u}_t + \delta^0 + S^\top(-\delta^0).$$

On vérifie que :

$$\sum_{k=1}^K g_k(W, \delta) = 0,$$

ce qui garantit l'absence de création ou de destruction de masse électorale.

## 5 Vote utile

Le vote utile est modélisé comme une perturbation endogène non linéaire qui s'active en fin de campagne.

### 5.1 Activation

La fonction d'activation est :

$$a_t = \frac{1}{1 + \exp\left(\frac{(T-t) - \tau_0}{s_\tau}\right)}.$$

Le paramètre  $\tau_0$  représente une date pivot proche du dernier jour de campagne (fixée à  $\tau_0 = 10$ ). Le paramètre  $s_\tau = 7$  contrôle la vitesse d'activation.

Cette forme logistique assure une montée progressive et différentiable.

### 5.2 Seuil de qualification

Le vote utile est calibré par rapport à un seuil de qualification endogène. On définit

$$q_0 = v_{(2)}(t_0),$$

où  $v_{(2)}(t_0)$  désigne le score du deuxième candidat à la date  $t_0 = T - \tau_0$  dans la trajectoire simulée sans vote utile.

Ce seuil représente le niveau de vote à atteindre pour être compétitif dans la course à la qualification au second tour.

### 5.3 Enjeu

La fonction d'enjeu donne davantage de poids aux candidats proches du seuil de qualification  $q_0$  :

$$e_k(v) = \exp\left(-\frac{1}{2} \left(\frac{v_k - q_0}{\sigma_q}\right)^2\right).$$

Il s'agit d'un noyau gaussien centré sur  $q_0$ , avec  $\sigma_q = 0.05$ .

## 5.4 Rétention

Pour chaque candidat  $k$ , on note  $r_k(v_k)$  la part de ses électeurs qui peuvent changer d'avis par stratégie. Cette rétention dépend de la distance entre le score du candidat et le seuil  $q_0$ . On adopte une fonction logistique :

$$r_k(v_k) = r_{max} - (r_{max} - r_{min}) \text{logit}^{-1} \left( \text{logit}(0.05) + \frac{v_k - q_0}{\sigma_r} \right), \quad \text{logit}^{-1}(x) = \frac{1}{1 + e^{-x}}.$$

Le paramètre  $\sigma_r$  contrôle l'échelle de la transition et est fixé à  $\sigma_r = 0.01$  dans les simulations. Les paramètres  $r_{min}$  et  $r_{max}$  sont fixés à  $r_{min} = 0.5$  et  $r_{max} = 0.9$ .

Cette spécification implique que la rétention est élevée (proche de  $r_{max}$  pour les candidats proches ou au-dessus du seuil de qualification, et faible (proche de  $r_{min}$ ) pour les petits candidats nettement en dessous du seuil. La masse d'électeurs susceptible de se déplacer depuis le candidat  $k$  est alors donnée par

$$m_k = (1 - r_k) v_k.$$

Ces électeurs sont ensuite redistribués entre les autres candidats selon leur proximité idéologique et leur attractivité.

## 5.5 Attractivité

L'attractivité du candidat  $k$  pour le vote utile dépend de l'enjeu  $e_k$  et du paramètre de dynamique structurelle  $\delta_k$  spécifique à chaque candidat :

$$\text{att}_k = \exp((\alpha \delta_k + \beta) e_k).$$

Les paramètres d'échelle sont fixés à  $\alpha = 5$  et  $\beta = 10$ .

## 5.6 Redistribution

La redistribution du vote utile s'effectue vers les candidats idéologiquement proches :

$$A_{jk} = \frac{(W_{jk})^{\lambda_{cos}} \text{att}_k}{\sum_m (W_{jm})^{\lambda_{cos}} \text{att}_m},$$

avec

$$\lambda_{cos} = 5.$$

Le vote ajusté est :

$$v^{VU} = r \odot v + A^\top m.$$

## Interprétation

Ce mécanisme produit :

- une dynamique non linéaire,
- un renforcement des candidats proches du seuil,
- une attraction dépendante de la dynamique  $\delta_k$ .

## 6 Second tour

Les deux premiers candidats sont notés  $A, B$ .

Pour un électeur du segment  $s$ , les utilités sont :

$$\ell_A = \beta(W_{sA})^\lambda - \rho_A, \quad \ell_B = \beta(W_{sB})^\lambda - \rho_B, \quad \ell_{NE} = \alpha_{nonexpr}.$$

avec

$$\beta = 7, \quad \lambda = 3, \quad \alpha_{nonexpr} = -2.259515.$$

Le paramètre  $\alpha_{nonexpr}$  est estimé à partir des législatives 2024.

### 6.1 Effets de barrage

L'effet de barrage pour chaque candidat dépend de sa composante idéologique au bloc de droite  $W_{ideo,k,3}$  (resp. au bloc de gauche  $W_{ideo,k,1}$ ), et des paramètres de barrage sélectionnés par l'utilisateur `rejet_D` et `rejet_G`.

$$\rho_k = \text{rejet\_D} \cdot W_{ideo,k,3} + \text{rejet\_G} \cdot W_{ideo,k,1}.$$

Les valeurs par défaut sont estimées à l'aide des législatives 2024 :

$$\text{rejet\_D} = 4.909898, \quad \text{rejet\_G} = 2.240084.$$

Le choix final suit :

$$P(A, B, NE) = \text{softmax}(\ell_A, \ell_B, \ell_{NE}).$$

## 7 Monte Carlo

On simule  $S = 500$  trajectoires indépendantes.

Les probabilités estimées convergent vers leurs valeurs asymptotiques à vitesse  $O(S^{-1/2})$ .

On calcule :

- la probabilité de qualification au second tour,
- la probabilité de victoire,
- la médiane des trajectoires du premier tour ainsi que leurs intervalles de confiance.