**Spatial homogamy and online dating**

Leena Maaß 1 & Andreas Filser 2

1 SOCIUM?

2 Institute for Employment Research, Nuremberg, Germany

# **Abstract**

Spatial homogamy has gained increased attention in relation to mate choice in recent years. However, it is still an under-researched dimension in homogamy studies. People tend to choose a spatially proximate partner. Theoretically, in the context of digitalisation, it has become possible to find a partner anywhere in the world, so the importance of geographical distance may have diminished over time. The aim of the study is to show the development of geographical distances in partner choice between the online and offline meeting context in a longitudinal section for Germany. With the help of the pairfam dataset, the individual partnership histories of waves 2 to 12 are collected (2009 to 2019). In a second step, the pooled OLS regression results are tested for robustness using a fixed-effects model. The empirical results indicate a difference in distance between the two identification locations. Over time, the distances for partnerships initiated online have decreased, which means that the proportion of spatial homogamy has increased. In contrast, the distances between offline initiated partnerships have increased over time. This means that the proportion of spatial homogamy has decreased. The various contact opportunities therefore influence spatial homogamy. With a few exceptions, the current relevance of having a partner close by is confirmed.

# **Introduction**

Over the last two decades, the internet has profoundly changed dating behaviour (Finkel et al. 2012). Online dating has now replaced friends as the most common social mediator in partner choice (Rosenfeld and Thomas 2012; Rosenfeld et al. 2019, pp. 17754). The rise of online dating has shifted how couples meet, including a potential 'death of distance' (Cairncross 1997): while partner markets once were limited to individuals close-by, internet dating has expanded dating opportunities to more distant partners (Stoye et al. 2014, p. 91). Online daters now have the largest potential selection of partners at their disposal, and it is easy to make contact (Finkel et al. 2012). The elimination of space-time restrictions in online dating means that daters are theoretically no longer limited to the pool of individuals they physically meet (Bossard 1932; Skopek et al. 2009; Wellman 2001). On the internet, dating communication can span regions over even continents (Haandrikman 2019, p. 440). On the other hand, most online daters still prefer to meet their romantic counterpart in person. Thus, even online daters will still prefer find a partner from close-by and online dating might rather help to daters to better find a suitable partner from their region. Whether online dating has led to a shift in spatial partner markets has not yet been answered (Eckhard and Stauder 2019, p. 3).

In this study, we to study how online dating as an opportunity structure affects distance when choosing a partner.

Partner choice

People tend to choose a partner who is both similar to them and comes from their immediate geographical environment. The tendency to choose or marry a person who is similar to oneself in terms of various characteristics is known as homogamy (Eckhard and Stauder 2019, p. 2). The choice of a person within a nearby geographical area is referred to as spatial homogamy. Haandrikman and Hutter (2012, p. 241) refer to spatial homogamy as "similarity between partners concerning their geographical origin". At the beginning of research into the spatial proximity of partnerships in the 1930s, the pattern of 'distance decay' was repeatedly demonstrated (e.g. Bossard 1932; Coleman and Haskey 1986). Distance decay means that the greater the distance between potential partners, the less likely it is that they will choose a partner (Bossard 1932). Due to the lack of current literature, the question arises as to how distances have developed in partner choice in Germany.

Spatial homogamy may have diminished over time, due to increased mobility, growing prosperity and rising levels of education. Cairncross (1997) refers to this insignificance of distance as the 'death of distance'. Although space and time play a subordinate role in this perspective, the costs that an individual has to invest in a long-distance relationship in order to maintain it increase (Samers 2010). From the perspective of cost-benefit maximisation, this means that individuals prefer a partner close by, as the investment required to maintain the relationship is lower (Catton and Smircich 1964). This makes it less likely that partnerships will be formed from distant locations. Even if relationships begin in the virtual environment, the search for a partner is usually limited to a short distance, as most partnerships are lived in reality (Klein and Stauder 2016, p. 271).

Differences in spatial homogamy between online and offline dating have only received little attention in the literature so far (see Potarca 2020 for an exception). Due to this research gap and the question of what influence the opportunity of online dating has, this thesis deals with the significance of geographical distances in partner choice in Germany. The aim of this study is to determine the geographical trends in partner choice in a longitudinal section over the years 2009 to 2019. The resulting research question is: To what extent has the distance in partner selection between online and offline initiated partnerships changed over time?

The online dating term used in this study includes meeting people via dating sites as well as dating apps and social networks. This research question is investigated with the help of the pairfam family and relationship panel (Brüderl et al. 2020).

Description Structure of the paper

# **Theoretical considerations on the process of partner choice**

Theoretical perspectives on the spatial patterning of partner choice build on the concept of partner markets and Blau’s structural theory. …..is used to look at contact opportunities in the context of partner choice (Blau 1977b; 1977a; Blau and Schwartz 1984), which are pre-structured by the social structure (Stauder 2015, p. 402). As his consistently macro-sociological approach does not take individual action and social integration into action contexts into account, the theoretical perspective is extended by Scott Feld's focus theory (1981), which includes a smaller space and thus a concrete social environment.

Partner markets are understood as meeting contexts for individuals that predefine interactions with potential partners and in which they can get to know each other (Häring et al. 2014; Kalmijn and Flap 2001; Klein 2015; Stauder 2008). Klein and Stauder (2016, p. 256) further define the partner market as "a spatial-social delimitation and description of the space of possibilities".

On the one hand, this possibility space forms the basis for the formation of partnerships (Blau 1977a; 1977b; 1994; Feld 1981 and structures partner markets (Blau et al. 1982; 1984). In this perspective, people enter into partnerships with similar partners "because people in similar social positions have similar social experiences, assume similar roles and exhibit similar characteristics and attitudes that increase the likelihood of a relationship" (Stauder 2015, 403, quoted from Blau 1977a, p. 36). He also assumes "that the social differentiation of the population influences the patterns of people's relationships" (Stauder 2008, p. 266, quoted from Blau 1977a; 1977b). In the multidimensional space assumed by Blau, the "Blau space", individuals occupy a social position according to combinations of characteristics such as age, ethnicity or education. Social and spatial proximity within the "blue space" is an important criterion for whether individuals get to know each other and increases this probability (Blau 1977b, p. 31, p. 251). Depending on the social distribution of individual characteristics, it is more or less likely to meet a partner with the desired characteristics (Klein and Stauder 2016, p. 256). Similar people have an increased probability of interacting and developing a relationship due to the aforementioned structuring (Blau et al. 1984). Therefore, social interactions result from the opportunity structures of getting to know each other (Blau and Schwartz 1984, p. 29). Partner selection preferences can be adjusted by social structural restrictions if people with the desired characteristics are not available on the market.

It is critically noted that Blau's space of possibilities is very broad, as it is "defined by national borders and the social structure of society" (Klein 2015, p. 328). This wide-ranging view of the partner market does not reflect the actual meeting possibilities in the concrete environment of individuals, as the formation of partnerships is focussed on spatial proximity (Bossard 1932; Catton and Smircich 1964; Clarke 1952; Lengerer 2001). This means that a small-scale approach is required.

In contrast to Blau, this "small-scale social environment" (Zillmann 2017, p. 34) is taken into account in Scott Feld's (1981) focus theory. It has been shown that partnerships arise primarily in action spaces in which a continuous exchange takes place (Hill and Kopp 2006, p. 165). This allows the living conditions of individuals to be better mapped. According to Feld (1981), individuals get to know each other through so-called *focus of activity*. A focus is "a social, psychological, legal, or physical entity around which joint activities are organised" (Feld 1981, p.1016), which can be, for example, the workplace, the family or regular activities.

The probability of a relationship increases with the interaction in the context of activities in the same focus, as the foci pre-structure encounters with potential partners. There is an opportunity structure within these learning locations, which influences individuals in various ways or limits or expands their options for action (Zillmann 2017, p. 33f). Individuals are therefore integrated into action contexts through their workplace, for example. These realise various encounters (Feld 1981) and thus have an effect on the "field of the selectable" (Winch 1955), which pre-structures the opportunities for interaction. Depending on the partner selection preference, the action contexts are designed differently in order to come into contact with the preferred characteristics (Kalmijn and Flap 2001; Stauder 2008).

The two theories outlined above show how contacts arise during the search for a partner. The opportunity structures have an effect on individual behaviour. On the one hand, they open up various options for action, but on the other hand, they can also be accompanied by restrictions (Wirth 2000, p. 50). Furthermore, it is shown that people with similar characteristics find themselves in similar opportunity contexts, which in turn are differently accessible for certain groups. However, since the partnership formation does not only depend on the meeting context, Chapter 0 includes other factors that are important for spatial homogamy. "[W]hile meeting depends on opportunities, mating depends on both attraction and opportunities." (Verbrugge 1977, p. 577).

# **Partner choice**

There is a large body of literature on patterns of partner choice in both offline and online encounter contexts.

## **Empirical factors influencing partner choice**

As outlined in the theoretical considerations in Chapter 2, opportunity structures influence the possibility of meeting a partner, but partner choice preferences and cultural norms also influence partner choice in the sociological literature (Kalmijn 1991). In this context, Haandrikman and Hutter (2012, p. 241) state that "[e]ach of these factors may be influenced by geographical distance".

Individuals tend to choose a partner who is similar to them in terms of their characteristics and traits (e.g. Kalmijn 1994, 1998). These partnership preferences limit the supply of suitable partners and the preferences of the other person determine the individual's resources on the partner market (Klein 2015, 326, quoted from Bourdieu 1976). The choice of a (similar) partner also has far-reaching consequences. Patterns of partner choice reproduce social structures (Klein 2015, p. 326, quoted from Bourdieu 1976; Skopek et al. 2009). If partners belong to a different social class or have an unequal social background, this leads directly to social or professional advancement or relegation. Marriages between a native and a migrant promote integration (Klein 2015, p. 326).

While evidence on For Germany, homogamy patterns have been documented in terms of education (Blossfeld and Timm 1997; Timm 2004) and age (Klein 1996; Klein 2005).

The predominant educational similarity of couples is explained by the selective education system, which creates more homogeneous groups. Due to the opportunities for contact that arise in the education system, it is not surprising that partnerships are formed between people with the same or similar qualifications (Timm 2004). In addition, people with a similar socio-economic background tend to concentrate spatially. Furthermore, age has a compositional effect on the proportion of spatial homogamy. The availability of potential partners in certain age groups on the dating market influences the chances of meeting. If women or men do not come into contact with men or women of the same age on the partner market due to different birth rates, this extreme case is referred to as 'marriage squeeze' (Bhrolchain 2001).

* + 1. **Spatial distance in partner choice**

Much of the sociological literature dealing with the spatial proximity of spouses is more than 70 years old (Bell 1957; Bossard 1932; Clarke 1952; Coleman and Haskey 1986; Davie and Reeves 1939; Ellsworth 1948; Koller 1948). S Subsequently, for a long time, no attention was paid to geography in partner choice (Lengerer 2001, p. 134). A handful of recent articles are once again attaching greater importance to its role (Haandrikman et al. 2008; Haandrikman 2019; Lengerer 2001).

The pioneering study by Bossard (1932) highlighted the important role of spatial homogamy in partner choice among couples in Philadelphia. This is because proximity increases the likelihood of spontaneous encounters, meaning that couples are more likely to form in the neighbourhood. Replications for other cities in the United States confirm this concentration of marriage circles (e.g. Abrams 1943; Clarke 1952; Davie and Reeves 1939; Harris 1935).

More recent studies confirm these findings: ,cohabitation (Haandrikman et al. 2008). In Sweden, half of all couples live a maximum of 9.2 kilometres apart and the average distance to their partner is 57 kilometres (Haandrikman 2019). For Germany, one study reports that the majority of couples meet within a radius of 20 kilometres (Lengerer 2001).

However, spatial homogamy has been shown to vary across contextual and individual factors. Already early studies report that in rural areas, fewer couples live in the immediate vicinity of each other (Bell 1957; Ellsworth 1948; Schnepp and Roberts 1952). Similarly, population size and density is associated with spatial homogamy in the Netherlands, Sweden, and the Netherlands (Lengerer 2001 Haandrikman et al. 2008; Haandrikman 2019). Explanations for this urban-rural devide evoke differences in the opportunity structures: urban areas offer a wider variety of focal points such as jobs, clubs and educational institutions where people meet potential partners. In rural areas, on the other hand, there are fewer potential partners in the immediate vicinity. *Villagers consider a certain distance to their potential spouse within a certain radius, but only go as far as they have to (ibid., p. 446). Bell (1957, p. 74f) also shows that the greater the distance, the fewer local people marry a non-local person. In addition to the factors mentioned above, this effect is favoured by geographical location or weakened by geographical barriers such as rivers (ibid., p. 77). Geographical factors can therefore also influence spatial homogamy in partner choice.*

In addition, individual-level socio-demographic characteristics correlate with spatial homogamy. Harris (1935) and Koller (1948) showed that the geographical distance between spouses increases with a person's higher occupational position. As a rule, higher social classes and levels of education are associated with a greater distance to the spouse (e.g. Coleman and Haskey 1986; Clegg et al. 1998; Haandrikman et al. 2008; Lengerer 2001). Harris (1935, p. 259) explains the growing marriage radius with the increasing mobility resulting from the higher professional position and the associated larger geographical areas in which they can search for potential partners. On the other hand, Harris (1935, 258) shows a non-linear effect of age. On average, the distance to the partner decreases with increasing age (Clegg et al. 1998; Coleman and Haskey 1986; Harris 1935; Haandrikman et al. 2008). Partnerships "that are formed over a greater distance [...] [are] consistently more heterogamous in terms of age and education. [...] [Lengerer also assumes] that the action contexts in which potential partners from distant places of residence meet are less homogamously pre-structured" (Lengerer 2001, p. 159).

In addition to the factors mentioned above, such as age, level of education and place of residence, cultural factors also contribute to changes in spatial homogamy. People living with their parents and single parents travel a shorter distance to their partner (Haandrikman et al. 2008, p. 396).

However, no distinction is made here between where the partners met. Geographical as well as socio-economic and demographic factors explain the varying distances between couples in Sweden. People who live in the city, were born abroad, have a low degree of mobility or whose parents live nearby travel the shortest distances to their partner (Haandrikman 2019).

In addition, socio-economic and demographic factors and spatial conditions such as population density, degree of urbanisation and spatial barriers influence meeting opportunities, which can affect the distance to the partner in different ways. Spatial barriers make it more difficult to reach partners, which in turn leads to a longer distance. Thus, spatial determinants lead to regional differences in the proportion of spatial homogamy. It has also been shown that the distance covered has not increased significantly over time, despite the increase in mobility and participation in education, meaning that there has only been a slight change in the expansion of the geographical radius (Haandrikman et al. 2008; Haandrikman 2019; Lengerer 2001). Accordingly, partnerships from an easily accessible area are still favoured in the present. Looking back at the theoretical considerations, the foci of mate choice have been limited to local places since the beginning of research on spatial homogamy. In the context of online dating, the question arises as to whether and, if so, how the opportunity structures have changed. A theoretical approach to this question follows in the next chapter.

## **Online dating**

In recent years, the importance of the Internet in the search for a partner has steadily increased. This is reflected in the high number of online dating users (e.g. Cacioppo et al. 2013; Potarca 2017; 2020; Rosenfeld and Thomas 2012; Rosenfeld et al. 2019; Thomas 2020) and the economic success (Skopek 2012, p. 31f.). However, valid data on the use of online dating sites is rare (Schmitz 2017, p. 16). Due to its popularity and the high number of users, online dating has become an important context for getting to know people when looking for a partner (Klein and Stauder 2016, p. 254; Skopek 2012, p. 30). In line with Feld's focus theory, online dating also represents a focus in the search for a partner. Due to its characteristics, online dating can also be described as a hyper-focus (see Chapter 3.2.1) (Schmitz 2014, p. 114).

* + 1. **Structural changes in online dating**

The primary difference between online dating and everyday encounters lies in their functions. Online dating acts as an intermediary in the search for a potential partner (see Schmitz 2014, p. 114). This means that the focus of online dating is on finding a partner, as contact between users is limited to getting to know each other. Unlike in everyday life, the use of online dating is associated with "a certain degree of certainty of expectation" (Skopek 2012, p. 76). By comparison, in an offline dating context, it is uncertain whether a person is available to the dating market or not (Stauder 2006). Skopek (2012, p. 76) describes online dating in terms of "the search for a partner [as] more efficient and much more systematic and goal-oriented than partial marriage markets in everyday life, in which the search for a partner can also be based on chance encounters in socially pre-structured contact networks."

Unlike in traditional meeting contexts, where "the partnership formation [...] is an unplanned side effect of the context-specific practice" (Schmitz 2014, 114), online users specifically choose this context for their partner search (Schmitz 2017, p. 33; Skopek 2012, p. 76f.). By shifting the search for a partner to the digital space, searchers are able to meet strangers who are not in their circle of friends or with whom they have no social ties (Rosenfeld and Thomas 2012, p. 524) and would otherwise not have met (Dutton et al. 2009). Since online dating acquaintances are loose contacts, the context of the behaviour is less embedded in social structures (Schmitz 2014, p. 114f). There is a "detachment of social interaction structures from [everyday] life" (ibid. 2014, p. 116). In addition, the space-time barrier is removed (Finkel et al. 2012, p. 4; Skopek et al. 2009) and the scope of action for users is expanded (Skopek 2012, p. 82). As mentioned at the beginning, time-shifted communication allows them to meet people from different places. By removing themselves from their everyday lives and social environment, those looking for a partner encounter a heterogeneous partner market. With regard to work, it should be noted above all that the geographical factor that determines the likelihood of meeting potential partners locally is eliminated in online dating (ibid. 2012, p. 77).

With public access to the online partner market, a very large number of potential partners are available to those looking for a partner in online dating. Unlike in offline dating markets, there is no pre-sorting according to individual characteristics such as age at school or level of education at work (Klein 2015, 339). Regional or institutional meeting places, such as neighbourhoods or educational institutions (Kalmijn and Flap 2001), are also far more homogeneously organised than the internet. This means that users have a large heterogeneous pool of partner opportunities at their disposal. This strong heterogeneity is characteristic of the hyper-focus mentioned at the beginning. As users use exclusion principles to select which characteristics are important to them when getting to know someone, such as gender, proximity search, education or age, people who do not match these characteristics can be sorted out prematurely. Zillmann (2017, p. 73) refers to this as negative selection.

In conclusion, it can be stated that the virtual focus (Feld 1981) has a weaker influence on the social structure and the resulting encounters than real encounter contexts due to its general access. Due to the lower pre-structuring of the encounters, the participants in the virtual partner markets have socially heterogeneous characteristics compared to real encounter contexts. In addition, contact between potential partners is not determined by geographical location or time of day, which expands the opportunities for contact. Individuals make their own selections according to criteria that are relevant to them. This means that individual preferences will be the main deciding factor when choosing a partner. People from urban centres in particular benefit from online dating, as they have a large number of potential partners at their disposal (see Hobbs et al. 2017, p. 272).

* + 1. **Empirical findings on online partner selection**

Homogamy studies in the context of online dating come to mixed results. According to Dutton et al. (2009), the use of the internet is associated with a greater difference between partners in terms of age and educational level. On the other hand, couples are more similar in their interests and values (ibid. 2009, 16). Furthermore, partnerships initiated online are less similar in terms of ethnicity, religion and education (Potarca 2017; Thomas 2020). However, partners are more similar in age (ibid. 2020, 1281). Overall, it is shown that "[...] the Internet has had a more limited (but not necessarily unimportant) direct impact on exogamy in the population as a whole" (Thomas 2020, p. 1281f.). The study of gender-specific partner preferences shows that geographical proximity is seen by both men and women as a luxury good when choosing a partner (Jonason et al. 2017). Partner seekers accept travelling to satisfy more relevant partner preferences. In short-term relationships, men place a stronger focus on geographical proximity to their partner. Women, on the other hand, invest more in the search for a partner who is geographically close in the long term (ibid. 2017, p. 590f. & p. 593).

In the study of dating behaviour, interactions between people from different regions across America are observed, but these are limited. When it comes to finding a romantic partner, partners from their own region are preferred (Bruch and Newman 2019). The national dating market is divided into local units in which geographical proximity plays an important role (ibid. 2019, p. 222). Within cities, demographic characteristics create submarkets that are differentiated by age and ethnicity. More men use online dating at a young age, whereas women use the medium primarily at an older age. On average, men are slightly more than 1.5 years older than women in all submarkets. Most interactions take place within a submarket or between two neighbouring units (ibid. 2019, p. 224).

In Switzerland, friends are still the largest social mediator in partner selection, but this is declining (Potarca 2020). The internet has seen a sharp increase in matchmaking and is similar to the figures from Rosenfeld and Thomas (2012) in the USA. Potarca's findings show that there are positive effects between both dating apps and dating websites and the distance to the partner. Compared to offline meetings, partnerships via dating apps are associated with a greater geographical distance for both medium and long distances. The extended search options make it easier to establish relationships with people living further away geographically. In the case of dating websites, an effect is mainly observed for long distances, which suggests that people are looking further afield. This means that relationships are more likely to develop outside of their own living space. In contrast to online dating, dating apps promote contact with geographically neighbouring areas (Potarca 2020, p. 13ff.) and thus increase spatial homogamy.

In conclusion, it can be said that the structural changes in online meeting places enable meetings from geographically more distant locations (Bruch and Newman 2019). In addition, empirical findings confirm that partnerships initiated online are associated with a greater distance than everyday meeting contexts (Potarca 2020). However, variations can also be found within the individual forms of online dating. However, these are not considered further in this study. The significance of spatial homogamy in online dating thus appears to be changing. Accordingly, online encounters are more likely to be associated with greater distances to the partner and the proportion of spatial homogamy is lower than in everyday encounters. In addition, gender-specific preferences have been demonstrated (Jonason et al. 2017). As the comparative literature on spatial homogamy in partner selection in online and offline contexts is scarce and no study looks at a longitudinal comparison, the question arises as to how the internet affects spatial homogamy. Chapter 3.3 generates a framework of expectations for this question.

## **Framework of expectations regarding the development of distance when choosing a partner**

Based on the theoretical considerations outlined above and the results of Potarca (2020), it can initially be assumed that partnerships initiated online are associated with a greater distance than partnerships that have met in person. However, cost-benefit maximisation provides an argument as to why the distance between couples will not increase indefinitely. Due to the limited scope of this study, this aspect will not be discussed in depth.

Everyday opportunities for encounters are limited to local places in the neighbourhood. For example, individuals get to know a potential partner at school who comes from the same catchment area. At the discotheque, people meet people who come from an easily accessible area and the meeting context of the neighbourhood means that people are directly in the room. Since empirical results have shown that people also tend to form partnerships from their spatial environment (e.g. Bossard 1932; Haandrikman et al. 2008; Haandrikman 2019; Lengerer 2001), offline initiated partnerships will be associated with a small distance at the beginning of the study and the proportion of spatial homogamy should be higher. As there is no significant evidence in the literature of a widening of the partner selection radius in recent years, it is expected that the distances between partners have remained stable over the study period in Germany.

In recent years, online dating has become an important focus in the choice of a partner and implies more opportunities to encounter potential partners. In contrast to everyday encounters, individuals' contact networks are not pre-structured in online dating, making it more likely that they will meet people who are not embedded in their environment (Dutton et al. 2009; Rosenfeld and Thomas 2012; Schmitz 2014). The elimination of space-time barriers in particular creates new room for manoeuvre for those seeking a partner. The geographical factor, which implies who a person comes into contact with, no longer applies when getting to know someone via online dating (Skopek 2012). Users are free to decide the radius within which they want to search for potential partners using the proximity search. The radius will presumably differ between men and women (Jonason et al. 2017). Due to the theoretical reasons and the empirical evidence (Potarca 2020), partnerships initiated online will on average travel further distances to the partner.

At the beginning of the study period in 2009, it was therefore expected that relationships initiated online would cover a greater distance than couples who met in everyday life. At that time, online dating still played a subordinate role in dating in Germany. Accordingly, this opportunity structure was used by fewer people. At the end of the study in 2018, online dating was very popular. Since 2009, the number of online dating users has doubled.[[1]](#footnote-3) As a result of the increase in online dating users, more potential partners live in the local area and a larger pool of partners is available to the individual locally. There may be gender-specific differences in the relevance of geography when choosing a partner (Jonason et al. 2017). Over time, the average distances to the partner should therefore tend to decrease. This would result in an increase in spatial homogamy.

The following explorative expectations can be generated from the theoretical argumentation:

*Expectation 1:* At the beginning of the study period, online-initiated partnerships cover a greater distance to their partner than offline-initiated partnerships.

*Expectation 2a:* Over the course of time, the distance decreases for partnerships initiated online.

*Expectation 2b:* Over time, the distances travelled by partnerships initiated offline will not change significantly.

From expectations E2a and E2b, it is concluded that the closer it gets to the end of the study period, the more the distances between the online and offline groups will equalise. Overall, the distances travelled to the partner will vary depending on the place of residence and the level of education.

In the following chapter, the data and methods used will be discussed in order to answer these expectations.

# **Data and method**

After defining the theoretical framework, the next step is to test the expectations. Chapter 4 presents the methodological focus of the work. Firstly, the data used is presented in Chapter 4.1. The methodological approach is then outlined in section 4.2 and the data selection in section 4.3. Following on from this, the variables are operationalised in chapter 4.4. Finally, section 4.5 discusses the limitations of the work.

## **Data**

The German relationship and family panel called pairfam is used to test the expectations. Pairfam is a nationwide, randomised longitudinal study of relationships and families in which annual surveys have been conducted since 2008 with three birth cohorts (1971-73, 1981-83 and 1991-93). In the first wave, 12,402 respondents took part in the survey. In addition to the cohort approach, which ensures a representative initial sample, the pairfam design is characterised by the multi-actor approach. The respective partners of the ancestor as well as the parents or stepparents and the children living in the household (from wave 2) are also surveyed. The annual surveys are referred to as waves (Brüderl et al. 2020).

The pairfam's research interests include partnerships, parenthood and fertility as well as family and intergenerational relationships and behaviour. In addition, parenting and child development as well as social embedding are of interest. The data set also offers a broad spectrum of socio-demographic characteristics. Changes are recorded retrospectively for the last year in the survey. The partnership, employment and housing histories are available in an event-history format (Brüderl et al. 2021), so that all events in the life course of the individual are included. The data to be analysed are available in a panel structure, so that several "consecutive observations" are available for each individual (Schröder 2007, 261).

A detailed description of the relationship and family panel can be found in Huinink et al. (2011).

## **Method**

First, the data is analysed using a pooled OLS regression. A fixed-effects model is calculated to control for possible distortions. Due to gender differences, as discussed in the literature, the analysis considers models separately by gender.

* + 1. **Pooled OLS regression**

Regression analyses are the most frequently used statistical analysis method. On the one hand, they are used to quantitatively describe correlations and, on the other, to estimate the values of the dependent variables. In addition, they analyse causal relationships (Backhaus et al. 2016, p. 64), whereby a linear relationship is assumed. The ordinary least squares (OLS) regression method is used to calculate the best-fit regression line (see Diaz-Bone 2019, p. 96ff.).

In the pooled OLS regression, the existing panel character of the pairfam data is ignored (Stoetzer 2020, p. 237), which means that the data is considered cross-sectionally (Giesselmann and Windzio 2012, p.27). For the work, this means that individuals with several partnerships at different points in time are assumed to have relationships with different people and not with the same observation unit (Park 2011, p. 7). The analysis distinguishes between an online and an offline place of identification. The estimation of the coefficients is used to calculate whether and to what extent the two analysed groups differ from each other in the distances involved in pair formation (Stoetzer 2020, p. 238f.).

Since an empirical situation can never be fully modelled by the limited explanatory variables in an OLS model, unconsidered influences can affect the model. This problem is described in more detail in the limitations in section 4.5.2. As a result, the estimated distances are controlled by a fixed-effects regression.

* + 1. **Fixed-effects regression**

In the fixed-effects model (FE model), the panel character of the data is maintained in comparison to the pooled OLS regression. The data is transformed in order to depict the potential for longitudinal questions (Giesselmann and Windzio 2012, p. 28). The fixed effects models are used to analyse the intra-individual changes in the distances of all relationships of an individual over time (for more details on estimating fixed effects see Brüderl and Ludwig 2014; Giesselmann and Windzio 2012; Wooldridge 2010).

The advantage of FE models is that only intra-individual changes over time are used for the estimation. In FE models, time-invariant characteristics, such as gender or birth cohort, are not quantified due to the lack of variation in these time-constant variables (Giesselmann and Windzio 2012, p. 44).[[2]](#footnote-4) As a result, the unobserved heterogeneity or endogeneity of independent variables caused by time-constant variables does not distort the coefficient estimates in the FE model (Stoetzer 2020, p. 238).

This means for the work: The cross-sectional estimates of the effects of the mate learning location on the distance to the partner could be distorted by unobserved differences. Particularly in the context of analysing partner markets, the availability of individual groups can influence the estimates. Consequently, an OLS regression would lead to biased estimates of the effects of meeting location on distance to partner as long as factors that influence both meeting location and distance to partner are not controlled for. Estimates of regressions with fixed effects would not be biased, as they are not based on the comparison of offline and online initiated partnerships, but on the comparison of intra-individual differences.

The fixest package (0.9.0.) was used in RStudio to calculate the FE models.

## **Data selection**

Only the Ancor dataset is used for the analysis, as all relevant partnership events and variables for the research question can be generated from it. The unit of investigation to be analysed is the individual. In the work, the individuals are equated with their partnerships.

The analysis takes into account all new partnerships entered into during the period under investigation. Newly formed partnerships are partnerships that were newly entered into during the last wave. The newly formed partnerships are identified by the place where the person met their partner (pa3). A distinction is made in the study between meeting in an online and offline meeting context. Thus, the variable pa3 serves to identify all observations on the one hand and, on the other hand, the meeting location represents the independent variable in order to explain the object of investigation of distance.

As just described, only partnerships that were newly formed during the last wave are analysed. Accordingly, waves 2 to 11 (2009 to 2018) are included in the analysis as the period under investigation. Wave 1 is not included, as these partnerships were surveyed retrospectively over the course of a lifetime. The relationships surveyed therein were entered into in different years and not during the last year of the first survey. Due to the longitudinal approach, the top-up cohorts from wave 11 are also not taken into account. By not taking into account the top-up sample of wave 11 and wave 1, a distortion of the distances is avoided. The data set is therefore left-censored. A detailed description of all analysed partnerships can be found in chapter 5.1.

From the 11 waves of the panel, 32,615 observations of 2,965 individuals were initially counted. At the beginning of data preparation, the first wave was excluded. All cases where no new partnership could be identified were then excluded. No dating location was observed for these cases. In addition, all cases where no information on the distance to the partner was provided were excluded prior to the analysis. After data selection (see 4.4.1) and the exclusion of missing values, this results in a sample of 1,981 observations from 1,224 individuals. As the number of observation units is not constant for individual points in time, this results in an unbalanced panel.

## **Operationalisation**

* + 1. **Dependent variable**

The object of the study is the geographical distance between the partners. Due to the limited access to data, no complete statements can be made about the specific place of residence of the interviewee and their partner and therefore the distances in kilometres cannot be calculated independently. Due to this restriction, the geographical distance between the partners is calculated from the average journey time in hours (hcp3h) and minutes (hcp3m) to the partner's place of residence using the usual means of transport (pairfam Group 2020, p. 92).[[3]](#footnote-5) The two variables are summarised in a metric variable, which is the dependent variable of the study. Partnerships in which the partners live more than 1,000 minutes apart are considered outliers and excluded from the analysis.[[4]](#footnote-6) After excluding twelve cases, the longest journey time to the partner is 960 minutes (16 hours). The range is 960 minutes. Previously, the maximum journey time between partners was 2,880 minutes.

* + 1. **Independent variable**

The place where the couple met (pa3) is used as an independent variable for the subject of the study. In the first wave of pairfam, those who stated that they had a partner were asked to name the place where they met this partner. In the following waves, this question was only asked of those who had entered into a new partnership.

The study distinguishes between getting to know each other in an online and offline meeting context. The offline meeting places mentioned in the questionnaire include all meeting places that are not virtual. These include school or training, work, hobbies, clubs, pubs or discos, friends and acquaintances, relatives, printed advertisements, holidays and other. In addition to the Internet (waves 2 and 3), the online meeting context includes dating sites (from wave 4), dating via social networks, chat forums or similar (from wave 4) and mobile dating (from wave 10). As the detailed breakdown of the various online meeting places has only been carried out recently, the analysis does not differentiate between the individual online meeting places. If the meeting place is online, this is coded with 1 in the factor variable and with 0 for an offline meeting place.

* + 1. **Control variables**

As changes over time are relevant to the research question, the survey year (wave) is included as a metric variable both in the pooled OLS regressions and in the fixed-effects models. By including the time-varying variable, an initial trend can be derived with regard to the development of distances (Stoetzer 2020, p. 239). An interaction between the year and the location is modelled in order to work out a change in relation to the place of identification.

The complete pooled OLS model is extended by a number of individual and contextual control variables. At the individual level, the level of education (isced) is controlled for. The categorical variable distinguishes whether the person has a low, medium or high level of education. In the OLS regression, the low level of education serves as the reference category. The other two categories are interpreted in comparison to the low level of education. In addition, the categorical variable of the cohort, with the characteristics 1970, 1980 and 1990, is controlled for. 1970 represents the reference category. The context-related control variable of the size of the place of residence (type of region) (gkpol) in which the individual lives is also included in the study. The categorical variable differentiates between a small town (less than 20,000 inhabitants), a medium-sized town (20,000-100,000 inhabitants) and a large city (100,000 inhabitants or more). The small town serves as a reference category. Potential differences between the size of municipalities can be categorised as urban-rural differences (Lengerer 2001, p. 141).

Control variables are also added to the fixed-effects models (FE model). As only time-varying variables can be quantified, the constant variable of the birth cohort is not included in the model. In order to be able to calculate the predicted values for the FE model, a dummy variable with 0 and 1 was created for all characteristics of the size of place of residence and the level of education except for the two reference categories. These dummy variables thus reflect changes in status in the interpretation and not the actual status as in the OLS regression.

Table 1 describes the categorical variables of the total sample according to frequency and percentage share.

**Insert Table 1: Overall sample description of the categorical variables**

# **Results**

## **Descriptive results**

Figure 1 shows the development of the relative frequency of partnerships from 2009 to 2018. A total of 1,981 partnerships were observed during this period. Of these, 325 (16.4 %) partnerships were formed online and 1,656 (83.6 %) partnerships offline. At the beginning of the study period in 2009 (wave 2), 234 couples met offline and at the end of the study period in 2018 (wave 11), 102 couples met. 2009 is the year with the most new offline partnerships. In contrast, only few partnerships initiated online were observed. For example, 36 partnerships were formed online in 2009 and 28 in 2018. With 40 observations, the most online-initiated partnerships were formed in 2014 (wave 7). After a stable percentage of partnerships initiated online in the first five years, the proportion of these partnerships has been increasing since 2013. The relative share of partnerships initiated online has risen over time to just over 20 %. More detailed information on the relative and absolute frequencies of newly formed partnerships can be found in the appendix (Table 5).

Unlike Rosenfeld and Thomas (2012) for the USA and Potarca (2020) for Switzerland, this data does not confirm the trend of online meeting places replacing offline meeting places as social mediators in partner selection: In Germany, online relationships do not rise sharply during the study period and thus do not replace everyday encounters as the main mediator in partner selection.

**Figure 1: Relative frequencies of newly formed partnerships from online and offline encounters by wave (source: pairfam, wave 2-11, own illustration)**

The analysis of the time distances to the partner shows a difference between the partnerships initiated online and offline. Table 2 shows the average distance and the median distance to the partner in minutes in the online and offline meeting context. On average, couples who met offline need 54 minutes to reach their partner and half of all people live less than 20 minutes away from their partner. Partnerships initiated online take an average of 77 minutes to reach their partner and half of the couples live less than 40 minutes away from their partner. This means that both the median time and the average distance to the partner is greater for partnerships initiated online than for those initiated offline. On average, partnerships from the online meeting context travel 23 minutes longer to their partner (see Appendix Figure 7). 25 % of all couples live 10 minutes away from each other. The longest distance between two couples who met offline in both 2014 and 2015 was 960 minutes (16 hours).

Overall, it can be seen that the newly formed partnerships from the offline meeting context are on average associated with a shorter time distance than the partnerships from the online meeting context. The two-sided t-test shows that the differences in the average distances between the meeting locations are statistically significant (see Appendix Table 6).

**Table 2: Temporal distances for the online and offline meeting context**

The next step is to show how the temporal distances to the partner have developed. Figure 2 serves to illustrate the trends in geographical distance over time. It shows the development of the temporal distances of newly formed partnerships from 2009 to 2018. The online and offline partnerships are shown in separate figures.

The average time distance between couples who met offline rose from 40.3 minutes in 2009 to 63.6 minutes in 2018. However, this increase is not linear. From 2008 to 2015, the average time distance between partners initially increased. After the distances fall in 2016 and 2017, they rise again in 2018. Since 2014, the average time distance has been above the mean value of 54 minutes. With the exception of the years 2014 to 2016, the median distance of 20 minutes has remained constant over the ten years under review. In this time frame, the time distance increased by 5 or 10 minutes.

At the beginning of the 2009 study period, couples who met online spent a comparatively long 89.9 minutes away from their partner.

During the period under review, the average time distance for couples who met online fell. By 2018, this had fallen by almost a third to 60 minutes. In 2016, partnerships initiated online traveled the shortest distance to their partner at 48.1 minutes. When looking at the development over time, it should be noted that the distances vary greatly between years and that no clear trend can be identified. The same applies to the median. The range is between 25 and 55 minutes.

Figure 2: Development of the time distances to the partner for the offline and online meeting context. Representation of the mean and median values of travel times (source: pairfam, wave 2-11, own representation)

In conclusion, it can be said that partnerships initiated online are initially associated with a greater distance to the partner compared to those initiated offline. This applies to both the average duration and the median. At the beginning of the study period in 2008, the online partnerships commuted an average of 49.6 minutes longer to their partner than the couples who met offline. At the end of the study period in 2018, however, couples who met offline traveled 3.6 minutes longer to their partner than couples who met online (see Appendix Table 7). When looking at the partnerships initiated online, it can be seen that the average time distances have fallen overall. Over the course of time, partners are therefore more likely to be found in the immediate vicinity. In contrast, the average duration to the partner of partnerships from the offline encounter context has increased. This means that this group of people appears to be less spatially concentrated. However, half of these partnerships travel a maximum of 20 minutes to their partner. This means that the increasing average distances may be the result of extreme values. There is no comparable contemporary study for Germany that can be placed in relation to the time distances identified here. Only Lengerer (2001) was able to show that the majority of pairs occurred within a radius of 20 kilometers. Depending on the means of transportation and the size of the place of residence, 20 kilometers can be covered in 20 minutes.

Table 3 shows the distances in minutes according to the size of the place of residence. People who live in a small town travel an average of 48.4 minutes to their partner. People from large cities travel an average of 70.8 minutes. Half of all people living in a small town and half in a medium-sized city travel 20.0 minutes to see their partner. The median journey time for city dwellers is 30.0 minutes. Contrary to the assumption that the place of residence promotes the proportion of spatial homogamy with increasing size, this cannot be confirmed with the descriptive representation of duration. The average duration of partnerships initiated online in large cities has decreased over time. Whereas a person needed an average of 114 minutes to reach their partner in 2009, in 2018 it was only 40 minutes. In contrast, the time distance for partnerships initiated offline in the city increased from 56.5 minutes in 2009 to 103 minutes in 2018 (see Appendix Table 8).

**Table 3: Time distances by size of place of residence**

Table 4 shows the time distances to the partner by level of education. At 72.3 minutes, people with a high level of education cover the longest distance to their partner on average. People with a low level of education travel 57.4 minutes and people with a medium level of education 50.9 minutes. This does not clearly confirm that the distances increase with the level of education. There is also no clear trend in the time distances according to education level. What is striking is that the average time distance of relationships entered into offline by people with a low level of education has increased over time. In 2009, the distance for this group of people was 34.9 minutes and rose to 90.7 minutes in 2018. However, the corresponding median shows that half of these partnerships have a constant distance of no more than 20 minutes, except for the years 2014 to 2016. For people with a high level of education who met their partner online, there were major changes in the average distance to their partner during the period under investigation. Initially, this group of people needed an average of 170 minutes to reach their partner. In 2018, it was only 69.5 minutes. The distance to the partner has therefore decreased over time.

**Table 4: Time distances by level of education**

## **Regression results**

The following chapter looks at the pooled OLS regression results with regard to the development of distances in partner choice. For this purpose, a bivariate model was first calculated with the factor variable "place of identification" and then a second model with an interaction between the place of identification and the year, separated by gender. The model for the women contains 1,074 observations and that for the men 907.

* + 1. **OLS models for the women**

In the first step, a bivariate model was created for the women (see Appendix Table 10). Women who met their partner online drive 17.5 minutes longer to their partner compared to women who met their partner offline. This correlation between distance and meeting location is significant at the 0.05 level. However, only 0.4 percent of the variance in distance can be explained with the help of the meeting location. The fluctuations in the distances are therefore due to other influences that were not included in the regression equation.

In the second step, the bivariate model was expanded to include an interaction effect between the place of identification and the year (see Appendix Table 10). This shows a significant interaction effect at the 0.05 level between the place of identification and the year for the women. In addition, there is a significant main effect at the 0.01 level for both the meeting location and the year. Online dating users travel a 56.7 - 6.6 \* year longer distance to their partner than those who date offline. 1.3 percent of the variance can be explained by the model. This means that the fluctuations in distances can be attributed to other influences that were not included in the regression equation.

Figure 3 shows the adjusted predictions graphically, which depict the predicted distances for the interaction between the identifier location and the year. The respective year within the observation period is shown on the x-axis and the predicted distance in minutes to the partner is shown on the y-axis. The 95 % confidence interval is plotted for each pair of values.

In 2009, the predicted distance for women who met their partner online was just under 87 minutes. The negative slope shows that the predicted distances for online encounters have decreased over time. Every year, the distance to the partner decreases by around 2.5 minutes. In 2018, the time distance traveled to the partner was 64 minutes. In contrast, the offline encounter context shows a positive increase. While women who initiated their relationship offline drove just under 43 minutes to their partner in 2009, they drove 79 minutes in 2018. For partnerships initiated offline, the distance to the partner increases by around 4 minutes each year. In 2009, online-initiated partnerships took around 43.6 minutes longer to get to their partner than couples who met their partner in person, while in 2018, offline-initiated couples traveled 15.5 minutes longer. It is also clear that the distances from online and offline meeting contexts differed significantly by 2011. This is reflected in the non-overlap of the 95 % confidence intervals. From 2012 onwards, there is no longer a statistically significant difference between the distances covered by women in comparison to the online and offline meeting context.

**Figure 3: Predicted values for the distance with the interaction between the identification location and the year for the women. Error ranges are indicated by the 95 % CI. (Source: pairfam, wave 2-11)**

* + 1. **OLS models for the men**

A bivariate model was also calculated for the men in the first step (see Appendix Table 11). Men who met their partner online drive an average of 29.7 minutes longer compared to men who met their partner offline. This correlation between distance and meeting location is significant for men at the 0.001 level. Nevertheless, only 1.5 percent of the variance in distance can be explained with the help of the meeting location. Here too, the model does not provide an explanation of the variance. The fluctuations in the distances are also attributable to other influences for men that were not included in the regression equation.

In the second step, the bivariate model was again expanded to include an interaction effect between the place of identification and the year (see Appendix Table 11). This shows a significant interaction effect at the 0.05 level between the place of identification and the year for the men. In addition, there is a significant main effect of dating location at the 0.001 level and a significant main effect of year at the 0.05 level. Online dating users cover a 73.2 - 7.1 \* year longer distance to their partner than partnerships initiated offline. The model explains 2.4 percent of the variance.

To visualize the regression model, the adjusted predictions are shown graphically in Figure 4, with the years on the x-axis and the predicted temporal distances in minutes on the y-axis.

**Figure 4: Predicted values for the model with the interaction between the identifier location and the year for the men. Error ranges are represented by the 95 % CI. (Source: pairfam, wave 2-11)**

In 2009, the predicted distance for men who met their partner online was just under 99 minutes. The predicted time distances for online encounters have fallen by around 4.5 minutes per year. In 2018, the distance was reduced to 58 minutes. In contrast, the offline meeting context shows a positive increase. While men drove an average of just under 40 minutes to their partner in 2009, they drove 63 minutes in 2018. This means that the distance to their partner increases by around 2.6 minutes per year. At the end of the study period, online-initiated partnerships even drive a distance to their partner that is around 4.8 minutes shorter than offline-initiated partnerships. In 2009, the difference was still 59 minutes. Until 2014, the distances differed significantly between online and offline encounters. From 2015 onwards, the differences in distance are no longer significant.

## **Robustness test**

The OLS models from Chapter 5.2 compared the two identifier locations. Chapter 5.3 checks the OLS results using a fixed-effects regression (FE model). FE models determine the differences in distance between the intra-individual relationships.

The FE model for women includes 1,074 observations, of which only one relationship was observed for 656 people. Women entered into up to five new relationships during the study period. For men, 907 observations are included in the analysis. This includes 568 people for whom only one relationship was observed. In the case of men, up to four new partnerships were observed during the study period (see Appendix Table 12 for women and Table 13 for men). The complete models include control variables in order to uncover possible spurious correlations.

* + 1. **Fixed-effects model for women**

The coefficient for the full FE model of the identification learning locus also remains positive in comparison to the full OLS model (see Appendix Table 14). Thus, even when controlling for person-specific heterogeneity, there is a positive effect of the place of identification for women. A comparison between partnerships of a female person shows that the person who first entered into a partnership offline and then online has a 75.6 minute higher distance difference. This value is statistically significant at the 0.05 level. The FE model therefore shows that people with the variation in meeting location travel further to their partner when they meet online than at times when they have met a partner offline. The model also shows a statistically significant interaction effect between meeting location and year at the 0.05 level. The main effect of year is also statistically significant at the 0.05 level. If a change in an individual's educational status to the middle level is observed, this has a negative effect on distance.

The difference in explanatory power between the OLS and FE models is striking: while 2.9 percent of the variation in distance could be explained by the independent variables in the OLS regression, this figure is 66.2 percent based on the fixed-effects-transformed data. The independent variables therefore explain fewer differences between people than within people. In the FE model, only 2.8 percent of the variance within the individual is explained. This means that only a small proportion of the fluctuations in the distances between women can be attributed to other influences that were not included in the regression equation.

**Figure 5: Fixed-effects model for the women, taking into account the control variables (year, municipality size, level of education and an interaction between place of identification and year). Error ranges are indicated by the 95 % CI. (Source: pairfam, wave 2-11)**

Figure 5 illustrates the predicted distances of the FE model for women. The predicted intraindividual distances for partnerships of an individual from the online dating context have decreased over time. In 2009, the predicted distance for a person who initiated their partnerships online was around 67.6 minutes, whereas in 2018 it was only 31.3 minutes. This means that the intra-individual distance to a partner for online dating has decreased by more than 50 percent over the period. This represents an annual change of 4 minutes. In contrast, the predicted distance of an individual's partnerships from the offline encounter context has increased by around 6 minutes over time. In 2009, the predicted distance to the partner was around 12 minutes. For an individual's offline-initiated partnerships, the distance to the partner in 2018 was 66 minutes. Since 2015, an individual's offline-initiated partnerships have been associated with a higher distance. However, none of the predicted values for women are statistically significant.

* + 1. **Fixed-effects model for men**

The coefficient of the online encounter context from the FE model for men also remains positive. However, it is lower and no longer statistically significant (see Appendix Table 15). However, even when controlling for person-specific heterogeneity, there is a positive effect of the meeting location. The comparison between partnerships of a male person shows that the person who first entered into a partnership offline and then online has a 24.7 minute greater difference in distance. In the FE model for men, none of the control variables indicate a significant explanatory power.

The difference in explanatory power between the OLS and FE models is striking: while 4.5 percent of the variation in distance could be explained by the independent variables in the OLS regression, this figure is 75.8 percent on the basis of the fixed-effects-transformed data. Accordingly, the independent variables explain fewer differences between individuals than within individuals. But even in the FE model, only 3.2 percent of the variance within the individual is explained. This means that only a small proportion of the fluctuations in distances among men can be attributed to other influences that were not included in the regression equation.

Figure 6 visualizes the predicted distances for men. Unlike the previous models predicted, the predicted intraindividual distances for online partnerships increase slightly over time. While the predicted distance for offline-initiated partnerships of one person was 26.5 minutes in 2009, it was around 34.5 minutes in 2018. This represents a total intra-individual distance change of 6 minutes over the ten years. An individual's online relationships increase by 0.9 minutes per year. In contrast, the intra-individual differences in distance from offline encounters increase more strongly over time. Whereas in 2009 the value was 5.2 minutes to their partner, in 2018 men were already traveling 28.5 minutes to their partner. Even though there were no significant differences in distance between the partnerships of an individual at the beginning of the study period, these have steadily decreased over time. At the end of the study period, only slight intra-individual distance differences of around 6 minutes were observed between the partner locations. However, none of the predicted values for the men are statistically significant.

**Figure 6: Fixed-effects model for the men, taking into account the control variables (place of identification, year, size of place of residence, level of education and an interaction between place of identification and year). Error bars are indicated by the 95% CI. (Source: pairfam, wave 2-11)**

# **Discussion**

This article has examined the development of geographical distances in the choice of partners between those met online and offline for the German region over time. Taking into account the theoretical considerations and the empirical results, three expectations regarding the development of distances were formulated and tested using the pairfams relationship and family panel. The results confirm the first two expectations: At the beginning of the study period, couples initiated online cover a greater distance from their partner than partnerships initiated offline. Over time, the distance decreases in online-initiated partnerships. However, the third expectation is not confirmed: the average distance increases over time for partnerships initiated offline.

To generate the results, the first step involved a descriptive approximation of the development of the distances. In the second step, the distances were predicted using a pooled OLS regression. At the beginning of the study, significant differences were found between the online and offline encounter context for both women and men. At the end of the study period, it was found that both women and men who met their partner in an offline meeting context traveled further to their partner. Geographical distances have therefore decreased over time in online dating and partnerships are increasingly being formed from nearby locations. The increased distances in offline dating, on the other hand, have a negative effect on the proportion of spatial homogamy. After robustness checks of the regression results with the FE models, the third step also shows that the intra-individual distance differences between women have decreased over time in the online dating context. In contrast, both the offline and online initiated partnerships of a man show a slight increase in distances over time. It is striking that men cover a shorter distance overall compared to female partnerships. This means that men are more likely to enter into relationships from a nearby environment. Women, on the other hand, also enter into partnerships further away.

Despite the similar results after the robustness test, the lack of significant differences is due to weak statistical significance. A larger sample would be necessary to obtain more precise results. Despite this limitation, the study nevertheless makes a small exploratory contribution to spatial homogamy in mate choice. Previous studies have only rudimentarily worked out the differences between the online and offline encounter context in the longitudinal section. Consequently, it can be assumed for the contact opportunities in online dating that fewer contacts were spatially restricted at the beginning of the study and that online dating initially reduced the proportion of spatial homogamy. The offline meeting context resulted in a higher proportion of spatial homogamy at that time. Over the years, the differences in distance between the two meeting contexts have decreased. This means that the meeting context now only seems to make a small difference to the proportion of spatial homogamy.

The question remains as to why the distances in the online context have decreased over time. As the analysis did not produce any meaningful explanatory factors for the change in distance, it would be important to investigate this further and also include cultural factors (Haandrikman et al. 2008; Haandrikman 2019). On the theoretical side, the higher number of users could explain the change. If people have a larger local partner pool, they are more likely to search for a partner in a small radius, as the probability of finding a potential partner in this radius is higher. On the other hand, the popular use of dating apps could also have an impact on distance (Potarca 2020). Dating apps encourage contacts in the immediate vicinity. It would therefore be interesting for future studies to focus more on the differences between the individual forms of online dating and their effects on spatial homogamy.

## **Limitation**

* + 1. **Data**

Since the same people are interviewed every year in a panel, the number of people interviewed decreases over time. If the number of interviewees decreases over time, the observed partnerships will also decrease during the study period. The withdrawal or non-participation of a person in the panel is also referred to as panel mortality (see Schupp 2019, p. 1274). Another reason for the decline in newly formed partnerships is that actors tend to be in partnerships over the course of their lives (Eckhard and Stauder 2019, p. 3f.). The problem of low case numbers also applies to the online meeting context. In total, only 325 online-initiated partnerships were observed in the data set. If the number of cases is too small, it may not be possible to detect effects, which also results in uncertain estimates. The distances can also be distorted by homosexual partnerships and partnerships where the partner lives abroad. Studies that analyse geographical proximity in partner selection only include heterosexual couples in the analysis, as the partner market is distributed differently for homosexual people. Long-distance relationships with a person living abroad are also associated with greater distances. Due to the resulting data exclusion, it is assumed that the couples are heterosexual and live in Germany. This represents a severe limitation and, as indicated, can lead to a distortion of the distances.

Another limitation of the study is that the data basis does not allow us to go into more in-depth differences between the specific types of online dating. The individual online meeting contexts were only implemented in pairfam over the course of the waves. In particular, mobile dating has only been included for two waves, meaning that no development can be shown. However, since empirical evidence has emerged of differences between online dating forms (Potarca 2020), it is necessary not to ignore these effects on spatial homogamy in the future.

For a more meaningful and comparable analysis, it would be important to know where the respondents live. This would allow the exact distances in kilometres to the partner to be determined. This would lead to exact measurements and thus also better statements about the proportion of spatial homogamy. In the context of this study, however, only the distances between partners in minutes are available. In addition, measuring the distance in minutes has a further disadvantage, as the pairfam does not specify which means of transport a person uses to cover the distance. This can lead to a strong distortion if the journey time between the two routes is identical. Within 45 minutes, a person can cycle 10 kilometres to their partner or the same person can drive 45 kilometres by car in the same time. This leaves open how far the distance between the partnerships really is. The time distance can therefore only be used to derive tendencies regarding geographical preference. It is therefore challenging to categorise the results, as there is no comparable data.

* + 1. **Methodological limitations**

As indicated in section 4.2.2, the unobserved heterogeneity in the pooled OLS regression can prove problematic. In this case, the estimated coefficients are distorted and not reliable, meaning that they are not random. As a result of unobserved heterogeneity or time, autocorrelation and heteroscedasticity of the residuals can lead to further problems (Stoetzer 2020, p. 238f.). Heteroscedasticity leads to inefficient estimators and affects the standard error. Linear models also assume that the residuals are uncorrelated. If there is autocorrelation in the data, the residuals are no longer randomly distributed, which also leads to distorted standard errors (Backhaus et al. 2016, p. 103ff.). To avoid this, robust standard errors should be estimated in the pooled OLS regression. The cluster-robust standard errors produce reliable t-values, which in turn result in significant coefficients (Stoetzer 2020, p. 239). A major advantage of FE models is that unobserved heterogeneity can be eliminated. For this reason, the pooled OLS regression results are tested for robustness by the FE models (see section 5.3).

1. While there were 61 million users (6.9 million active users) in 2009, there were 135.7 million users (8.6 million active users) in 2017. This discrepancy between active and passive users can prove to be problematic (Langbein et al. n.d., 7). [↑](#footnote-ref-3)
2. Time-invariant refers to all variables that remain the same over time for an individual but take on different values between individuals (Stoetzer 2020, S. 234). [↑](#footnote-ref-4)
3. The usual means of transport is not defined more precisely in the codebook. However, the means of transport used influences the journey time to the partner People who live further apart are more likely to use a car or train than a bicycle. This is a comment that cannot be discussed further in the paper.

   cannot be discussed further in this study. [↑](#footnote-ref-5)
4. The exclusion of the higher values is based on the duration between the northernmost and southernmost point in Germany. The northernmost point in Germany is List on Sylt and the southernmost place is Oberstdorf in Bavaria. As the crow flies, the distance between the two places is 855 km and the driving distance is 1,060 km (https://www.luftlinie.org/, accessed on 02/06/2021, 10:23). According to Google Maps, this distance takes 13 hours and 15 minutes by car and 13 hours and 35 minutes by train (accessed on 02/06/2021, 10:25). [↑](#footnote-ref-6)