Banquier d'affaires - quel parcours !

Arthur Miguel et Madalina Olteanu

Université Paris Dauphine PSL

Travail réalisé en collaboration avec Rosalie Milner

14 octobre 2024

Introduction

Constats et objectif de la thèse

- **Des histoires similaires** :
  - Deux dynasties financières juives qui se développent au XIXe siècle et dont l'implantation est internationale.
- **Divergence dans les trajectoires suivies** :
  - Alors que Lazard voit les descendants de la famille fondatrice évincés de la direction dans les années 2000, la famille Rothschild, quant à elle, maintient un contrôle familial.
- **Objectif** :
  - En s'appuyant sur l'étude des sociétés de conseil Lazard Frères et Rothschild & Cie, comprendre comment chaque banque a répondu aux pressions de la concentration et de la mondialisation.

Thèse

- La contradiction entre Lazard et Rothschild est **partielle**.
- **Rothschild** a suivi une trajectoire qui rejoint celle de **Lazard** sous beaucoup d'aspects :
  - Les profils sociologiques des cadres dirigeants des deux banques convergent.

Hypothèse concernant les secteurs d'activité

- On s'attend à observer **deux phénomènes** en étudiant les carrières professionnelles de nos enquêtés sous l'angle des **secteurs d'activité** :
- **Une spécialisation dans le secteur financier** :
  - La spécialisation dans un secteur est typique des carrières des cadres dirigeants contemporains, caractérisées par une stabilité et une mobilité réduite.
  - Cette spécialisation permet de développer une expertise sectorielle, de bâtir un réseau solide et de maximiser les opportunités de promotion interne.
  - (Koch, Forgues, & Monties, 2017)
- **Le passage par la haute administration** :
  - En France, un certain nombre de cadres dirigeants commencent leur carrière dans la haute fonction publique avant de rejoindre le secteur privé, un phénomène connu sous le nom de **pantouflage**.
  - (Dudouet et al., 2024, p. 46-63)

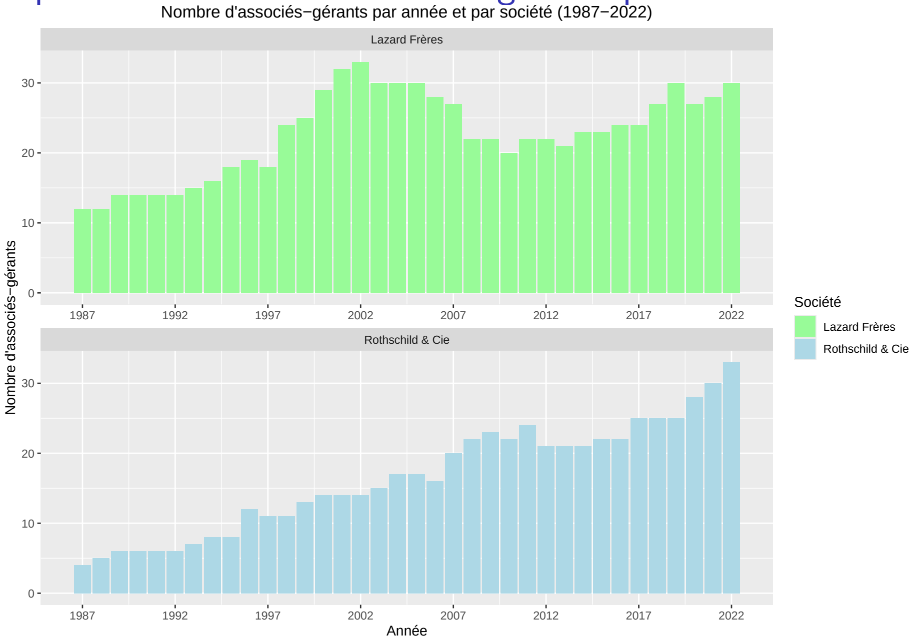
Construction de la base de données

- **Construction d'une base de données prosopographiques** :
  - Recueil d'informations biographiques détaillées dans une base de données relationnelles, notamment sur les trajectoires professionnelles.
- **Sources principales** :
  - Who's who, LinkedIn, presse, sites internet des entreprises, Geneanet (origines sociales).

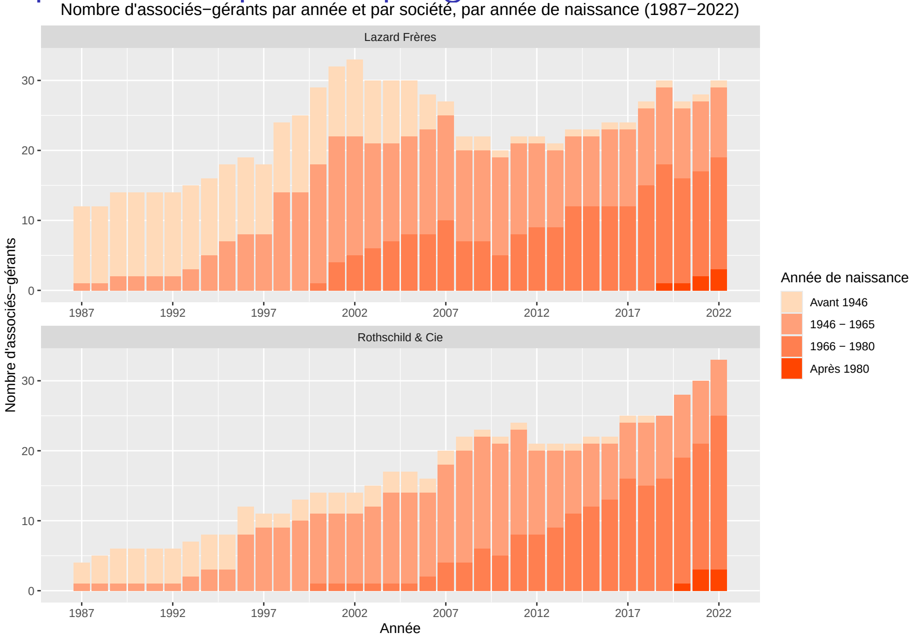
Population - Informations générales

- **Total des individus** : 139
- **Critères de sélection** :
  - Les individus ont occupé le poste d'**associé-gérant** durant la période étudiée au sein des sociétés de conseil **Rothschild & Cie** et **Lazard Frères**.
  - On a inclus ceux ayant le **statut formel** d'associé-gérant au sein d'une société en commandite simple, ainsi que ceux portant le **titre** après la transformation de Lazard Frères en société anonyme.
- **Précisions supplémentaires** :
  - **Exhaustivité** : Toutes les personnes répondant aux critères sont incluses.
  - **Augmentation très forte** des effectifs sur la période étudiée.

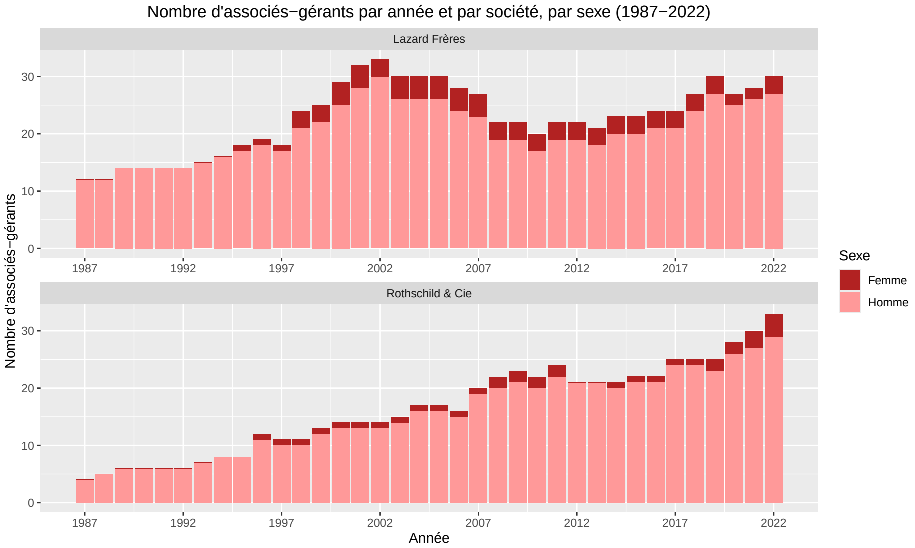
Population - Nombre d'associés-gérants par année



Population - Répartition par génération

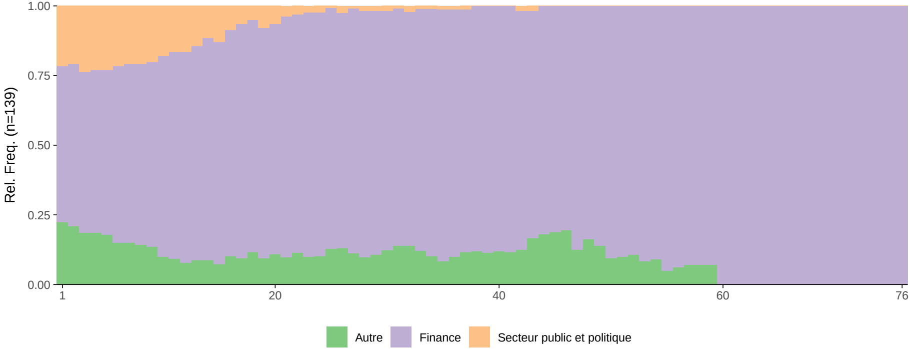


Population - Répartition par sexe



Quels outils pour l'analyse quantitative ?

- On se focalise dans la suite sur les trajectoires du point de vue du secteur d'activité.



- Analyse longitudinale ?
- Modèles probabilistes ?

Un court passage par l'analyse longitudinale

- Approche exploratoire et holiste (Abbott, 1986).
- Les *clusters* sont extraits après calcul d'une matrice de dissimilarité entre les trajectoires et une étape de classification hiérarchique.
- **Quel choix pour la dissimilarité** ?
  - de très nombreuses propositions dans la littérature (Hamming, LCS, χ^2, optimal matching, ...).
  - chaque distance est sensible à un ou plusieurs aspects : enchaînement des états, chronologie, durées de séjour.
  - certaines dissimilarités nécessitent des choix arbitraire de coûts (optimal matching).
  - en général, les trajectoires doivent être de même longueur. Dans notre cas, *les données sont censurées à droite*.

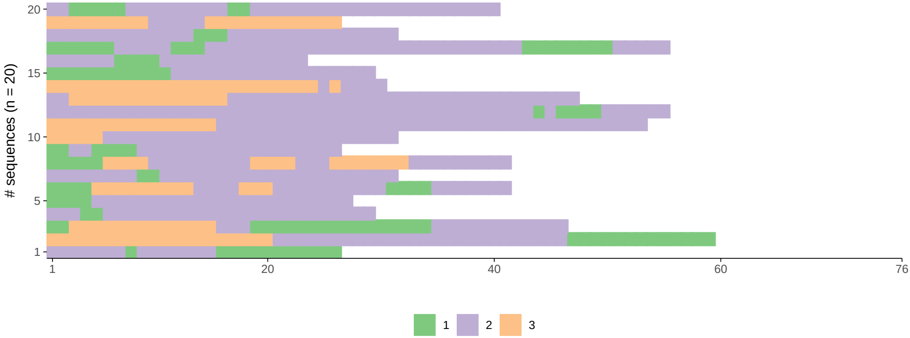
L'approche probabiliste via des modèles markoviens

- Changer de point de vue et se focaliser sur la dynamique des trajectoires.
- Passer d'une approche qui considère la séquence comme une entité à analyser vers une approche qui décrit un processus pas à pas, et au cours duquel le passé a une influence sur la séquence.
- Partiellement intégré dans les matrices de coûts pour les distances OM.
- Modèle probabiliste, *génératif*.
- Permettent des valeurs manquantes, les trajectoires peuvent avoir des longueurs différentes.
- Permettent même l'imputation de données manquantes.

Les processus semi-Markoviens pour le clustering de séquences

Processus de renouvellement Markovien

- On souhaite modéliser des séquences de données **catégorielles**, **ordonnées** dans le temps, possiblement de **longueurs inégales**.



- Comment modéliser les transitions entre les différents états ?
- Comment modéliser les temps de séjour dans chaque état ?
- Chaines de Markov, semi-Markov, modèles de mélange, ...

Processus de renouvellement Markovien - définition

- On considère un ensemble fini d'états $\mathcal{S} = \{1, ..., D\}$.
- On peut caractériser une trajectoire via la suite $(S_p, X_p)_{p \geq 1}$, où les S_p désignent les états visités, et X_p le temps de séjour associé à chaque visite.
- **Pour modéliser les transitions entre les états**, on introduit une chaîne de Markov *homogène* $(S_p)_{p \geq 1}$ sur $\mathcal{S}$, définie par sa matrice de transition $\mathbf{P} = (p_{ij})_{i,j} \in [0,1]^{D \times D}$

$$p_{ij} = \mathbb{P}(S_{p+1} = j | S_p = i) \ , \forall i, j = 1, ..., D,$$

et par les probabilités de l'état initial S_1, $\alpha = (\alpha_1, ...\alpha_D)$,

$$\alpha_i = \mathbb{P}(S_1 = i) \ , \forall i = 1, ...D.$$

Processus de renouvellement Markovien - définition

- **Pour modéliser les temps de séjour dans chaque état**, on introduit leur distribution (paramétrique) conditionnellement aux états visités :

$$\phi_{ij}(t) = (X_p \leq t | S_p = i, S_{p+1} = j) \ , \forall t \in [0, \infty), i \neq j.$$

- Plusieurs choix parmi des distributions usuelles : Gamma, exponentielle, ...
- $(S_p, X_p)_p$ est un **processus de renouvellement Markovien** (Pyke, 1961).
- $(X_t)_t$ est un **semi-processus de Markov** (Barbu et al., 2008).

Processus de renouvellement Markovien - définition

- On supposera $p_{ii} = 0$, $\forall i = 1, ..., D$ pour des questions d'identifiabilité.

- **Propriété de Markov** : pour tout $t \in [0, \infty)$, $i, j = 1, ...D$,

$$\mathbb{P}(X_p \leq t, S_{p+1} = j | S_p = i, S_{p-1}, ..., S_1, X_{p-1}, ..., X_1) = p_{ij}\phi_{ij}(t).$$

- Modèle simplifié :
  - le temps de séjour dépend uniquement de l'état actuellement visité (*pas d'anticipation*) :

$$\mathbb{P}(X_p \leq t | S_p = i, S_{p+1} = j) = \mathbb{P}(X_p \leq t | S_p = i) = \phi_i(t).$$

  - hypothèse réaliste dans de nombreux cas, et aussi moins de paramètres à estimer.

- Paramètre du modèle :

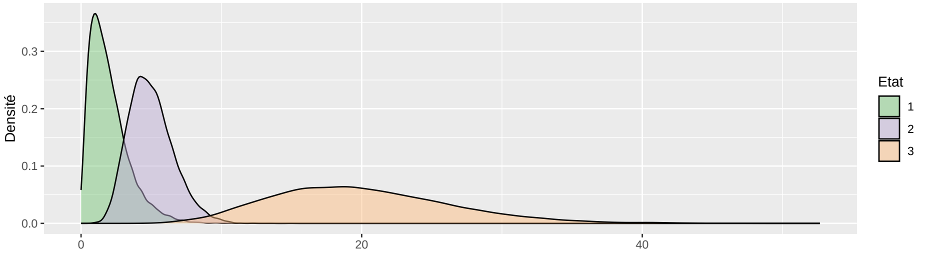
$$\theta = (\alpha, \mathbf{P}, (\gamma_i)_{i=1,...,D}) \in [0,1]^D \times [0,1]^{D \times D} \times \mathbb{R}^{DxL}.$$

Exemple sur des données simulées

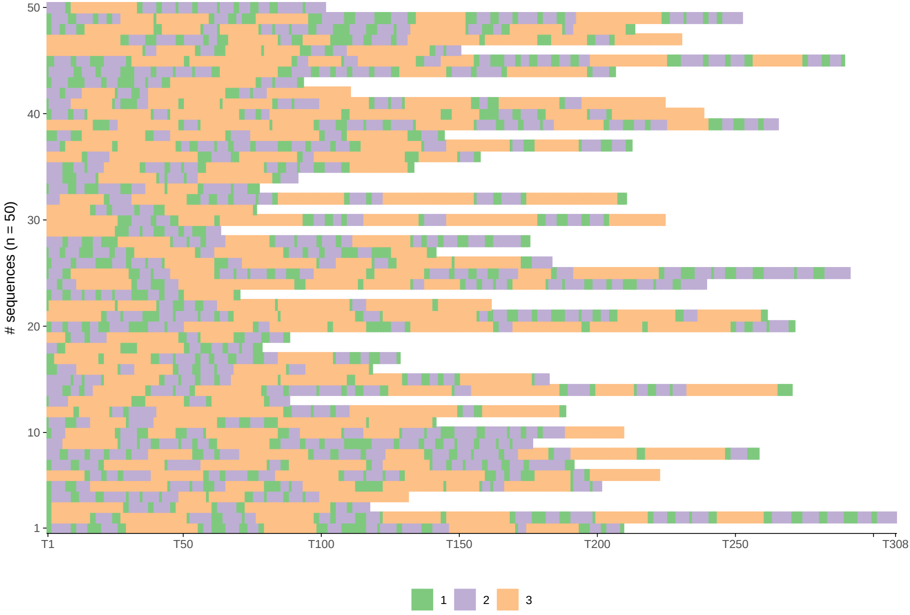
- Probabilités initiales : $\alpha = (0.5, 0.2, 0.3)$.
- Matrice de transition

$$\mathbf{P} = \begin{pmatrix} 0 & 0.8 & 0.2 \\ 0.8 & 0 & 0.2 \\ 1 & 0 & 0 \end{pmatrix}$$

- Distribution des temps de séjour (famille Gamma) :



Exemple sur des données simulées



Et si on a plusieurs populations ?

- Dans un *contexte de données hétérogènes*, on peut considérer que les trajectoires observées ont été engendrées par plusieurs processus semi-Markoviens, chacun étant déterminé par un jeu de paramètres différent et étant responsable pour un sous-ensemble des trajectoires observées (**une sous-population**).
- L'un peut ainsi définir des **mélanges de processus semi-Markov**, (Cardot et. al, 2019).
- On considère G sous-populations, avec des poids théoriques $\pi = (\pi_1, ..., \pi_G)$.
- Chaque trajectoire est la réalisation d'un processus semi-Markovien associé à une sous-population.

Mélange de processus semi-Markoviens

Pour générer une nouvelle trajectoire :

- choisir une sous-population $Z \in \{1, ..., G\}$ selon une loi multinomiale $\mathcal{M}(\pi, 1)$.
- selon la valeur $Z = g$, choisir le jeu de paramètres associé $\theta^g = (\alpha^g, \mathbf{P}^g, (\gamma_i^g)_{i=1,\dots,D})$.
- engendrer un trajectoire selon le processus semi-Markovien de paramètre θ^g, $(S_p^g, X_p^g)_{p\geq 1}$.

Paramètres du modèle global :

$$\pi = (\pi_1, ...\pi_G), (\alpha^g)_{g=1,\dots,G}, (\mathbf{P}^g)_{g=1,\dots,G}, (\gamma_i^g)_{i=1,\dots,D;g=1,\dots,G}$$

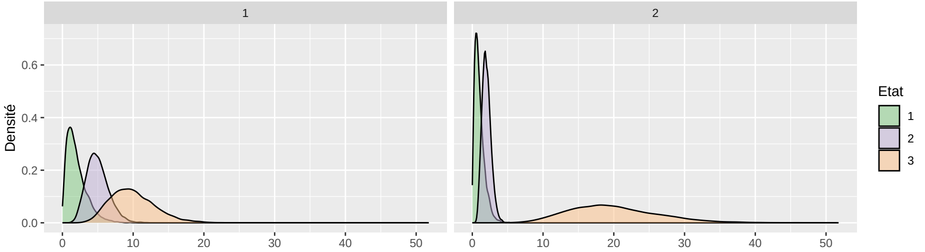
.

Exemple sur des données simulées

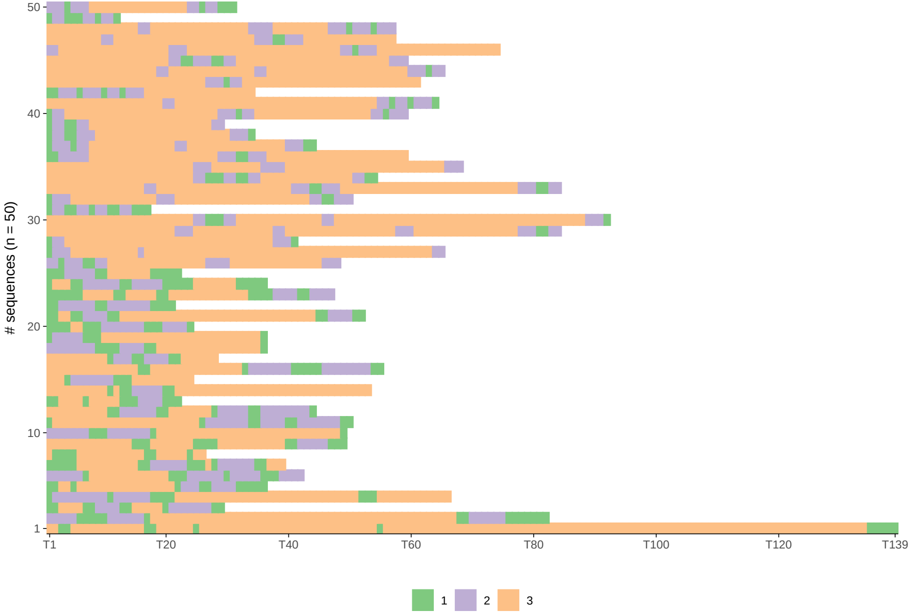
- Poids des sous-populations : $\pi_1 = \pi_2 = 0.5$.
- Probabilités initiales : $\alpha^1 = \alpha^2 = (0.5, 0.2, 0.3)$.
- Matrice de transition

$$\mathbf{P^1} = \begin{pmatrix} 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; \mathbf{P^2} = \begin{pmatrix} 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{pmatrix}$$

- Distribution des temps de séjour (famille Gamma) :



Exemple sur des données simulées



Et dans la vraie vie ?

- **On ne dispose que des données !**
- On souhaite inférer et calculer :
  - le nombre de sous-populations G ;
  - les poids des sous-populations π_g ;
  - les probabilités initiales pour chaque sous-population α^g ;
  - les matrices de transition pour chaque sous-population $\mathbf{P}^g$;
  - les paramètres des distributions des temps de séjour dans chaque sous-population γ^g ;
  - **mais aussi** la sous-population la plus vraisemblable pour chaque trajectoire.

Et dans la vraie vie ?

- **Comment ?**
  - maximisation de la vraisemblance basée sur un algorithme EM ;
  - le nombre de sous-populations G doit être choisi avec un critère de vraisemblance pénalisée (BIC, ...).
- **Quel intérêt ?**
  - **clustering** implicite par l'affectation de chaque trajectoire à la sous-population la plus probable ;
  - interprétabilité des clusters grâce aux paramètres estimés ;
  - pas de contrainte sur la longueur des trajectoires ;
  - modèle génératif : on peut engendrer de nouvelles trajectoires, simuler la suite, ...
  - possibilité d'inclure des covariables.

Population étudiée - trois clusters

Trois clusters - Paramètres estimés

- Poids des classes $\hat{\pi}_1$, $\hat{\pi}_2$, $\hat{\pi}_3$:

Classe	
1	0.19
2	0.46
3	0.35

- Probabilités initiales des 3 états dans chaque classe, $(\hat{\alpha}_\ell^g)_{g=1,2,3,\ell=1,2,3}$:

	Autre	Finance	Sect. public & politique
1	0.05	0.82	0.13
2	0.34	0.24	0.41
3	0.15	0.85	0.00

Trois clusters - Paramètres estimés - Cluster 1

- Matrice de transition $\hat{\mathbf{P}}$:

	Autre	Finance	Sect. public & politique
Autre	0.0000	1.0000	0
Finance	1.0000	0.0000	0
Sect. public & politique	0.1227	0.8773	0

- Distribution des temps de séjour $\Gamma(\hat{a}, \hat{\lambda})$:

Etat			Moyenne	Variance
Autre	0.11	0.06	1.73	28.12
Finance	28.51	0.58	49.27	85.14
Sect. public & politique	0.21	0.15	1.36	9.00

Trois clusters - Paramètres estimés - Cluster 2

- Matrice de transition $\hat{\mathbf{P}}$:

	Autre	Finance	Sect. public & politique
Autre	0.0000	0.8123	0.1877
Finance	0.7907	0.0000	0.2093
Sect. public & politique	0.2738	0.7262	0.0000

- Distribution des temps de séjour $\Gamma(\hat{a}, \hat{\lambda})$:

Etat			Moyenne	Variance
Autre	0.93	0.13	7.02	52.77
Finance	1.70	0.14	12.41	90.68
Sect. public & politique	0.69	0.13	5.19	39.28

Trois clusters - Paramètres estimés - Cluster 3

- Matrice de transition $\hat{\mathbf{P}}$:

	Autre	Finance	Sect. public & politique
Autre	0.0000	1.0000	0
Finance	1.0000	0.0000	0
Sect. public & politique	0.9992	0.0008	0

- Distribution des temps de séjour $\Gamma(\hat{a}, \hat{\lambda})$:

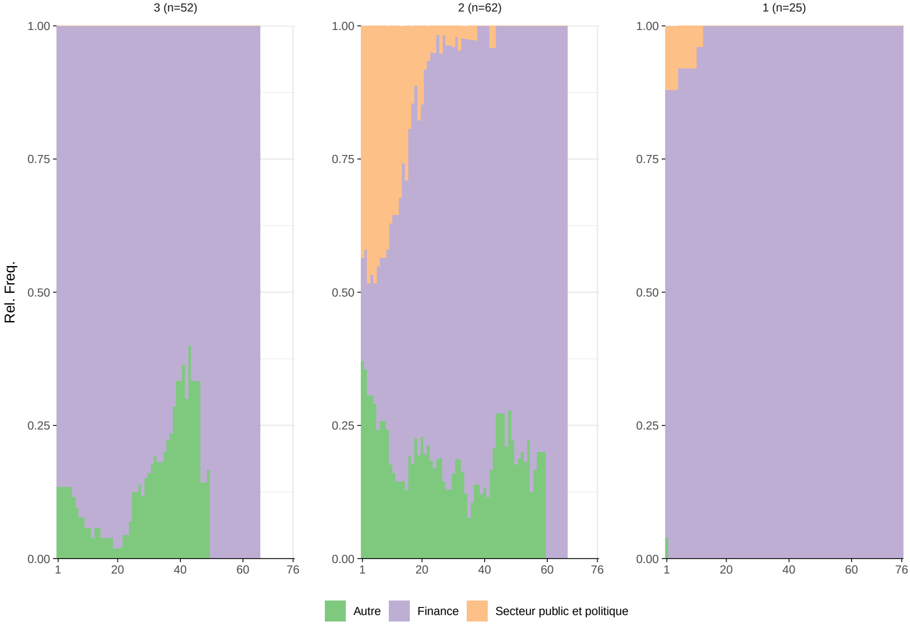
Etat			Moyenne	Variance
Autre	3.86	0.36	10.79	30.13
Finance	18.77	0.67	28.03	41.87
Sect. public & politique	0.05	0.05	1.17	25.45

Trois clusters - composition

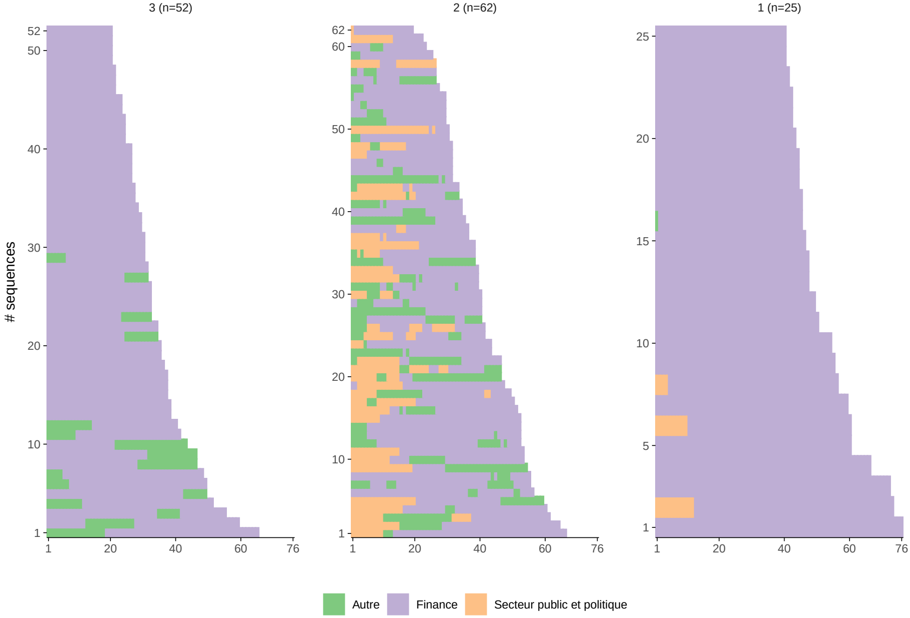
- Informations générales sur les trois clusters :

Cluster	Nombre d'individus	Proportion (%)
1	25	18.0
2	62	44.6
3	52	37.4

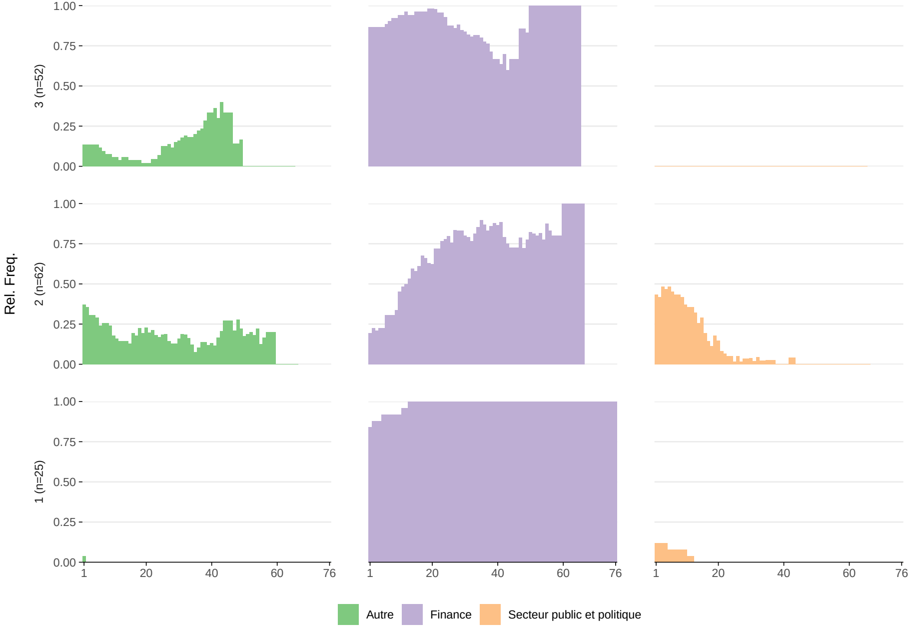
Trois clusters - composition



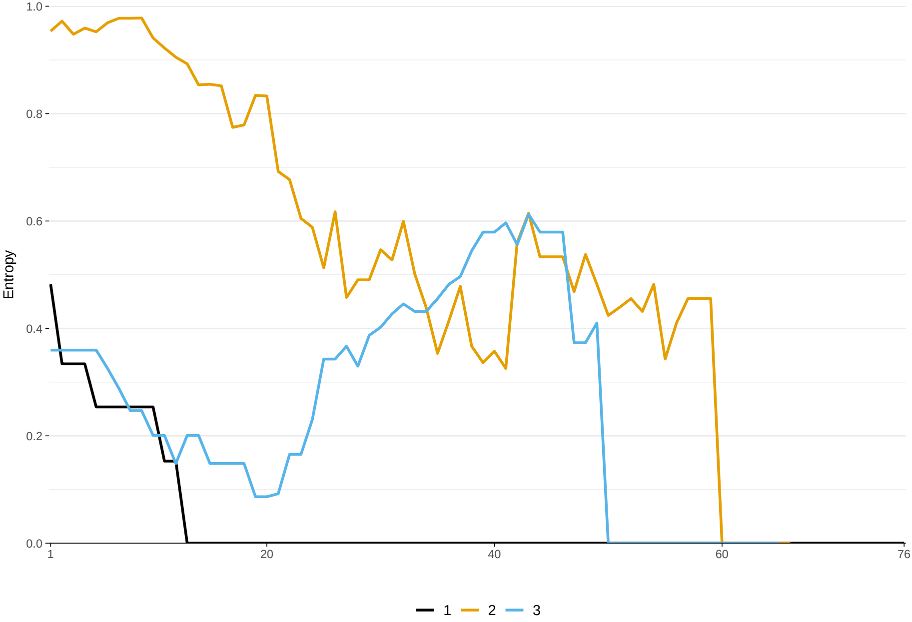
Trois clusters - composition



Trois clusters - composition



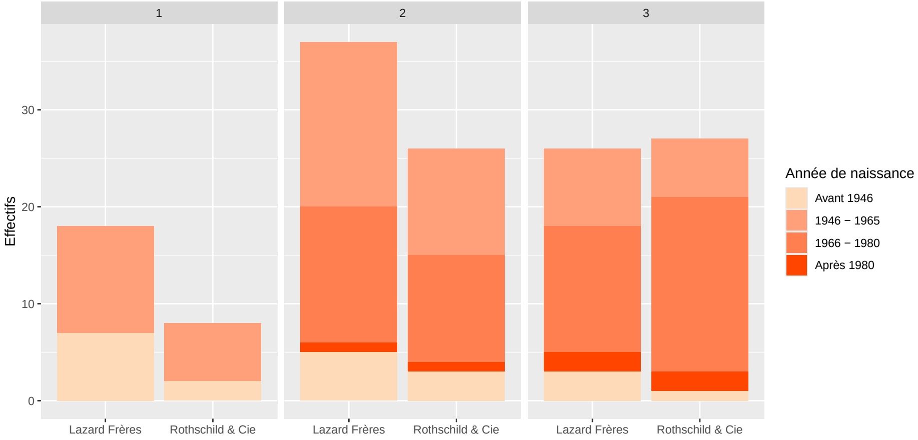
Trois clusters - composition



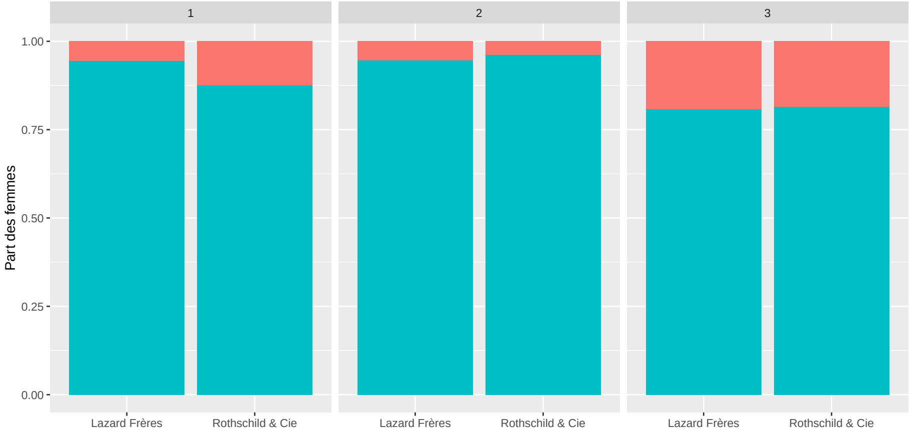
Interprétation générale des trois clusters

- **Cluster 1** :
  Ce cluster regroupe des individus présentant une **spécialisation très forte dans la finance**, avec peu ou pas de transitions vers d'autres secteurs.
- **Cluster 2** :
  Ce cluster est le plus diversifié en termes de **transitions entre les secteurs**. Les individus peuvent débuter dans le **secteur public et politique** ou dans un **autre secteur**, puis passer à la **finance**, avec des durées de séjour variables.
- **Cluster 3** :
  Bien que ce cluster soit également **très concentré sur la finance**, il se distingue par des **transitions notables** vers les **autres secteurs**, souvent à un stade plus avancé de la carrière.

Distribution des clusters par génération



Distribution des clusters par sexe



Conclusion et perspectives

- **Spécialisation d'une grande partie de la population dans le secteur financier** : La finance domine les trois clusters, avec une prédominance écrasante dans le cluster 1 et une importance marquée dans le cluster 3. Cette spécialisation se manifeste dès le début des carrières, indiquant un alignement rapide vers le secteur financier pour une grande partie des associés-gérants.
- **Persistance du pantouflage** : Le cluster 2 illustre la persistance du 'pantouflage', c'est-à-dire des passages par la haute administration française en début de carrière avant un passage dans le secteur privé.
- **La finance comme tremplin** : Le cluster 3 suggère que la finance peut également constituer un tremplin pour accéder à d'autres secteurs. Bien que cela reste à démontrer, on suppose qu'il s'agit de l'accès à des postes de premier plan.

Bibliographie

- ABBOTT, A., & FORREST, John. Optimal matching methods for historical sequences. *The Journal of Interdisciplinary History*, 1986, Vol. 16(3), pp. 471-494.
- BARBU, V. S., & LIMNIOS, N. (2008). *Semi-Markov Chains and Hidden Semi-Markov Models Toward Applications: Their Use in Reliability and DNA Analysis*. New York: Springer Science and Business Media.
- CARDOT, H., LECUELLE, G., SCHLICH, P., & VISALLI, M. (2019). Estimating Finite Mixtures of Semi-Markov Chains: An Application to the Segmentation of Temporal Sensory Data, *Journal of the Royal Statistical Society Series C: Applied Statistics*, Vol. 68, pp. 1281–1303.
- DUDOUET, F., & VION, A. (2024). III / Les carrières de dirigeants. Sociologie des dirigeants de grandes entreprises. La Découverte, pp. 46-63.
- KOCH, M., FORGUES, B., & MONTIES, V. (2017). The way to the top: Career patterns of Fortune 100 CEOs. *Human Resource Management*, Vol. 56(2), pp. 267-285.
- PYKE, R. (1961) Markov renewal processes: definitions and preliminary properties. *Ann. Math. Statist.*, Vol. 32, pp. 1231–1242.