
(Re)faire la sociologie du couple avec des données massives

Marie Bergström, Ined

marie.bergstrom@ined.fr

Les rencontres de statistiques appliquées

Données du Web – Questions et pratiques en SHS, 12 février 2021



La présentation



- Introduction
- **Apport** des données massives (à la sociologie du couple)
- **Illustrations** d'usage
- **Limites** & difficultés
- Conclusion

Introduction



L'usage de données du web pour étudier les rencontres

- **Sites et applications de rencontres** : profils & interactions

- **Un corpus croissant de recherches** :

Processus d'appariement : Hitsch, Hortacsu et Ariely, 2010; Bergström, 2018; Bruch et Newman, 2018, 2019.

Homogamie sociale : Skopek, Schulz et Blossfeld, 2011; Schmitz, 2012, 2016; Bergström, 2016.

Homogamie ethnoraciale : Feliciano, Robnett et Komaie, 2009; Feliciano, Lee et Robnett, 2011; Lin et Lundquist, 2013; Curington, Lin et Lundquist, 2015; Lundquist et Lin, 2015; Potârcă et Mills, 2015.

- **Des partenariats Universités-Entreprises**

Apport des données massives



Les sources habituelles pour étudier la formation des couples

Données administratives : état civil & données fiscales

Enquêtes

- *Le choix du conjoint*, 1959, Ined.
- *La formation des couples*, 1983-1984, Ined.
- *L'étude des parcours individuels et conjugaux*, 2013-2014, Ined-Insee.

Sources mobilisées pour :

- Cerner la *composition* des unions (ex : l'homogamie)
- Cerner les logiques *du choix du conjoint* (ex : l'homophilie)

Apport des données massives



Les sources habituelles pour étudier la formation des couples

- **Une approche *ex post facto*** : étude des couples déjà constitués
- **Implications théoriques** : les couples comme « préférences révélées » (Samuelson, 1938)

Apport des données massives



Les sources habituelles pour étudier la formation des couples

- Une approche *ex post facto* : étude des couples déjà constitués
- Implications théoriques : les couples comme « préférences révélées » (Samuelson, 1938)

Problème empirique : les choix ne reflètent pas toujours les préférences

Problème méthodologique : confusion entre les niveaux d'analyse

Problème théorique : « *scholastic fallacy* »

« [Erreur] très commune dans les sciences sociales, (...) par laquelle le savant place dans la tête des agents qu'il étudie (...) les considérations et les constructions théoriques qu'il a dû élaborer pour rendre compte de leurs pratiques »

- Bourdieu, « Le champ économique », 1997, p. 61-62.

Apport des données massives



L'originalité des données de sites & d'applications

- Étudier les **rencontres en train de se faire** (*input* et non *output*)
- Saisir autant la **sélection** que le **rejet** de partenaires potentiels
- **L'observation** plutôt que la déclaration de pratiques
- Enregistrement de **flux** et possibilités d'analyses dynamiques

Illustrations d'usage



Une recherche en partenariat avec Meetic Group

- **2 types d'informations**
 - **Profils anonymisés** : informations renseignées (sexe, âge, taille etc.) + métadonnées (date d'inscription, dernier login, support etc.)
 - **Comportements de contact** : métadonnées (identifiants, date, statut)
 - Navigation : « qui regarde qui »
 - Sollicitations : « *winks* » envoyés
 - Messages : « qui contacte qui »
 - Ratings : « *swipe* » droite ou gauche
- **15 pays européens** (FR, UK, DE, ES, IT, NL, BE, SWE...)
- **7 plateformes** (Match, Meetic, Neu, Lexa, DatingDirect, Friendscout, Ourtime)

Illustrations d'usage



L'homogamie sociale : qui contacte qui ?

- **Objectif** : saisir le *modus operandi* de l'homogamie
- **Échantillon** : profils (400 000) et interactions (25M) en 2014
- **Analyse** : interactions en fonction du diplôme déclaré

Illustrations d'usage

L'homogamie sociale : qui contacte qui ?

PRISE DE CONTACT (1^e EMAIL)

Expéditeur	Destinataire	Niveau					
		lycée et inférieur	Bac	Bac +2	Bac +3	Bac +4	Bac +5 et plus
Niveau lycée et inférieur		1,52	1,23	1,04	0,93	0,83	0,67
Bac		1,21	1,14	1,05	0,98	0,90	0,79
Bac+2		0,98	1,02	1,04	1,02	0,98	0,94
Bac+3		0,90	0,96	1,00	1,02	1,03	1,05
Bac+4		0,87	0,92	0,97	1,00	1,10	1,12
Bac+5 et plus		0,66	0,78	0,91	1,02	1,14	1,36

ÉCHANGE PROLONGÉ (10^e EMAIL)

Expéditeur	Destinataire	Niveau					
		lycée et inférieur	Bac	Bac +2	Bac +3	Bac +4	Bac +5 et plus
Niveau lycée et inférieur		1,93	1,36	0,97	0,82	0,67	0,52
Bac		1,35	1,24	1,04	0,93	0,82	0,68
Bac+2		0,97	1,04	1,08	1,02	0,98	0,89
Bac+3		0,78	0,93	1,03	1,07	1,06	1,06
Bac+4		0,62	0,77	0,96	1,07	1,24	1,28
Bac+5 et plus		0,46	0,61	0,87	1,07	1,26	1,62

<0,9 Sous-représentation
 0,9-1,1 Proche du taux attendu
 >1,1 Surreprésentation

Illustrations d'usage



L'homogamie sociale : qui contacte qui ?

- Une **tendance** d'homophilie...
- ...particulièrement **marquée aux « pôles »** de l'échelle des diplômes...
- ... qui découle à la fois d'un **choix** initial et d'un **processus**

Cf. Bergström, 2016, « (Se) correspondre en ligne. L'homogamie à l'épreuve des sites de rencontres », *Sociétés contemporaines*, 4(104), p. 13-40.

Illustrations d'usage

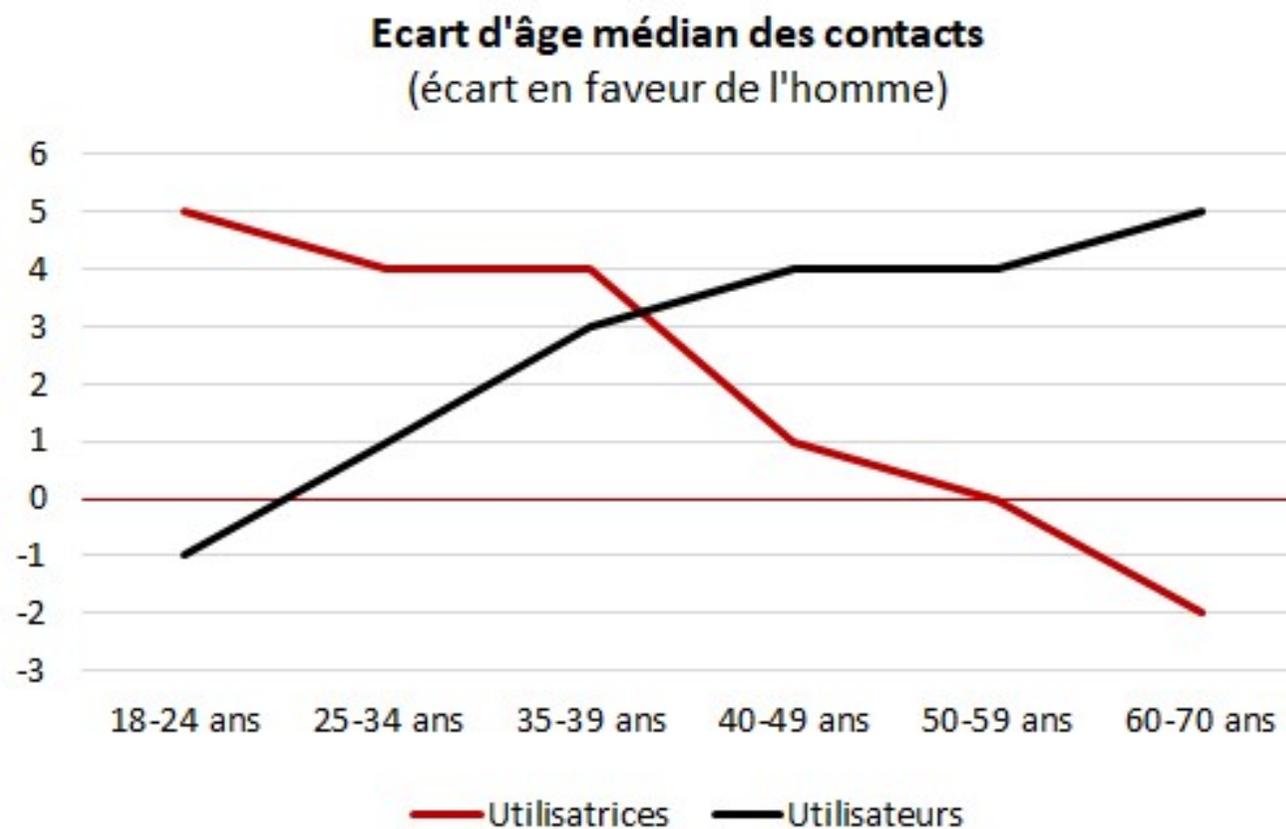


L'écart d'âge : une préférence féminine ou masculine ?

- **Objectif** : ouvrir la « boîte noire » de l'écart d'âge
- **Échantillon** : profils (400 000) et interactions (25M) en 2014
- **Analyse** : interactions en fonction du sexe et de l'âge

Illustrations d'usage

L'écart d'âge : une préférence féminine ou masculine ?



Illustrations d'usage

L'écart d'âge : une préférence féminine ou masculine ?

- Une **tendance statistique forte**...
- ... qui découle de **processus distincts** :
 - En début de parcours, l'écart d'âge est désiré par **les femmes**
 - À des âges + avancés, l'écart d'âge est désiré par **les hommes**

Cf. Bergström, 2018, « De quoi l'écart d'âge est-il le nombre ? L'apport des big data à l'étude de la différence d'âge au sein des couples », *Revue Française de Sociologie*, 59(3), p. 395-422.

Limites & difficultés



Les limites :

- **La représentativité** : données non représentative de la population FR
- **La finalité** : des rencontres amoureuses mais surtout sexuelles
- **La véracité** : on ne dit pas toujours vrai sur Internet (on arrondit)

Les difficultés :

- **L'accès** : négociations et contreparties
- **La voluminosité** : des milliards d'observations ; des bases multiples
- **La fiabilité** : une moindre qualité ; une moindre maîtrise

Conclusion

L'apport des données du web

- À la sociologie du couple :
un nouveau regard sur le processus d'appariement des partenaires

Données d'enquête	Données de <i>Meetic</i>
Couples formés	Couples formés
↓	↑
Appariement	Appariement
↓	↑
Préférences	Préférences

- Aux sciences sociales :
des possibilités nouvelles d'**observations quantifiées**
des données complémentaires (points aveugles des enquêtes)

Merci de votre attention !

Références bibliographiques



- Bergström** Marie, 2016, « (Se) correspondre en ligne. L'homogamie à l'épreuve des sites de rencontres », *Sociétés contemporaines*, 4(104), p. 13-40.
- Bergström** Marie, 2018, « De quoi l'écart d'âge est-il le nombre ? L'apport des big data à l'étude de la différence d'âge au sein des couples », *Revue Française de Sociologie*, 59(3), p. 395-422.
- Bruch** Elizabeth E., **Newman** M.E.J., 2018, « Aspirational pursuit of mates in online dating markets », *Science Advances*, 4(8) [en ligne].
- Bruch** Elizabeth E., **Newman** M.E.J., 2019, « Structure of online dating markets in US cities », *Sociological science*, 6, p. 219-234.
- Curington** Celeste Vaughan, **Lin** Ken-Hou, **Lundquist** Jennifer Hikes, 2015, « Positioning multiraciality in cyberspace: Treatment of multiracial daters in an online dating website », *American Sociological Review*, 80(4), p. 764-788.
- Feliciano** Cynthia, **Robnett** Belinda, **Komaie** Golnaz, 2009, « Gendered racial exclusion among white internet daters », *Social Science Research*, 38(1), p. 39-54.
- Feliciano** Cynthia, **Lee** Rennie, **Robnett** Belinda, 2011, « Racial Boundaries among Latinos: Evidence from Internet Daters' Racial Preferences », *Social Problems*, 58(2), p. 189-212.
- Hitsch** Günter, **Hortacsu** Ali, **Ariely** Dan, 2010, « Matching and Sorting in Online Dating », *The American Economic Review*, 100(1), p. 130-163.
- Lin** Ken-Hou, **Lundquist** Jennifer, 2013, « Mate Selection in Cyberspace: The Intersection of Race, Gender, and Education », *American Journal of Sociology*, 119(1), p. 183-215.
- Lundquist** Jennifer H., **Lin** Ken-Hou, 2015, « Is love (color) blind? The economy of race among gay and straight daters », *Social Forces*, 93(4), p. 1423-1449.
- Potârcă** Gina, **Mills** Melinda, 2015, « Racial Preferences in Online Dating across European Countries », *European Sociological Review*, 31(3), p. 326-341.
- Schmitz** Andreas, 2012, « Elective Affinities 2.0? Bourdieu's Approach to Partnership Formation in the Light of E-Dating Contact Patterns », *RESET. Recherches en sciences sociales sur Internet*, 1(1) [en ligne].
- Schmitz** Andreas, 2016, *The Structure of Digital Partner Choice. A Bourdieusian perspective*, New-York, Springer.
- Skopek** Jan, **Schulz** Florian, **Blossfeld** Hans-Peter, 2011, « Who Contacts Whom? Educational Homophily in Online Mate Selection », *European Sociological Review*, 27(2), p. 180-195.