What to choose

The influence of social network structure on fertility intentions

L. J. Venema

Student number: 3453634

Study: Master Sociologie

Supervisor: dr. Gert Stulp

Referent: Jaap Nieuwenhuis, PhD

university of groningen

faculty of behavioural and social sciences

18th March 2024

1 Abstract

Recent literature has shown that fertility rates in the Netherlands have been in decline. Social networks play an important role in the formation of peoples' fertility intentions. They exert social influence which can be enhanced or hindered by the structure of the network. Therefore, this thesis studies the effects of network structure on fertility intentions. This is done by using the Girvan-Newman method to identify clusters within the network. For this research data from the LISS panel has been used. The LISS panel is a representative sample of Dutch individuals who participate in monthly Internet surveys. The analyses show that while there is an effect of the opinion of the personal network on fertility intentions, no significant effect of network polarisation on certainty about fertility intentions was found. A recommendation for further research would be to look into the role of social pressure as mediator for the effect of polarisation within the network, as this could have a large effect on the influence of the network on fertility intentions.

Contents

2 Introduction

Over the past 50 years a few notable trends in fertility in the Netherlands have emerged. Firstly, birth rates have been in decline since the 1970's, and the fertility rate of 1.62 children per woman is lower than what is needed to replace the population (CBS, 2023). Secondly, the number of women who are voluntarily childless has been increasing among women born after the second world war, which has resulted in a larger percentage of women born around 1965 to be childless than the generations before them (CBS, n.d.). These trends show that there is a shift in fertility behaviour and suggest that fertility intentions, while a personal choice, can have an impact on a societal scale. Low societal fertility can result in an ageing society, which brings several problems. In terms of the labour market, an ageing workforce will reduce overall labour participation, and physical labour could result in negative health effects if people are forced to continue working (Liu et al., 2021). It will result in a strain on the healthcare system, because older people tend to have more healthcare needs than young people (Dallmeyer et al., 2017; Kim et al., 2018). If there is a large increase in people that need healthcare, then that would be a strain on the resources available (Tang & Li, 2021). Furthermore, social welfare could get too expensive to be maintained (de Albuquerque, 2018). The social welfare system relies on the people that are working to pay out pensions. If there are more people that receive pensions than there are working, then that will jeopardise the social welfare system (Han, 2013).

Fertility intentions seem to be an individuals' or a couples' choice, but that choice, like all choices, can be heavily influenced by the social environment a person is embedded in (Bernardi et al., 2007; Lazer et al., 2010). Several theories, such as the theory of planned behaviour (Ajzen, 1991), and the social influence theory (Kelman, 1974), attempt to explain how personal decisions can be influenced by others. The theory of planned behaviour states that behaviour can be predicted by intentions, attitudes, and perceived behavioural control (Ajzen, 1991). The social influence theory complements the theory of planned behaviour by explaining the processes through which intentions, behaviours and attitudes can be influenced (Kelman, 1974). Following these theories, there have been many empirical studies that investigated the effects of social influence on fertility intentions (e.g. Buyukkececi et al., 2020; Pink et al., 2014; Lyngstad & Prskawetz, 2010; Bernardi, 2003). These studies have found that social pressure, social support, social learning and social contagion have an effect on the formation of fertility intentions.

Social influence is most likely to be effective when the influencer is someone known and close to the individual (Latané, 1981). It is therefore important to study the people that have the most influence on an individual, i.e. their personal networks. The people that are part of these networks are often also connected to each other, forming network structures. Studies have found that aspects of network structure, such as density or composition, can enable or hinder the effects of social influence (e.g. Bernardi, 2003; Bernardi & Klärner, 2014; Bühler & Fratczak, 2007; Vacca, 2020). Other studies have focussed more specifically on the different effects of network structure on fertility intentions (e.g. Madhavan et al., 2003; Kohler et al., 2001; Stulp & Barrett, 2021).

Many of the findings on how network structure relates to fertility intentions are based on qualitative research (e.g. Keim, 2011; Kavas & De Jong, 2020; Keim et al., 2009). These studies, which are often based on highly selective samples, are useful in determining some of the processes that play a role in the formation of fertility intentions, but are less appropriate for determining which processes are most important in the larger population, nor can they establish the magnitude of the effects of social influence. The reason that most of these studies are qualitative in nature is because of the difficulties connected with collecting a large sample of large personal networks. It is difficult to collect large personal networks, because of the burden to respondents in having to fill out many questions about the people in their networks and the ties between these people (Robins, 2015). This burden can cause motivational loss within the respondents, which can lead to a decrease in data quality (Stadel & Stulp, 2022). Many network studies will therefore ask the respondents to name up to only five network partners, or other approximations are made to discover network structure (e.g. Colleran, 2020; Kohler et al., 2001; Mönkediek & Bras, 2014; Mathews & Sear, 2013). However, a smaller network sample can lead to other biases. For example, a small network will most likely not contain any weak ties, and the information about the density or composition of the network will be unreliable because there is information missing (Stadel & Stulp, 2022). The inclusion of this information could generate insight into the effects of more realistic networks. It is therefore necessary to collect data about larger networks to identify the network structure.

There are different types of network structure that can influence fertility intentions. One way to examine the effects of network structure on fertility intentions is by dividing networks into different types of networks. Through a series of in depth interviews, Keim (2011) identified six types of personal networks and the effects that they have on fertility intentions: the family centred network, the supportive network, the polarised network, the family remote network, the non-supportive network, and the childless by choice network. These networks are mostly characterised by their composition, the number of relatives and friends or colleagues, and by structural characteristics such as density and tie strength. A polarised network, however, differs from the other types of network, as this is a network in which there are multiple subgroups present that have opposing opinions about a certain subject (Interian et al., 2023). This thesis will attempt to discover if these findings concerning polarisation can also be found in the Netherlands.

Polarisation is a much discussed topic, both in the media and in the academic world. In particular political polarisation is a concern of many (Liu et al., 2021; Interian et al., 2023). For example, the increase in harsh statements in both political and public debates, and the decrease in manners in these discussions have raised concern (Ministerie van Volksgezondheid, 2023). The increase in political polarisation can threaten democracies (Liu et al., 2021). On a smaller scale the consequences of a polarised network are less drastic, but on a personal level it can create tension between people that have opposing views on certain subjects. If one network member has a strong desire to have children, while another strongly opposes, tension can form between them or between them and the person who has them in their network. Keim (2011) has found that people with a polarised network tend to be ambivalent about their fertility intentions. It is therefore relevant to study cases of polarisation in small networks, in order to understand how the processes of polarisation and reconciliation work. Additionally, this thesis will concentrate on how network structure influences the formation of fertility intentions, which is a current topic for sociologists (Biondi et al., 2023). This thesis will try to discover what the effects of network composition and structure are on fertility intentions and to see if a polarised network structure has an effect on fertility intentions.

This leads to the following research question: *"How does the polarity or opinion diversity of personal social networks shape fertility intentions?"* To answer this question we need to look into the ways social networks can influence opinions or behaviour.

3 Theory

There are multiple ways for a social network to change behaviour or attitudes of people, but the mechanisms through which it does so are always related to social influence. Social influence is the process by which a person changes their behaviour based on social interactions (Kelman, 1974; Montgomery & Casterline, 1996). The theory of planned behaviour (Ajzen, 1991), and social influence theory (Kelman, 1974) are theories that explain how social influence can affect behaviour. This section will first explain what the different types of social influence are and then focus on the theories to explain behaviour. The section will close with the hypotheses that will be tested in the analyses.

3.1 Types of social influence

There are four types of social influence: social support, social pressure, social learning, and social or emotional contagion (e.g. Lois, 2016; Keim, 2011; Bernardi & Klärner, 2014; Kavas & De Jong, 2020). In this section I will explain each of these mechanisms of social influence.

Social support is the help that network members can provide, such as child care or advice (Kavas & De Jong, 2020; Lois, 2016). Social support is an important mechanism, because its presence is necessary for people to make big life changes, such as having children. Depending on the culture, the presence of social support is necessary for a couple to decide to have children, and network members can exert a large amount of influence by granting or withholding support (Kavas & De Jong, 2020). Generally speaking, the presence of social support is what allows people to realise their fertility intentions (Hank & Kreyenfeld, 2003). Additionally, social support is the one mechanism that is dependent on the actions of the network members for it to be put into effect; it takes conscious action for a person to decide whether or not to extend help to someone else. Social support is therefore not necessarily a type of social influence that has an effect on the formation of fertility intentions, but it can have an effect on the realisation of fertility intentions if the network gives indication about whether or not they will provide support.

The second type of social influence, social pressure, is a force that the network can exert to make individuals conform to social norms (Kuhnt & Trappe, 2016; Balbo & Mills, 2011;

Balbo & Barban, 2014). Social norms are related to the desirability of certain behaviours, the characteristics of a group of actors, and are enforced by social sanctioning or rewards by the network members (Liefbroer & Billari, 2010; Axelrod, 1986; Lois & Becker, 2014). In contrast to social support, social pressure can be put into effect both actively and passively by the network. An active way to exert social pressure is by reminding someone of the norms that are present, such as asking a couple when they are planning on having children, or asking for grandchildren. Social pressure can also be passively present in a social environment or network. In this case people observe and conform to the norms they see around them without any direct action from the people in their network. Studies have found that social pressure has caused people to change their fertility intentions (Balbo & Mills, 2011; Mönkediek & Bras, 2018; Kuhnt & Trappe, 2016; Lois & Becker, 2014).

The last two types of social influence, social learning and social contagion, are similar concepts and will therefore be discussed simultaneously. Social learning happens when interacting with a network member and learning about their experiences with parenthood and using these experiences to form fertility intentions (Lois, 2016; Keim et al., 2009). Social contagion tends to be an emotional response to interactions with young children. This emotional response can then influence the decision this person makes when it comes to their own fertility intentions (Keim et al., 2009; Lois, 2016; Bernardi & Klärner, 2014). Both social learning and social contagion mostly take place without the network members actively trying to exercise social influence. Social learning is something that can happen in casual conversation, while social contagion is an emotional response, be it positive or negative, to small children.

The kind of effect that any type of social influence has, whether it is positive or negative, depends on the content of the interaction. The presence of social pressure and social learning does not necessarily result in positive fertility intentions (Bernardi & Klärner, 2014; Keim et al., 2009). For example, if the network has a negative opinion about reproduction, or if a person learns about the negative aspects of raising children, they might be influenced against having children. Additionally, network members can through either granting or withholding social support, control the timing and even the number of children being born (Kavas & De Jong, 2020).

Different network structures and ties can enable or impede certain types of social influence. Social pressure and social support are most effective when the network has a high density and the recipient has strong ties with the network partners that exert this influence (Bernardi, 2003; Bernardi & Klärner, 2014; Bühler & Fratczak, 2007; Vacca, 2020). A high density in the network means that the influence from one person is more likely to be enforced by others, because the network members are more interconnected. Strong ties not only allow for social pressure or social support to take place, it also enables social learning and social contagion (Keim, 2011). It is after all easier to listen to, and consider, the opinion of a close friend than that of a stranger or acquaintance (Latané, 1981). Social learning is effective amongst both weak and strong ties; whereas in the case of social learning, it is not necessary for two people to be close to each other for this mechanism to be effective (Lyngstad & Prskawetz, 2010; Pink et al., 2014; Buyukkececi et al., 2020). Social contagion is again more likely to occur through strong ties, as these enable contact between adults and young children.

Social pressure has a large effect on fertility intentions (Balbo & Mills, 2011; Liefbroer & Billari, 2010). Social pressure is, as mentioned earlier, the force that makes individuals conform to social norms, and is the main mechanism by which two theories explain human behaviour. These theories are the theory of planned behaviour (Ajzen, 1991), and social influence theory (Kelman, 1974). These will be further discussed in the following section.

3.2 Theory of planned behaviour

One way to explain how behavioural intentions are formed and how they influence actual behaviour is described in the theory of planned behaviour, developed by (Ajzen, 1991). This theory has often been used to explain fertility behaviour (Kuhnt & Trappe, 2016). The theory of planned behaviour explains how behaviour can be predicted by intentions and perceived behavioural control. Intentions are formed by attitudes, the perceived behavioural control, and subjective norms (Ajzen, 1991). Attitudes are the assumptions people hold about the consequences, positive or negative, of a behaviour like having children or not (Mönkediek & Bras, 2018). Perceived behavioural control is the perceived ability to perform a certain behaviour (Ajzen, 1991), which will both influence someone's intentions, as well as the ability to perform that behaviour. For example, if someone feels physically or emotionally unable to raise a child, they feel as if they don't have the ability to perform that behaviour. This can influence whether that person tries to have children or not, even if they might have a desire to have children. Subjective norms are norms that an individual believes are present (i.e., they are perceived norms).

Behavioural intentions are heavily influenced by the social environment, and in particular by social pressure. Both the attitudes people have toward a certain behaviour and subjective norms are dependent on the people that surround them, i.e. their networks. When people are exposed to pressure to have children, they will believe this is the norm in that social environment. If the presence of this norm causes the desire to have children, then their attitudes have also changed. If a person enters an environment where all other people are having children by a certain age, they might feel that is an expectation in that environment. They believe it is the norm in that environment to have children by that age. If they then attempt to have children at that age, their attitudes have changed in response to the environment. When the people in the network have strong attitudes towards a subject, this can influence the attitudes and behaviour of the people they are in contact with (Ajzen, 1991).

The theory of planned behaviour provides a framework on how people can be influenced by others and how social pressure can form different attitudes towards certain subjects. This can be seen in Kuhnt & Trappe (2016), who applied this theory to explain the effects of social influence on the realisation of fertility intentions. They found that people in the network have their own norms, and will use social pressure to enforce these norms (Kuhnt & Trappe, 2016). Contradictory norms in the network can however cause uncertainty about which norm to uphold (Kmetty & Tardos, 2022).

3.3 Social influence theory

The social influence theory complements the theory of planned behaviour by further elaborating on the mechanisms that can influence behavioural intentions. Kelman (1974) states that in order for social influence to take place, three requirements need to be met. The first requirement is that social influence must be about a goal that is important for a person to be met. It is therefore not possible to influence someone to do something that they are not interested in. The second requirement is that the influencer must be considered to be relevant to the achievement of the goal. The third requirement is that there must be some evident path. This means that the goal cannot be met by following another course of action. In terms of fertility behaviour, these requirements are easily met. People tend to have to make a choice about their fertility intentions at one point in their lives, there are a lot of people that can be relevant to the decision that is made, e.g. by being able to provide social support, and the choice is a clear one that can not be achieved through another course of action.

According to the social influence theory, attitudes, behaviours, and beliefs are influenced through three processes: the compliance process, the internalisation process, and the identification process (Kelman, 1974). The compliance process entails that a person's behaviour is based on the expectation of reward or punishment. This means that someone accepts influence in order to gain a reward or avoid punishment from others. This is similar to the influence of social pressure, where people will use sanctions or rewards to get others to change their behaviour and conform to their norms. An example of this is when people choose to have children out of fear of being stigmatised by their network partners.

The internalisation process is related to how the beliefs and values of others are received in an individual. It occurs when an individual integrates the norms and values of others into their own goals. This often happens based on the perception of someone's social norms (Yang, 2018). This process does not necessarily change the goals of an individual, but integrates their motivation for those goals into that of their social environment. In the case of internalisation, the influence is less of a direct process, but more an adaptation of the viewpoints of the social environment. For example, a person might not have strong fertility intentions and form them to conform with those they believe their network members hold.

The third process, the identification process, is related to the satisfaction of others (Kelman, 1974). Through this process, a person accepts influence in order to maintain the relationship they have with that person or group. In contrast with the compliance process, the individual is not actively seeking a reward or trying to avoid social sanctions, but is mainly concerned with their relationship with the group (Kelman, 1974). An example of this is when someone holds of of having children when their network members are strongly against having them in order to maintain their relationships with these network members.

The theory of planned behaviour and the social influence theory show how behaviour can be influenced by the social environment people are embedded in. These theories provide complimentary explanations about how social influence can change an individual's behaviour. The mechanism through which both of these theories work is social pressure. Through enforcement of norms, social pressure influences people to change their attitudes, intentions or behaviour.

3.4 Social influence and polarised networks

Historically speaking, there was a norm for people to have children, and those who did not were stigmatised (Agrillo & Nelini, 2008). In the last few decades however, this norm has shifted and fertility rates have declined throughout the world (Munshi & Myaux, 2006). This shift in fertility norms can have multiple reasons: economic decline, a decline in child mortality rates, and the rise of contraceptives all result in people having smaller families (Bhattacharya & Chakraborty, 2012). Nowadays more people are voluntarily childless than in previous decades (CBS, n.d.). The increase in voluntary childless couples is a possible effect of the second demographic transition, which resulted in a shift in social norms. As more women went to higher education and joined the workforce, it became normal to start having children later, or not to have children at all (Agrillo & Nelini, 2008). This resulted in more people being in contact with voluntarily childless couples and the acceptance of voluntarily childless couples. The larger spread acceptance of voluntarily childless couples has created the possibility for them to exert social influence on their network partners in the same way that couples that do want to have children can.

Contact with people that have different social norms or opinions will reduce the likelihood of having polarising opinions (Facciani et al., 2023). This can be explained by opinion convergence, as stated in opinion dynamics literature (Mueller & Tan, 2018). This entails that exposure to other viewpoints will move two people closer together in opinions about a certain issue (Baumgaertner et al., 2016). In other words, when an individual is in contact with people that have opposing opinions to each other, this will result in that person having less strong opinions of their own. This is in line with the findings of Keim (2011), who found that people with polarised networks tend to be ambivalent about their fertility intentions.

3.5 Hypotheses

Following the reasoning of the theory of planned behaviour, a network that is mostly in agreement about children, will only exert pressure in one direction. In this type of network there is one clear perceived norm about having children, which is the opinion of the network. This will then influence a person to adopt the same norm as the network (Ajzen, 1991). The mechanism through which this person adopts this norm is (perceived) social pressure or social learning through the interactions with the network members.

This is in line with the social influence theory, where the identification process will cause someone to adapt their own behaviour in order to be liked by a group. If the network is not polarised and in agreement on a certain issue, then that will lead to a behavioural outcome similar to that of the network. These arguments lead to the following hypothesis:

H1: A network that is more pronatal will lead to more positive fertility intentions

Similarly, according to the theory of planned behaviour, behavioural intentions are formed based on the attitudes, perceived behavioural control, and subjective norms (Ajzen, 1991). Both attitudes and behavioural norms are receptive to social pressure, as one can be pressured into changing one's opinions and norms. Polarised networks are characterised by the presence of multiple subgroups with contrasting opinions (Interian et al., 2023). When someone has a polarised network, they will receive social pressure from both groups in the network to conform to the norms belonging to each group. In the context of fertility intentions, this means that an individual will receive pressure both to have and not to have children. Exposure to opposing standpoints can lead to ambivalence about that topic (Kmetty & Tardos, 2022).

This can also be found when applying the theory of social influence to a situation of fertility intentions. The process of compliance is the most easily identified form of social influence, but the processes of internalisation and identification can be more effective. In the internalisation process, an individual will take on the norms they perceive others around them to have, and make them their own. When this happens, it can be hard to notice that a person has been influenced at all. Applied to networks, this means that a person will adapt to the opinions or norms of the others in the network. This is more difficult to do in polarised networks, which could lead to someone not knowing which side to choose. This leads to the following hypothesis:

H2: a polarised network will result in ambivalence in fertility intentions

3.6 In this study

This study is aimed at increasing the understanding of the influence of network structures on fertility intentions. To achieve this, I will use cluster analysis to detect communities (or clusters) in the networks. Cluster analysis is a group of techniques that can be used to assign items into different groups based on the similarities and distances between them (Borgatti et al., 2018). This will allow me to determine what the different subgroups or clusters are, which is necessary to determine network polarisation. A polarised network can be defined as a network in which there are multiple groups present that have differing opinions (Keim, 2011; Interian et al., 2023). Network polarisation is a difficult measure, as personal networks are rarely completely polarised. It is likely that multiple subgroups will have varying degrees of opinion similarity. To be able to deal with this variation, a scale was created from -1 to 1, both on the level of the clusters to measure opinion, and on the level of the network to measure polarisation. Keim (2011) has found that people with polarised networks tend to be ambivalent about their own fertility intentions. This study will try to determine whether this effect can be found in the population and try to discover the size of this effect.

4 Methods

4.1 Description of data & methods of data collection

In this paper I make use of data from the LISS (Longitudinal Internet studies for the Social Sciences) panel administered by Centerdata (Tilburg University, The Netherlands). The LISS panel is a representative sample of Dutch individuals who participate in monthly Internet surveys. The panel is based on a true probability sample of households drawn from the population register. Households that could not otherwise participate are provided with a computer and internet connection. A longitudinal survey is fielded in the panel every year, covering a large variety of domains including health, work, education, income, housing, time use, political views, values and personality.

4.2 Data collection

The data was collected from a sample of the people who participated in the LISS panel, which allows researchers to submit their own surveys. For the *Social networks and fertility* survey all women between the ages of 18 and 40 who participated in the LISS panel were invited to participate between February 20 and March 27, 2018 (Stulp, 2021). Of the 1332 people who were approached for the survey, 758 people responded. The respondents were comparable on multiple background variables to the women who did not participate (Stulp, 2020, 2021). The respondents were informed that the survey would take 25-30 minutes to be completed and they received €12.50 for completing the survey. The first part of the survey contained questions concerning the fertility intentions of the respondents and their partners, if they had one. The network data was collected through the use of the program GENSI, which creates visualisations of the network and the alters. This aids the respondents in filling in questions about the alters and alter-alter ties. The respondents were asked to name 25 people with whom they were in contact in the last year. They were then asked a series of questions about their relationship with these people, before they were asked who of these people knew each other.

The dataset measures the fertility intentions of 758 women and their personal networks.

The number of alters that was collected is large enough to allow for the network structure to be observed, while remaining manageable for the respondents to fill in the survey and alter-alter ties (Robins, 2015).

Seven respondents did not fill in any alter-alter ties; those were excluded from the analysis. Additionally, there were people for whom it was not possible to calculate the level of polarisation in their networks. This was either because of missing network data, or because these people only had one community of three or more people in their network (for further explanation, see below). These people have also been excluded from the analysis. Lastly, I will restrict my analyses to women that do not have children (yet), in order to exclude the influence that parenthood might have on the composition and structure of the network, and because the decision process to have a first child is different than for a second child (Balbo & Mills, 2011). This leaves a dataset with 478 respondents. The average age of the women in this sample is around 26 years old.

4.3 Measurement of fertility intentions

The main question that will be used as the outcome variable to determine fertility intentions is: "Do you plan on having (more) children in the future?". The answer options to this question are: "Absolutely not" , "Probably not" , "I don't know", "Probably" , and "Definitely".

Because there is no variable that measures the certainty the respondents have about their fertility intentions, the same question is used to describe the respondents' certainty in fertility intentions. In order to create the best of certainty, three different operationalisations were constructed. For the first operationalisation I transformed the fertility intention question into a binary variable that only includes those who answered "I don't know" as uncertain, and the rest as certain. For the second operationalisation, I also included those who answered "Probably yes/no" as uncertain. Both of these variables have a value of 0 for people that are uncertain, and a value of 1 for people who are certain about their fertility intentions.

Given that the people who answered "probably" when asked about their fertility intentions can be argued to be more certain than the people who answered "I don't know", I created a variable with three outcomes: people that are certain ("Absolutely yes/no"), people that are less certain ("Probably yes/no"), and people that are uncertain ("I don't know"). The code for this variable and all others used in this thesis can be found in Appendix A.

4.4 Fertility intentions of network members

The hypotheses, *H1: A network that is more pronatal will lead to more positive fertility intentions,* and *H2: a polarised network will result in ambivalence in fertility intentions*, are all related to the fertility intentions of the network members. The first relates to the overall network opinion, the second to how these opinions are related to the network structure. The fertility intentions of the alters were measured in three questions to the respondent about their network members. These questions are: "Which of these people have children or are currently expecting a child?", "From which individuals do you know that they would like to have children?", and "From which individuals do you know that they would not like to have children?". The answer options to these questions are: "Has a child or is expecting a child"/"Does not have a child"; "Would like to have children"/"Don't know whether individual wants children or not"; "Would not like to have children"/"Don't know whether individual does not want children". I then combined these three questions into one variable where for each alter it was determined whether s/he preferred to have children (coded as 1), whether the fertility preference was unknown (coded as 0), or whether s/he preferred to not have children (coded as -1). Network members with children are included in the group with positive intentions, based on the assumption that their children were intended.

For the first hypothesis I created a variable that measures the overall fertility intentions of the network through summing the intentions of the network members to each other, but excluding those network members whose opinion is unknown from the calculation. A high number means that there are more network members with positive fertility intentions, a low number means that the network generally has negative fertility intentions. If there is only one opinion in the entire network, then the number would be either positive or negative 25. The number of network members whose fertility intentions are unknown are included in all analyses as a control variable.

In order to see what the effect of a network with a clear opinion was on the certainty of the respondents about their fertility intentions, I also created an absolute value of this variable. I did this by first calculating the square of the variable that measured the overall network opinion, and then taking the root of the outcome. This resulted in a new variable with values from 0 to 25, where a high number means that the majority of the network shares the same fertility intentions.

4.5 Polarisation

Next to fertility intentions, hypothesis 2 also relates to the structure of the network, in particular its polarisation. Polarisation can be measured in different ways. One way to determine whether a network is polarised or not is to look at the number of different opinions in the network. Another way to determine whether a network is polarised or not, is to determine whether opinions differ across different groups in the networks; if different groups have very different opinions, the network can be said to be polarised. To determine these groups, clusters will be determined. I will use the variation in opinion between clusters as a method to determine polarisation. There are multiple methods that can be used to determine clusters in a network. One of these methods is hierarchical clustering. This method can be applied in two ways, agglomerative and divisive (Newman & Girvan, 2004). The agglomerative method finds nodes in the network that have the highest similarity and detects clusters based on these similarities. This process is however not the most successful, and has a tendency to overlook nodes that are peripheral to the cluster they belong to. In contrast, the divisive method attempts to find the least similar nodes in the network and removes the edge between them (Newman & Girvan, 2004). The Girvan-Newman algorithm, used in this thesis to identify clusters, uses the divisive method to remove the edges that have the highest betweenness centrality. Nodes with high betweenness centrality are generally connected to nodes outside of a cluster. Nodes with low betweenness centrality are more central in a cluster. This method will enable me to determine the clusters within the network.

After determining what the clusters in the network are, I calculated the average fertility intentions of the network members per cluster, by calculating the means within clusters and then assess the variation in these means by calculating the standard deviation in the means across these clusters. The final polarisation value is on the level of the entire network and consists of the variation between the average intentions between clusters. This means that the total polarisation of the network would be equal to 0 if the average intentions across clusters were similar. Cluster analysis also identifies isolates and dyads as their own clusters. Because these isolates will only have one opinion, their influence will be bigger than that of an alter that is part of a larger cluster. To negate that, I only included clusters of size 3 or larger. Figure 1 shows a network where the intentions of all the alters are known and there are similar numbers of pro-natal and child-free people in the network. In this example four clusters were detected, not including the isolate. The upper two clusters are connected by one node, but they are two separate clusters. The interconnecting node is part of the upper left cluster. In three out of these four clusters we see a dominant opinion within the cluster. The polarisation of this network is 0.768, making this one of the more polarised networks in the data. As mentioned earlier, the polarisation of a network is determined by the differences between the means in the clusters (through calculating a standard deviation of means). The fertility intentions of the network have been assigned the number -1 if they have negative fertility intentions, and 1 if they have positive fertility intentions. This means that the mean of the bottom left cluster will be 0.2 ((3-2)/5), while the mean of the bottom right cluster will be -1. The node that is connected to both of the upper clusters is part of the left cluster, but the node on the left with only a single tie to that cluster has been excluded, making the mean of that cluster -2/3 (-4/6). The mean of the last cluster is therefore (4/6). The standard deviation of all of these numbers is 0.768 which is the measure of polarisation. A second example of a network in this data can be seen in figure 2. This is a network in which the fertility intentions of most network members are unknown, and the overall level of polarisation is low. This network consists of four clusters, two isolates and a dyad. The isolates and the dyad are excluded from the calculation of polarisation. The clusters are the group of four nodes on the right side of the figure; the group of five nodes on the bottom of the figure; the nine nodes on the top left; and the three nodes on the top right. The last two clusters are again connected by one node, which is part of the cluster on the left. For model 2 the calculation of polarisation is relatively easy, three out of four clusters have a mean of 0, the last cluster is the one on the top left. This cluster includes one person with negative fertility intentions, 3 with positive fertility intentions, and 5 whose fertility intentions are unknown. The mean of this cluster is 0.222 (2/9). The standard deviation of these four clusters, and thereby the value of polarisation for this network, is 0.111.

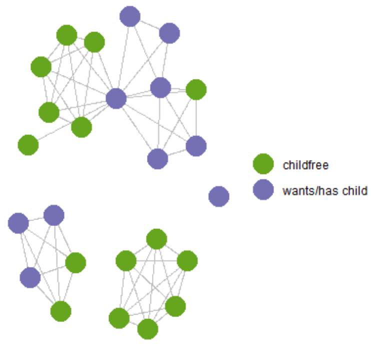


Figure 1: Network visualisation of a network with 25 alters, where all fertility intentions are known (15 childfree people, and 10 people that either want to or have children) and a relatively high measure of polarisation. Polarisation = 0.768

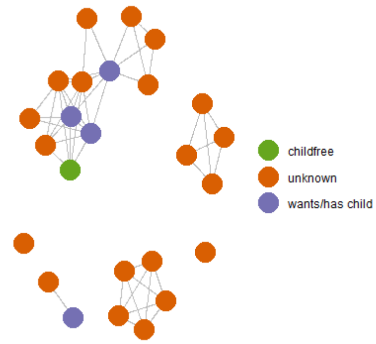


Figure 2: Network visualisation of a network with 25 alters, where most fertility intentions are unknown and with a low amount of polarisation. Polarisation = 0.111

The average polarisation across networks in this sample is small, but there is substantial variation (m = 0.33, SD = 0.19). Figure 3 shows that most observations of polarisation are close to 0.2. This means that there is not a lot of variation in the different scores between clusters. There are also 20 respondents with a value of 0 for polarisation. This means that these respondents have a network in which there is no difference in averages between clusters. Due to the operationalisation of fertility intentions of the network, this can mean different things: either a network that is entirely pro- or anti-natal, or a network in which the fertility intentions of the network members are completely unknown, or where fertility intentions are spread out equally across the clusters.

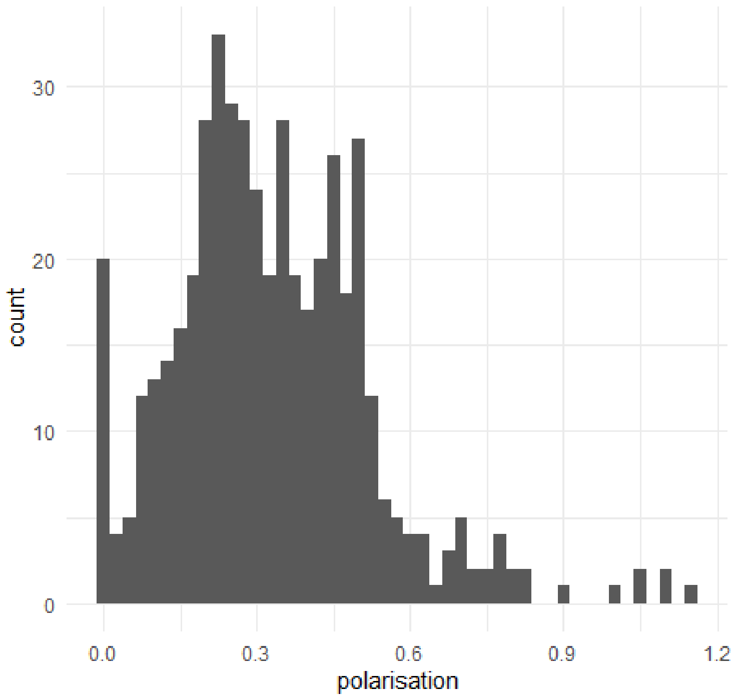


Figure 3: Distribution of polarisation.

4.6 Control variables

Apart from my predictor variables, I will also include a number of control variables. By adding control variables to a study, the influence of confounding and extraneous variables will be limited, which increases the internal validity of the study. The control variables in this study are age, education level, whether or not the respondent is in a serious relationship, density, and the number of alters whose fertility intentions are unknown. Age has an effect on fertility intentions, as people who are older tend to be more certain about their fertility intentions (Sobotka, 2009).

Education mostly has effects on fertility outcomes, but a higher education level can cause ambivalence about fertility intentions (Berrington & Pattaro, 2014).

The type of relationship the respondents have with their partner is included, because people in more serious relationships can be considered to be more open to having children (Berrington & Pattaro, 2014).

I used cohabitation as a measurement of the strength of the relationship, because my expectation is that people do not move in together if they are not serious about their relationship. Furthermore, the density of the network can have an effect on the amount and effectiveness of social pressure (Latané, 1981). This can result in networks with a higher density having more influence than networks with a lower density.

Lastly the number of alters whose fertility intentions are unknown are included because these are included in the measurement of polarisation, but can be argued not to have an effect on the fertility intentions of the respondents.

The variables that measured education and cohabitation are categorical variables. Cohabitation was measured as a yes/no question. I transformed these values into a binary variable, including those without a partner in the group that was not cohabiting with their partner at the time of the survey. Educational degree measured the highest education level the respondent had achieved. This was measured in 8 categories: primary school, vmbo (intermediate secondary education, US: junior high school, havo/vwo (higher secondary education/preparatory university education, US: senior high school), mbo (intermediate vocational education, US: junior college), hbo (higher vocational education, US: college), wo (university), Other, and Not (yet) completed any education. I assigned a numeric value to all of these, where a higher number means a higher level of education. The group "other" was assigned the lowest number (0), because there is no clear other classification for that group, and it is more likely that they haven't finished an education than otherwise.

The density of a network is calculated by dividing the number of ties, or connections, in the network by the maximum number of connections. In this dataset the maximum number of connections is 300. This means that density is the proportion of ties in the network, based on the total number of ties. A network in which all network members are connected to each other will therefore have a density of 1, while a network in which no one is connected to anyone will have a density of 0.

4.7 Analysis strategy

I will test my hypotheses using different kinds of regression analyses. The first hypothesis *"A network that is more pronatal will lead to more positive fertility intentions"*, will be tested by making use of an ordinal regression analysis. An ordinal regression analysis is generally used when the outcome variable is an ordinal variable. In this case the outcome variable has five values. An ordinal regression model divides the different outcome categories into binary groups and calculates the difference between them. The interpretation of an ordinal regression model is similar to a logistic regression model. A logistic regression model calculates the odds that a case falls within one of two groups based on the predictor variables, whereas an ordinal regression model calculates the odds that a case is more or less likely to fall within a category, as opposed to the lower categories. Although it is common to treat an ordinal variable with five categories as a linear variable, and thus using a linear regression for the analysis; in this case, many assumptions of a linear regression were violated, making a linear regression analysis unsuitable for the testing of this model (see Appendix C for further details).

The second hypothesis *"A polarised network will result in ambivalence in fertility intentions"* will be tested by first analysing the influence of the opinion of the network on the certainty of fertility intentions, and then to add the effect of polarisation on certainty. This is done in order to determine whether the effects found in this study are due to overall network opinion or whether network structure plays an important role. The analyses to test this hypothesis will consist of two logistic regression analyses and one ordinal regression analysis, due to the different operationalisations of certainty. A logistic regression is a regression where the outcome variable is a binary or dichotomous variable. The outcome of the regression analysis will be the change in the (log-)odds of the outcome variable.

Some additional analyses will be conducted that could delve into the mechanisms of social influence on fertility intentions. Before testing the second hypothesis it is first important to see if the overall opinion of the network has an effect on the certainty of the fertility intentions of the respondents. This will be done by testing the effect of the overall opinion of the network on certainty, and therefore will include all measurements of certainty. Furthermore, I will do multiple robustness checks on how opinions across clusters are associated with fertility intentions. This will be done in three ways, the outcomes of which can be found in Appendix D. The first is to try to determine the influence of networks selecting only those with only three clusters (n=160), which are the most common in the data. Having the number of clusters in the network fixed will facilitate the interpretation of the measurement of polarisation. Furthermore, reducing the number of clusters in the analysis can create a better understanding of the influence of individual clusters. The fewer clusters present in the network, the larger the influence of each individual cluster. The second is to try to determine the influence of large clusters. This was done by only including clusters which contained five people or more in the calculation of polarisation. The calculation using only large clusters is included because larger clusters within the network will be able to assert more social pressure than smaller clusters (Latané, 1981).

Finally, I will do an additional analysis to see what the influence of the network members whose fertility intentions are unknown on the effect of polarisation is. This will be done by removing them from the calculation of polarisation. The variable of polarisation will then be calculated only on the basis of the network members whose fertility intentions are known.

The second hypothesis will therefore be tested through a series of analyses that will not only include the direct effect of network polarisation on certainty, but also the influence of clusters and cluster size in determining the effects of network structure on the certainty of fertility intentions. The results of the analyses will be presented in terms of the odds ratio. All of the tables report the odds ratio for the variables, the confidence interval of the odds ratio, and the p-value. The odds ratio can take on any value above zero, any value smaller than one means that the effect is negative, any value larger than one means that the effect is positive.

4.8 Reliability

A regression analysis is considered to be reliable when its assumptions are met. The testing of the assumptions can be found in Appendix C, a short summary and explanation will be provided here. The assumptions for both logistic regression models were met; the outcome variables are binary and the cases are independent. The main assumption for an ordinal regression model is that of parallel slopes. This assumption means that the slopes for the different categories of the outcome variable are parallel to each other. If this is the case, then the estimate created by the regression analysis is applicable to all categories of the outcome variable. This means that when this assumption does not hold, the estimate predicted by the regression analysis does not match the actual relationship between the concepts tested. There is no standardised test for this in R, but an approximation can be made by determining whether the slopes for variables of interest from the models are roughly similar for each comparison within the dependent variable being the level of certainty. These calculations determine the slopes for the different groups by running multiple logistic regression models for the different groups and comparing the slopes to each other. The results of these assumption analyses can be seen in Appendix C, which shows that for different values of polarisation the difference between the different outcome categories stays relatively constant. This means that the assumption of parallel slopes holds for the predictor variable of polarisation. This is however not the case for the calculation of polarisation that only includes three clusters. It is therefore important to be more careful in interpreting the results of this analysis.

Apart from the assumptions of the regression models, it is important to check the data for influential points. This can be done through calculating the leverage or Cook's distance for the different models. The leverage measures how far the independent variable values of a particular observation is from the other observations. A high leverage means that this distance is large and that this observation could be an outlier. Similarly, Cook's distance is a tool that can be used to identify outliers. The outcomes of these calculations can be found in Appendix 2. There are a few points that were identified as possible outliers, but because these points were not very influential they were kept in the analyses.

5 Results

5.1 Descriptives

In this chapter, I will give a description of the variables used in the analyses, starting with the dependent variables followed by the independent variables. This description will be given by first showing the distributions of the dependent and independent variables, and then by discussing the bivariate statistics and correlations between all of the variables in the models. The distribution of polarisation can be seen in figure 3 and was discussed in the methods chapter, and will therefore not be discussed further here.

The distribution of the fertility intentions of the respondents shows that most of the respondents have a positive attitude towards having children. This is illustrated in figure 4, which shows the distribution of the fertility intentions of the respondents.

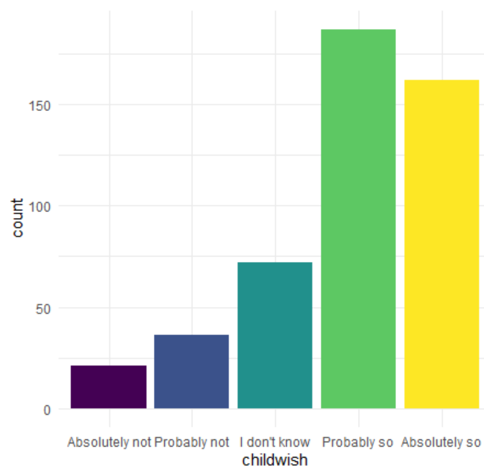
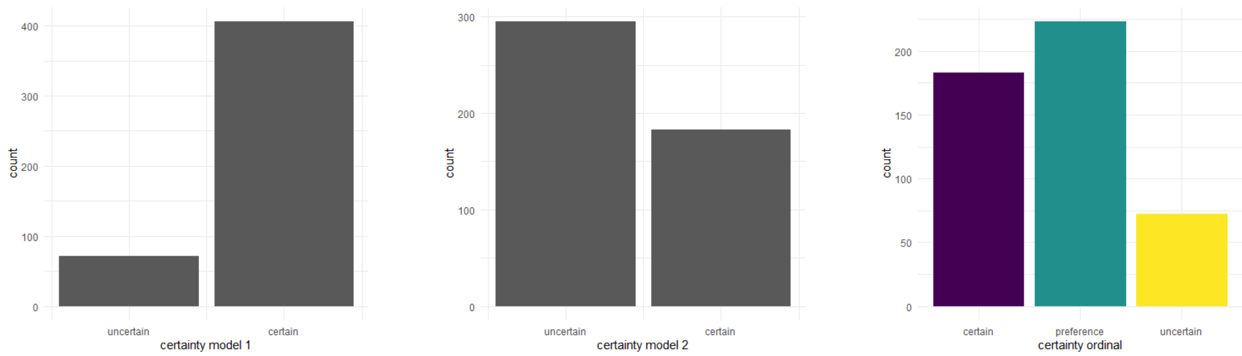


Figure 4: Distribution of the fertility intentions of the respondents (n=478).

As seen in figure 4, 4.39% of the respondents did not want to have a child, and 7.53% probably not. In contrast, 33.9% definitely wanted a child, and 39.1% probably. The remaining 15.1% did not know (see appendix B for coding). This meant for the first operationalisation of uncertainty 84.9% was uncertain, and 15.1% was certain. In the second operationalisation, respondents who indicated "probably" were also included into the uncertain category, meaning that 61.7% was uncertain, and 38.3% was certain. In the third operationalisation, which separates the outcome variable into certain, preference and uncertain, 38.3% of the respondents were certain, 46.7% had a preference, and 15.1% were uncertain. These distributions can also be seen in figures 5a, 5b, and 5c.



(a) First operationalisation of certainty; uncertain (n=72), certain (n=406).

(b) Second operationalisation of certainty; uncertain (n=295), certain (n=183).

(c) Ordinal operationalisation of certainty; certain (n=183), preference (n=223), uncertain (n= 72)

Figure 5: Distribution of operationalisation of certainty.

The majority of respondents have a network that generally has a positive attitude towards having children (m = 11.93, SD = 5.97), but there is a lot of variation in the number of people with positive fertility intentions (see figure 6). There are however some respondents (n=6) who have a network where the majority of the network members have a negative opinion about having children.

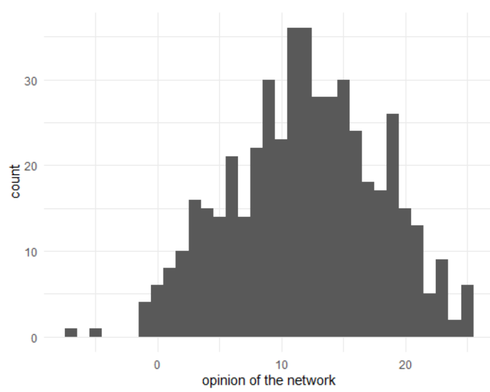


Figure 6: The distribution of the opinion on having children across the network of 478 respondents. Twenty-five refers to the maximum number of alters in the network with positive fertility intentions.

Due to the operationalisation of polarisation, the network members whose fertility intentions are unknown are included in the calculation of polarisation. It could be argued however that these people have less influence on the fertility intentions of the respondents than those network members whose fertility intentions are known. As can be seen in figure 7, there are 17 respondents who know the fertility intentions of all of their network members. The average number of unknown fertility intentions of the network members is around 10 (m = 10.1, median = 9.5).

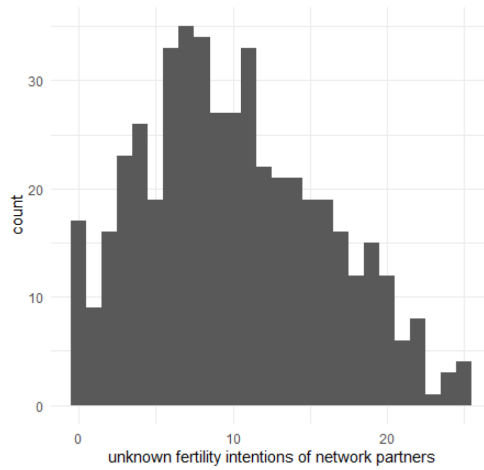


Figure 7: Distribution of network members whose fertility intentions are unknown

Table 1 reports the descriptive statistics for all variables used in the analyses. Apart from the distributions of these variables, it is also important to see how they are related to each other. Table 2 shows the correlations between the continuous variables in the models. The correlation between the opinion of the network and the number of unknown opinions is very strong (-0.797), and between the number of unknown opinions and the absolute value of the opinion of the network the correlation is stronger (-0.828). The reason for this is because most network members have positive fertility intentions (see figure 6). This creates a higher value for the variable that measures the opinion of the network, while lowering the number of unknown opinions. Even though these variables are in the same model, these strong correlations are not very problematic, because their multicollinearity is still relatively low (see appendix C).

Table 1: Bivariate statistics of all variables in all models

variable	mean/ %(n)	sd	min	median	max
Age	26.275	5.687	18	25	41
Educational degree	4.363	1.451	1	4	8
Cohabiting: No	61.5				
Yes	38.5				
density	0.238	0.108	00.20	0.223	0.670
Unknown opinion network	10.105	5.845	0	9.5	25
Opinion network	11.929	5.968	-7	12	25
Opinion network absolute values	11.996	5.832	0	12	25
Polarisation	0.330	0.193	0	0.306	1.154
Childwish	3.906	1.086	1	4	5
Certainty 1:					
Certain	84.9				
Uncertain	15.1				
Certainty 2:					
Certain	38.3				
Uncertain	61.7				
Certainty 3:					
Certain	38.3				
Preference	46.7				
Uncertain	15.1				

Table 2: Correlations between the variables

	age	Educational degree	density	Opinion network	absolute value opinon network	Unknown opinion network	polarisation
age	1						
Educational degree	0.332	1					
density	-0.173	-0.039	1				
Opinion network	0.218	0.157	0.197	1			
absolute value opinon network	0.223	0.153	0.200	0.990	1		
Unknown opinion network	-0.368	-0.117	-0.096	-0.797	-0.828	1	
polarisation	-0.024	-0.014	-0.010	-0.122	0.112	-0.086	1

5.2 Network influence on fertility intentions

To test the first hypothesis *"A network that is more pronatal will lead to more positive fertility intentions"*. An ordinal regression was run to test the effects of the opinion of the network on the fertility intentions of the respondents. This analysis showed that the overall effect of the network on fertility intentions seems to be a positive one (see table 3). In line with hypothesis 1, a positive effect on the number of positive fertility intentions on the desire to have children was found, with an odds ratio of 1.114 ($p<0.001$). This effect is rather large, as the addition of a single person with positive fertility intentions increases the odds of the respondent being in the next higher group of positive fertility intentions by approximately 11%. In other words, all other variables being constant, the odds that a respondent has positive fertility intentions increases by 15 (exp(0.108*25)), if that person has a network in which all network partners have positive fertility intentions compared to a network where the opinions are perfectly divided.

Table 3: Estimates from an ordinal regression model with childwish as dependent variable n= 477

coefficients	OR	CI OR Lower - Upper	Pr (>|z|)
Age	0.860	0.829 - 0.893	<0.001
Educational degree	1.103	0.972 - 1.250	0.127
Cohabiting	2.313	1.598 - 3.352	<0.001
Density	1.756	0.324 - 9.493	0.513
Unknown opinion network	1.051	0.993 - 1.113	0.087
Opinion network	1.114	1.052 - 1.179	<0.001
Log Likelihood ratio	103.17		
df	6		
p-value	<0.001		

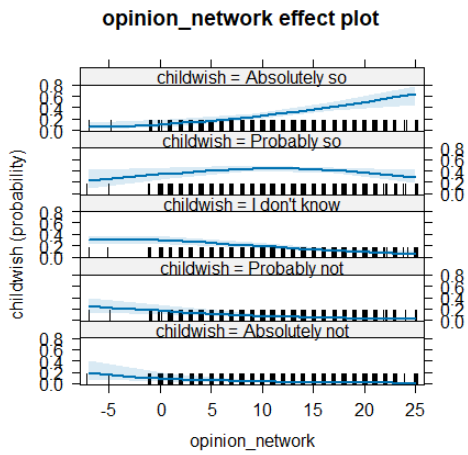


Figure 8: Influence of network opinion on fertility intentions

Figure 8 shows the effect of the opinion of the network on the fertility intentions of the respondents. It does so by showing the slopes predicted probabilities for each category of the outcome variable. As seen in this figure, and consistent with the hypothesis, the effect of network opinion is negative for negative fertility intentions (and for people who do not know), and strongly positive for the most positive intentions.

5.3 Influence of network opinion on certainty

The second hypothesis *"A polarised network will result in ambivalence in fertility intentions"* will be tested in multiple steps, and through multiple analyses. The first analysis will test the influence of the opinion of the network on certainty, in the next part I will test the effects of polarisation on certainty. The effect of the overall opinion of the network is determined by creating the absolute value of the opinion of the network variable used in the previous analysis. A higher value for the overall opinion of the network means that most of the network is in agreement about their fertility intentions. A low value means that the network is more divided or that the fertility intentions of the network are unknown. Table 4 reports the effects of the analyses that test the effects of the overall network opinion on certainty. The models are numbered based on the operationalisation of certainty they use, and this numbering is used across all analyses. As mentioned before, the first model only includes people who do not know if they want to have children as uncertain. The second model includes those with a preference in the group of uncertain people. The third model separates these groups into people who are certain, people with a preference, and people who are uncertain.

Table 4 shows that there does not seem to be a strong effect of the overall opinion of the network on certainty about fertility intentions. The direction of the effect varies across the different models, and none of them are significant. The lack of significance in all of the models and the large confidence intervals, which in all cases include 1, show that these effects could very likely be different due to chance.

Table 4: effects of the overall opinion of the network on certainty about fertility intentions

Model 1 dependent variable:"I don't know" is uncertain, all others are certain.

Model 2: dependent variable: "i don't know" and "probably yes/no" are uncertain, "absolutely yes/no" is certain

Model 3: dependent variable: "i don't know" is uncertain, "probably yes/no" is preference, " absolutely yes/no" is certain.

n=477

	Model 1			Model 2			Model 3		
	OR	CI OR Lower - Upper	p-value	OR	CI OR Lower - Upper	p-value	OR	CI OR Lower - Upper	p-value
age	0.901	0.856 - 0.948	$<$0.001	0.971	0.934 - 1.010	0.157	1.056	1.017 - 1.095	0.004
Educational degree	1.339	1.090 - 1.645	0.005	1.011	0.880 - 1.162	0.874	0.915	0.805 - 1.039	0.173
cohabiting	1.724	0.958 - 3.105	0.069	1.609	1.066 - 2.401	0.021	0.588	0.404 - 0.857	0.006
density	7.365	0.499 - 108.858	0.146	3.411	0.561 - 20.742	0.182	0.225	0.042 - 1.206	0.082
Unknown opinion network	0.971	0.898 - 1.051	0.465	0.957	0.901 - 1.017	0.161	1.037	0.978 - 1.101	0.219
Overall opinion network	1.038	0.960 - 1.123	0.343	0.985	0.929 - 1.045	0.615	1	0.944 - 1.058	0.991
Log likelihood ratio	33.29			11.75			23.39		
df	6			6			6		
p-value	$<$0.001			0.068			0.001		

5.4 Influence of polarisation

The next section will test whether the addition of network structure to the analysis has an effect on certainty about fertility intentions. The following analysis calculates the polarisation for the entire network and tests its effects across the three models discussed earlier.

All three models show that there is a positive effect of polarisation on fertility intentions (see table 5). The odds ratios (ORmodel1 = 1.679, ORmodel2= 1.499) mean that, keeping all other variables constant, the odds of a person with a value of 1 for polarisation are approximately 1.6 times as likely to be certain about their fertility intentions than a person with a value of 0 for polarisation. However, the large confidence intervals show that there is a lot of dispersion within this prediction, which means that the effect of polarisation could be very different with a different sample. The lack of significance in both the predictor and the control variables suggest that these models are not good predictors of certainty about fertility intentions. The odds ratio of 0.607 for polarisation in the third model means that the odds of a person with a polarised network with the value of polarisation of 1 to be uncertain about their fertility intentions is about 0.6 times that of a person without a polarised network who will have a polarisation value of 0. This means that it is about twice as likely that someone with a polarised network is more certain about their fertility intentions than someone with a polarised network. Figure 9 shows visualisations of the effects of the first two models. The x-axis represents the different variables that are included in the models, the y-axis represents the outcome variable. The slope for each variable is the effect of that variable on the certainty of the respondents, while holding all other variables constant. The visualisation of the third model can be found in figure 10, which shows the effect of polarisation on all three categories for certainty. In this figure you can see that the certainty of the respondents increases with polarisation. The regression tables which give the value for each slope can be found in Appendix D.

Table 5: effects of polarisation on certainty about fertility intentions

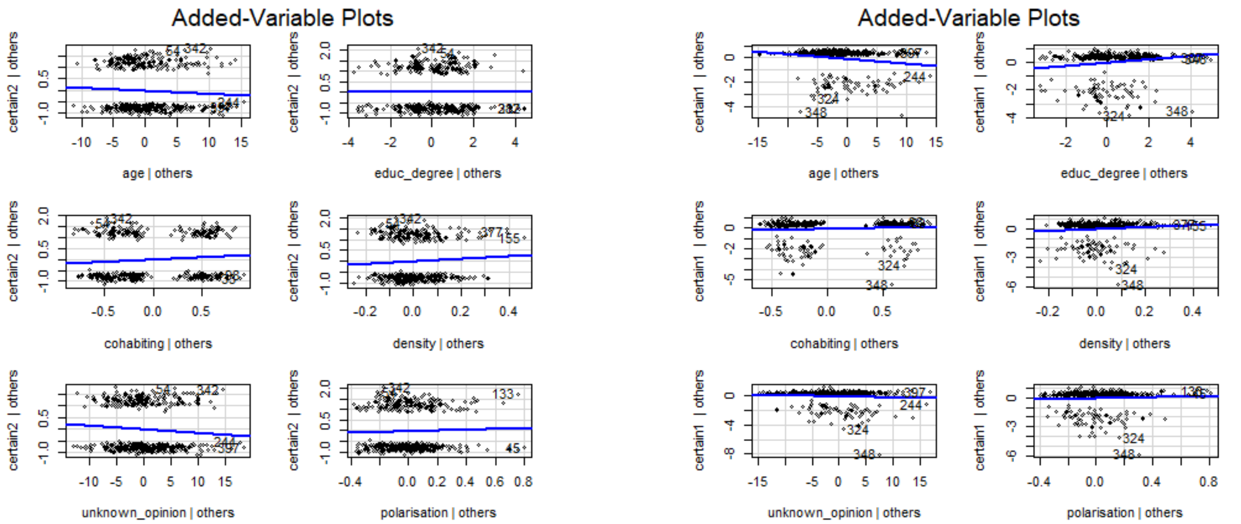
Model 1 dependent variable:“I don’t know” is uncertain, all others are certain.

Model 2: dependent variable: “i don’t know” and “probably yes/no” are uncertain, “absolutely yes/no” is certain

Model 3: dependent variable: “i don’t know” is uncertain, “probably yes/no” is preference, “ absolutely yes/no” is certain.

n=477

	Model 1			Model 2			Model 3		
	OR	CI OR Lower - Upper	p-value	OR	CI OR Lower - Upper	p-value	OR	CI OR Lower - Upper	p-value
age	0.899	0.855 - 0.946	<0.001	0.974	0.936 - 1.012	0.191	1.055	1.016 - 1.094	0.004
Educational degree	1.359	1.107 - 1.665	0.003	1.007	0.878 - 1.155	0.926	0.913	0.806 - 1.036	0.160
cohabiting	1.728	0.960 - 3.111	0.068	1.601	1.067 - 2.403	0.023	0.590	0.405 - 0.859	0.006
density	9.552	0.659 - 138.440	0.098	3.218	0.550 - 18.802	0.198	0.214	0.041 - 1.128	0.069
Unknown opinion network	0.946	0.900 - 0.993	0.025	0.971	0.937 - 1.005	0.104	1.036	1.002 - 1.071	0.035
Polarisation	1.679	0.417 - 6.750	0.466	1.499	0.572 - 3.933	0.411	0.607	0.249 - 1.480	0.272
Log Likelihood ratio	32.950			12.170			24.600		
df	6			6			6		
p-value	<0.001			0.058			<0.001		



(a) Added variable plot of model 1

(b) Added variable plot of model 2

Figure 9: Effects of polarisation on certainty for models 1 and 2

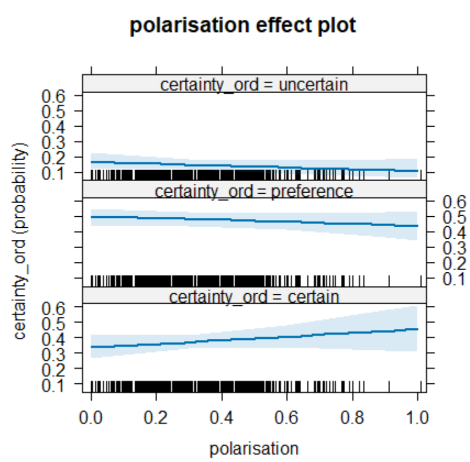


Figure 10: Influence of polarisation on certainty

Additionally, I will do three robustness checks to increase the validity of these analyses. These checks will be done to determine whether the effects of polarisation are due to the predictive ability of the variable, or due to the choices made in the operationalisation of this variable. The checks will include a calculation of polarisation where only networks with three clusters are included; a calculation of polarisation where only large clusters are included; and a calculation of polarisation where only known fertility intentions of the network members are included in the calculation. These checks will then allow me to determine whether there is an effect of polarisation on certainty about fertility intentions. The results of these robustness tests can be found in Appendix D and were similar to the outcomes found in the analyses presented above.

6 Conclusion

The goal of this thesis was to investigate what the effects of network structure are on fertility intentions through quantitative data analysis. The LISS data enabled me to study personal networks of a large sample of Dutch women and their fertility intentions. The research question "How does the polarity or opinion diversity of personal social networks shape fertility intentions?" was tested with the hypotheses H1: A network that is more pronatal will lead to more positive fertility intentions and H2: a polarised network will result in ambivalence in fertility intentions. Through cluster analysis the polarisation of the networks was calculated. Support was found for the first hypothesis. There is a positive relationship between the opinion of the network and the fertility intentions of the respondents. Additionally, it was expected that the opinion of the network would shape how (un)certain people were about having children. However, little support for this expectation was found, even across different operationalisations of uncertainty. This means that there is not an effect of polarisation.

7 Discussion

7.1 Reflection on findings

The first analysis shows that the fertility intentions of network members influence the fertility intentions of the respondents. The opinion of the network had a positive effect on the fertility intentions of the respondents, meaning that the more network members there were with certain fertility intentions, the more likely it is that the respondent has the same fertility intentions. This is in line with the theory of planned behaviour (Ajzen, 1991), and the social influence theory (Kelman, 1974), and is also consistent with the findings of several other studies (c.f. e.g. Balbo & Barban, 2014; Bernardi & Klärner, 2014; Madhavan et al., 2003; Lois, 2016).

However, the analyses that tried to determine the relationship between the polarisation of the network did not yield significant results. While the overall opinion of the network has an influence on the fertility intentions of the respondents, it does not have any effect on the certainty the respondents have. This finding may be related to the lack of respondents with a high level of polarisation in their networks (see figure 3). A more even distribution of people with a high and low level of polarisation in their networks could have provided a better insight into the effects of polarisation.

Furthermore, the direction of the effects that were found, significant only in one of the robustness checks (see Appendix D), were contradictory to the hypothesis. People who had higher levels of polarisation in their networks were more certain about their fertility intentions than those with lower levels of polarisation in their networks. This is not in line with the findings of Keim (2011). It is however more similar to the findings of Lois (2016), who found that those with a polarised network have an average transition rate to family formation compared to other types of networks as identified by Keim (2011). A possible explanation for these findings is that exposure to contrasting viewpoints (e.g. in a highly polarised network) can help someone to better understand their own, or make them reject the viewpoint that they disagree with (Keijzer et al., 2024).

Finally, it could be argued that people choose their networks based on shared opinions. People tend to enjoy the company of those that are similar to them over that of those who are different. This principle, called homophily, could mean that the opinions of the network members is determined by the opinions of the respondents, rather than the other way around (Steglich et al., 2012). However, it is not possible to test whether the network is formed based on the opinions of the respondents, or if the opinions of the respondents are formed based on the network due to the cross-sectional design of this study. It is however more likely that the network shapes the opinion, because a longitudinal study has shown that the network does not change that much after experiencing parenthood, and changes to the network are more likely to be in the form of new ties than in the loss of old ties (Klärner et al., 2016).

7.2 Limitations

A limitation of this research is that it was not possible to test for the effects of social pressure, due to the way the data was collected. Social pressure was measured on a network level, making it impossible to use on the level of the clusters used in the analysis. Additionally social pressure was only measured as the pressure to have children. It could be interesting to see if there is as much pressure not to have children as there is to have children. The theories used to develop the hypotheses uses social pressure as the main mechanism for how the social influence affects the respondents. For the second hypothesis concerning the effects of network structure on certainty about fertility intentions, it operates on the assumption that both pro-natal and anti-natal network members are asserting social pressure. This was however not measured in the data and could have a large influence on the results. A study to determine whether the inclusion of social pressure, including the pressure not to have children, on the level of the individual clusters has an effect on the influence of polarisation could therefore provide more insight into how these processes work.

Another limitation is related to the measurement of certainty about fertility intentions. There was no question directly related to the level of certainty of the respondents. It is therefore possible that the variable used to create the measurements of certainty is not a good representation of the actual certainty of the respondents. A possible solution for this could be to use another variable or a combination of variables related to ideal family size and the certainty about that family size to create an approximation of certainty with this data. Another solution could lie in the collection of new data that includes this information.

A third limitation is the lack of inclusion of tie strength in the analyses. The emotional closeness of people is an important factor in the effectiveness of social influence (Latané, 1981). The inclusion of tie strength could show which clusters within the network are exerting more influence, which could directly influence the results. If the network seems to have a high level of polarisation, but one the respondent is much closer to one group in the network, then it is more likely that the respondent is more influenced by that group than by the rest of the network. The measurement of polarisation used in this research would only be accurate if the different groups in the network have the same influence on the respondent. Group size was controlled for in one of the robustness checks (see appendix D). Tie strength was however excluded from the analyses because the operationalisation of polarisation is based on the network clusters, whereas the tie strength is measured on an individual level.

7.3 Insights

This study provides insight in how a persons social environment can affect the formation of their attitudes and behaviours. This large dataset containing network data of 758 Dutch women, enables researchers to study several kinds of connections and effects of social networks on fertility. Social networks do influence people's decisions about their fertility intentions. It is however clear that the polarisation of a personal network does not have an effect on the certainty about fertility intentions. In order to create further insights into how demographic changes come to be, it is interesting to study the moderating effects of social pressure and tie strength on the effects of network structures on opinion formation.

References

Agrillo, C., & Nelini, C. (2008). Childfree by choice: a review. *Journal of cultural geography*, *25*(3), 347–363.

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179–211.

Axelrod, R. (1986). An evolutionary approach to norms. *American political science review*, *80*(4), 1095–1111.

Balbo, N., & Barban, N. (2014). Does fertility behavior spread among friends? *American sociological review*, *79*(3), 412–431.

Balbo, N., & Mills, M. (2011). The effects of social capital and social pressure on the intention to have a second or third child in france, germany, and bulgaria, 2004–05. *Population Studies*, *65*(3), 335–351.

Baumgaertner, B. O., Tyson, R. C., & Krone, S. M. (2016). Opinion strength influences the spatial dynamics of opinion formation. *The Journal of mathematical sociology*, *40*(4), 207–218.

Bernardi, L. (2003). Channels of social influence on reproduction. *Population research and policy review*, *22*, 427–555.

Bernardi, L., Keim, S., & Von der Lippe, H. (2007). Social influences on fertility: A comparative mixed methods study in eastern and western germany. *Journal of mixed methods research*, *1*(1), 23–47.

Bernardi, L., & Klärner, A. (2014). Social networks and fertility. *Demographic research*, *30*, 641–670.

Berrington, A., & Pattaro, S. (2014). Educational differences in fertility desires, intentions and behaviour: A life course perspective. *Advances in life course research*, *21*, 10–27.

Bhattacharya, J., & Chakraborty, S. (2012). Fertility choice under child mortality and social norms. *Economics Letters*, *115*(3), 338–341.

Biondi, E., Boldrini, C., Passarella, A., & Conti, M. (2023). Dynamics of opinion polarization. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*.

Borgatti, S. P., Everett, M. G., & Johnson, J. C. (2018). *Analyzing social networks*. Sage.

Bühler, C., & Fratczak, E. (2007). Learning from others and receiving support: The impact of personal networks on fertility intentions in poland. *European societies*, *9*(3), 359–382.

Buyukkececi, Z., Leopold, T., van Gaalen, R., & Engelhardt, H. (2020). Family, firms, and fertility: A study of social interaction effects. *Demography*, *57*(1), 243–266.

CBS. (n.d.). *Kinderen krijgen.* Retrieved from https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/levensloop/kinderen-krijgen.

CBS. (2023). *Geboorte.* Retrieved from https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/bevolkingsgroei/geboren-kinderen

Colleran, H. (2020). Market integration reduces kin density in women's ego-networks in rural poland. *Nature communications*, *11*(1), 266.

Dallmeyer, S., Wicker, P., & Breuer, C. (2017). How an aging society affects the economic costs of inactivity in germany: empirical evidence and projections. *European review of aging and physical activity*, *14*, 1–9.

de Albuquerque, P. C. A. M. (2018). Population ageing and the labour market during the recent crisis in portugal. *Portuguese Journal of Social Science*, *17*(1), 105–119.

Facciani, M., Lazić, A., Viggiano, G., & McKay, T. (2023). Political network composition predicts vaccination attitudes. *Social Science & Medicine*, 116004.

Han, K. J. (2013). Saving public pensions: Labor migration effects on pension systems in european countries. *The Social Science Journal*, *50*(2), 152–161.

Hank, K., & Kreyenfeld, M. (2003). A multilevel analysis of child care and women's fertility decisions in western germany. *Journal of marriage and family*, *65*(3), 584–596.

Interian, R., G. Marzo, R., Mendoza, I., & Ribeiro, C. C. (2023). Network polarization, filter bubbles, and echo chambers: an annotated review of measures and reduction methods. *International Transactions in Operational Research*, *30*(6), 3122–3158.

Kavas, S., & De Jong, J. (2020). Exploring the mechanisms through which social ties affect fertility decisions in turkey. *Journal of Marriage and Family*, *82*(4), 1250–1269.

Keijzer, M., Mäs, M., & Flache, A. (2024). Polarization on social media: Micro-level evidence and macro-level implications. *Journal of Artificial Societies and Social Simulation*, *27*(1).

Keim, S. (2011). *Social networks and family formation processes.* Springer.

Keim, S., Klärner, A., & Bernardi, L. (2009). Qualifying social influence on fertility intentions: Composition, structure and meaning of fertility-relevant social networks in western germany. *Current sociology*, *57*(6), 888–907.

Kelman, H. C. (1974). Social influence and linkages between the individual and the social system: Further thoughts on the processes of compliance, identification, and internalization. In *Perspectives on social power* (p. 125–171). Albine.

Kim, Y.-S., Lee, J., Moon, Y., Kim, K. J., Lee, K., Choi, J., & Han, S.-H. (2018). Unmet healthcare needs of elderly people in korea. *BMC geriatrics*, *18*, 1–9.

Klärner, A., Keim, S., & von der Lippe, H. (2016). Social network dynamics in the course of family formation: Results from a mixed-methods longitudinal study. *International Review of Social Research*, *6*(4).

Kmetty, Z., & Tardos, R. (2022, Jul). Party nexus position generator. *Social Networks*, *70*, 112–125.

Kohler, H.-P., Behrman, J. R., & Watkins, S. C. (2001). The density of social networks and fertility decisions: Evidence from south nyanza district, kenya. *Demography*, *38*(1), 43–58.

Kuhnt, A.-K., & Trappe, H. (2016). Channels of social influence on the realization of short-term fertility intentions in germany. *Advances in Life Course Research*, *27*, 16–29.

Latané, B. (1981). The psychology of social impact. *American psychologist*, *36*(4), 343.

Lazer, D., Rubineau, B., Chetkovich, C., Katz, N., & Neblo, M. (2010). The coevolution of networks and political attitudes. *Political communication*, *27*(3), 248–274.

Liefbroer, A. C., & Billari, F. C. (2010). Bringing norms back in: A theoretical and empirical discussion of their importance for understanding demographic behaviour. *Population, space and place*, *16*(4), 287–305.

Liu, L., Wang, X., Chen, X., Tang, S., & Zheng, Z. (2021). Modeling confirmation bias and peer pressure in opinion dynamics. *Frontiers in Physics*, *9*.

Lois, D. (2016). Types of social networks and the transition to parenthood. *Demographic Research*, *34*, 657–688.

Lois, D., & Becker, O. A. (2014). Is fertility contagious? using panel data to disentangle mechanisms of social network influences on fertility decisions. *Advances in life course research*, *21*, 123–134.

Lyngstad, T. H., & Prskawetz, A. (2010). Do siblings' fertility decisions influence each other? *Demography*, *47*(4), 923–934.

Madhavan, S., Adams, A., & Simon, D. (2003). Women's networks and the social world of fertility behavior. *International family planning perspectives*, 58–68.

Mathews, P., & Sear, R. (2013). Does the kin orientation of a british woman's social network influence her entry into motherhood? *Demographic Research*, *28*, 313–340.

Ministerie van Volksgezondheid, W. e. S. (2023). *Burgerperspectieven bericht 2 2022.* Author. Retrieved from https://www.scp.nl/publicaties/publicaties/2022/12/29/continu-onderzoek-burgerperspectieven---bericht-2-2022

Mönkediek, B., & Bras, H. (2014). Strong and weak family ties revisited: Reconsidering european family structures from a network perspective. *The History of the Family*, *19*(2), 235–259.

Mönkediek, B., & Bras, H. (2018). Family systems and fertility intentions: Exploring the pathways of influence. *European Journal of Population*, *34*, 33–57.

Montgomery, M. R., & Casterline, J. B. (1996). Social learning, social influence, and new models of fertility. *Population and development review*, *22*, 151–175.

Mueller, S. T., & Tan, Y.-Y. S. (2018). Cognitive perspectives on opinion dynamics: The role of knowledge in consensus formation, opinion divergence, and group polarization. *Journal of Computational Social Science*, *1*, 15–48.

Munshi, K., & Myaux, J. (2006). Social norms and the fertility transition. *Journal of development Economics*, *80*(1), 1–38.

Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, *69*(2), 026113.

Pink, S., Leopold, T., & Engelhardt, H. (2014). Fertility and social interaction at the workplace: Does childbearing spread among colleagues? *Advances in life course research*, *21*, 113–122.

Robins, G. (2015). *Doing social network research: Network-based research design for social scientists*. SAGE.

Sobotka, T. (2009). Intentions de fécondité inférieures au seuil de remplacement en autriche. *European Journal of Population/Revue européenne de Démographie*, *25*, 387–412.

Stadel, M., & Stulp, G. (2022). Balancing bias and burden in personal network studies. *Social Networks*, *70*, 16–24.

Steglich, C., Sinclair, P., Holliday, J., & Moore, L. (2012). Actor-based analysis of peer influence in a stop smoking in schools trial (assist). *Social Networks*, *34*(3), 359–369.

Stulp, G. (2020). Methods and materials of the social networks and fertility survey (sociale relaties en kinderkeuzes). *DataverseNL.*, *10*.

Stulp, G. (2021). Collecting large personal networks in a representative sample of dutch women. *Social Networks*, *64*, 63–71.

Stulp, G., & Barrett, L. (2021). Do data from large personal networks support cultural evolutionary ideas about kin and fertility? *Social Sciences*, *10*(5), 177.

Tang, B., & Li, Z. (2021). A country-level empirical study on the fiscal effect of elderly population health: the mediating role of healthcare resources. In *Healthcare* (Vol. 10, p. 30).

Vacca, R. (2020). Structure in personal networks: Constructing and comparing typologies. *Network Science*, *8*(2), 142–167.

Yang, X. (2018). Social influence or personal attitudes? understanding users' social network sites continuance intention. *Kybernetes*, *48*(3), 424–437.

8 Appendix A

Preparation of data and operationalisations

activating packages

library(tidyverse)

library(FertNet)

library(sna)

library(ggplot2)

library(tidygraph)

library(ggraph)

library(purrr)

library(igraph)

library(broom)

library(car)

library(MASS)

library(Hmisc)

library(foreign)

library(effects)

library(rms)

activating dataset

```
\newline data <- produce_data()

\newline data<- produce_data(background_vars = TRUE)
```

remove parents from data

```
data<- data |>

filter(has_children == "No")
```

create variable for alters intent

```
alters_intentions <- function(alter_attr_data) {

alter_attr_data$alter_intent = case_when(
```

```
alter_attr_data$childfree_a == "Prefers to remain childless" &
    alter_attr_data$childwish_a == "I don't know whether person wishes to have
    children" & alter_attr_data$has_child_a == "Does not have (a) child(ren) and
    is not expecting a child"~ -1,
   alter_attr_data$childfree_a == "I don't know whether person wishes to remain
       childless" & alter_attr_data$childwish_a == "I don't know whether person
       wishes to have children" & alter_attr_data$has_child_a == "Does not have
       (a) child(ren) and is not expecting a child" ~ 0,
   alter_attr_data$childfree_a == "I don't know whether person wishes to remain
       childless" & alter_attr_data$childwish_a == "Wishes to have children" &
       alter_attr_data$has_child_a == "Does not have (a) child(ren) and is not
       expecting a child" ~ 1,
   alter_attr_data$has_child_a == "Does have (a) child(ren) or is expecting a
       child" ~ 1,
   TRUE ~ NA)
  return(alter_attr_data)
}
```

variable alter_intent for visualisation

```
alters_intentions_categorical <- function(alter_attr_data) {
  alter_attr_data$alter_intent_categorical = case_when(
   alter_attr_data$childfree_a == "Prefers to remain childless" &
       alter_attr_data$childwish_a == "I don't know whether person wishes to have
       children" & alter_attr_data$has_child_a == "Does not have (a) child(ren)
       and is not expecting a child"~ "childfree",
   alter_attr_data$childfree_a == "I don't know whether person wishes to remain
       childless" & alter_attr_data$childwish_a == "I don't know whether person
       wishes to have children" & alter_attr_data$has_child_a == "Does not have
       (a) child(ren) and is not expecting a child" ~ "unknown",
   alter_attr_data$childfree_a == "I don't know whether person wishes to remain
       childless" & alter_attr_data$childwish_a == "Wishes to have children" &
```

```
      alter_attr_data$has_child_a == "Does not have (a) child(ren) and is not
      expecting a child" ~ "wants/has child",
   alter_attr_data$has_child_a == "Does have (a) child(ren) or is expecting a
      child" ~ "wants/has child",
   TRUE ~ NA)
  return(alter_attr_data)
}
```

adds alter intent into alter_attr

```
data <- data |>
  mutate(
   alter_attr = map(alter_attr, function(x) alters_intentions(x))
  )

data <- data |>
  mutate(
   alter_attr = map(alter_attr, function(x) alters_intentions_categorical(x))
  )

includes alter intent in tidygraph

data <- data |>
  filter(!is.na(edgelist)) |>
  mutate(
   tidygraph = map2(alter_attr, edgelist,
                function(x, y) tbl_graph(x, as.data.frame(y), directed = FALSE))
  )
```

Identifying clusters

```
add_membership <- function(alter_attr, membership) {
  alter_attr[, "membership"] <- membership
```

```
  return(alter_attr)
}

data <- data |>
  mutate(
   communities = map(tidygraph,
                    function(x) edge.betweenness.community(x)),
   membership = map(communities, function(x) x$membership),
   alter_attr = map2(alter_attr, membership,
                    function(x, y) add_membership(x, y) )
  )
```

allow to filter isolates and dyads

```
add_connections <- function(alter_attr, connections) {
  alter_attr[, "connections"] <- connections
  return(alter_attr)
}

data<- data |>
  mutate(
   connections = map(tidygraph,
                    function(x) degree(x)),
   alter_attr = map2(alter_attr, connections, function(x, y) add_connections(x,
       y))
  )
```

Function to calculate polarisation

```
calculate_polarisation <- function(alter_attr) {
  means <- alter_attr |> group_by(membership) |>
   filter(connections > 1) |>
   summarise(means = mean(alter_intent, na.rm = TRUE))
```

```
  sd(means$means)
}
```

adds polarisation for all respondents

```
data <- data |>
  mutate(
   polarisation = map_dbl(alter_attr,
                        function(x) calculate_polarisation(x))
  )
```

filters out those with missing values on polarisation

```
data<- data |>
  filter(
   !is.na(polarisation)
  )
```

calculate number of clusters (excepting isolates/dyads)

```
data <- data |> mutate(
  community_detection = map(tidygraph, function(x)
      igraph::cluster_edge_betweenness(x, directed = FALSE) ),
  community_sizes = map(community_detection, function(x) c(table(x$membership))),
  comm_2orhigher = map_dbl(community_sizes, function(x) sum(x >= 2)),
)
```

calculate polarisation for those with 3 clusters in the network

```
data <- data |> mutate(polaris_comm3 = if_else(comm_2orhigher == 3,
    polarisation, NA))
```

calculates polarisation for large clusters

```
calculate_polarisation_large <- function(alter_attr) {
```

```
  means <- alter_attr |>
   group_by(membership) |>
   mutate(community_size = n()) |>
   filter(community_size >= 5) |>
   summarise(means = mean(alter_intent, na.rm = TRUE))
  sd(means$means)
}

data <- data |>
  mutate(
   polarisation_large = map_dbl(alter_attr,
                            function(x) calculate_polarisation_large(x))
  )
```

calculate polarisation for all known opinons in the networks

```
calculate_polarisation_allknown <- function(alter_attr) {
  means <- alter_attr |>
   group_by(membership) |>
   filter(connections >1) |>
   filter(alter_intent != 0) |>
   summarise(means = mean(alter_intent, na.rm = TRUE))
  sd(means$means)
}

data <- data |>
  mutate(
   polarisation_allknown = map_dbl(alter_attr,
                              function(x) calculate_polarisation_allknown(x))
  )
```

outcome variable certainty ordinal

```
data<- data |>
  mutate(
   certainty_ord = fct_collapse(childwish,
                              "uncertain" = "I don't know",
                              "preference" = "Probably so",
                              "preference" = "Probably not",
                              "certain" = "Absolutely not",
                              "certain" = "Absolutely so"
                                 )
  )



data <- data |>
  mutate(
   certainty_collapsed = fct_collapse(childwish,
                                 "1" = "Absolutely not",
                                 "1" = "Absolutely so",
                                 "2" = "Probably not",
                                 "2" = "Probably so",
                                 "3" = "I don't know" )
  )

data$certainty_collapsed <- as.numeric(data$certainty_collapsed)
```

outcome variable certainty logistic model 1

```
data <- data |>
  mutate(
   certain1 = if_else(childwish == "I don't know", 0, 1)
  )
```

outcome variable certainty logistic model 2

```
data <- data |>
  mutate(
   certain2 = if_else(childwish == "Probably not"|childwish == "I don't know"|
       childwish == "Probably so", 0, 1)
  )
```

overall network opinion

```
calculate_opinion <- function(alter_attr){
  alter_attr |>
    filter(alter_attr$alter_intent !=0) |>
    summarise(sum(alter_intent, na.rm = TRUE))
}

data<- data |>
  mutate(
    opinion_network = map(alter_attr,
                            function(x) calculate_opinion(x)))


data$opinion_network <- unlist(data$opinion_network)
```

absolute value opinion network

```
data<- data |>
  mutate(
   opinon_network_absval = sqrt(I(opinion_network)^2)
  )
```

calculate number of unknown opinions

```
data <- data |>
  mutate(
   unknown_opinion = map(alter_attr, function(x) sum(ifelse(x$alter_intent ==
```

```
    "0", 1, 0), na.rm = TRUE) )
  )

data$unknown_opinion <- unlist(data$unknown_opinion)
```

creating numeric continuous control variables

```
data<- data |>
  mutate(
   educ_degree = fct_recode(educ_degree,
                          "2" = "primary school",
                          "3" = "vmbo (intermediate secondary education, US:
                             junior high school)",
                          "4" = "havo/vwo (higher secondary education/preparatory
                             university education, US: senior high school)",
                          "5" = "mbo (intermediate vocational education, US:
                             junior college)",
                          "6" = "hbo (higher vocational education, US: college)",
                          "7" = "wo (university)",
                          "0" = "other",
                          "1" = "Not (yet) completed any education",
                          "<NA>" = "<NA>" )
  )


data<- data |>
  mutate(
   cohabiting = fct_recode(cohabiting,
                          "0" = "No",
                          "1" = "Yes"


   )
```

```
  )
```

adds people without a partner to the group that is not cohabiting with their partner

```
data$cohabiting[is.na(data$cohabiting)]<- 0


data$educ_degree <- as.numeric(data$educ_degree)

data$cohabiting <- as.numeric(data$cohabiting)
```

calculate density

```
data <- data |> mutate(density = map_dbl(tidygraph, function(x) edge_density(x)))
```

9 Appendix B

Descriptive statistics and visualisations

Age:

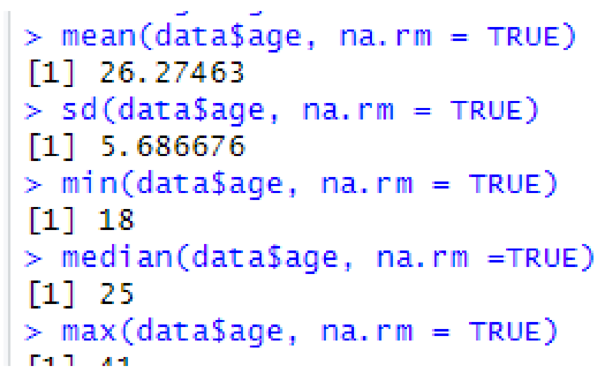
```
mean(data$age, na.rm = TRUE)

sd(data$age, na.rm = TRUE)

min(data$age, na.rm = TRUE)

median(data$age, na.rm =TRUE)

max(data$age, na.rm = TRUE)
```



```
> mean(data$age, na.rm = TRUE)
[1] 26.27463
> sd(data$age, na.rm = TRUE)
[1] 5.686676
> min(data$age, na.rm = TRUE)
[1] 18
> median(data$age, na.rm =TRUE)
[1] 25
> max(data$age, na.rm = TRUE)
```

Figure B1: descriptives age

Educational degree:

```
mean(data$educ_degree, na.rm = TRUE)

sd(data$educ_degree, na.rm = TRUE)

min(data$educ_degree, na.rm = TRUE)

median(data$educ_degree, na.rm =TRUE)

max(data$educ_degree, na.rm = TRUE)
```

```
> mean(data$educ_degree, na.rm = TRUE)
[1] 4.362683
> sd(data$educ_degree, na.rm = TRUE)
[1] 1.451061
> min(data$educ_degree, na.rm = TRUE)
[1] 1
> median(data$educ_degree, na.rm =TRUE)
[1] 4
> max(data$educ_degree, na.rm = TRUE)
[1] 8
```

Figure B2: Descriptives educational degree

Cohabitation:

```
data |>
  group_by( cohabiting ) |>
  summarise( percent = 100 * n() / nrow( data ) )
```

```
# A tibble: 2 × 2
  cohabiting percent
       <dbl>   <dbl>
1          1    61.5
2          2    38.5
```

Figure B3: Percentage cohabiting

1= no, 2=yes

Density:

```
mean(data$density, na.rm = TRUE)
sd(data$density, na.rm = TRUE)
min(data$density, na.rm = TRUE)
median(data$density, na.rm =TRUE)
max(data$density, na.rm = TRUE)
```

```
> mean(data$density, na.rm = TRUE)
[1] 0.2379637
> sd(data$density, na.rm = TRUE)
[1] 0.1078457
> min(data$density, na.rm = TRUE)
[1] 0.02
> median(data$density, na.rm =TRUE)
[1] 0.2233333
> max(data$density, na.rm = TRUE)
[1] 0.67
```

Figure B4: Descriptives density

Unkown opinion:

```
mean(data$unknown_opinion, na.rm = TRUE)

sd(data$unknown_opinion, na.rm = TRUE)

min(data$unknown_opinion, na.rm = TRUE)

median(data$unknown_opinion, na.rm =TRUE)

max(data$unknown_opinion, na.rm = TRUE)
```

```
> mean(data$unknown_opinion, na.rm = TRUE)
[1] 10.1046
> sd(data$unknown_opinion, na.rm = TRUE)
[1] 5.845454
> min(data$unknown_opinion, na.rm = TRUE)
[1] 0
> median(data$unknown_opinion, na.rm =TRUE)
[1] 9.5
> max(data$unknown_opinion, na.rm = TRUE)
[1] 25
```

Figure B5: Descriptives unknown opinion

Opinion network:

```
mean(data$opinion_network, na.rm = TRUE)

sd(data$opinion_network, na.rm = TRUE)

min(data$opinion_network, na.rm = TRUE)

median(data$opinion_network, na.rm =TRUE)

max(data$opinion_network, na.rm = TRUE)
```

```
> mean(data$opinion_network, na.rm = TRUE)
[1] 11.92887
> sd(data$opinion_network, na.rm = TRUE)
[1] 5.967695
> min(data$opinion_network, na.rm = TRUE)
[1] -7
> median(data$opinion_network, na.rm =TRUE)
[1] 12
> max(data$opinion_network, na.rm = TRUE)
[1] 25
```

Figure B6: descriptives opinion network

Absolute value opinion network:

```
mean(data$opinon_network_absval, na.rm = TRUE)
sd(data$opinon_network_absval, na.rm = TRUE)
min(data$opinon_network_absval, na.rm = TRUE)
median(data$opinon_network_absval, na.rm =TRUE)
max(data$opinon_network_absval, na.rm = TRUE)
```

```
> mean(data$opinon_network_absval, na.rm = TRUE)
[1] 11.99582
> sd(data$opinon_network_absval, na.rm = TRUE)
[1] 5.831669
> min(data$opinon_network_absval, na.rm = TRUE)
[1] 0
> median(data$opinon_network_absval, na.rm =TRUE)
[1] 12
> max(data$opinon_network_absval, na.rm = TRUE)
[1] 25
```

Figure B7: Descriptives absolute value opinion network

Polarisation

```
mean(data$polarisation, na.rm = TRUE)
sd(data$polarisation, na.rm = TRUE)
min(data$polarisation, na.rm = TRUE)
median(data$polarisation, na.rm = TRUE)
max(data$polarisation, na.rm = TRUE)
```

```
> mean(data$polarisation, na.rm = TRUE)
[1] 0.3298073
> sd(data$polarisation, na.rm = TRUE)
[1] 0.1933994
> min(data$polarisation, na.rm = TRUE)
[1] 0
> median(data$polarisation, na.rm = TRUE)
[1] 0.3062512
> max(data$polarisation, na.rm = TRUE)
[1] 1.154701
```

Figure B8: Descriptives polarisation

Childwish:

```
mean(data$childwish_numerical, na.rm = TRUE)
sd(data$childwish_numerical, na.rm = TRUE)
min(data$childwish_numerical, na.rm = TRUE)
median(data$childwish_numerical, na.rm = TRUE)
max(data$childwish_numerical, na.rm = TRUE)
```

```
> mean(data$childwish_numerical, na.rm = TRUE)
[1] 3.905858
> sd(data$childwish_numerical, na.rm = TRUE)
[1] 1.086185
> min(data$childwish_numerical, na.rm = TRUE)
[1] 1
> median(data$childwish_numerical, na.rm = TRUE)
[1] 4
> max(data$childwish_numerical, na.rm = TRUE)
[1] 5
```

Figure B9: Descriptives childwish

```
data |>
  group_by( childwish ) |>
  summarise( percent = 100 * n() / nrow( data ) )
```

```
# A tibble: 5 × 2
  childwish        percent
  <ord>              <dbl>
1 Absolutely not      4.39
2 Probably not        7.53
3 I don't know       15.1
4 Probably so        39.1
5 Absolutely so      33.9
```

Figure B10: Percentages childwish

Certainty; logistic model 1

```
data |>
  group_by( certain1 ) |>
  summarise( percent = 100 * n() / nrow( data ) )
```

```
# A tibble: 2 × 2
  certain1 percent
     <dbl>   <dbl>
1        0    15.1
2        1    84.9
```

Figure B11: Percentages certain 1

0=uncertain, 1= certain

Certainty; logistic model 2:

```
data |>
  group_by( certain2 ) |>
  summarise( percent = 100 * n() / nrow( data ) )
```

```
# A tibble: 2 x 2
  certain2 percent
     <dbl>   <dbl>
1        0    61.7
2        1    38.3
```

Figure B12: Percentages certain 2

0=uncertain, 1= certain

Certainty; ordinal model:

```
data |>
  group_by( certainty_collapsed ) |>
  summarise( percent = 100 * n() / nrow( data ) )
```

```
# A tibble: 3 x 2
  certainty_collapsed percent
                <dbl>   <dbl>
1                   1    38.3
2                   2    46.7
3                   3    15.1
```

Figure B13: Percentages certain 3

1= certain, 2 = preference, 3 = certain

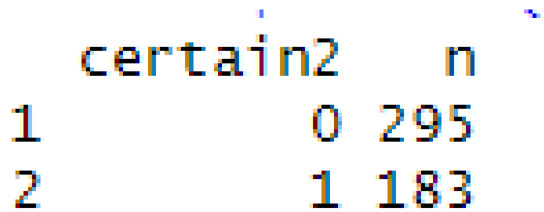
Counts of people for each group in all operationalisation of certainty

```
Model 1
data |> count(certain1)
```

```
> data |> count(certain1)
  certain1   n
1        0  72
2        1 406
```

Figure B14: Count certain1

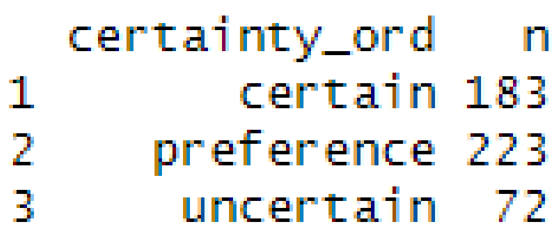
```
Model 2
data |> count(certain2)
```



	certain2	n
1	0	295
2	1	183

Figure B15: Count certain2

```
Model 3
data |> count(certainty_ord)
```



	certainty_ord	n
1	certain	183
2	preference	223
3	uncertain	72

Figure B16: Count certain3

calculate correlations

```
datacorrelations <-
  dplyr::select(data, age, educ_degree, density, opinion_network,
      opinon_network_absval, unknown_opinion, polarisation)

cor(datacorrelations, use = "pairwise.complete.obs")
```

	age	educ_degree	density	opinion_network	opinon_network_absval	unknown_opinion	polarisation
age	1.00000000	0.33160620	-0.172856697	0.2175013	0.2230312	-0.36832624	-0.024040520
educ_degree	0.33160620	1.00000000	-0.038975221	0.1567040	0.1530148	-0.11725789	-0.014376421
density	-0.17285670	-0.03897522	1.000000000	0.1974883	0.1997786	-0.09615716	-0.009624173
opinion_network	0.21750133	0.15670401	0.197488328	1.0000000	0.9909330	-0.79734092	-0.121731655
opinon_network_absval	0.22303124	0.15301484	0.199778623	0.9909330	1.0000000	-0.82782983	-0.110750265
unknown_opinion	-0.36832624	-0.11725789	-0.096157156	-0.7973409	-0.8278298	1.00000000	-0.086143019
polarisation	-0.02404052	-0.01437642	-0.009624173	-0.1217317	-0.1107503	-0.08614302	1.000000000

Figure B17: Correlations

bar chart outcome variables

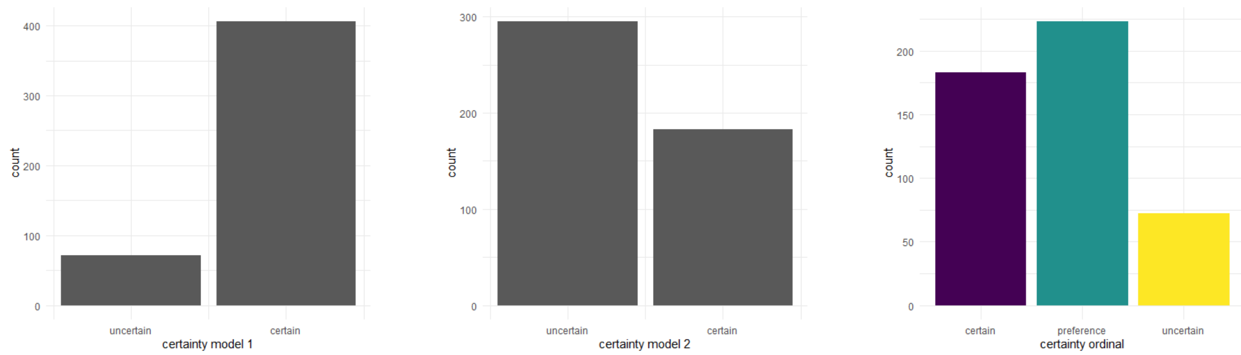
```
bar chart certainty1
data |> ggplot(aes(x = certain1))+
  geom_bar()+
  scale_x_continuous(breaks = seq(0, 1, by = 1), labels = c("uncertain",
      "certain")) +
  labs(x = "certainty model 1")+
  theme_minimal()


bar chart certainty2
data |> ggplot(aes(x = certain2))+
  geom_bar()+
  scale_x_continuous(breaks = seq(0, 1, by = 1), labels = c("uncertain",
      "certain")) +
  labs(x = "certainty model 2")+
  theme_minimal()


bar chart certainty_ord
data |> ggplot(aes(x = certainty_ord))+
  geom_bar(aes(fill = certainty_ord), show.legend = FALSE)+
  labs(x = "certainty ordinal")+
  theme_minimal()


bar chart childwish
data |> ggplot(aes(x = childwish))+
  geom_bar(aes(fill = childwish), show.legend = FALSE)+
```

```
  theme_minimal()
```



(a) First operationalisation of certainty; uncertain (n=72), certain (n=406).

(b) Second operationalisation of certainty; uncertain (n=295), certain (n=183).

(c) Ordinal operationalisation of certainty; certain (n=183), preference (n=223), uncertain (n= 72)

Figure B18: Distribution of operationalisation of certaint.

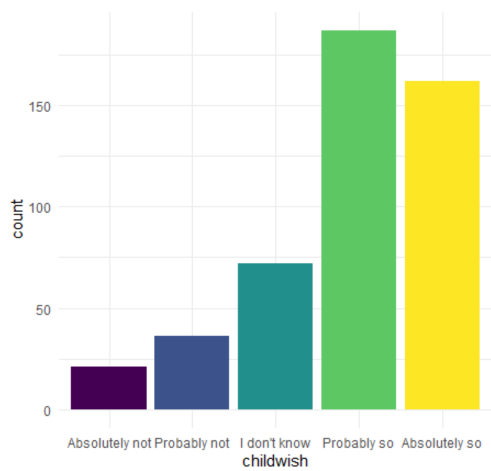


Figure B19: Distribution of the fertility intentions of the respondents (n=478).

distribution number of clusters

```
data |> ggplot(aes(x = comm_2orhigher))+
  geom_histogram(binwidth = 1)+
  labs(x = "number of communities")+
  theme_minimal()
```

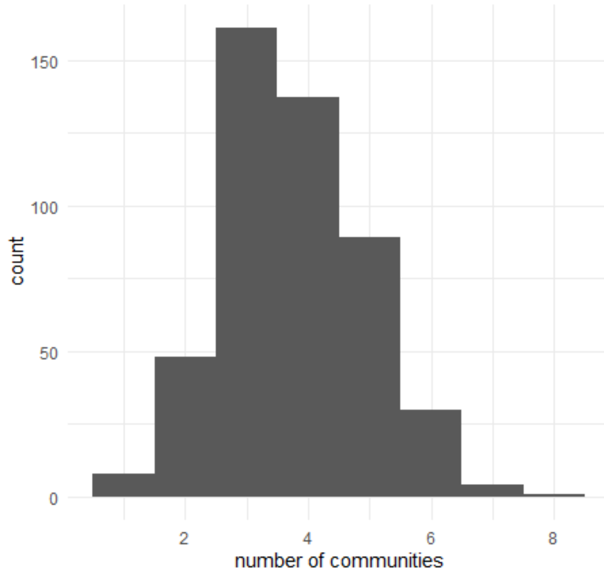


Figure B20: Histogram number of communities

distribution number of unknown opinions

```
data |> ggplot(aes(x = unknown_opinion))+
  geom_histogram(binwidth = 1)+
  labs(x = "unknown fertility intentions of network partners")+
  theme_minimal()
```

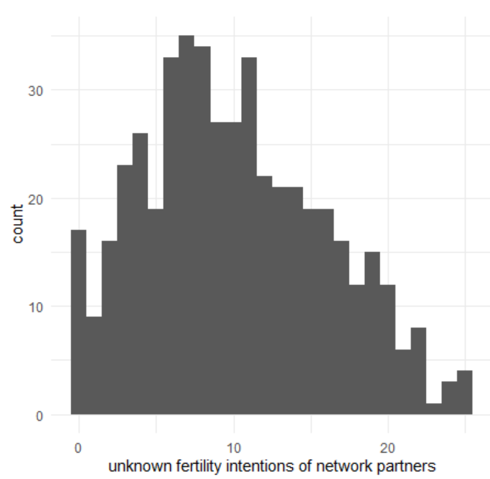


Figure B21: Distribution of network members whose fertility intentions are unknown visualisation of a polarised network

```
ggraph(data$tidygraph[[344]], layout = "kk") +
  geom_edge_link(colour = "grey") +
  geom_node_point(aes(colour = alter_intent_categorical), size = 7) +
  scale_color_manual(breaks = c("childfree", "unknown", "wants/has child"),
                values=c("#66A61E", "#D95F02", "#7570B3"))+
  labs(colour = NULL) +
  theme_graph()
```

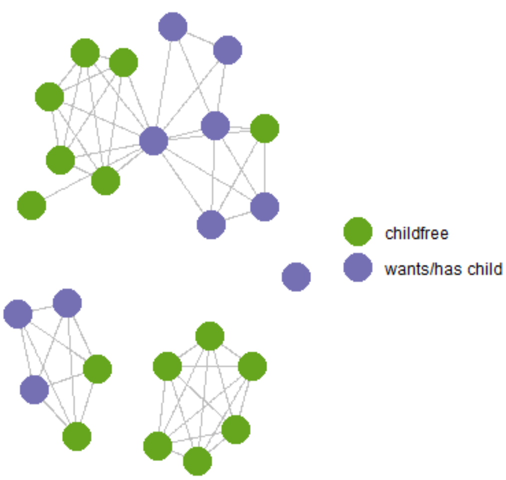


Figure B22: Network visualisation

visualisation of a non-polarised network

```
ggraph(data$tidygraph[[17]], layout = "kk") +
  geom_edge_link(colour = "grey") +
  geom_node_point(aes(colour = alter_intent_categorical), size = 7) +
  scale_color_manual(breaks = c("childfree", "unknown", "wants/has child"),
                values=c("#66A61E", "#D95F02", "#7570B3"))+
  labs(colour = NULL) +
  theme_graph()
```

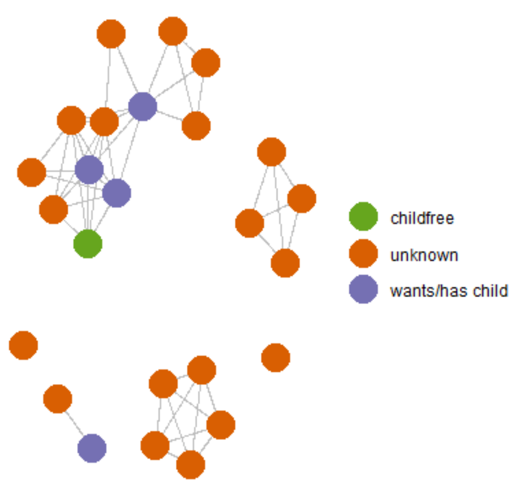


Figure B23: Network visualisation

historgram polarisation

```
data |> ggplot(aes(x = polarisation))+
  geom_histogram(binwidth = 0.025)+
  theme_minimal()
```

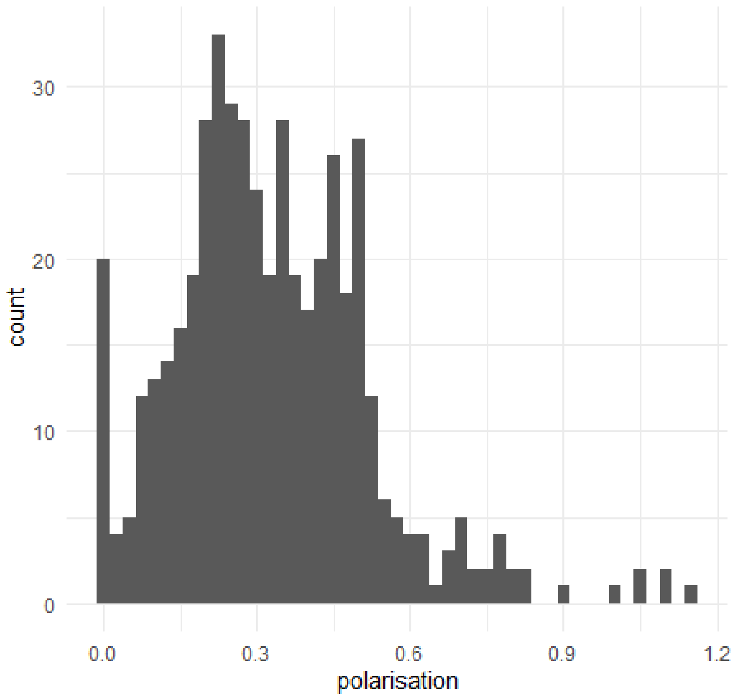


Figure B24: Distribution of polarisation.

histogram opinion network

```
data |> ggplot(aes(x = opinion_network))+
  geom_histogram(binwidth = 1)+
   labs(x = "opinion of the network")+
  theme_minimal()
```

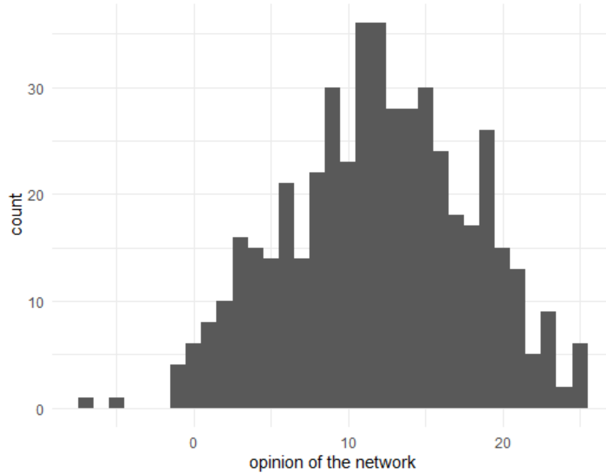


Figure B25: The distribution of the opinion on having children across the network of 478 respondents. Twenty-five refers to the maximum number of alters in the network with positive fertility intentions.

10 Appendix C

Ordinal regression analysis on the effects of network opinion on fertility intentions

Ordinal model

```
general_ordmodel <- polr(childwish~
                         age+
                         educ_degree+
                         cohabiting+
                         density+
                         unknown_opinion+
                         opinion_network,
                         data = data,
                         Hess=TRUE)
```

Proportional odds

```
glm(I(as.numeric(childwish) >= 2) ~ opinion_network, family="binomial", data =
    data)
```

```
Call:  glm(formula = I(as.numeric(childwish) >= 2) ~ opinion_network,
    family = "binomial", data = data)

Coefficients:
    (Intercept)  opinion_network
        2.63293          0.03964

Degrees of Freedom: 477 Total (i.e. Null);  476 Residual
Null Deviance:      172.3
Residual Deviance: 171.2        AIC: 175.2
```

Figure C1: Proportional odds

```
glm(I(as.numeric(childwish) >= 3) ~ opinion_network, family="binomial", data
    = data)
```

```
Call:  glm(formula = I(as.numeric(childwish) >= 3) ~ opinion_network,
    family = "binomial", data = data)

Coefficients:
    (Intercept)  opinion_network
        1.39179          0.05428

Degrees of Freedom: 477 Total (i.e. Null);  476 Residual
Null Deviance:      349.3
Residual Deviance: 344.1        AIC: 348.1
```

Figure C2: Proportional odds

```
glm(I(as.numeric(childwish) >= 4) ~ opinion_network, family="binomial", data
    = data)
```

```
glm(I(as.numeric(childwish) >= 5) ~ opinion_network, family="binomial", data =
    data)
```

```
Call:  glm(formula = I(as.numeric(childwish) >= 4) ~ opinion_network,
    family = "binomial", data = data)

Coefficients:
    (Intercept)  opinion_network
        0.27470          0.06306

Degrees of Freedom: 477 Total (i.e. Null);  476 Residual
Null Deviance:      557.5
Residual Deviance: 544.5        AIC: 548.5
```

Figure C3: Proportional odds

```
Call:  glm(formula = I(as.numeric(childwish) >= 5) ~ opinion_network,
    family = "binomial", data = data)

Coefficients:
    (Intercept)  opinion_network
       -0.99970          0.02743

Degrees of Freedom: 477 Total (i.e. Null);  476 Residual
Null Deviance:      612.1
Residual Deviance: 609.3        AIC: 613.3
```

Figure C4: Proportional odds

VIF

```
vif(general_ordmodel)
```

```
            age     educ_degree      cohabiting         density unknown_opinion opinion_network
       1.456573        1.163322        1.149823        1.097440        3.898247        3.778326
```

Figure C5: VIF score

Analyses to test the effects of overall network opinion on certainty

Model 1

```
log_model_netop1 <- glm(certain1~age +
                    educ_degree +
                    cohabiting +
                    density +
                    unknown_opinion +
                    opinon_network_absval,
                 family="binomial",
                 data=data)
```

checking linear relationship model 1

```
plot(log_model_netop1, 1)
```

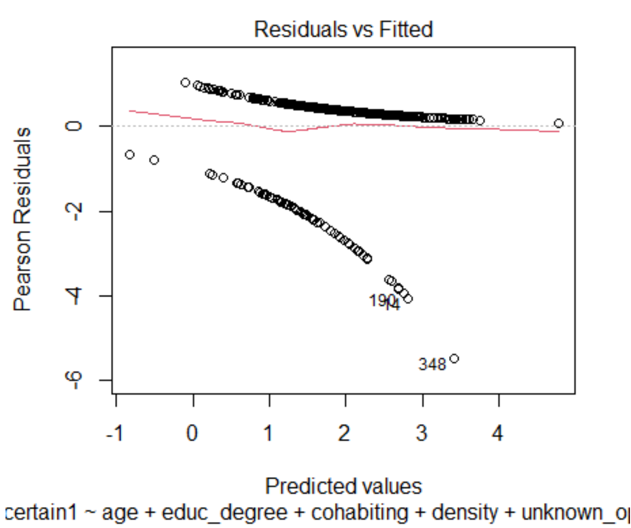


Figure C6: Linearity model 1

checking multicollinearity model 1

```
vif(log_model_netop1)
```

```
            age        educ_degree         cohabiting            density
       1.554486           1.214845           1.088409           1.086282
unknown_opinion opinon_network_absval
       3.295402           2.963164
```

Figure C7: VIF model 1

influential values model 1

```
plot(log_model_netop1, 4)

plot(log_model_netop1, 5)
```

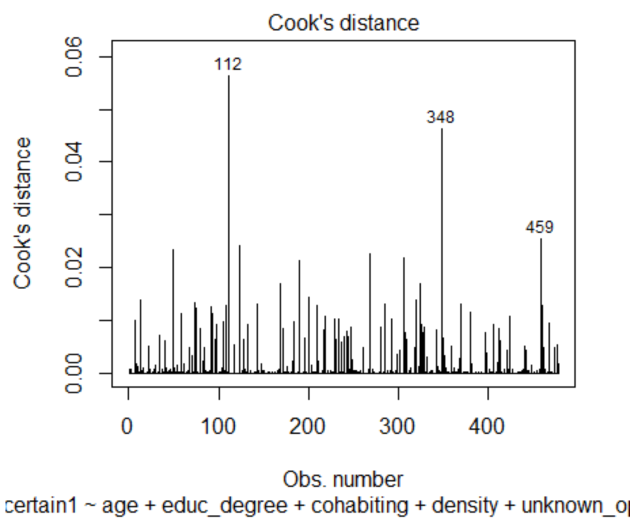


Figure C8: Cook's distance

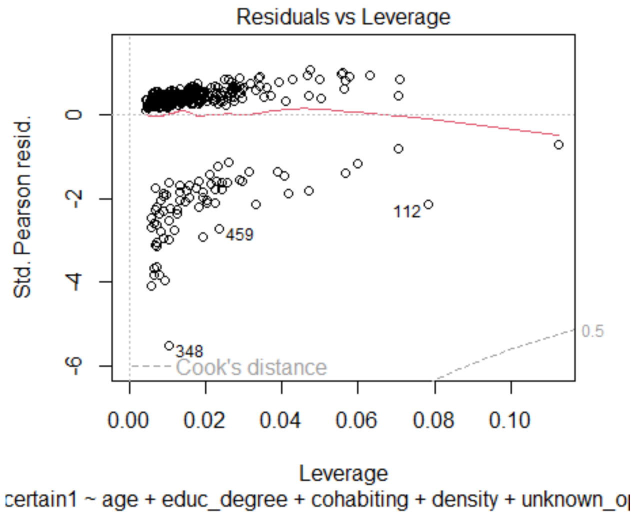


Figure C9: Leverage

Model 2

```
log_model_netop2 <- glm(certain2~age +
                    educ_degree +
                    cohabiting +
                    density +
                    unknown_opinion +
                    opinon_network_absval,
                family="binomial",
                data=data)
```

checking linear relationship model 2

```
plot(log_model_netop2, 1)
```

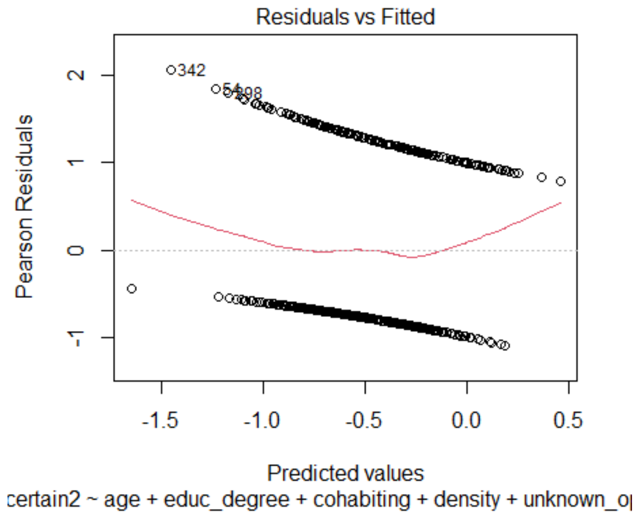


Figure C10: Linearity

checking multicollinearity model 2

```
vif(log_model_netop2)
```

```
            age         educ_degree          cohabiting             density
       1.455758            1.181446            1.135223            1.096832
unknown_opinion opinon_network_absval
       3.575299            3.371473
```

Figure C11: VIF model 2

influential values model 2

```
plot(log_model_netop2, 4)
plot(log_model_netop2, 5)
```

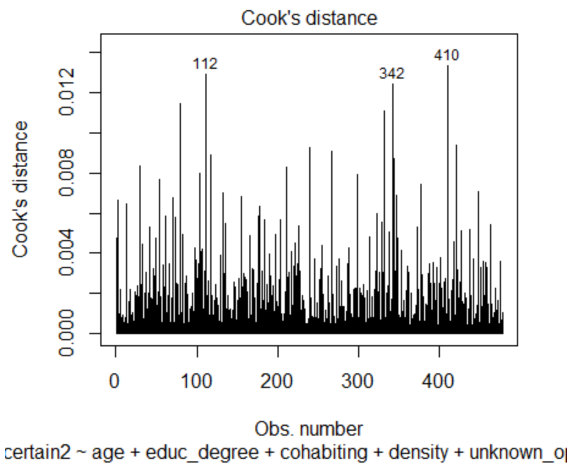


Figure C12: Cook's distance

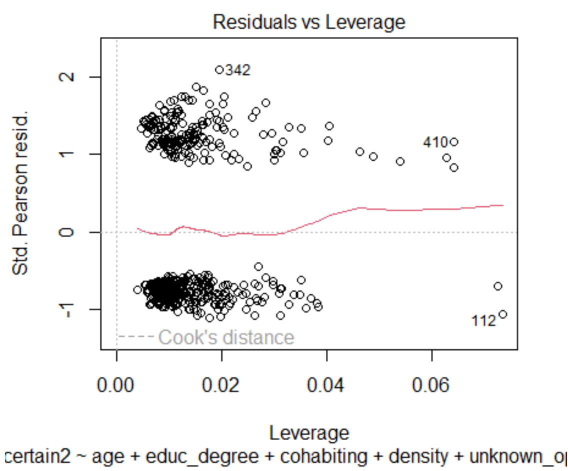


Figure C13: Leverage

Model 3

```
ord_model_netop <- polr(certainty_ord~
                     age+
                     educ_degree+
                     cohabiting+
                     density+
                     unknown_opinion+
                     opinon_network_absval,
                  data = data,
                  Hess=TRUE)
```

proportional odds

```
glm(I(as.numeric(certainty_ord) >= 2) ~ opinion_network, family="binomial", data
    = data)
glm(I(as.numeric(certainty_ord) >= 3) ~ opinion_network, family="binomial", data
    = data)
```

```
Call:  glm(formula = I(as.numeric(certainty_ord) >= 2) ~ opinion_network,
    family = "binomial", data = data)

Coefficients:
    (Intercept)  opinion_network
        0.70285         -0.01877

Degrees of Freedom: 477 Total (i.e. Null);  476 Residual
Null Deviance:      636.2
Residual Deviance: 634.8        AIC: 638.8
> glm(I(as.numeric(certainty_ord) >= 3) ~ opinion_network, family="binomial", data = data)

Call:  glm(formula = I(as.numeric(certainty_ord) >= 3) ~ opinion_network,
    family = "binomial", data = data)

Coefficients:
    (Intercept)  opinion_network
       -1.15290         -0.05105

Degrees of Freedom: 477 Total (i.e. Null);  476 Residual
Null Deviance:      405.1
Residual Deviance: 399.5        AIC: 403.5
```

Figure C14: Proportional odds

checking multicollinearity model 3

```
vif(ord_model_netop)
```

```
              age          educ_degree           cohabiting              density
         1.460447             1.187820             1.137842             1.088648
   unknown_opinion opinon_network_absval
         3.881089             3.674381
```

Figure C15: VIF

Influence of polarisation on certainty

Model 1

```
log_model1 <- glm(certain1~age +
             educ_degree +
             cohabiting +
             density+
             unknown_opinion+
             polarisation,
          family="binomial",
          data=data)
```

checking linear relationship model 1

```
plot(log_model1, 1)
```

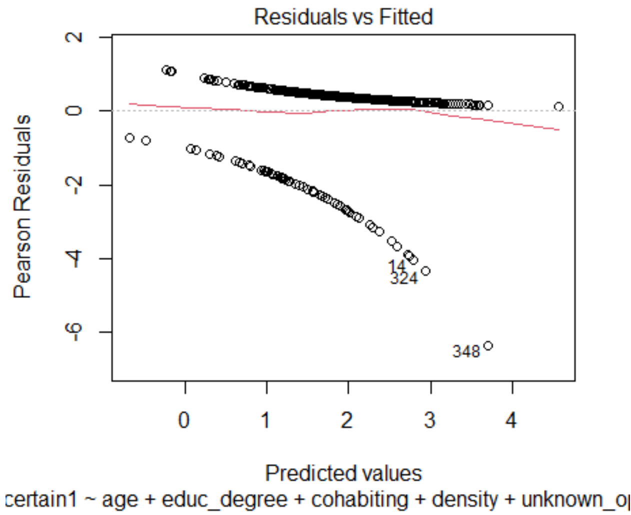


Figure C16: Linearity model 1

checking multicollinearity model 1

```
vif(log_model1)
```

```
           age      educ_degree       cohabiting          density unknown_opinion     polarisation
      1.536933         1.190243         1.089791         1.054566        1.306855         1.031393
```

Figure C17: VIF model 1

influential values model 1

```
plot(log_model1, 4)

plot(log_model1, 5)
```

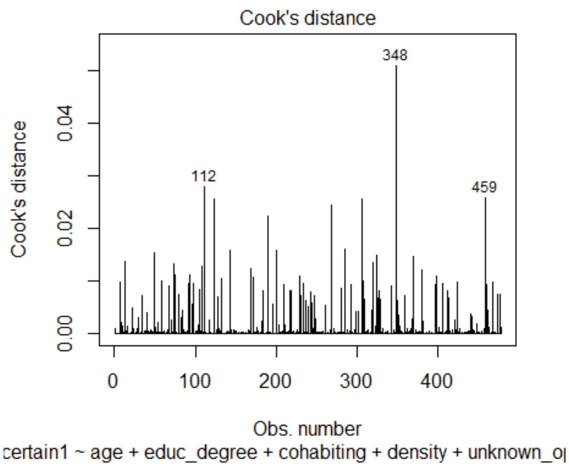


Figure C18: Cook's distance model 1

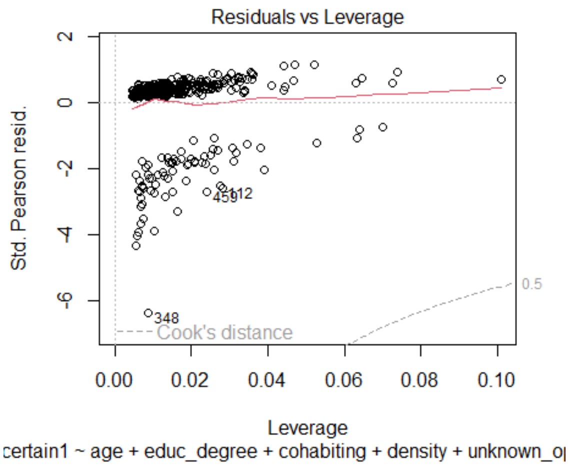


Figure C19: Leverage model 1

Model 2

```
log_model2 <- glm(certain2~age +
             educ_degree +
             cohabiting +
             density+
             unknown_opinion+
             polarisation,
          family="binomial",
          data=data)
```

checking linear relationship model 2

```
plot(log_model2, 1)
```

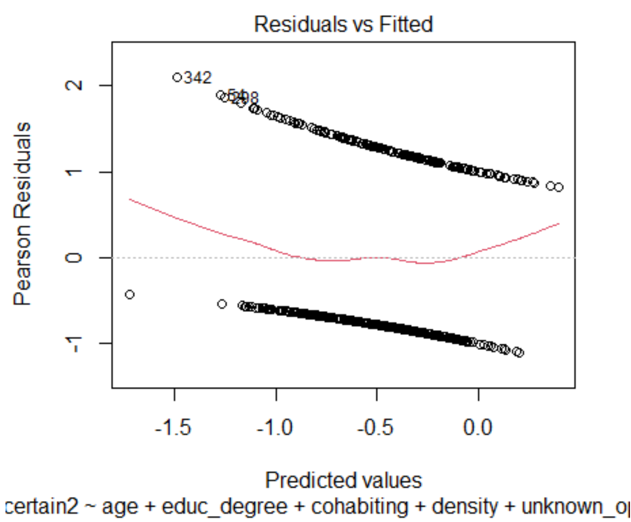


Figure C20: Linearity model 2

checking multicollinearity model 2

```
vif(log_model2)
```

```
     age      educ_degree      cohabiting      density unknown_opinion    polarisation
1.421499         1.152880        1.135999     1.059406        1.204358        1.009094
```

Figure C21: VIF model 2

influential values model 2

```
plot(log_model2, 4)
plot(log_model2, 5)
```

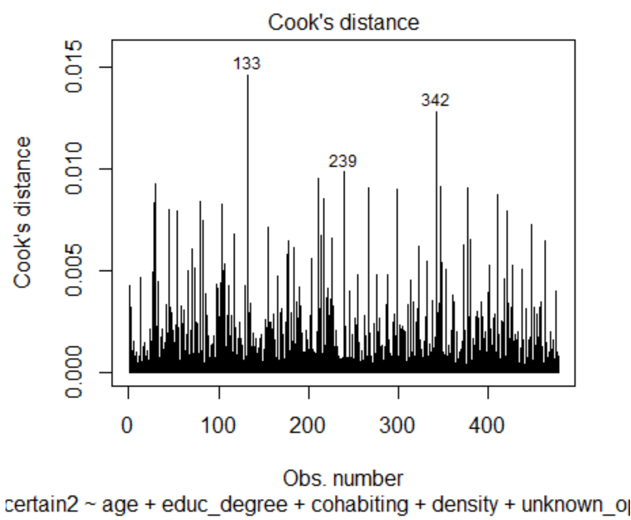


Figure C22: Cook's distance model 2

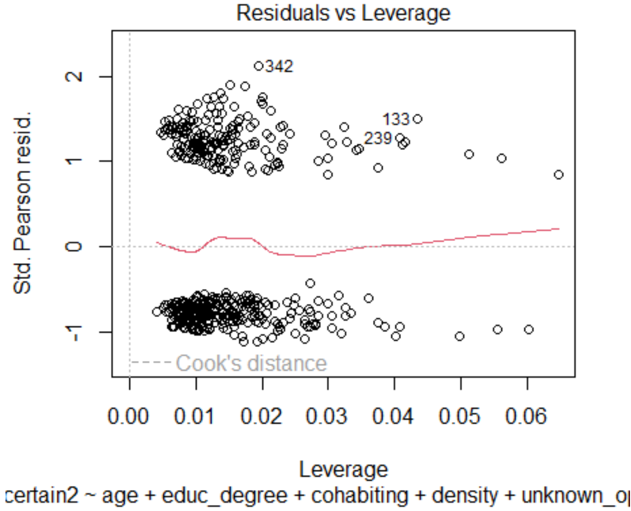


Figure C23: Leverage model 2

Model 3

```
ord_model <- polr(certainty_ord~
              age+
              educ_degree+
              cohabiting+
              density+
              unknown_opinion +
              polarisation,
           data = data,
           Hess=TRUE)
```

proportional odds

```
glm(I(as.numeric(certainty_ord) >= 2) ~ polarisation, family="binomial", data =
    data)
glm(I(as.numeric(certainty_ord) >= 3) ~ polarisation, family="binomial", data =
    data)
```

```
Call:  glm(formula = I(as.numeric(certainty_ord) >= 2) ~ polarisation,
    family = "binomial", data = data)

Coefficients:
 (Intercept)  polarisation
      0.6450       -0.5046

Degrees of Freedom: 477 Total (i.e. Null);  476 Residual
Null Deviance:      636.2
Residual Deviance: 635.1        AIC: 639.1
> glm(I(as.numeric(certainty_ord) >= 3) ~ polarisation, family="binomial", data = data)

Call:  glm(formula = I(as.numeric(certainty_ord) >= 3) ~ polarisation,
    family = "binomial", data = data)

Coefficients:
 (Intercept)  polarisation
     -1.4687       -0.8172

Degrees of Freedom: 477 Total (i.e. Null);  476 Residual
Null Deviance:      405.1
Residual Deviance: 403.7        AIC: 407.7
```

Figure C24: Proportional odds

checking multicollinearity model 3

```
vif(ord_model)
```

```
         age     educ_degree      cohabiting         density unknown_opinion    polarisation
    1.424131        1.156222        1.135691        1.061521        1.198484        1.012818
>
```

Figure C25: VIF model 3

Influence of polarisation in only 3 clusters on certainty

Model 1

```
log_model_polarisation3 <- glm(certain1~age +
                    educ_degree +
                    cohabiting +
                    density+
                    unknown_opinion+
                    polaris_comm3,
                family="binomial",
                data=data)
```

checking linear relationship model 1

```
plot(log_model_polarisation3, 1)
```

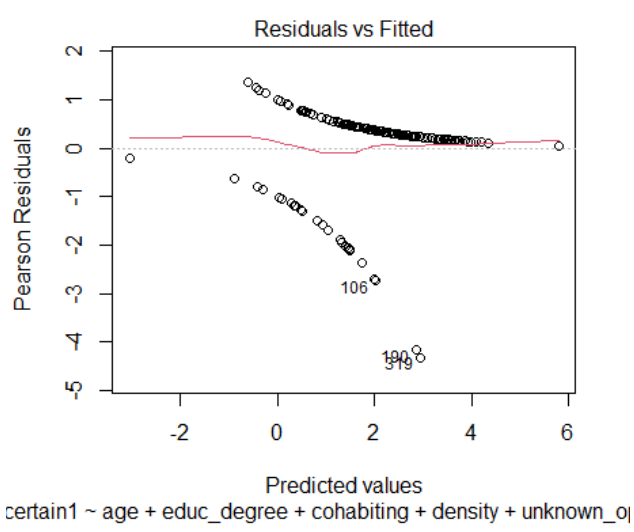


Figure C26: Linearity

checking multicollinearity model 1

```
vif(log_model_polarisation3)
```

```
          age   educ_degree    cohabiting       density unknown_opinion    polaris_comm3
     1.621555      1.337692      1.170494      1.029664        1.194841         1.098170
```

Figure C27: VIF model 1

checking influential values model 1

```
plot(log_model_polarisation3, 4)
plot(log_model_polarisation3, 5)
```

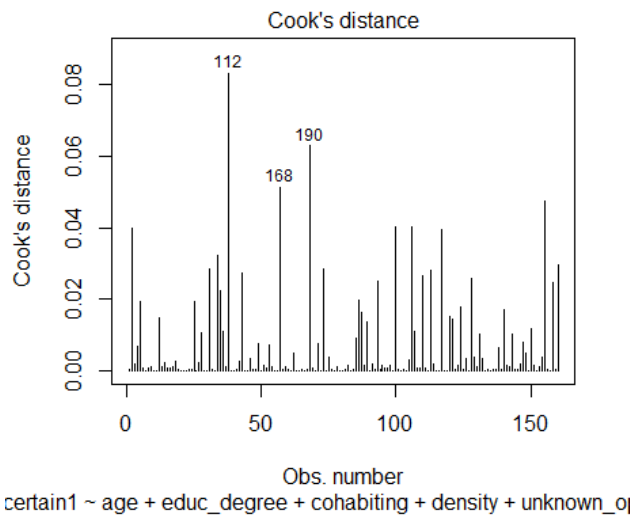


Figure C28: Cook's distance model 1

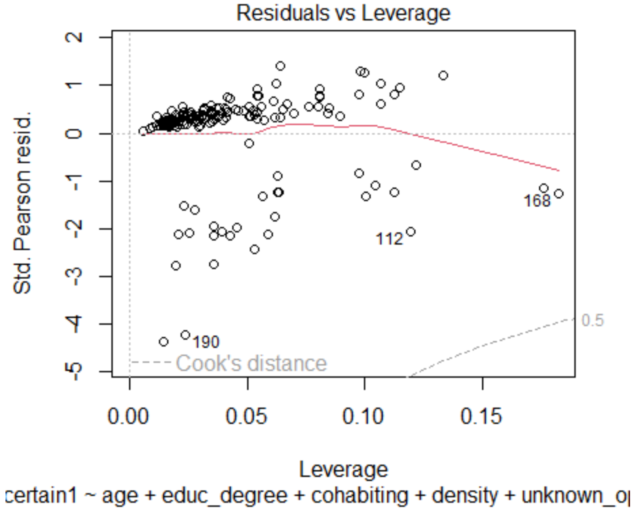


Figure C29: Leverage model 1

Model 2

```
log_model2_polarisation3 <- glm(certain2~age +
                                  educ_degree +
                                  cohabiting +
                                  density +
                                  unknown_opinion +
                                  polaris_comm3,
                               family="binomial",
                               data=data)
```

checking linear relationship model 2

```
plot(log_model2_polarisation3, 1)
```

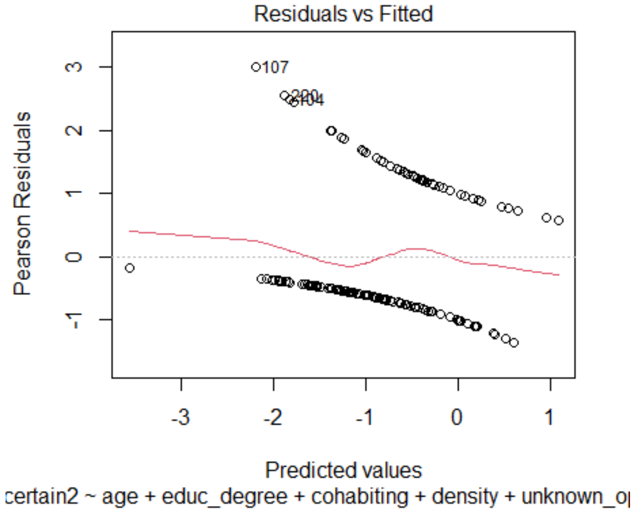


Figure C30: Linearity model 2

checking multicollinearity model 2

```
vif(log_model2_polarisation3)
```

age	educ_degree	cohabiting	density	unknown_opinion	polaris_comm3
1.491595	1.212190	1.206717	1.043094	1.133118	1.039400

Figure C31: VIF mdoel 2

checking influential values model 2

```
plot(log_model2_polarisation3, 4)
plot(log_model2_polarisation3, 5)
```

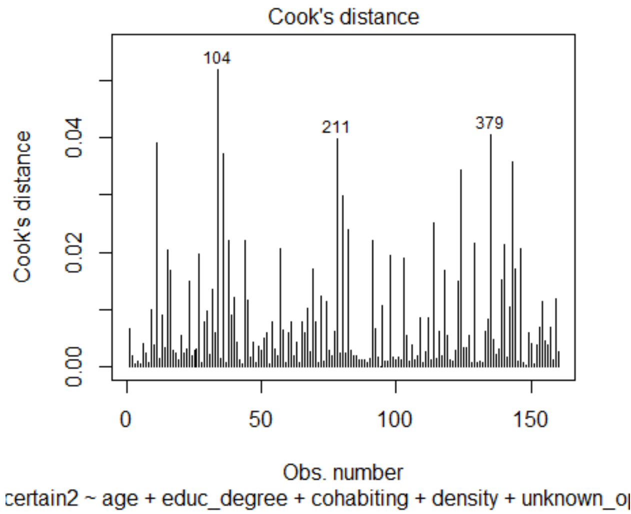


Figure C32: Cook's distance model 2

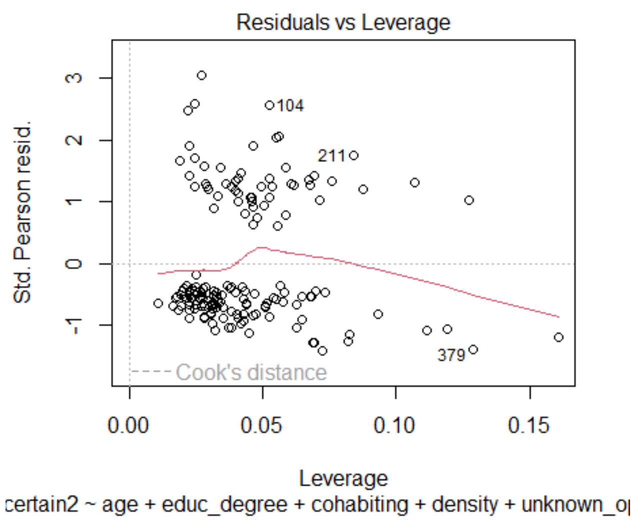


Figure C33: Leverage model 2

Model 3

```
ord_model_polarisation3 <- polr(certainty_ord~
                            age+
                            educ_degree+
                            cohabiting+
                            density+
                            unknown_opinion+
                            polaris_comm3,
                        data = data,
                        Hess=TRUE)
```

proportional odds

```
glm(I(as.numeric(certainty_ord) >= 2) ~ polaris_comm3, family="binomial", data =
    data)
glm(I(as.numeric(certainty_ord) >= 3) ~ polaris_comm3, family="binomial", data =
    data)
```

```
Call:  glm(formula = I(as.numeric(certainty_ord) >= 2) ~ polaris_comm3,
    family = "binomial", data = data)

Coefficients:
  (Intercept)  polaris_comm3
       0.9500        -0.7066

Degrees of Freedom: 160 Total (i.e. Null);  159 Residual
  (317 observations deleted due to missingness)
Null Deviance:      204
Residual Deviance: 203.3        AIC: 207.3
> glm(I(as.numeric(certainty_ord) >= 3) ~ polaris_comm3, family="binomial", data = data)

Call:  glm(formula = I(as.numeric(certainty_ord) >= 3) ~ polaris_comm3,
    family = "binomial", data = data)

Coefficients:
  (Intercept)  polaris_comm3
      -0.7852        -2.4095

Degrees of Freedom: 160 Total (i.e. Null);  159 Residual
  (317 observations deleted due to missingness)
Null Deviance:      151.8
Residual Deviance: 147.3        AIC: 151.3
```

Figure C34: Proportional odds

checking multicollinearity model 3

```
vif(ord_model_polarisation3)
```

```
         age     educ_degree      cohabiting         density unknown_opinion   polaris_comm3
    1.419822        1.196608        1.200007        1.035772        1.089071        1.053656
```

Figure C35: VIF model 3

Influence of polarisation in large clusters on certainty

Model 1

```
log_model_large <- glm(certain1~age +
                  educ_degree +
                  cohabiting +
                  density +
                  unknown_opinion+
                  polarisation_large,
                family="binomial",
                data=data)
```

checking linear relationship model 1

```
plot(log_model_large, 1)
```

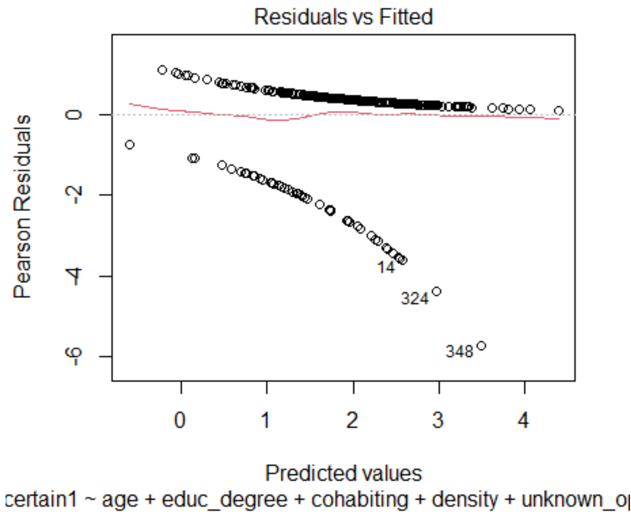


Figure C36: Linearity model 1

checking multicollinearity model 1

```
vif(log_model_large)
```

```
      age       educ_degree        cohabiting           density    unknown_opinion polarisation_large
 1.557448          1.164445          1.110078          1.087976          1.282777           1.046830
```

Figure C37: VIF model 1

checking influential values model 1

```
plot(log_model_large, 4)
  plot(log_model_large, 5)
```

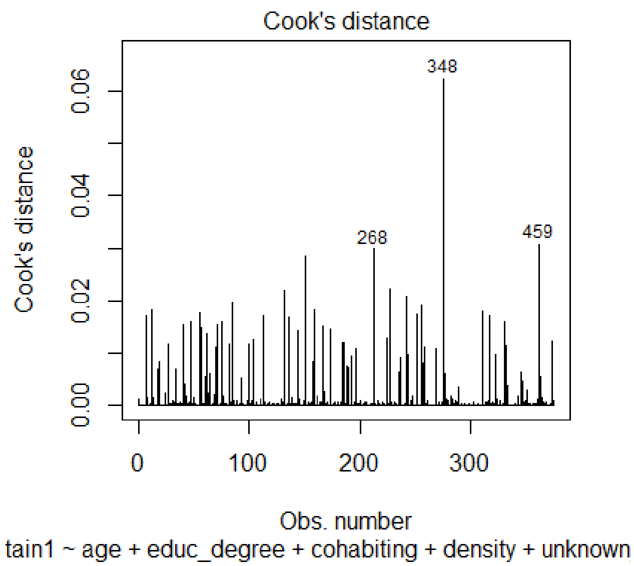


Figure C38: Cook's distance model 1

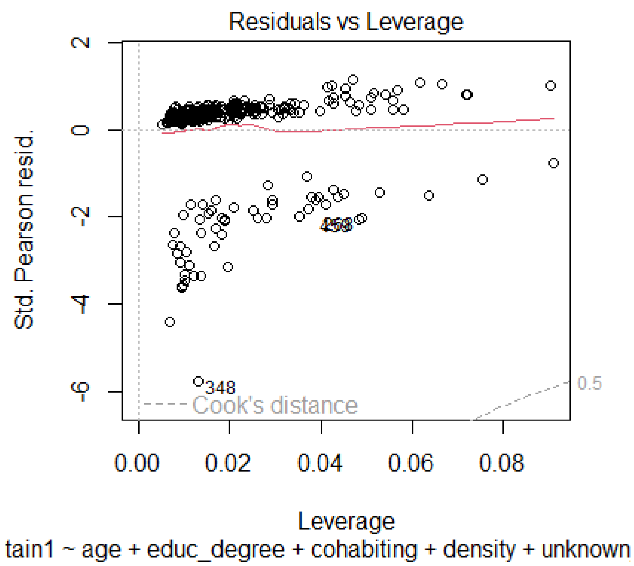


Figure C39: Leverage model 1

Model 2

```
log_model2_large <- glm(certain2~age +
                    educ_degree +
                    cohabiting +
                    density +
                    unknown_opinion+
                    polarisation_large,
                  family="binomial",
                  data=data)
```

checking linear relationship model 2

```
plot(log_model2_large, 1)
```

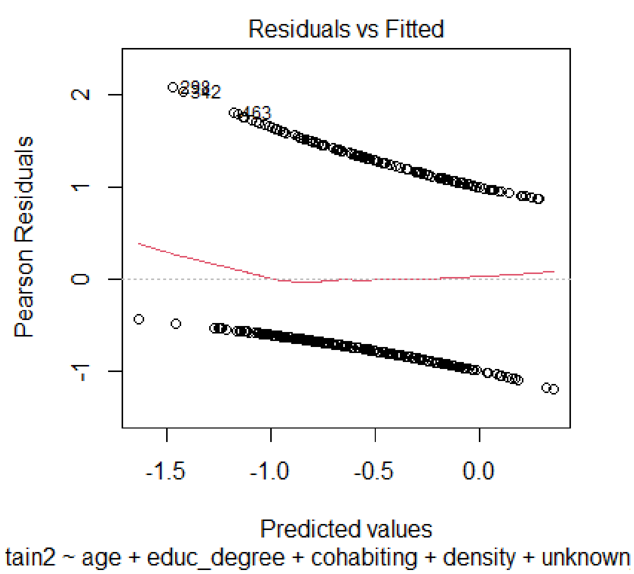


Figure C40: Linearity model 2

checking multicollinearity model 2

```
vif(log_model2_large)
```

checking influential values model 2

```
           age         educ_degree          cohabiting             density    unknown_opinion polarisation_large
      1.417244            1.133900            1.147324            1.077470            1.187908           1.034803
```

Figure C41: VIF model 2

```
plot(log_model2_large, 4)
plot(log_model2_large, 5)
```

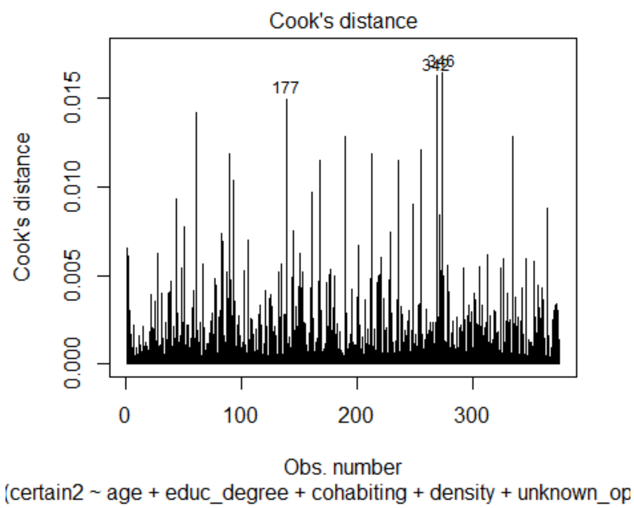


Figure C42: Cook's distance model 2

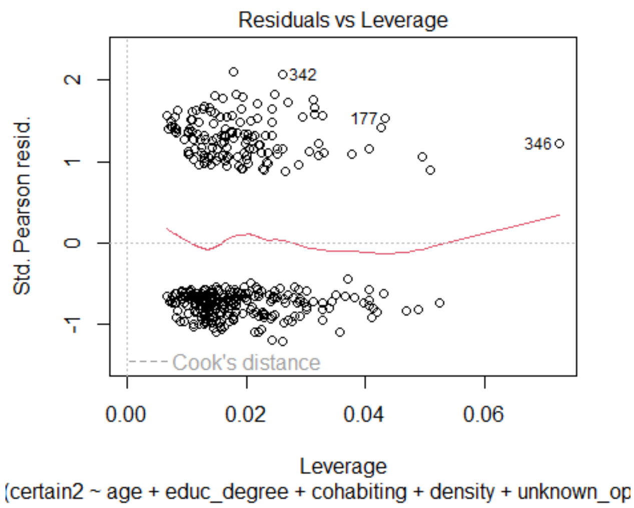


Figure C43: Leverage model 2

Model 3

```
ord_model_large <- polr(certainty_ord~
                          age+
                          educ_degree+
                          cohabiting+
                          density +
                          unknown_opinion+
                          polarisation_large,
                        data = data,
                        Hess=TRUE)
```

proportional odds

```
glm(I(as.numeric(certainty_ord) >= 2) ~ polarisation_large, family="binomial",
    data = data)
glm(I(as.numeric(certainty_ord) >= 3) ~ polarisation_large, family="binomial",
    data = data)
```

```
Call:  glm(formula = I(as.numeric(certainty_ord) >= 2) ~ polarisation_large,
    family = "binomial", data = data)

Coefficients:
       (Intercept)  polarisation_large
             0.427               0.486

Degrees of Freedom: 375 Total (i.e. Null);  374 Residual
  (102 observations deleted due to missingness)
Null Deviance:      494.3
Residual Deviance: 493.8        AIC: 497.8
> glm(I(as.numeric(certainty_ord) >= 3) ~ polarisation_large, family="binomial", data = data)

Call:  glm(formula = I(as.numeric(certainty_ord) >= 3) ~ polarisation_large,
    family = "binomial", data = data)

Coefficients:
       (Intercept)  polarisation_large
           -1.8979              0.4514

Degrees of Freedom: 375 Total (i.e. Null);  374 Residual
  (102 observations deleted due to missingness)
Null Deviance:      309.4
Residual Deviance: 309.2        AIC: 313.2
```

Figure C44: Proportional odds

checking multicollinearity model 3

```
vif(ord_model_large)
```

```
            age     educ_degree       cohabiting          density    unknown_opinion polarisation_large
       1.435847        1.143634         1.152089         1.076882           1.165643           1.036213
```

Figure C45: VIF model 3

Influence of polarisation on certainty when all intentions are known

Model 1

```
log_model_allknown <- glm(certain1~age +
                        educ_degree +
                        cohabiting +
                        density+
                        unknown_opinion+
                        polarisation_allknown,
                        family="binomial",
                        data=data)
```

checking linear relationship model 1

```
plot(log_model_allknown, 1)
```

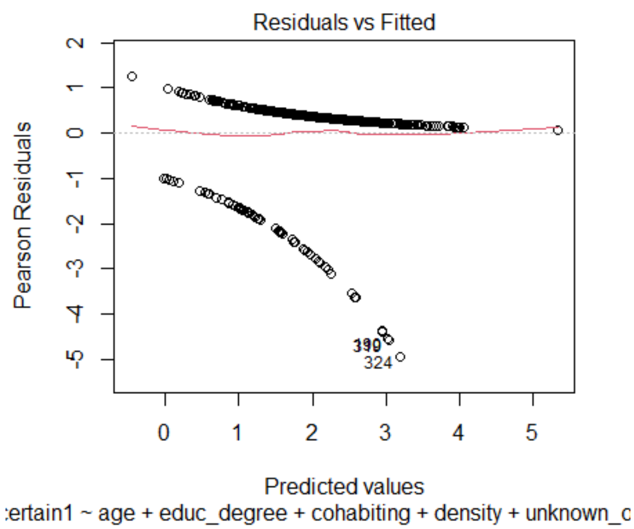


Figure C46: Linearity model 1

checking multicollinearity model 1

```
vif(log_model_allknown)
```

```
           age          educ_degree           cohabiting              density      unknown_opinion polarisation_allknown
      1.654905             1.238199             1.083965             1.059962             1.394932             1.075774
```

Figure C47: VIF model 1

checking influential values model 1

```
plot(log_model_allknown, 4)
plot(log_model_allknown, 5)
```

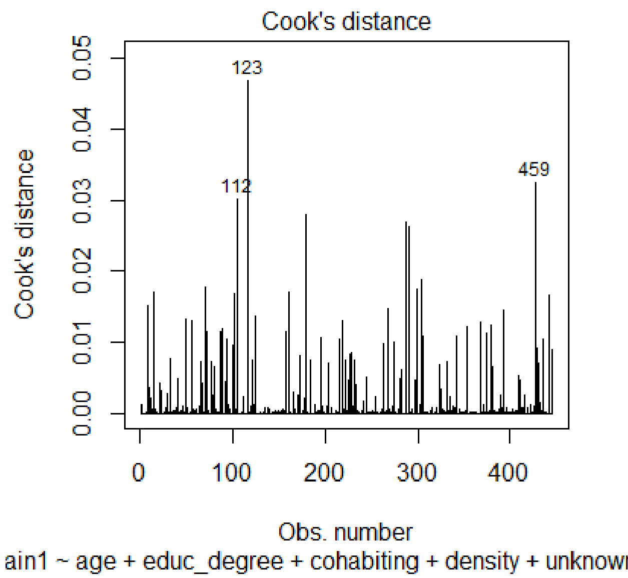


Figure C48: Cook's distance model 1

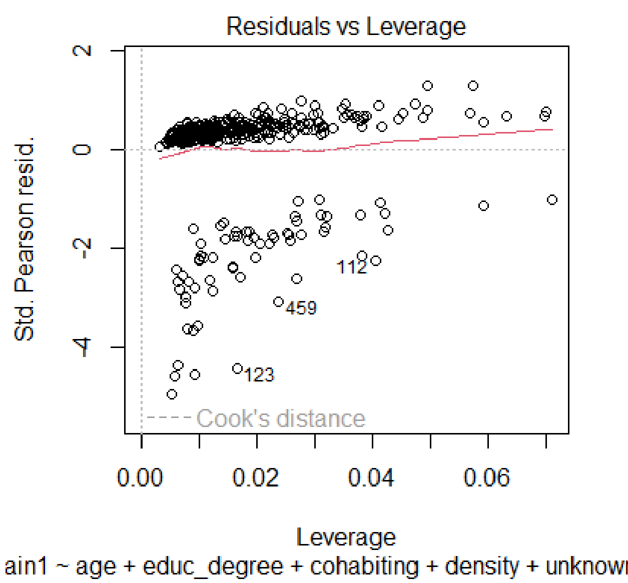


Figure C49: Leverage model 1

Model 2

```
log_model2_allknown <- glm(certain2~age +
                                 educ_degree +
                                 cohabiting +
                                 density +
                                 unknown_opinion +
                                 polarisation_allknown,
                              family="binomial",
                              data=data)
```

checking linear relationship model 2

```
plot(log_model2_allknown, 1)
```

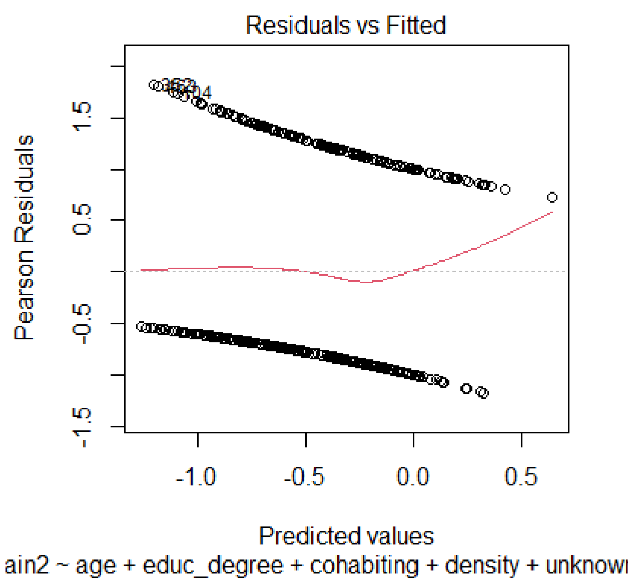


Figure C50: Linearity model 2

checking multicollinearity model 2

```
vif(log_model2_allknown)
```

```
age        educ_degree         cohabiting            density   unknown_opinion polarisation_allknown
1.451729      1.164913           1.140662           1.066152          1.266977              1.076105
```

Figure C51: VIF model 2

checking influential values model 2

```r
plot(log_model2_allknown, 4)

plot(log_model2_allknown, 5)
```

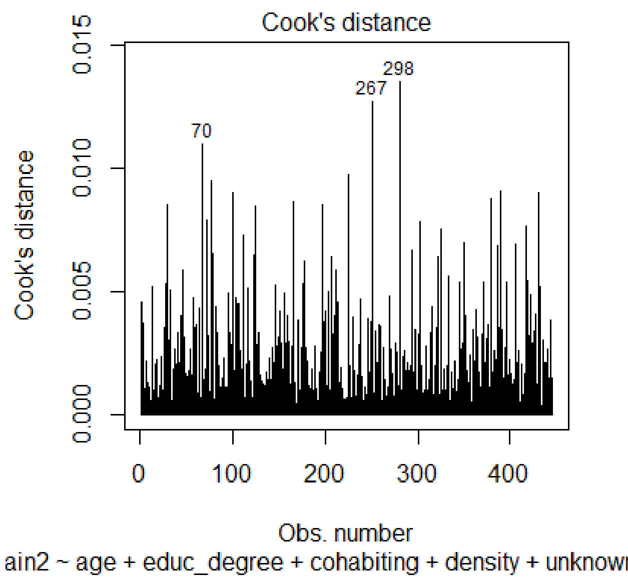


Figure C52: Cook's distance model 2

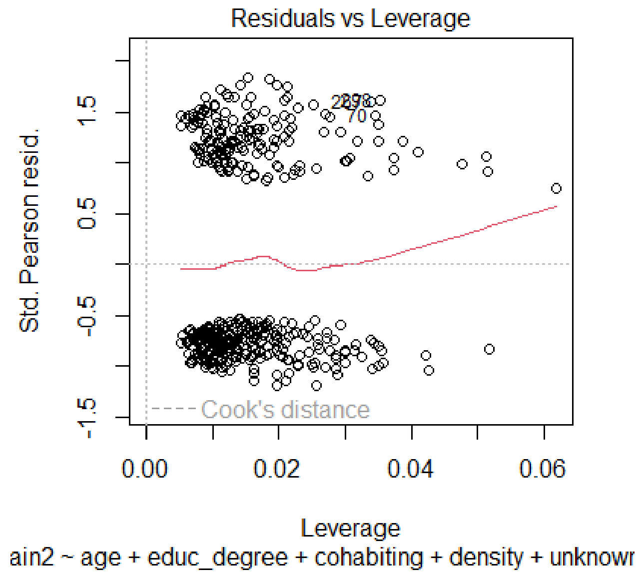


Figure C53: Leverage model 2

Model 3

```
ord_model_allknown <- polr(certainty_ord~
                          age+
                          educ_degree+
                          cohabiting+
                          density+
                          unknown_opinion+
                          polarisation_allknown,
                        data = data,
                        Hess=TRUE)
```

proportional odds

```
glm(I(as.numeric(certainty_ord) >= 2) ~ polarisation_allknown,
    family="binomial", data = data)
glm(I(as.numeric(certainty_ord) >= 3) ~ polarisation_allknown,
    family="binomial", data = data)
```

```
Call:  glm(formula = I(as.numeric(certainty_ord) >= 2) ~ polarisation_allknown,
    family = "binomial", data = data)

Coefficients:
          (Intercept)  polarisation_allknown
              0.46504               -0.04745

Degrees of Freedom: 446 Total (i.e. Null);  445 Residual
  (31 observations deleted due to missingness)
Null Deviance:      597.6
Residual Deviance: 597.5        AIC: 601.5
> glm(I(as.numeric(certainty_ord) >= 3) ~ polarisation_allknown, family="binomial", data = data)

Call:  glm(formula = I(as.numeric(certainty_ord) >= 3) ~ polarisation_allknown,
    family = "binomial", data = data)

Coefficients:
          (Intercept)  polarisation_allknown
             -1.75969               -0.03712

Degrees of Freedom: 446 Total (i.e. Null);  445 Residual
  (31 observations deleted due to missingness)
Null Deviance:      370.7
Residual Deviance: 370.7        AIC: 374.7
```

Figure C54: Proportional odds

checking multicollinearity model 3

```
vif(ord_model_allknown)
```

```
              age           educ_degree            cohabiting               density        unknown_opinion polarisation_allknown
         1.448672              1.167252              1.133433              1.066491              1.267408              1.077036
```

Figure C55: VIF model 3

Linear model to test the effects of the opinion of the network on fertility intentions; Assumptions do not hold

linear regression model

```
lin_model <- lm(childwish_numerical~
           age+
           educ_degree +
           cohabiting +
           opinion_network,
           data= data)
```

checking for linearity

```
plot(lin_model, 1)
```

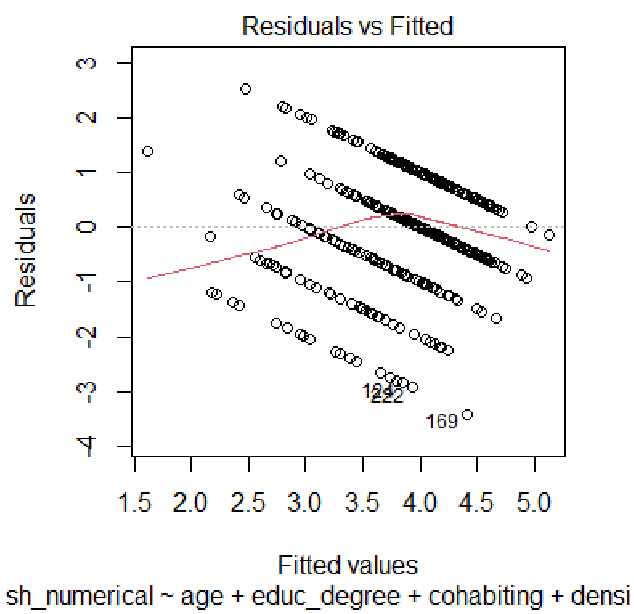


Figure C56: Linearity linear model

checking for normality

```
plot(lin_model, 2)
```

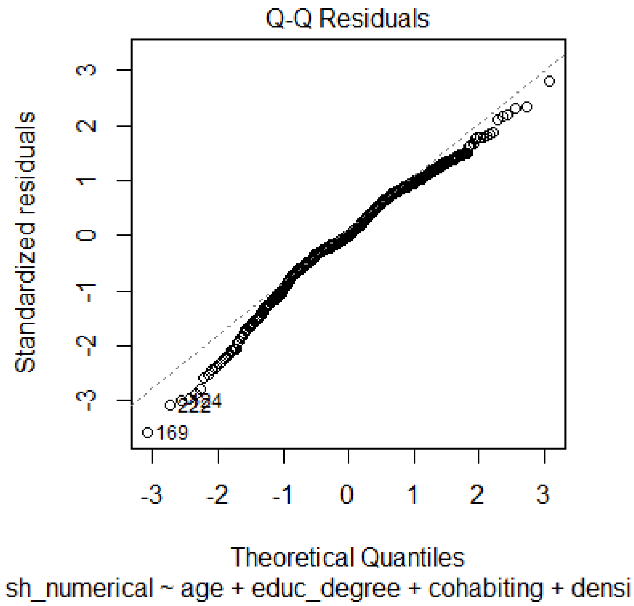


Figure C57: Normality linear model

checking for homoscedasticity

```
plot(lin_model, 3)
```

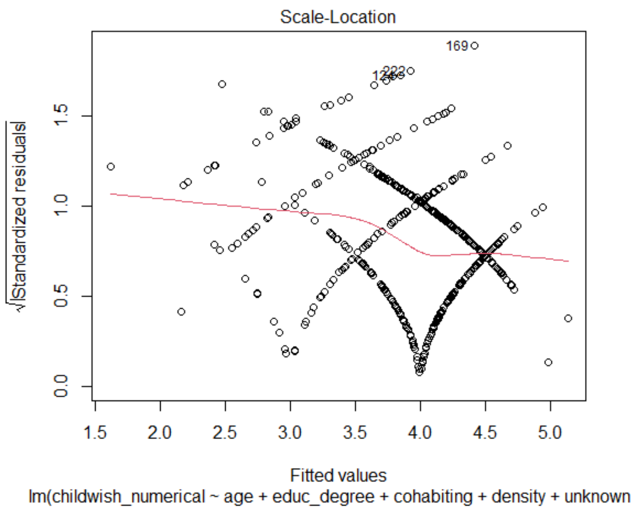


Figure C58: homoscedasticity linear model

checking for influential points

```
plot(lin_model, 4)
plot(lin_model, 5)
```

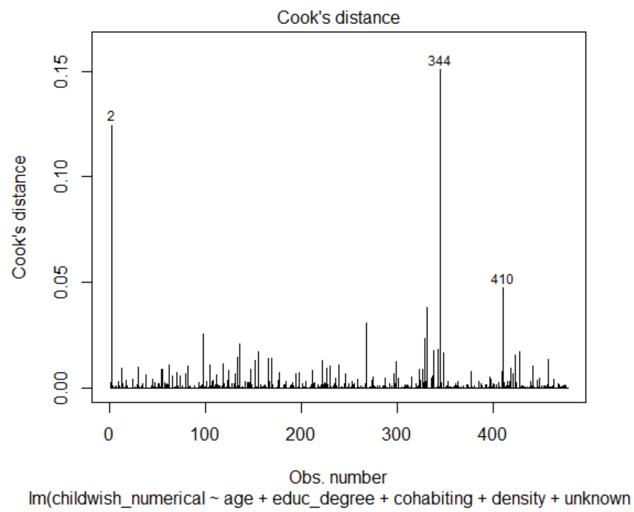


Figure C59: Cook's distance linear model

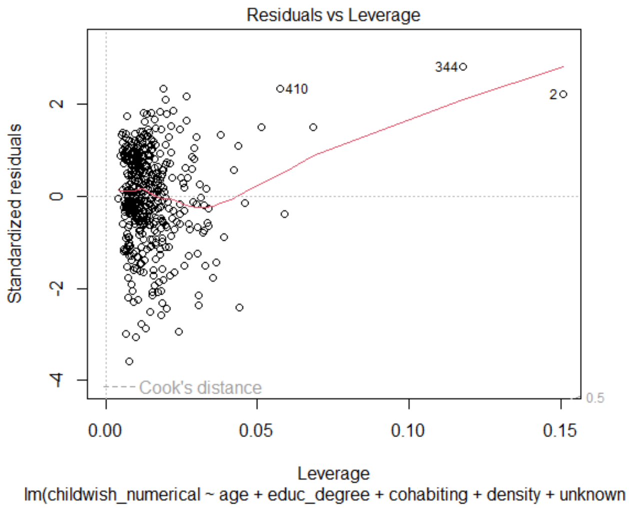


Figure C60: Leverage linear model

checking for multicollinearity

```
vif(lin_model)
```

```
     age    educ_degree     cohabiting        density unknown_opinion opinion_network
1.428662       1.179984       1.121632       1.097202        3.140620        2.957948
```

Figure C61: VIF linear model

11 Appendix D

Regression analyses

Influence network on fertility intentions

Model

```
general_ordmodel <- polr(childwish~
                           age+
                           educ_degree+
                           cohabiting+
                           density+
                           unknown_opinion+
                           opinion_network,
                           data = data,
                           Hess=TRUE)

(ctable1 <- coef(summary(general_ordmodel)))


p1 <- pnorm(abs(ctable1[, "t value"]), lower.tail = FALSE) * 2
```

combined table

```
(ctable1 <- cbind(ctable1, "p value" = p1))
```

```
                               Value Std. Error    t value      p value
age                      -0.15044202 0.01894475 -7.9410912 2.004103e-15
educ_degree               0.09771333 0.06396878  1.5275160 1.266328e-01
cohabiting                0.83871409 0.18947933  4.4264146 9.581231e-06
density                   0.56328477 0.86083674  0.6543456 5.128891e-01
unknown_opinion           0.04992899 0.02914538  1.7131012 8.669393e-02
opinion_network           0.10807410 0.02872011  3.7630116 1.678794e-04
Absolutely not|Probably not -3.97778772 0.87600810 -4.5408116 5.603810e-06
Probably not|I don't know -2.73668543 0.85669354 -3.1944742 1.400858e-03
I don't know|Probably so  -1.49294941 0.84693945 -1.7627581 7.794129e-02
Probably so|Absolutely so  0.42801827 0.84348190  0.5074422 6.118446e-01
```

Figure D1: output ordinal regression analysis

model fit

```
lrm(formula = childwish~
  age+
  educ_degree+
  cohabiting+
  density+
  unknown_opinion+
  opinion_network,
  data = data)
```

```
Frequencies of Responses

Absolutely not   Probably not   I don't know    Probably so  Absolutely so
            21             36             72            186            162


                       Model Likelihood      Discrimination    Rank Discrim.
                             Ratio Test             Indexes          Indexes
Obs             477    LR chi2     103.17      R2       0.208    C       0.680
max |deriv| 2e-06    d.f.            6      R2(6,477)0.184    Dxy     0.361
                     Pr(> chi2) <0.0001    R2(6,428.1)0.203    gamma   0.361
                                             Brier    0.085    tau-a   0.254

                  Coef    S.E.   Wald Z Pr(>|Z|)
y>=Probably not    3.9774 0.8760  4.54  <0.0001
y>=I don't know    2.7366 0.8567  3.19  0.0014
y>=Probably so     1.4929 0.8469  1.76  0.0779
y>=Absolutely so -0.4281 0.8435 -0.51  0.6118
age              -0.1504 0.0189 -7.94  <0.0001
educ_degree       0.0977 0.0640  1.53  0.1266
cohabiting        0.8387 0.1895  4.43  <0.0001
density           0.5632 0.8608  0.65  0.5130
unknown_opinion   0.0499 0.0291  1.71  0.0867
opinion_network   0.1081 0.0287  3.76  0.0002
```

Figure D2: model fit

OR

```
exp(cbind(OR = coef(general_ordmodel)))
```

```
                       OR
age             0.8603276
educ_degree     1.1026466
cohabiting      2.3133902
density         1.7564325
unknown_opinion 1.0511964
opinion_network 1.1141303
```

Figure D3: Odds ratio

visualisation

```
plot(Effect(focal.predictors = "opinion_network",general_ordmodel))
```

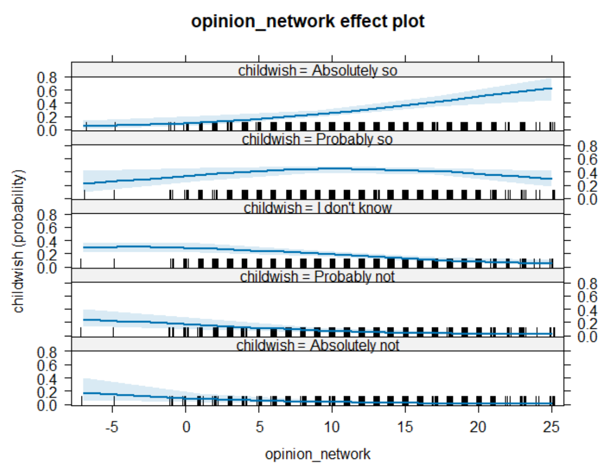


Figure D4: visualisation effect opinion network

Effect overall network opinion on certainty

Model 1

```
log_model_netop1 <- glm(certain1~age +
                          educ_degree +
                          cohabiting +
                          density+
                          unknown_opinion+
                          opinon_network_absval,
                        family="binomial",
                        data=data)

summary(log_model_netop1)
```

```
Call:
glm(formula = certain1 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + opinon_network_absval, family = "binomial",
    data = data)

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)            1.99475    1.21653   1.640   0.1011
age                   -0.10429    0.02620  -3.980 6.89e-05 ***
educ_degree            0.29224    0.10481   2.788   0.0053 **
cohabiting             0.54476    0.29987   1.817   0.0693 .
density                1.99669    1.37433   1.453   0.1463
unknown_opinion       -0.02909    0.03979  -0.731   0.4646
opinon_network_absval  0.03769    0.03973   0.949   0.3428
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 404.82  on 476  degrees of freedom
Residual deviance: 371.53  on 470  degrees of freedom
  (1 observation deleted due to missingness)
AIC: 385.53

Number of Fisher Scoring iterations: 5
```

Figure D5: estimates logistic regression analysis model 1

model fit

```
lrm(formula = certain1 ~ age +
  educ_degree +
  cohabiting +
  density+
  unknown_opinion+
  opinon_network_absval,
  data=data)
```

```
Frequencies of Missing Values Due to Each Variable
             certain1                   age           educ_degree             cohabiting
                    0                     1                     1                      0
              density        unknown_opinion opinon_network_absval
                    0                     0                     0

Logistic Regression Model

lrm(formula = certain1 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + opinon_network_absval, data = data)


                       Model Likelihood      Discrimination     Rank Discrim.
                             Ratio Test             Indexes           Indexes
Obs            477     LR chi2      33.29      R2        0.118     C       0.715
 0              72     d.f.             6      R2(6,477)0.056     Dxy     0.430
 1             405     Pr(> chi2) <0.0001    R2(6,183.4)0.138     gamma   0.431
max |deriv| 7e-06                              Brier     0.119     tau-a   0.111

                      Coef    S.E.   Wald Z Pr(>|Z|)
Intercept              1.9948 1.2165  1.64  0.1011
age                   -0.1043 0.0262 -3.98  <0.0001
educ_degree            0.2922 0.1048  2.79  0.0053
cohabiting             0.5448 0.2999  1.82  0.0693
density                1.9967 1.3743  1.45  0.1463
unknown_opinion       -0.0291 0.0398 -0.73  0.4646
opinon_network_absval  0.0377 0.0397  0.95  0.3428
```

Figure D6: Model fit model 1

calculate odds ratio

```
exp(cbind(OR = coef(log_model_netop1)))
```

```
                             OR
(Intercept)           7.3503906
age                   0.9009626
educ_degree           1.3394206
cohabiting            1.7241865
density               7.3646529
unknown_opinion       0.9713252
opinon_network_absval 1.0384063
```

Figure D7: Odds ratio model 1

Model 2

```
log_model_netop2 <- glm(certain2~age +
                          educ_degree +
                          cohabiting +
                          density+
                          unknown_opinion+
                          opinon_network_absval,
                        family="binomial",
                        data=data)


summary(log_model_netop2)
```

```
Call:
glm(formula = certain2 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + opinon_network_absval, family = "binomial",
    data = data)

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)           -0.09917    0.93285  -0.106   0.9153
age                   -0.02901    0.02050  -1.415   0.1571
educ_degree            0.01130    0.07128   0.159   0.8740
cohabiting             0.47557    0.20669   2.301   0.0214 *
density                1.22699    0.92134   1.332   0.1829
unknown_opinion       -0.04383    0.03129  -1.401   0.1613
opinon_network_absval -0.01520    0.03018  -0.504   0.6145
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 635.19  on 476  degrees of freedom
Residual deviance: 623.44  on 470  degrees of freedom
  (1 observation deleted due to missingness)
AIC: 637.44

Number of Fisher Scoring iterations: 4
```

Figure D8: estimates logistic regression analysis model 2

Model fit

```
lrm(formula = certain2 ~ age +
          educ_degree +
          cohabiting +
          density+
          unknown_opinion+
          opinon_network_absval,
        data=data)
```

```
                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs            477    LR chi2      11.75    R2         0.033    C       0.588
 0             294    d.f.             6    R2(6,477)0.012    Dxy     0.176
 1             183    Pr(> chi2) 0.0678   R2(6,338.4)0.017    gamma   0.176
max |deriv| 2e-06                           Brier      0.230    tau-a   0.084

                      Coef    S.E.   Wald Z Pr(>|Z|)
Intercept             -0.0992 0.9329 -0.11  0.9153
age                   -0.0290 0.0205 -1.42  0.1571
educ_degree            0.0113 0.0713  0.16  0.8740
cohabiting             0.4756 0.2067  2.30  0.0214
density                1.2270 0.9213  1.33  0.1829
unknown_opinion       -0.0438 0.0313 -1.40  0.1613
opinon_network_absval -0.0152 0.0302 -0.50  0.6145
```

Figure D9: Model fit model 2

odds ratio

```
exp(cbind(OR = coef(log_model_netop2)))
```

```
                             OR
(Intercept)           0.9055887
age                   0.9714106
educ_degree           1.0113675
cohabiting            1.6089324
density               3.4109435
unknown_opinion       0.9571178
opinon_network_absval 0.9849173
```

Figure D10: Odds ratio model 2

Model 3

```r
ord_model_netop <- polr(certainty_ord~
                            age+
                            educ_degree+
                            cohabiting+
                            density+
                            unknown_opinion+
                            opinon_network_absval,
                        data = data,
                        Hess=TRUE)
```

adding p-values to the table

```r
# Store table
table_ordreg_netop <- coef(summary(ord_model_netop))

# Calculate and store p values
p_netop <- pnorm(abs(table_ordreg_netop[, "t value"]), lower.tail = FALSE) * 2

# Combined table
(table_ordreg_netop <- cbind(table_ordreg_netop, "p value" = p_netop))
```

	Value	Std. Error	t value	p value
age	0.0541680651	0.01887120	2.87040960	0.004099404
educ_degree	-0.0889128276	0.06523842	-1.36289057	0.172917001
cohabiting	-0.5314477895	0.19239580	-2.76226290	0.005740224
density	-1.4920114563	0.85719590	-1.74057231	0.081758574
unknown_opinion	0.0367432172	0.02991504	1.22825251	0.219352192
opinon_network_absval	-0.0003306219	0.02916785	-0.01133515	0.990956053
certain|preference	-0.2018077667	0.88174247	-0.22887382	0.818966989
preference|uncertain	2.0828964862	0.88866851	2.34383966	0.019086375

Figure D11: Estimates ordinal regression model

```
lrm(formula = certainty_ord ~ age +
          educ_degree +
          cohabiting +
          density+
          unknown_opinion+
          opinon_network_absval,
        data=data)
```

```
                      Model Likelihood      Discrimination     Rank Discrim.
                            Ratio Test             Indexes           Indexes
Obs           477    LR chi2      23.39      R2         0.055    C         0.606
 certain      183    d.f.             6      R2(6,477)0.036      Dxy       0.212
 preference   222    Pr(> chi2) 0.0007     R2(6,400.3)0.043      gamma     0.212
 uncertain     72                            Brier      0.232    tau-a     0.130
max |deriv| 7e-12

                      Coef    S.E.   Wald Z Pr(>|Z|)
y>=preference          0.2018 0.8817  0.23  0.8189
y>=uncertain          -2.0829 0.8887 -2.34  0.0191
age                    0.0542 0.0189  2.87  0.0041
educ_degree           -0.0889 0.0652 -1.36  0.1729
cohabiting            -0.5315 0.1924 -2.76  0.0057
density               -1.4920 0.8572 -1.74  0.0818
unknown_opinion        0.0367 0.0299  1.23  0.2194
opinon_network_absval -0.0003 0.0292 -0.01  0.9909
```

Figure D12: Model fit model 3

odds ratio

```
exp(cbind(OR = coef(ord_model_netop)))
```

```
                             OR
age                   1.0556620
educ_degree           0.9149253
cohabiting            0.5877534
density               0.2249198
unknown_opinion       1.0374266
opinon_network_absval 0.9996694
```

Figure D13: Odds ratio model 3

Effects of polarisation of certainty

Logistic regression model 1

```
log_model1 <- glm(certain1~age +
                    educ_degree +
                    cohabiting +
                    density+
                    unknown_opinion+
                    polarisation,
                  family="binomial",
                  data=data)


summary(log_model1)
```

```
Call:
glm(formula = certain1 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation, family = "binomial", data = data)

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      2.45937    1.02284   2.404  0.01620 *
age             -0.10597    0.02614  -4.054 5.03e-05 ***
educ_degree      0.30646    0.10402   2.946  0.00322 **
cohabiting       0.54697    0.29993   1.824  0.06821 .
density          2.25672    1.36414   1.654  0.09806 .
unknown_opinion -0.05582    0.02494  -2.238  0.02519 *
polarisation     0.51833    0.71039   0.730  0.46561
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 404.82  on 476  degrees of freedom
Residual deviance: 371.87  on 470  degrees of freedom
  (1 observation deleted due to missingness)
AIC: 385.87

Number of Fisher Scoring iterations: 5
```

Figure D14: Estimates logistic regression model 1

Model fit

```
lrm(formula = certain1~age +
  educ_degree +
  cohabiting +
  density+
  unknown_opinion+
  polarisation,
data=data)
```

```
lrm(formula = certain1 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation, data = data)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs           477    LR chi2      32.95    R2        0.117    C       0.710
 0             72    d.f.             6    R2(6,477)0.055     Dxy     0.419
 1            405    Pr(> chi2) <0.0001  R2(6,183.4)0.137     gamma   0.419
max |deriv| 6e-06                          Brier     0.119    tau-a   0.108

                Coef    S.E.   Wald Z Pr(>|Z|)
Intercept        2.4594 1.0228  2.40  0.0162
age             -0.1060 0.0261 -4.05  <0.0001
educ_degree      0.3065 0.1040  2.95  0.0032
cohabiting       0.5470 0.2999  1.82  0.0682
density          2.2567 1.3641  1.65  0.0981
unknown_opinion -0.0558 0.0249 -2.24  0.0252
polarisation     0.5183 0.7104  0.73  0.4656
```

Figure D15: Model fit model 1

calculate odds ratio

```
exp(cbind(OR = coef(log_model1)))
```

```
                          OR
(Intercept)       11.6974119
age                0.8994488
educ_degree        1.3586133
cohabiting         1.7280153
density            9.5517264
unknown_opinion    0.9457121
polarisation       1.6792221
```

Figure D16: Odds ratio model 1

visualisation

```
avPlots(log_model1)
```

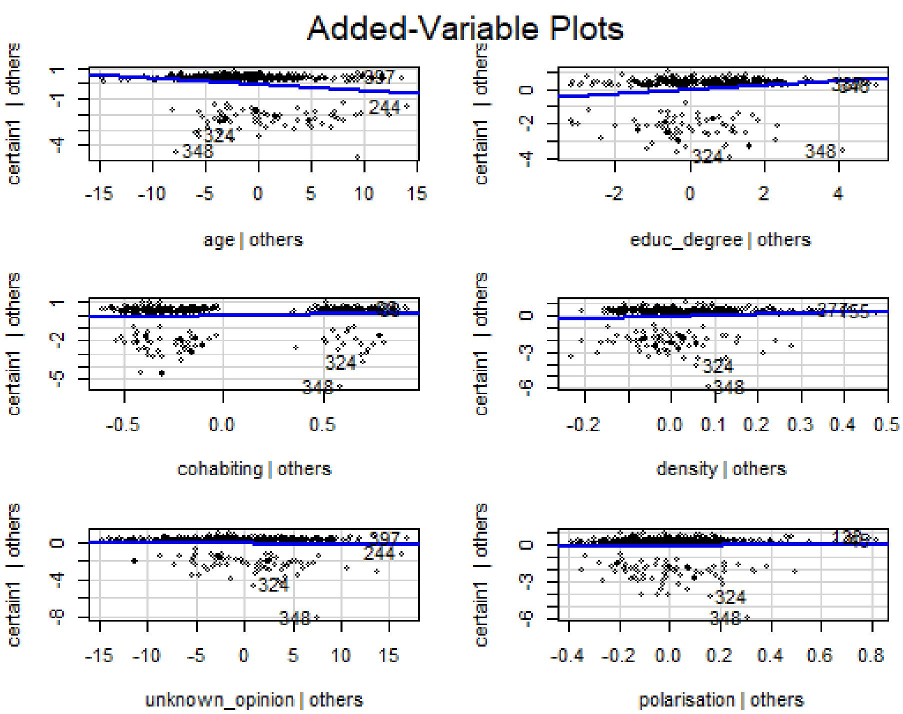


Figure D17: Added variable plot model 1

logistic regression model 2

```
log_model2 <- glm(certain2~age +
                educ_degree +
                cohabiting +
                density+
                unknown_opinion+
                polarisation,
             family="binomial",
             data=data)

summary(log_model2)
```

```
Call:
glm(formula = certain2 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation, family = "binomial", data = data)

Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)     -0.582844   0.737879  -0.790   0.4296
age             -0.026522   0.020279  -1.308   0.1909
educ_degree      0.006571   0.070459   0.093   0.9257
cohabiting       0.470725   0.206847   2.276   0.0229 *
density          1.168640   0.907059   1.288   0.1976
unknown_opinion -0.029657   0.018220  -1.628   0.1036
polarisation     0.404577   0.491925   0.822   0.4108
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 635.19  on 476  degrees of freedom
Residual deviance: 623.02  on 470  degrees of freedom
  (1 observation deleted due to missingness)
AIC: 637.02

Number of Fisher Scoring iterations: 4
```

Figure D18: Estimates of logistic regression model 2

Model fit

```
lrm(formula = certain2~age +
  educ_degree +
  cohabiting +
  density+
  unknown_opinion+
  polarisation,
data=data)
```

```
lrm(formula = certain2 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation, data = data)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs           477    LR chi2     12.17     R2       0.034    C       0.592
 0            294    d.f.            6     R2(6,477)0.013    Dxy     0.185
 1            183    Pr(> chi2) 0.0582   R2(6,338.4)0.018    gamma   0.185
max |deriv| 3e-13                          Brier    0.230    tau-a   0.087

                 Coef    S.E.   Wald Z Pr(>|Z|)
Intercept       -0.5828 0.7379 -0.79  0.4296
age             -0.0265 0.0203 -1.31  0.1909
educ_degree      0.0066 0.0705  0.09  0.9257
cohabiting       0.4707 0.2068  2.28  0.0229
density          1.1686 0.9071  1.29  0.1976
unknown_opinion -0.0297 0.0182 -1.63  0.1036
polarisation     0.4046 0.4919  0.82  0.4108
```

Figure D19: Model fit model 2

calculate odds ratio

```
exp(cbind(OR = coef(log_model2)))
```

```
                       OR
(Intercept)     0.5583082
age             0.9738266
educ_degree     1.0065930
cohabiting      1.6011549
density         3.2176133
unknown_opinion 0.9707789
polarisation    1.4986679
```

Figure D20: Odds ratio model 2

visualisation

```
avPlots(log_model2)
```

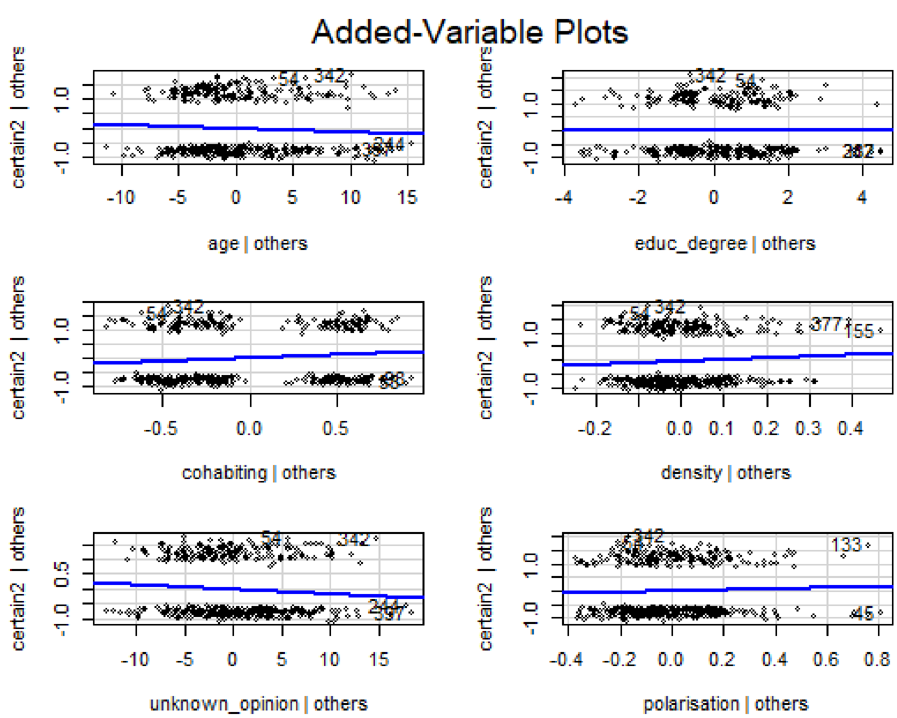


Figure D21: Added variable plot model 2

Model 3

```
ord_model <- polr(certainty_ord~
                    age+
                    educ_degree+
                    cohabiting+
                    density+
                    unknown_opinion +
                    polarisation,
                data = data,
                Hess=TRUE)
```

adding p-values to the table

```
# store table
table_ordreg <- coef(summary(ord_model))

# calculate and store p values
p <- pnorm(abs(table_ordreg[, "t value"]), lower.tail = FALSE) * 2

# combined table
(table_ordreg <- cbind(table_ordreg, "p value" = p))
```

	Value	Std. Error	t value	p value
age	0.05316208	0.01861798	2.8554155	0.004298056
educ_degree	-0.09054201	0.06440232	-1.4058812	0.159759414
cohabiting	-0.52764297	0.19232752	-2.7434606	0.006079535
density	-1.54086638	0.84669487	-1.8198603	0.068780282
unknown_opinion	0.03509998	0.01662448	2.1113434	0.034742804
polarisation	-0.49959578	0.45457052	-1.0990501	0.271746213
certain|preference	-0.42047962	0.67962472	-0.6186938	0.536118079
preference|uncertain	1.86902671	0.68675603	2.7215294	0.006498061

Figure D22: Estimates of the ordinal regression model

Model fit

```
lrm(formula = certainty_ord~age +
     educ_degree +
     cohabiting +
     density+
     unknown_opinion+
     polarisation,
   data=data)
```

```
lrm(formula = certainty_ord ~ age + educ_degree + cohabiting +
    density + unknown_opinion + polarisation, data = data)


                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs           477    LR chi2      24.60     R2       0.058    C       0.605
 certain      183    d.f.             6     R2(6,477)0.038    Dxy     0.211
 preference   222    Pr(> chi2) 0.0004    R2(6,400.3)0.045    gamma   0.211
 uncertain     72                           Brier    0.232    tau-a   0.129
max |deriv| 1e-11

                Coef    S.E.   Wald Z Pr(>|Z|)
y>=preference    0.4205 0.6796  0.62  0.5361
y>=uncertain    -1.8691 0.6867 -2.72  0.0065
age              0.0532 0.0186  2.86  0.0043
educ_degree     -0.0905 0.0644 -1.41  0.1598
cohabiting      -0.5277 0.1923 -2.74  0.0061
density         -1.5409 0.8467 -1.82  0.0688
unknown_opinion  0.0351 0.0166  2.11  0.0347
polarisation    -0.4996 0.4546 -1.10  0.2718
```

Figure D23: Model fit model 3

odds ratio

```
exp(cbind(OR = coef(ord_model)))
```

```
                       OR
age             1.0546006
educ_degree     0.9134360
cohabiting      0.5899940
density         0.2141954
unknown_opinion 1.0357233
polarisation    0.6067759
```

Figure D24: Odds ratio model 3

visualisation

```
plot(Effect(focal.predictors = "polarisation",ord_model))
```

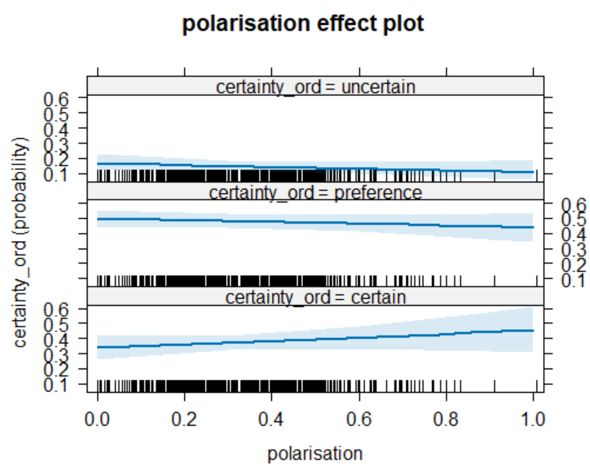


Figure D25: visualisation of the effect of polarisation

Influence of polarisation in only 3 clusters on certainty

Model 1

```
log_model_polarisation3 <- glm(certain1~age +
                           educ_degree +
                           cohabiting +
                           density+
                           unknown_opinion+
                           polaris_comm3,
                        family="binomial",
                        data=data)


summary(log_model_polarisation3)
```

```
Call:
glm(formula = certain1 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polaris_comm3, family = "binomial", data = data)

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)      0.69288    1.61197   0.430   0.6673
age             -0.11167    0.04543  -2.458   0.0140 *
educ_degree      0.30030    0.18451   1.628   0.1036
cohabiting       0.94957    0.58846   1.614   0.1066
density          6.06561    2.55798   2.371   0.0177 *
unknown_opinion -0.10065    0.04271  -2.357   0.0184 *
polaris_comm3    2.91894    1.38251   2.111   0.0347 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 151.45  on 159  degrees of freedom
Residual deviance: 124.57  on 153  degrees of freedom
  (318 observations deleted due to missingness)
AIC: 138.57

Number of Fisher Scoring iterations: 5
```

Figure D26: Estimates of logitisc regression model 1

Model fit

```
lrm(formula = certain1~age +
  educ_degree +
  cohabiting +
  density+
  unknown_opinion+
  polaris_comm3,
data=data)
```

```
lrm(formula = certain1 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polaris_comm3, data = data)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs           160    LR chi2      26.88    R2       0.253    C       0.795
 0             29    d.f.             6    R2(6,160)0.122    Dxy     0.590
 1            131    Pr(> chi2) 0.0002   R2(6,71.2)0.254    gamma   0.590
max |deriv| 6e-10                          Brier    0.124    tau-a   0.176

                 Coef    S.E.   Wald Z Pr(>|Z|)
Intercept         0.6929 1.6120  0.43  0.6673
age              -0.1117 0.0454 -2.46  0.0140
educ_degree       0.3003 0.1845  1.63  0.1036
cohabiting        0.9496 0.5885  1.61  0.1066
density           6.0656 2.5580  2.37  0.0177
unknown_opinion -0.1007 0.0427 -2.36  0.0184
polaris_comm3     2.9189 1.3825  2.11  0.0347
```

Figure D27: Model 1

Odds ratio

```r
exp(cbind(OR = coef(log_model_polarisation3)))
```

```
                           OR
(Intercept)         1.9994711
age                 0.8943416
educ_degree         1.3502664
cohabiting          2.5846076
density           430.7843642
unknown_opinion     0.9042486
polaris_comm3      18.5216346
```

Figure D28: Odds ratio model 1

Model 2

```r
log_model2_polarisation3 <- glm(certain2~age +
                      educ_degree +
                      cohabiting +
                      density +
                      unknown_opinion +
                      polaris_comm3,
                   family="binomial",
                   data=data)

summary(log_model2_polarisation3)
```

```
Call:
glm(formula = certain2 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polaris_comm3, family = "binomial", data = data)

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)     -0.60679    1.31381  -0.462   0.6442
age             -0.03385    0.03680  -0.920   0.3576
educ_degree     -0.14739    0.13069  -1.128   0.2594
cohabiting       0.91317    0.41132   2.220   0.0264 *
density          2.75420    1.97468   1.395   0.1631
unknown_opinion -0.08415    0.03680  -2.287   0.0222 *
polaris_comm3    0.74213    0.86739   0.856   0.3922
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 203.22  on 159  degrees of freedom
Residual deviance: 187.35  on 153  degrees of freedom
  (318 observations deleted due to missingness)
AIC: 201.35

Number of Fisher Scoring iterations: 4
```

Figure D29: Estimates of logistic regression model 2

Model fit

```
lrm(formula = certain2~age +
  educ_degree +
  cohabiting +
  density+
  unknown_opinion+
  polaris_comm3,
data=data)
```

```
lrm(formula = certain2 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polaris_comm3, data = data)


                      Model Likelihood     Discrimination    Rank Discrim.
                            Ratio Test            Indexes          Indexes
Obs           160    LR chi2      15.87     R2       0.131    C       0.693
 0            107    d.f.             6     R2(6,160)0.060    Dxy     0.386
 1             53    Pr(> chi2) 0.0145   R2(6,106.3)0.089    gamma   0.386
max |deriv| 5e-08                           Brier    0.200    tau-a   0.172

                Coef    S.E.   Wald Z Pr(>|Z|)
Intercept       -0.6068 1.3138 -0.46  0.6442
age             -0.0338 0.0368 -0.92  0.3576
educ_degree     -0.1474 0.1307 -1.13  0.2594
cohabiting       0.9132 0.4113  2.22  0.0264
density          2.7542 1.9747  1.39  0.1631
unknown_opinion -0.0842 0.0368 -2.29  0.0222
polaris_comm3    0.7421 0.8674  0.86  0.3922
```

Figure D30: Model fit model 2

Odds ratio

```
exp(cbind(OR = coef(log_model2_polarisation3)))
```

```
                        OR
(Intercept)      0.5451001
age              0.9667194
educ_degree      0.8629611
cohabiting       2.4922099
density         15.7084623
unknown_opinion  0.9192890
polaris_comm3    2.1004081
```

Figure D31: Odds ratio model 2

Model 3

```
ord_model_polarisation3 <- polr(certainty_ord~
                         age+
                         educ_degree+
                         cohabiting+
                         density+
                         unknown_opinion+
                         polaris_comm3,
                      data = data,
                      Hess=TRUE)
```

Adding p values to the table

```
# store table
table_ordreg_polarisation3 <- coef(summary(ord_model_polarisation3))

# calculate and store p values
p_polarisaion3 <- pnorm(abs(table_ordreg_polarisation3[, "t value"]), lower.tail
    = FALSE) * 2

# combined table
(table_ordreg_polarisation3 <- cbind(table_ordreg_polarisation3, "p value" =
    p_polarisaion3))
```

	Value	Std. Error	t value	p value
age	0.0581987122	0.03152032	1.846387057	0.06483603
educ_degree	-0.0003209381	0.10967899	-0.002926159	0.99766527
cohabiting	-0.9049714653	0.37399616	-2.419734649	0.01553184
density	-4.0139051110	1.77270573	-2.264281687	0.02355680
unknown_opinion	0.0873693340	0.03053874	2.860934739	0.00422394
polaris_comm3	-1.4214584301	0.79862863	-1.779874121	0.07509656
certain|preference	-1.1697345630	1.15457811	-1.013127262	0.31099938
preference|uncertain	1.3415953517	1.15733823	1.159207665	0.24637155

Figure D32: Estimates of the ordinal regression model

Model fit

```
lrm(formula = certainty_ord~age +
  educ_degree +
  cohabiting +
  density+
  unknown_opinion+
  polaris_comm3,
data=data)
```

```
lrm(formula = certainty_ord ~ age + educ_degree + cohabiting +
    density + unknown_opinion + polaris_comm3, data = data)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes           Indexes
Obs           160    LR chi2      26.27    R2         0.174    C       0.699
 certain       53    d.f.             6    R2(6,160)0.119    Dxy     0.399
 preference    78    Pr(> chi2) 0.0002  R2(6,134.7)0.140    gamma   0.399
 uncertain     29                          Brier      0.203    tau-a   0.249
max |deriv| 3e-08

                 Coef    S.E.   Wald Z Pr(>|Z|)
y>=preference     1.1697 1.1546  1.01  0.3110
y>=uncertain     -1.3416 1.1573 -1.16  0.2463
age               0.0582 0.0315  1.85  0.0648
educ_degree      -0.0003 0.1097  0.00  0.9977
cohabiting       -0.9050 0.3740 -2.42  0.0155
density          -4.0138 1.7727 -2.26  0.0236
unknown_opinion   0.0874 0.0305  2.86  0.0042
polaris_comm3    -1.4214 0.7986 -1.78  0.0751
```

Figure D33: Model fit model 3

```
exp(cbind(OR = coef(ord_model_polarisation3)))
```

```
                       OR
age            1.05992559
educ_degree    0.99967911
cohabiting     0.40455343
density        0.01806272
unknown_opinion 1.09129966
polaris_comm3  0.24136175
```

Figure D34: Odds ratio model 3

Table D1: effects of polarisation in only 3 clusters on certainty

Model 1 dependent variable:"I don't know" is uncertain, all others are certain.

Model 2: dependent variable: "i don't know" and "probably yes/no" are uncertain, "absolutely yes/no" is certain

Model 3: dependent variable: "i don't know" is uncertain, "probably yes/no" is preference, " absolutely yes/no" is certain.

n=160

	Model 1			Model 2			Model 3		
	OR	CI OR Lower Upper	p-value	OR	CI OR Lower Upper	p-value	OR	CI OR Lower Upper	p-value
Age	0.894	0.819 0.976	0.014	0.967	0.899 1.039	0.358	1.060	0.995 1.128	0.065
Educational degree	1.350	0.939 1.940	0.104	0.863	0.668 1.116	0.259	1	0.806 1.240	0.998
Cohabiting	2.585	0.817 8.186	0.107	2.492	1.113 5.576	0.026	0.405	0.194 0.842	0.016
Density	403.784	2.864 64840.131	0.018	15.708	0.327 753.704	0.163	0.018	0.001 0.583	0.024
Unknown opinion network	0.904	0.831 0.983	0.018	0.919	0.855 0.989	0.022	1.091	1.027 1.159	0.004
Polarisation	18.522	1.233 278.852	0.035	2.100	0.384 11.488	0.392	0.241	0.050 1.156	0.075

Influence of polarisation in large clusters on certainty

Model 1

```
log_model_large <- glm(certain1~age +
                         educ_degree +
                         cohabiting +
                         density +
                         unknown_opinion+
                         polarisation_large,
                       family="binomial",
                       data=data)
summary(log_model_large)
```

```
Call:
glm(formula = certain1 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation_large, family = "binomial",
    data = data)

Coefficients:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)         4.33432    1.23796   3.501 0.000463 ***
age                -0.13147    0.03004  -4.377  1.2e-05 ***
educ_degree         0.26796    0.11809   2.269 0.023257 *
cohabiting          0.69135    0.35399   1.953 0.050816 .
density            -0.34584    1.75249  -0.197 0.843561
unknown_opinion    -0.07782    0.02817  -2.762 0.005740 **
polarisation_large -0.68048    0.94755  -0.718 0.472667
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 309.12  on 374  degrees of freedom
Residual deviance: 282.14  on 368  degrees of freedom
  (103 observations deleted due to missingness)
AIC: 296.14

Number of Fisher Scoring iterations: 5
```

Figure D35: Estimates of logistic regression model 1

Model fit

```
lrm(formula = certain1~age +
  educ_degree +
  cohabiting +
  density+
  unknown_opinion+
  polarisation_large,
data=data)
```

```
lrm(formula = certain1 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation_large, data = data)

                      Model Likelihood     Discrimination    Rank Discrim.
                         Ratio Test           Indexes          Indexes
Obs            375    LR chi2     26.97    R2         0.124    C       0.725
 0              54    d.f.            6    R2(6,375)0.054    Dxy     0.451
 1             321    Pr(> chi2) 0.0001  R2(6,138.7)0.140    gamma   0.451
max |deriv| 1e-05                          Brier      0.114    tau-a   0.111

                   Coef    S.E.   Wald Z Pr(>|Z|)
Intercept           4.3343 1.2380  3.50  0.0005
age                -0.1315 0.0300 -4.38  <0.0001
educ_degree         0.2680 0.1181  2.27  0.0233
cohabiting          0.6913 0.3540  1.95  0.0508
density            -0.3458 1.7525 -0.20  0.8436
unknown_opinion    -0.0778 0.0282 -2.76  0.0057
polarisation_large -0.6805 0.9475 -0.72  0.4727
```

Figure D36: Model fit model 1

odds ratio

```
exp(cbind(OR = coef(log_model_large)))
```

```
                            OR
(Intercept)         76.2733540
age                  0.8768035
educ_degree          1.3072989
cohabiting           1.9964052
density              0.7076273
unknown_opinion      0.9251341
polarisation_large   0.5063760
```

Figure D37: Enter Caption

Model 2

```
log_model2_large <- glm(certain2~age +
                      educ_degree +
                      cohabiting +
                      density +
                      unknown_opinion+
                      polarisation_large,
                    family="binomial",
                    data=data)

summary(log_model2_large)
```

```
Call:
glm(formula = certain2 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation_large, family = "binomial",
    data = data)

Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)         0.248238   0.874106   0.284   0.7764
age                -0.033079   0.023399  -1.414   0.1575
educ_degree        -0.004885   0.079437  -0.061   0.9510
cohabiting          0.580894   0.236439   2.457   0.0140 *
density            -0.715324   1.256476  -0.569   0.5691
unknown_opinion    -0.044575   0.020048  -2.223   0.0262 *
polarisation_large -0.392016   0.674248  -0.581   0.5610
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 493.41  on 374  degrees of freedom
Residual deviance: 481.07  on 368  degrees of freedom
  (103 observations deleted due to missingness)
AIC: 495.07

Number of Fisher Scoring iterations: 4
```

Figure D38: Estimates of logistic regression model 2

Model fit

```
lrm(formula = certain2~age +
  educ_degree +
  cohabiting +
  density+
  unknown_opinion+
  polarisation_large,
data=data)
```

```
lrm(formula = certain2 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation_large, data = data)


                      Model Likelihood     Discrimination    Rank Discrim.
                            Ratio Test            Indexes          Indexes
Obs            375    LR chi2      12.34     R2         0.044    C        0.609
 0             237    d.f.             6     R2(6,375)0.017    Dxy      0.218
 1             138    Pr(> chi2) 0.0547    R2(6,261.6)0.024    gamma    0.218
max |deriv| 7e-14                            Brier      0.225    tau-a    0.102

                   Coef    S.E.   Wald Z Pr(>|Z|)
Intercept           0.2482 0.8741  0.28  0.7764
age                -0.0331 0.0234 -1.41  0.1575
educ_degree        -0.0049 0.0794 -0.06  0.9510
cohabiting          0.5809 0.2364  2.46  0.0140
density            -0.7153 1.2565 -0.57  0.5691
unknown_opinion    -0.0446 0.0200 -2.22  0.0262
polarisation_large -0.3920 0.6742 -0.58  0.5610
```

Figure D39: Model fit model 2

odds ratio

```
exp(cbind(OR = coef(log_model2_large)))
```

```
                          OR
(Intercept)        1.2817653
age                0.9674620
educ_degree        0.9951273
cohabiting         1.7876366
density            0.4890337
unknown_opinion    0.9564042
polarisation_large 0.6756935
```

Figure D40: Odds ratio model 2

Model 3

```
ord_model_large <- polr(certainty_ord~
                          age+
                          educ_degree+
                          cohabiting+
                          density +
                          unknown_opinion+
                          polarisation_large,
                       data = data,
                       Hess=TRUE)
```

Adding p-values to the table

```
# store table
table_ordreg_large <- coef(summary(ord_model_large))


# calculate and store p values
p_large <- pnorm(abs(table_ordreg_large[, "t value"]), lower.tail = FALSE) * 2


# combined table
(table_ordreg_large <- cbind(table_ordreg_large, "p value" = p_large))
```

	Value	Std. Error	t value	p value
age	0.06645154	0.02173320	3.0576046	0.0022311379
educ_degree	-0.07127789	0.07217942	-0.9875099	0.3233927506
cohabiting	-0.64994568	0.22049343	-2.9476873	0.0032016074
density	0.49390108	1.12113639	0.4405361	0.6595488546
unknown_opinion	0.05226122	0.01828484	2.8581716	0.0042608986
polarisation_large	0.39465079	0.61163962	0.6452342	0.5187754354
certain|preference	0.68722301	0.79331903	0.8662631	0.3863458891
preference|uncertain	3.11615137	0.81248059	3.8353548	0.0001253832

Figure D41: Estimates of the ordinal regression model

Model fit

```
lrm(formula = certainty_ord~age +
  educ_degree +
  cohabiting +
  density+
  unknown_opinion+
  polarisation_large,
data=data)
```

```
lrm(formula = certainty_ord ~ age + educ_degree + cohabiting +
    density + unknown_opinion + polarisation_large, data = data)


                      Model Likelihood      Discrimination    Rank Discrim.
                         Ratio Test            Indexes           Indexes
obs             375    LR chi2     21.80    R2        0.065    C       0.624
 certain        138    d.f.            6    R2(6,375)0.041    Dxy     0.248
 preference     183    Pr(> chi2) 0.0013  R2(6,311.6)0.049    gamma   0.248
 uncertain       54                         Brier     0.226    tau-a   0.150
max |deriv| 4e-11

                   Coef    S.E.   Wald Z Pr(>|Z|)
y>=preference      -0.6873 0.7933 -0.87  0.3863
y>=uncertain       -3.1163 0.8125 -3.84  0.0001
age                 0.0665 0.0217  3.06  0.0022
educ_degree        -0.0713 0.0722 -0.99  0.3234
cohabiting         -0.6499 0.2205 -2.95  0.0032
density             0.4941 1.1211  0.44  0.6594
unknown_opinion     0.0523 0.0183  2.86  0.0043
polarisation_large  0.3946 0.6116  0.65  0.5188
```

Figure D42: Model fit model 3

odds ratio

```
exp(cbind(OR = coef(ord_model_large)))
```

```
                         OR
age               1.0687092
educ_degree       0.9312031
cohabiting        0.5220741
density           1.6386965
unknown_opinion   1.0536509
polarisation_large 1.4838659
```

Figure D43: Odds ratio model 3

Table D2: effects of polarisation in large clusters on certainty

Model 1 dependent variable:"I don't know" is uncertain, all others are certain.

Model 2: dependent variable: "i don't know" and "probably yes/no" are uncertain, "absolutely yes/no" is certain

Model 3: dependent variable: "i don't know" is uncertain, "probably yes/no" is preference, " absolutely yes/no" is certain.

n=375

	Model 1			Model 2			Model 3		
	OR	CI OR Lower Upper	p-value	OR	CI OR Lower Upper	p-value	OR	CI OR Lower Upper	p-value
Age	0.877	0.827 0.930	<0.001	0.967	0.925 1.012	0.158	1.069	1.023 1.115	0.002
Educational degree	1.307	1.037 1.648	0.023	0.995	0.852 1.162	0.951	0.931	0.809 1.073	0.323
Cohabiting	1.996	0.997 3.994	0.051	1.788	1.126 2.839	0.014	0.522	0.339 0.803	0.003
Density	0.708	0.023 21.931	0.844	0.489	0.042- 5.736	0.569	1.638	0.182 14.749	0.660
Unknown opinion network	0.925	0.876 0.977	0.006	0.956	0.919 0.994	0.026	1.054	1.017 1.091	0.004
Polarisation	0.506	0.079 3.248	0.473	0.676	0.180 2.532	0.561	1.484	0.447 4.926	0.519

Influence of polarisation on certainty when all intentions are known

Model 1

```
log_model_allknown <- glm(certain1~age +
                          educ_degree +
                          cohabiting +
                          density+
                          unknown_opinion+
                          polarisation_allknown,
                          family="binomial",
                          data=data)

summary(log_model_allknown)
```

```
Call:
glm(formula = certain1 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation_allknown, family = "binomial",
    data = data)

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)            2.04559    1.07816   1.897 0.057788 .
age                   -0.11251    0.02856  -3.939 8.17e-05 ***
educ_degree            0.39067    0.11253   3.472 0.000517 ***
cohabiting             0.60712    0.31559   1.924 0.054384 .
density                3.42930    1.51832   2.259 0.023907 *
unknown_opinion       -0.06404    0.03008  -2.129 0.033252 *
polarisation_allknown  0.53405    0.39424   1.355 0.175536
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 370.40  on 445  degrees of freedom
Residual deviance: 334.79  on 439  degrees of freedom
  (32 observations deleted due to missingness)
AIC: 348.79

Number of Fisher Scoring iterations: 5
```

Figure D44: Estimates of logistic model 1

Model fit

```
lrm(formula = certain1~age +
educ_degree +
cohabiting +
density+
unknown_opinion+
polarisation_allknown,
data=data)
```

```
lrm(formula = certain1 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation_allknown, data = data)

                        Model Likelihood      Discrimination    Rank Discrim.
                           Ratio Test             Indexes           Indexes
Obs             446    LR chi2      35.62     R2        0.136    C       0.727
 0               65    d.f.             6     R2(6,446)0.064     Dxy     0.455
 1              381    Pr(> chi2) <0.0001   R2(6,166.6)0.163     gamma   0.455
max |deriv| 3e-12                             Brier     0.114    tau-a   0.113

                       Coef    S.E.   Wald Z Pr(>|Z|)
Intercept               2.0456 1.0782  1.90  0.0578
age                    -0.1125 0.0286 -3.94  <0.0001
educ_degree             0.3907 0.1125  3.47  0.0005
cohabiting              0.6071 0.3156  1.92  0.0544
density                 3.4293 1.5183  2.26  0.0239
unknown_opinion        -0.0640 0.0301 -2.13  0.0333
polarisation_allknown   0.5340 0.3942  1.35  0.1755
```

Figure D45: Model fit model 1

odds ratio

```
exp(cbind(OR = coef(log_model_allknown)))
```

```
                              OR
(Intercept)            7.7337249
age                    0.8935892
educ_degree            1.4779670
cohabiting             1.8351412
density               30.8549303
unknown_opinion        0.9379634
polarisation_allknown  1.7058250
```

Figure D46: Odds ratio model 1

Model 2

```
log_model2_allknown <- glm(certain2~age +
                              educ_degree +
                              cohabiting +
                              density +
                              unknown_opinion +
                              polarisation_allknown,
                           family="binomial",
                           data=data)

summary(log_model2_allknown)
```

```
Call:
glm(formula = certain2 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation_allknown, family = "binomial",
    data = data)

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)           -0.85803    0.75458  -1.137  0.25550
age                   -0.02881    0.02128  -1.354  0.17571
educ_degree            0.03506    0.07306   0.480  0.63129
cohabiting             0.57272    0.21359   2.681  0.00733 **
density                1.65097    0.94569   1.746  0.08085 .
unknown_opinion       -0.02870    0.02066  -1.389  0.16473
polarisation_allknown  0.26419    0.27899   0.947  0.34367
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 596.58  on 445  degrees of freedom
Residual deviance: 582.75  on 439  degrees of freedom
  (32 observations deleted due to missingness)
AIC: 596.75

Number of Fisher Scoring iterations: 4
```

Figure D47: Estimates of logistic model 2

Model fit

```
lrm(formula = certain2~age +

educ_degree +

cohabiting +

density+

unknown_opinion+

polarisation_allknown,

data=data)
```

```
lrm(formula = certain2 ~ age + educ_degree + cohabiting + density +
    unknown_opinion + polarisation_allknown, data = data)

                      Model Likelihood      Discrimination    Rank Discrim.
                         Ratio Test                Indexes          Indexes
Obs           446    LR chi2      13.83     R2          0.041    C       0.593
 0            272    d.f.             6     R2(6,446)0.017       Dxy     0.187
 1            174    Pr(> chi2) 0.0316    R2(6,318.3)0.024       gamma   0.187
max |deriv| 1e-13                           Brier       0.230    tau-a   0.089

                      Coef    S.E.   Wald Z Pr(>|Z|)
Intercept             -0.8580 0.7546 -1.14  0.2555
age                   -0.0288 0.0213 -1.35  0.1757
educ_degree            0.0351 0.0731  0.48  0.6313
cohabiting             0.5727 0.2136  2.68  0.0073
density                1.6510 0.9457  1.75  0.0808
unknown_opinion       -0.0287 0.0207 -1.39  0.1647
polarisation_allknown  0.2642 0.2790  0.95  0.3437
```

Figure D48: Model fit model 2

odds ratio

```
exp(cbind(OR = coef(log_model2_allknown)))
```

```
                             OR
(Intercept)           0.4239980
age                   0.9715973
educ_degree           1.0356823
cohabiting            1.7730921
density               5.2120576
unknown_opinion       0.9717056
polarisation_allknown 1.3023703
```

Figure D49: Odds ratio model 2

Model 3

```
ord_model_allknown <- polr(certainty_ord~
                              age+
                              educ_degree+
                              cohabiting+
                              density+
                              unknown_opinion+
                              polarisation_allknown,
                           data = data,
                           Hess=TRUE)
```

adding p-values to the table

```
# store table
table_ordreg_allknown <- coef(summary(ord_model_allknown))

# calculate and store p values
p_allknown <- pnorm(abs(table_ordreg_allknown[, "t value"]), lower.tail = FALSE)
    * 2

# combined table
(table_ordreg_allknown <- cbind(table_ordreg_allknown, "p value" = p_allknown))
```

	Value	Std. Error	t value	p value
age	0.05318480	0.01954708	2.720857	0.006511293
educ_degree	-0.12356866	0.06676314	-1.850851	0.064190931
cohabiting	-0.60871336	0.19913071	-3.056853	0.002236737
density	-2.13886157	0.88957625	-2.404360	0.016200810
unknown_opinion	0.03482196	0.01910469	1.822691	0.068350174
polarisation_allknown	-0.32406436	0.25801053	-1.256012	0.209111570
certain|preference	-0.75820816	0.69913980	-1.084487	0.278148846
preference|uncertain	1.55784494	0.70416507	2.212329	0.026943926

Figure D50: Estimates of the ordinal regression model

Model fit

```
lrm(formula = certainty_ord~age +
educ_degree +
cohabiting +
density+
unknown_opinion+
polarisation_allknown,
data=data)
```

```
lrm(formula = certainty_ord ~ age + educ_degree + cohabiting +
    density + unknown_opinion + polarisation_allknown, data = data)


                      Model Likelihood      Discrimination      Rank Discrim.
                            Ratio Test             Indexes            Indexes
Obs             446    LR chi2     26.31     R2          0.066   C         0.610
 certain        174    d.f.            6     R2(6,446)0.045     Dxy       0.220
 preference     207    Pr(> chi2) 0.0002   R2(6,373.5)0.053     gamma     0.221
 uncertain       65                          Brier      0.232   tau-a     0.135
max |deriv| 3e-11

                      Coef    S.E.   Wald Z Pr(>|Z|)
y>=preference          0.7582 0.6991  1.08  0.2782
y>=uncertain          -1.5579 0.7041 -2.21  0.0269
age                    0.0532 0.0195  2.72  0.0065
educ_degree           -0.1236 0.0668 -1.85  0.0642
cohabiting            -0.6087 0.1991 -3.06  0.0022
density               -2.1388 0.8896 -2.40  0.0162
unknown_opinion        0.0348 0.0191  1.82  0.0683
polarisation_allknown -0.3240 0.2580 -1.26  0.2092
```

Figure D51: model fit model 3

odds ratio

```
exp(cbind(OR = coef(ord_model_allknown)))
```

```
                              OR
age                    1.0546245
educ_degree            0.8837610
cohabiting             0.5440504
density                0.1177889
unknown_opinion        1.0354353
polarisation_allknown  0.7232037
```

Figure D52: Odds ratio model 3

Table D3: effects of polarisation when all opinions are known on certainty

Model 1 dependent variable:"I don't know" is uncertain, all others are certain.

Model 2: dependent variable: "i don't know" and "probably yes/no" are uncertain, "absolutely yes/no" is certain

Model 3: dependent variable: "i don't know" is uncertain, "probably yes/no" is preference, " absolutely yes/no" is certain.

n=446

	Model 1			Model 2			Model 3		
	OR	CI OR Lower Upper	p-value	OR	CI OR Lower Upper	p-value	OR	CI OR Lower Upper	p-value
Age	0.893	0.844 0.945	<0.001	0.972	0.932 1.012	0.176	1.055	1.014 1.097	0.007
Educational degree	1.478	1.185 1.845	0.001	1.036	0.894 1.120	0.631	0.884	0.775 1.007	0.064
Cohabiting	1.835	0.988 3.409	0.054	1.773	1.166 2.698	0.007	0.544	0.368 0.803	0.002
Density	30.855	1.574 604.426	0.024	5.212	0.816 33.287	0.081	0.118	0.021 0.674	0.016
Unknown opinion network	0.938	0.884 0.995	0.033	0.972	0.932 1.012	0.165	1.035	0.998 1.075	0.068
Polarisation	1.706	0.788 3.692	0.176	1.302	0.754 2.250	0.344	0.723	0.436 1.199	0.209