What’s behind the age gap between spouses?
An empirical assessment of women’s and men’s age preferences using online dating data

Marie Bergström
marie.bergstrom@ined.fr
French Institute for Demographic Studies (INED), Paris, France

Session 215: Qualitative and Quantitative Innovations in Family Research
(First draft not to be cited or quoted without permission)
Abstract

There is an abundant literature in Demography and Sociology on the age gap between spouses and the explanatory factors behind this phenomenon. While different, the studies often share a methodological premise: the analysis is based on survey data on couples already formed, i.e. it looks at partner choice ex post facto. This approach presents important limits when the aim is to explain how the age gap comes to be. It leads the authors to infer conjugal behavior and preferences from existing couples rather than to assess these empirically. The paper offers an alternative approach. Taking opportunity of some features of online dating, it looks at age preferences as they are expressed by users in their profiles. In doing so, it suggests that – against the common idea of a strong preference for age asymmetry (men being older) – the age gap is the result of a negotiation between the sexes.

Key words
Couple formation – Age Gap – Gender – Organic data – Online dating
1. Introduction

In the majority of heterosexual couples, the woman is younger than the man. Far from being specific to the Western World or to the contemporary era, this age gap between spouses is observed throughout time and space with few exceptions. Only the magnitude of the gap differs between countries and time periods and seems to reduce with socio-economic development (Mignot, 2010). What are the mechanisms that produce this male superiority in age? In the existing literature, the age gap is generally presented as a strong preference among both women and men. It is considered as a result of gender specific criteria in partner selection, which lead women to value maturity and men to value youth (Bergstrom & Bagnoli, 1993; Collins & Coltrane, 1991; Elder, 1969; Taylor & Glenn, 1976). This common hypothesis is however rarely tested. Due to the limitations of available data, few studies have tried to empirically grasp female and male age preferences and see how they come into play in couple formation. In this paper, I use data from an online dating site to address women’s and men’s partner criteria with regard to age.

The aim of the paper is twofold. A first objective is to illustrate how “organic data” can be useful to the study of couple formation. The article starts with a methodological discussion on the limits of surveys when trying to assess partner preferences and choice. It presents alternative sources that have been used to sidestep these limitations, and makes a case for using data from online dating sites. In a second part of the paper, I apply this approach to the study of age preferences. Using data from the French site Meetic.fr (equivalent of Match.com), I suggest that the age gap between spouses should be understood as the result of a “bargaining” between the sexes rather than as a product of two matching wills. The second and more general aim of the paper is thus to give new insights into the modus operandi of couple formation, and to gain a better understanding of how this age asymmetry between the sexes comes to be.
2. **Methodological discussion: the case for using online dating data**

Recent years have seen a growing interest in the social sciences for “organic data”, understood as quantifiable information on human behavior that is collected for other purposes than the scientific observation of these very behaviors (Groves, 2011). The term is used in opposition to “designed data”, referring to surveys and censuses that are specifically produced for quantitative analysis. The amount and the scope of this type of data have dramatically increased over the last decades, notably because of the digitalization of society. At the same time, its use in social scientific research has become more important. One reason for this is the declining response rates to surveys which encourages researchers to resort to other sources of information. Another reason lies in the new vantage points offered by this material. The nature of so called “organic data” is often very different from the type of data habitually used in Sociology and Demography. This comes with problems, but also with truly original perspectives for research. This belief is the starting point of the paper. In the following pages, I present the opportunities offered by a specific type of “organic data” collected from online dating sites. I briefly review the traditional methodological approach to couple formation before turning to the limits and interest of studying these matchmaking services.

a. **Assessing couple formation: conventions and inventions**

The study of couple formation is largely based on census and survey data on married or cohabiting couples. This type of data is used not only to characterize existing relationships (degree of homogamy and endogamy, size of the age gap etc.) but also to seize partner preferences and the organization of the “marriage market”. In the latter case, the choice of data implies a retrospective approach. Rather than observing partner matching as such, this process is addressed indirectly through the study of couples already formed. Such an approach necessarily implies the use of hypotheses. Since the data only describes an output (the couple) and not the process leading up to this result (dating and matching), the researcher is forced to apply theoretical models in order to describe conjugal behavior.

The literature on the age gap between spouses illustrates well this retrospective approach. Using national censuses, survey data or marriage registries, researchers observe an average age gap in favor of the man. The interpretation of this observation is generally that of strong preference for age asymmetry: women look to older men while men are interested in younger women. These gender specific partner criteria are not observed but inferred from the characteristics of existing couples. The underlying assumption is that couples are “revealed preferences”, *i.e.* they reflect women’s and men’s expectations with regard to a spouse.
However, it is possible that the age gap primarily reflects the will of one partner that either has more pronounced age preferences than the other, or greater opportunities to see these preferences satisfied. Such discrepancies cannot be observed ex post facto. Our understanding of how the age asymmetry comes to be – and more generally how couples are formed – thus relies on behavioral models rather than on empiric research.

To remedy this lack of knowledge on the premises of couple formation, two main methods have been used. First, some surveys include questions on partner preferences in addition to questions on actual partner characteristics. A French study conducted in 2013, for example, asked respondents what qualities they valued in their partner when they first met, and if they would have easily accepted a spouse that was younger or older than themselves (cf. Rault & Régnier-Loilier 2015). This data gives empirical information on partner criteria but it still relies on a retrospective approach. Respondents were asked to recollect first impressions and to imagine alternative scenarios of couple formation. The apparent risk with such an approach is that former and present opinions get conflated, and that the survey registers current attitudes (that of a married or cohabiting person) rather than past ones (that of a younger and single self involved in dating).

Another attempt to empirically grasp partner preferences has been to study personal ads or data from marriage agencies (Goode, 1996; Ní Bhrolcháin, 2000; Singly, 1984). This method has the advantage of focusing on single individuals presently looking for a spouse. It allows researchers to apprehend partner preferences in the state of preferences, on the one hand, and “in action” on the other hand: e.g., partner criteria registered in a personal ad are done so to identify desirable partners and are not sheer declarations. The disadvantage of the method is the difficulty to generalize from the observations. Contrarily to survey data, the studies are not based on representative samples. Moreover, we know that personal ads and marriage agencies count for a small minority of couples (Laumann, Gagnon, Michael, & Michaels, 1994) and that they are mainly used for repartnering (Bozon & Héran, 1987). These features limit the possibilities to draw conclusion with regard to the general population.

However, as this paper will try to demonstrate, today’s online dating sites offer new opportunities to study couple formation. Since ten years back, researchers have used data from these sites to assess assortative matching (Hitsch, Hortacsu, & Ariely, 2010), educational homophily (Skopek, Schulz, & Blossfeld, 2011) and racial preferences across countries (Potârcâ & Mills, 2015). Although similar to older types of matchmaking services, online dating is a more promising field of investigation as it gathers more users than any other dating service, and because it relies on interactive platforms.
b. The opportunities offered by online dating

Unlike personal ads or marriage agencies, online dating is a widespread practice in most Western countries. In 2013, 11% of all American adults had used this type of service, 38% of those who declared to be single and looking for a spouse at the time of the survey (Smith & Duggan, 2013). The figures are similar for France: 14% of all individuals aged 26-65 years old had used an online dating site in 2013, 42% of those who were single and desirous to be in a relationship at that date (Bergström, 2016). These high rates of usage make online dating an interesting entry to the study of couple formation. It does not elude the problem of generalization: data from these sites is not representative of the general population, and although users are diverse with regard to age, social class and marital background, usage is still predominant in certain groups. However, the growing number of users strengthens the empirical salience of this type of data as the observations are indicative of behaviors of a growing part of the population. When the aim is to go beyond mere descriptions of existing couples, and to gain knowledge on the mechanisms of partner selection, this data can thus provide valuable information.

First, online dating is informative of partner preferences. Users are asked to give detailed sociodemographic information on themselves (gender, age, place of residence, marital status, occupation, education, race, etc.) as well as of partners of interest. Researchers cannot guarantee the accuracy of this information. Yet, studies conducted in order to cross-check the attributes displayed online with the actual characteristics of the users tend to indicate that deceiving behaviors are in fact limited (Toma, Hancock, & Ellison, 2008). This observation, as well as the mass of data registered, gives the researcher a good reason to believe that the tendencies observed are not random or misleading, but more likely informative of actual tendencies in partner criteria. Online dating then offers unprecedented opportunities to study people’s expectations in dating. This is all the more the case as the large number of users allows for detailed analysis of variations in preferences across numerous subgroups.

Second, online dating allows for the study of partner matching. Unlike personal ads, these sites are interactive and thus make it possible to not only identify partner preference but observe how these come into play on the dating “market”. Metadata on contact behaviors – who contacts whom, who responds to whom? – gives indications on how individuals behave with regard to their preferences, and how these are received (accepted or rejected) by potential

---

1 In France for example, 16% of executives and persons in higher-level occupations had used an online dating site in 2013, versus 13% of manual workers (Bergström, 2016).
partners. This is also an advantage with regard to surveys on already formed couples. Unlike this habitual approach, online dating captures the very process of matching and registers both rejection and selection in dating. It provides the missing link between stated preferences (intentions) and couple formation (outcome) by registering interactions (coupling).

The argument of this paper is not that online dating data – or other types of “organic data” – can or should replace census or survey data. These remain crucial to the study of couple formation. Because they focus on existing relationships, however, these types of data do not tell us how couples are actually formed and oblige researchers to resort to theoretical models of partner choice. Online dating can give empirical strength to these models. Without being based on representative samples, it can give indications about conjugal behavior that help consolidate hypothesis on partner selection. I join British demographer Maire Ní Bhrolcháin in her conclusion that the “theoretical analysis of the marriage market is in need of a solid empirical foundation” (Ní Bhrolcháin, 2000, p. 936). The argument of this paper is that online dating data can contribute to this. Using data from a major dating site, I hope to show the merits of this approach that I apply to the study of age preferences.

3. Data & Methods

The data used in this study comes from the online dating site Meetic.fr which is the French version of Match.com. The site was chosen for its numerous users. Meetic.fr is the biggest online dating site in France with more than 16 million user accounts registered between 2002 and 2014. The data was acquired from the company Meetic Group in 2014 and comprises two sources of information.

The first set of data contains all information displayed in the user profile, except pseudonyms and photographs. It also contains information regarding the use of the profile: when it was created, using what device (computer or mobile), whether the profile belongs to a paying member or not, and when the user last logged onto the site. In this paper I present results regarding a subset of profiles all created in 2014 and attributed to heterosexual individuals aged 18 to 70 years old (N : 2 339 032). More specifically, I look at age criteria reported by users in these profiles. Providing one’s age preferences is mandatory on Meetic.fr: to register an account and access the platform, the future users must indicate an age bracket for potential partners ranging from 18 to 99 years age. Within these limits, users freely indicate the interval of their choice and they can reduce it to 0 (45 to 45 years old for

---

2 The French company Meetic bought the European versions of the site Match.com in 2009. Since then, Match.com and Meetic.fr are based on the same platform in France.
example). This information allows me to study age preferences among women and men and across different age groups.

The second set of data contains information regarding user interactions. It comprises metadata regarding all emails exchanges on the platform. No content is registered in the database, only characteristics of the email sent: user identifiers of the sender and the receiver, date and time of the email sent and information on whether the sender and/or the receiver were paying members at the time of the interaction. In this paper I focus on emails exchanged during 2014 by individuals aged 18 to 70 years old, and I only look at first contacts and replies (N: 28 376 707). Matched with the profile information, this interactional data make it possible to see who contacts whom and who responds to whom. I use this information to study age differences between senders and receivers engaged in heterosexual interactions.

4. Results: The two-sided story behind the age gap

The age difference between spouses is a strong tendency of couple formation throughout the world. It characterizes a majority of heterosexual relationships in the countries where these partner characteristics have been observed (Mignot, 2010). In the case of France, this age gap has narrowed during the twentieth century, though without disappearing. In 2013, men were on average 2.3 years older than their female partners (Rault & Régnier-Loilier, 2015).

The strength of this demographic tendency has lead researchers to consider it as an organizing principle of the marriage market. In other words, the age gap is regarded as a pronounced preference among both women and men – one that determines mate selection. This interpretation is common in the field of Demography and is central to the literature on what is known as the “marriage squeeze” (Akers, 1967; Fraboni, 2004; Glick, Heer, & Beresford, 1963; Schoen, 1983). The idea that the age gap is central to partner choice is also common in Sociology. French sociologist Jean-François Mignot resumes this idea, claiming that “the fact that there is an average age difference between spouses in favor of the man in (almost) all known human societies seems to fundamentally derive from the fact – apparently also universal – that men prefer young women and that women prefer mature men” (Mignot, 2010, p. 316).

---

3 The age bracket by default indicated by the site is 25 to 45 years old. Among the profiles studied, only around 6% indicated this specific age bracket. Thus, a large majority of the members have changed and personalised the information regarding their age preferences.
Drawn from data on existing couples, this conclusion remains however a hypothesis. Female and male age preferences have in fact rarely been studied as such. In the following section, I test this hypothesis by studying age preferences as they are expressed by online dating sites users, in their profiles and through their contact behaviors. The analysis indicates that the age gap between spouses can hardly be regarded as a strong preference among women and men. Depending on their own age, users state different preferences that do not always promote an age gap in favor of the man. I first give a brief outline of the nature of male and female partner criteria and then show how these come together to produce an age asymmetry between the sexes.

a. Contrasting attitudes towards male superiority in age

Among the female users of *Meetic.fr*, young women stand out as having the strongest preference for a traditional age gap (see chart 1). Only a little more than one forth (26%) of the women between 18 and 24 years old consider dating a younger man. This reluctance towards younger partners is however specific to this group of women. Between ages 40 and 49, a vast majority of women accept the idea of a younger man (80%). As of age 60, the proportion is even higher (89%). At the same time, the exclusive preference for older men is rare among older women over 60 (4%), while the exclusive preference for a younger partner is rather common (18%).

<table>
<thead>
<tr>
<th>Female users</th>
<th>Accept younger partners</th>
<th>Exclusive preference for younger partners</th>
<th>Accept older partners</th>
<th>Exclusive preference for older partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-24 years old</td>
<td>26</td>
<td>1</td>
<td>98</td>
<td>42</td>
</tr>
<tr>
<td>25-34 years old</td>
<td>62</td>
<td>1</td>
<td>99</td>
<td>21</td>
</tr>
<tr>
<td>35-39 years old</td>
<td>69</td>
<td>1</td>
<td>98</td>
<td>19</td>
</tr>
<tr>
<td>40-49 years old</td>
<td>80</td>
<td>5</td>
<td>91</td>
<td>10</td>
</tr>
<tr>
<td>50-59 years old</td>
<td>84</td>
<td>12</td>
<td>84</td>
<td>7</td>
</tr>
<tr>
<td>60-70 years old</td>
<td>89</td>
<td>18</td>
<td>75</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Male users</th>
<th>Accept younger partners</th>
<th>Exclusive preference for younger partners</th>
<th>Accept older partners</th>
<th>Exclusive preference for older partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-24 years old</td>
<td>69</td>
<td>1</td>
<td>95</td>
<td>12</td>
</tr>
<tr>
<td>25-34 years old</td>
<td>96</td>
<td>4</td>
<td>92</td>
<td>1</td>
</tr>
<tr>
<td>35-39 years old</td>
<td>98</td>
<td>9</td>
<td>85</td>
<td>1</td>
</tr>
<tr>
<td>40-49 years old</td>
<td>98</td>
<td>21</td>
<td>70</td>
<td>1</td>
</tr>
<tr>
<td>50-59 years old</td>
<td>99</td>
<td>38</td>
<td>53</td>
<td>1</td>
</tr>
<tr>
<td>60-70 years old</td>
<td>99</td>
<td>51</td>
<td>38</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: the proportions are calculated from the age preferences (minimum and maximum age) stated in the user profiles.
Among male users, it is again the young ones who differ from the others, but this time by a relative indifference regarding the traditional age gap in favor of the man. More than nine men out of ten, aged 18-34, indicate a maximum age above their own. However, the proportion of men considering an older partner progressively decreases with age, going from 95% amongst the 18-24 year olds to only 38% amongst the 60-70 year-olds. As men age, their desire for a younger partner becomes stronger.

The first conclusion that can be made from this online dating data is that age preferences appear more flexible than what is often suggested (Ní Bhrolcháin, 1992, 2000). The analysis goes against the idea of an unequivocal preference for an age gap in favor of the man. Men’s and women’s attitudes towards this age difference vary according to their own age. This is all the more clear when age preferences are no longer deduced ad hoc from the age gap between spouses, but observed in the state of preferences. Moreover, age criteria evolve in a similar manner for both sexes. With increasing age, the desire for a younger partner becomes stronger for both women and men. This means that men’s and women’s age preferences are more similar than is often assumed.

b. The age gap as the result of a negotiation between the sexes

The diverging nature of male and female partner preferences suggests that the production of the age gap is a complex process. In order to shed light on this process, we need look at the interplay between female and male age criteria (instead of studying them apart). Chart 2 maps the age preferences of men (indicated in blue on the y-axis) and of women (indicated in red on the x-axis). The stronger the preferences are, the darker the colors are: the dark blue fields correspond to age criteria often indicated by men and the dark red fields correspond to age criteria often indicated by women. The chart also plots (for each gender and each age) the mean maximum age indicated by men (blue line) and women (red line), as well as the mean minimum age indicated by men (blue dotted line) and women (red dotted line). At last, the chart maps the area of intersection of men’s and women’s age criteria: the purple field indicates where the preference of one sex meets that of the other.
The chart gives us valuable indications about how the age gap between spouses is produced. First, the area of intersection of male and female age criteria (in purple) indicates that the age gap is the result of a negotiation between the sexes. Rather than the product of a common preference, it is the meeting point between two wills that tend to diverge. The clearest example of this is found among the older users: male and female preferences pull in different directions, but coincide in favor of a small age gap in favor of the man.

Second, the mean maximum and minimum ages (indicated in full and dotted lines), tell us that one sex is often more “responsible” for the age gap than the other. In early adulthood, it is women's preferences that underlie the age gap (the minimum age they indicate promotes this gap). As of age 40, it is men instead who push for this age difference in their favor (the maximum age they indicate underpins this gap). The age gap thus seems to result from the desire of one partner that is more or less imposed onto the other – the gender being dependent on the time of couple formation.

This interpretation is strengthened by the analysis of male and female contact patterns. Chart 3 shows the proportion of emails sent (as a first contact or a first reply), by each gender.
and by age groups, to partners younger, same age or older than the user. The results confirm
the conclusions drawn from the analysis of the stated age preferences. The contact behaviors
indicate that at younger ages – corresponding to first unions – women are the main actors in
creating an age asymmetry in favor of the man. On Meetic.fr, between 70% and 90% of the
emails sent by females aged 18 to 34 years old are sent to older users. The tendency is then
progressively inversed. At older ages – corresponding to repartnering – men appear as the
main initiators of the age gap. At ages 40-59 years old, around 60% to 75% of the emails sent
are done so to younger women. Once again the study reveals women’s and men’s contrasting
attitudes towards male superiority in age, and a converging interest for younger partners with
increasing age.

Chart 3: Contact patterns on Meetic.fr by gender and age

5. Discussion

Online dating offers a unique opportunity to study couple formation “in the making”. It makes it possible to capture not only partner selection but also rejection in dating. As such, it gives us a better understanding of how the “marriage market” actually works. Regarding the production of the age gap between spouses, the study shows that male and female preferences, far from always being compatible, are often in tension with each other. Couples – and the age gap that characterizes them – are the product of a negotiation between the sexes whose aspirations are potentially divergent and regarding which concessions are often necessary. Thus, the age asymmetry in heterosexual relationships can hardly be considered as a "revealed preferences", i.e., as indicative of the desire of both sexes. It must rather be understood as the meeting point between two accommodated aspirations. Accordingly, the age gap does not appear to be an organizing principle of the marriage market. This asymmetry between the sexes is in fact unequally desired by women and men at different periods in life.

References


