ISER Working Paper Series

Peer Groups, Social Support, and Well-being: Evidence from a Large Online Maternity Community

Lingqing Jiang
University of Essex

Zhen Zhu
University of Kent

No. 2021-1
January 2021

INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH

Economic and Social Research Council

University of Essex

www.iser.essex.ac.uk

Non-technical Summary

WHO reports that worldwide about 10% of pregnant women and 13% of women after childbirth experience depression. Moreover, studies show that maternal depression is related to several adverse child outcomes. One well-known intervention to improve maternal well-being and child development is the Home Visiting Program (HVP) that offers prenatal support to disadvantageous first-time mothers. However, its effects on mother and child well-being are mixed across countries. In this project we study an alternative source of prenatal social support that comes from virtual peer groups where mothers exchange personal experiences and knowledge about maternal caring. It has several advantages: 1) Creating and joining virtual peer groups is of minimal costs. 2) Virtual peer groups are open to all mothers with various socio-demographic background. Therefore, we can generalise its implication to a more representative population. 3) Mothers are both provider and recipient of the support in the peer groups. To some extent, this is a self-sufficient source of prenatal social support. 4) Virtual peer support enables mothers to overcome physical and time constraints embedded within their offline counterparts. Under certain circumstances such as COVID-19 where offline physical support is restricted, virtual support might be the only available source of social support. Moreover, the option of maintaining anonymity in online settings provides users with more convenience when discussing sensitive and controversial topics.

We investigate i) how quasi-randomly assigned peer groups affect the observable social support in terms of information exchange among pregnant women in the peer groups, and ii) how perceived social support in these peer groups relates to well-being of the mothers and the newborns. We combine an observational data set with a survey data set from a large online maternity community to answer these two questions.

Our study is based on an online community that assigns pregnant users who are at the same stage of the pregnancy to the same peer groups. Besides the default peer group, users can also manually join other peer groups. All peer groups are created ten months before the delivery month, and are active since the creation till the end of the corresponding delivery month. The information and knowledge exchange among pregnant women *themselves* in the peer groups constitutes the essence of prenatal social support.

Besides the observable social support in terms of information exchange in the peer groups from the observational dataset, we also use a survey questionnaire to measure 1) user's perceived social support in the peer groups using a modified version of the *Multidimensional Scale of Perceived Social Support*, 2) mental health both during pregnancy and after childbirth using the *Edinburgh Depression Scale*, and 3) newborn's birth outcomes (weight, whether planned natural birth, and whether breastfeeding) among users who gave birth within one year. (Notice that the sample of survey users is different from the sample of the users in the observational dataset as there are two years apart.)

We find that while having more peer groups reduces observable social support in terms of information exchange, it does not undermine perceived social support. Higher perceived social support from online peer groups is positively associated with both prenatal and postnatal mental health of the mothers as well as newborns' birth weight.

Peer Groups, Social Support, and Well-being: Evidence from a Large Online Maternity Community*†

Lingqing Jiang[1] and Zhen Zhu[2]

[1]University of Essex
[2]University of Kent

January 19, 2021

Abstract

Increasing attention is being paid to social determinants of health. We study how quasi-randomly assigned peer groups affect social support among over 24,000 pregnant women, and how social support in these peer groups relates to mothers' well-being both during and after pregnancy as well as birth outcomes. We find that while having more peer groups reduces observable social support in terms of information exchange, it does not seem to undermine perceived social support. Higher perceived social support from online peer groups is positively associated with both prenatal and postnatal mental health of the mothers as well as newborns' birth weight.

Keywords: *Social support, Peer groups, Mental health, IV, Survey*

JEL Classification: I10, I31, J13, C36

*Acknowledgement: The authors would like to thank Bo Honore, Sonia Bhalotra, Jordi Brandt, Gabriella Conti, Thomas Cornelissen, Monica Costa Dias, Subhasish Chowdhury, Francesco Drago, Marco Francesconi, Lorenz Goette, Angus Holford, Marco Manacorda, Friederike Mengel, Marie Claire Villeval, Christian Zehnder, and participants at the conferences of Royal Economic Society 2019, Reading behavioural Workshop 2019, ISER internal seminar, General Online Research 2019, NetMob 2019, and IC2S2 2019 for their insightful comments.

†We obtained ethical approval for this study from the University of Essex.

1 Introduction

Human development starts in the womb (Kail & Cavanaugh 2018). Childbearing is one of the most significant life events that brings not only joy but also anxiety and stress. The World Health Organization (WHO) reports that worldwide about 10% of pregnant women and 13% of women who have just given birth experience a mental disorder, primarily depression. In developing countries this is even higher, i.e. 15.6% during pregnancy and 19.8% after childbirth.[1] Moreover, maternal depression and stress are also related to several adverse child outcomes (Carlson 2015, Aizer et al. 2016, Persson & Rossin-Slater 2018).

One well-known intervention to improve maternal well-being and child development is the Home Visiting Program (HVP) that offers prenatal and parental support to disadvantageous first-time mothers. It has been implement in the US (Olds et al. 2019), UK (Robling et al. 2016), Ireland (Doyle et al. 2015), and Germany (Sandner et al. 2018). However, the findings on mother and child health outcomes are mixed. Nevertheless, studies that found positive effects suggest the benefits are driven by compensating the lack of social support (Sandner et al. 2018).[2]

In this paper we study an alternative source of prenatal social support that comes from peer groups where pregnant women exchange personal experiences and knowledge about maternal caring. In particular, we investigate i) how quasi-randomly assigned peer groups affect the observable social support in terms of information exchange, and ii) how perceived social support from these peer groups relates to the well-being of mothers and newborns. We combine an observational data set with a survey data set from a large online maternity community to answer these two questions.

The economic literature has long recognised that peers and peer groups influence individuals' behaviour in various contexts.[3] With the development of digital technologies, more and more traditional offline social communities have found an increasing popularity of their online counterparts. The formation of peer groups in online communities is one of the examples. However, whether peer groups in online settings provide effective social support and how virtual social support relates to individual well-being are underexplored.

[1] The figures are from the WHO's article on maternal mental health that can be accessed at `https://www.who.int/mental_health/maternal-child/maternal_mental_health/en/`

[2] Medical literature shows that prenatal exposure to stress and lack of prenatal social support presents a strong risk factor for depression during pregnancy and the postpartum period, as well as the mental health of the next generation (Collins et al. 1993, Elsenbruch et al. 2006).

[3] For example, peers or peer groups play an important role in productivity at work (Mas & Moretti 2009, Bandiera et al. 2009, Cornelissen et al. 2017), academic performance (Sacerdote 2001, Carrell et al. 2009), retirement planning (Duflo & Saez 2002, Brown & Laschever 2012), prosocial behaviour (Gächter et al. 2013, Bruhin et al. 2020), and sports competition (Yamane & Hayashi 2015, Jiang 2020).

The main feature of this online maternity community is that it assigns pregnant users into peer groups based on their estimated due date (EDD). The creation of peer groups provides a common space for users who are at the same stage of pregnancy to exchange relevant information among themselves. Users typically exchange information about how to deal with a particular symptom or reaction (e.g. morning sickness, edema), how to read a particular check-up and screening test report, what diet plan to follow during pregnancy etc. Such information exchange among pregnant women constitutes the essence of social support, as defined by *interpersonal exchange of potentially useful information or things* (Cohen & Syme 1985). Peer groups are labelled in the format of *Month Year* and are created ten months before the estimated delivery month. For example, a user who is expecting to give birth on the 1st March 2018 is assigned to the peer group *March 2018* which was created in May 2017. Each peer group is active since the creation till the end of the corresponding delivery month. The observational data set includes the full cycle of three peer groups with over 24,000 users.

We first look at the observable social support in terms of information exchange in the peer groups according to the definition by Cohen & Syme (1985). A user can exchange information in the peer groups by either initiating a post or responding to a post. We consider both posts and responses as information exchange. A post typically initiates information exchange by either sharing own experience or asking a question; A response typically exchanges information by either answering a question or asking a follow-up question. Therefore, we measure the observable social support by the sum of individual posts and responses.

The feature is designed such that each pregnant user has only one peer group, namely, the default peer group determined by the EDD. However, we observe that almost half (47.8%) of the users join more than one peer groups in practice.[4] There could be various motives for that: users may want to share information with a greater number of users, access information from more users, or compare information across different peer groups.[5] Another reason is the natural uncertainty of EDD.[6] In particular, there is a clear pattern that users whose EDD falling in the beginning (end) of the month tend to additionally join the peer group of the previous (next) month. Among those who joined multiple peer groups, almost 30% of them joined one of the adjacent peer groups, i.e. the peer groups of the previous or the next month.

[4]The group enrollment works as follows: users are required to put in their EDD at the registration step. After the registration, they will be automatically assigned to the default peer group corresponding to the month of their EDD. Users may join other peer groups manually by first clicking the other peer groups in the menu and then clicking "join". In principle, a user can search for and join any existing peer groups. The range of such peer groups featured on the homepage is one year before and ten months after the current month.

[5]We explicitly asked about the motives of joining additional peer groups in the survey (see details in Table 5).

[6]Only about 4-5% of the babies are born on their estimated due date. Childbirth has a standard deviation of 14-16 days (Hoffman et al. 2008).

We investigate the case of having multiple peer groups as it can have ambiguous effects on the information exchange. On the one hand, information consumes attention. Herbert A. Simon sharply pointed out that "a wealth of information creates a poverty of attention" (Simon 1957). Limited attention literature (e.g. Kahneman 1973, Camerer 2005) emphasises that processing massive amount of information is often beyond human beings' capacity of attention. Consequently, individuals who joined an additional peer group – that doubles the number of peers and the amount of available information – face a trade-off in allocating the time and effort in exchanging information between peer groups. On the other hand, we have seen remarkable benefits of increasingly connected social networks. Social network literature (e.g. Jackson 2010) suggests that the exposure to a wider social network promotes information diffusion and knowledge spillovers. Therefore, users who are members of two peer groups may have more information exchange, thanks to the positive spillovers between peer groups.

The empirical challenges of identifying the causal effect of having multiple peer groups lie in users' self-selection in i) which peer group to join and ii) how many peer groups to join.[7] Two desirable features of this setup allow us to overcome the challenges. First, the default peer groups can be seen as *homogeneous* as all users are exposed to peers in similar maternal status within the groups. Second, whether users join additional peer groups is partially quasi-random. One can think of the day of EDD as a natural lottery ticket: each pregnant user quasi-randomly draws a number ranging from 1 to 30/31. Therefore, we can use the day of EDD as an instrumental variable to predict the probability of having multiple peer groups as i) it directly affects the probability of joining one of the adjacent peer groups; and ii) it is a quasi-random variable that is uncertain due to both natural and measurement errors.

We perform the main empirical analysis using the observations that have exactly two peer groups, in particular, those who additionally join either the previous-month peer group or the next-month peer group.[8] We then generalise the analysis to multiple peer groups. The results are robust and qualitatively similar (see Table A4). The empirical analysis yields two findings. First, joining an additional peer group reduces users' information exchange in their default peer group. Second, the sum of the information exchange generated by the users in two peer groups is less than the sum generated by the users in only one (default) peer group.

We interpret our first finding as a substitution effect between the additional

[7] It is important to notice that since we are only comparing two peer groups vs one (default) peer group, our identification challenge is the endogenous group enrollment rather than the *reflection problem* in Manski (1993). In other words, we are interested in the $\hat{\beta}$ in $Y_i = \beta(\text{Two PGs})$ instead of $Y_i = \beta \bar{Y}_{-i}$.

[8] Among the multiple-peer-group users, 55.4% have exactly two peer groups. Among the two-peer-group users, 49.5% joined one of the adjacent peer groups (i.e. the previous or the next month). In our main analysis, we consider the latter case as the treatment group and the single-peer-group users as the control group.

peer group and the default peer group, which is in line with the limited attention literature. Our second finding, if assuming linear production function of information exchange, suggests that there could be other effects besides the substitution effects. One possible explanation is that having two peer groups weakens the group identity of either group, which can reduce contribution or encourage "lurking" behaviour in either group.[9]

Next, we turn to perceived social support in the peer groups and investigate how it relates to users' mental health and newborn's birth outcomes. We conducted a survey in 2020, two years after the time period of the observational data. The survey includes 500 participants who were currently or have been users of the peer groups in this online community, but were not in the observational data.[10] Among the survey participants, 300 were pregnant and 200 gave birth within one year at the time of the survey.[11] The survey questionnaire measures i) users' perceived social support in the peer groups using a modified version of the *Multidimensional Scale of Perceived Social Support* (MSPSS) (Zimet et al. 1988), ii) users' mental health using the *Edinburgh Depression Scale* (EDS) (Murray & Cox 1990), and iii) the newborn's birth outcomes (weight, whether planned natural birth, and whether breastfeeding) among users who gave birth within one year.

The analysis of the survey data yields the following three findings. First, users' perceived social support in the peer group(s) does not seem to be affected by whether having single or multiple peer group. Second, higher perceived social support in the peer group(s) during pregnancy is associated with better mental health not only during pregnancy but also after childbirth. Third, higher perceived social support in the peer groups is positively associated with newborn's weight.

Taking the results from both empirical analyses together, it suggests that while having more peer group reduces the amount of information exchange, it does not necessarily undermine perceived social support in the peer groups. There can be several explanations for this observation. First, methodologically speaking, while we identify the causal effect of having one additional peer group by using the day of EDD as IV in the observational data, we use OLS regressions in the survey data as the instrument is weak in the survey data. Second, the users participating in the survey data are likely to be more active than the average users in the large observational data set. This can be seen in the fact that the fraction of users that report having multiple peer groups is

[9]Lurking is very common in online communities. It is related to the free riding behaviour in the sense that lurking users solely receive information without generating any information. Research has shown that stronger feelings about group identity increase the level of contribution in the public goods games (Charness et al. 2014).

[10]We invite users of this online community to our anonymous survey on a voluntary basis. We commissioned Zhongyan Technology, a major online survey and research company in China, to implement our survey and collect the data. The two data sets are not linked with personal identifiers.

[11]Notice that these are two different sets of users, rather than the same set of users during and after pregnancy.

much higher than the one we actually observe in the large data. The positive correlation between the overall activeness and joining multiple peer groups may offset the potential negative effects of having multiple peer groups on perceived social support, if any. Last but not least, perceived social support is not purely determined by the quantity of the information exchange. Despite the reduction in the quantity of information, users may still feel socially supported in the peer groups. The role of peer groups is similar to the role of friends. Although the number *per se* plays a role, the fact that one has a peer group (or a friend) to interact with is probably more important than the number of peer groups (or friends).

Our findings have important policy implications. First, comparing to the traditional third-party support such as home visiting program, online peer groups is a new source of social support that involves minimal costs and proven to be effective both during and after pregnancy. Moreover, the peer groups are open to all pregnant women rather than a small group of disadvantaged women, and therefore, we can generalise our results to a more representative population. Furthermore, under certain circumstances where offline social support is restricted, online peer groups can still operate. Thus, pregnant women should be encouraged to adopt this form of prenatal social support. Second, regarding the design of peer groups – from the perspective of the users – as long as having multiple peer groups does not harm their well-being, they may preserve the freedom of joining multiple groups, despite the fact that it may dilute their attention to some extent. From the perspective of the online community, our results suggest that communities and organisations with similar features be cautious when designing peer groups. Allowing users to join multiple peer groups may not necessarily boost engagement.[12]

Our paper contributes to three strands of literature. First, we contribute to the vibrant literature that investigates the link between prenatal environments and the well-being of mothers as well as child development.[13] There are two novelties of our study with respect to the existing literature. First, we exclusively focus on well-defined peer groups where users are not only the recipients but also the providers of the support. This differs from the third-party support where users are solely the recipients. Second, we study a virtual form social support which provides users more flexibility and anonymity security. This differs from traditional offline settings where users face more physical and temporal constraints. Therefore, it is worth investigating this form of social support based on virtual peer groups which has become increasingly

[12]This can also apply beyond our context, especially in the domain of social media. Today, our online social media status is defined by numbers: number of "friends", "followers", "likes", and "subscribers". Users are often encouraged to create multiple social media accounts, make more connections and friends, attract more followers etc. The face value of the costs of doing so is extremely low if not for free. However, as the numbers rocket, to process all the information is clearly beyond our capacities.

[13]See the empirical evidence in Cunha et al. (2006) and comprehensive reviews on human development in Heckman & Mosso (2014) and on child development in Francesconi & Heckman (2016).

popular.

Second, we contribute to the literature on peer groups. The importance of peer groups has been studied extensively in the economic literature. Adding to the literature mentioned earlier in the introduction, a great deal of studies show that group membership and group identity have a great impact on individual preferences (Goette et al. 2006, Charness et al. 2007, Chen & Li 2009). Angrist & Lavy (1999) and Zhang & Zhu (2011) show that the size of the peer groups also matters. However, to our knowledge, no study has investigated the causal effect of having multiple peer groups (versus single peer group) on individual outcomes given the empirical challenges. Our study fills this gap as the setup allows us to exploit the natural uncertainty of the EDD as an instrument to identify the causal effect of well-defined multiple peer groups on the information exchange in the peer groups.

Third, the evidence for substituting information exchange between peer groups in our setup is consistent with the literature on limited attention across several disciplines. The anthropology literature predicts that the limit of social group size is a hundred and fifty given the size of our brain and anything beyond that would be too complicated to handle at optimal processing levels (Dunbar 1998, Dunbar & Shultz 2007). The management literature on multiple team membership often sees a negative relationship between multiple teams and the time devoted to team tasks in the workplace (O'leary et al. 2011, Cummings & Haas 2012, Pluut et al. 2014). The economic literature further investigates the substitution effects beyond the same domains. For example, a number of studies show that the time spent on social media crowds out a wide range of activities in education, family, labour market, health etc (see a summary in DellaVigna & La Ferrara 2015).

The remainder of the paper is organised as follows, Section 2 describes the observational data, Section 3 illustrates the identification challenges and strategy, Section 4 discusses behavioural predictions, Section 5 presents the results of the empirical analysis, Section 6 describes the survey data and analysis, and finally, Section 7 concludes.

2 Observational data

We collect the observational data set from one of the largest Chinese online maternity and parenting communities. As stated on its website, it aims to establish a platform where positive and meaningful experiences about maternal caring and parenting can be obtained. We first explain the creation and enrollment of the peer groups. Subsequently, we present the descriptive statistics of the peer groups.

2.1 Creation and enrollment of peer groups

The main feature of this online maternity community is that it assigns pregnant users into peer groups based on their estimated due date (EDD). Assigning pregnant users who are at the same stage of the pregnancy to the same peer groups helps them to exchange information more efficiently in three ways. First, it keeps certain time-sensitive information timely. Second, since users in the peer groups are both information givers and recipients, they have a high level of responsibility and trust among themselves to ensure the quality of information. Last but not least, it keeps the number of users in each peer group more or less balanced.

Peer groups are named in the format of *Month Year* – the month and the year of the EDD. As illustrated in Figure 1, for example, users who are expecting to give birth on any day in March 2018 would enroll in the peer group *March 2018*. All the peer groups are created ten months before the delivery month, e.g. the peer group *March 2018* was created in May 2017. Each peer group is active since the creation till the end of the corresponding delivery month.

Figure 1: Creation and enrollment of peer groups

Note: Peer groups are labelled in the format of *Month Year* – the month and the year of the EDD. Users whose EDDs are in the *Month* and *Year* are assigned to the peer group *Month Year* by default. Peer groups are ordered by month along the timeline.

The feature is designed such that each user has only one peer group based on her EDD by default. However, in practice, some users have more than one peer groups. There could be various motives for that: users may want to share information with a greater number of users, access to information from more users, or compare information across different peer groups. Another reason is the natural uncertainty of EDD. In particular, we observe a clear pattern that users whose EDD falling in the beginning (end) of the month tend to additionally join the peer group of the previous (next) month. Among those who joined multiple peer groups, almost 30% of them joined one of the adjacent peer groups, i.e. the peer groups of the previous or the next month. This is in line with the fact that the standard deviation of EDD is about $\pm$ 15 days (Hoffman et al. 2008).

2.2 Descriptive statistics

Three core peer groups are tracked with complete cycle, i.e. from the creation of the peer group to the end of the delivery month. They are *March 2018*, *April 2018*, and *May 2018*. We have information about users' EDD, the enrollment of groups (peer groups and non-peer groups), the number of each user's posts and responses in the default peer group, the additional peer groups and non-peer groups.[14]

Table 1: Descriptive statistics of users in the three core peer groups

VARIABLES	(1) mean	(2) sd	(3) p10	(4) p50	(5) p90	(6) # Obs
Day of Expected Due Date (EDD) (1st-30th/31st)	15.10	8.626	3	15	27	24,705
Peer Group Enrollment						
Default PG_T	1	0	1	1	1	24,705
Previous month PG_{T-1}	1	0	1	1	1	2,692
Next month PG_{T+1}	1	0	1	1	1	1,485
Number of Posts in						
Default PG_T	2.079	1.992	1	1	4	24,705
Previous month PG_{T-1}	0.290	0.825	0	0	1	2,692
Next month PG_{T+1}	0.272	0.919	0	0	1	1,485
Number of Responses in						
Default PG_T	3.679	7.185	0	1	11	24,705
Previous month PG_{T-1}	1.105	2.765	0	0	3	2,692
Next month PG_{T+1}	0.882	2.558	0	0	2	1,485
Non Peer Groups						
Number of non PGs enrolled	9.865	3.371	4	10	13	24,705
Number of Post and Responses in non PGs	2.718	5.261	0	1	8	24,705

Table 1 presents the descriptive statistics of the three core peer groups pooled together. There are 24,705 unique users – who are either *single-default-peer-group* users or *two-peer-group* users additionally joining one of the adjacent peer groups – in these three core peer groups. The distribution of the day of EDD (1st-30th/31st) is approximately a uniform distribution with a mean of 15.1 and a standard deviation of 8.6. We first look at the peer group

[14]Our data collection process allows us to capture users who posted at least one message in her default peer group. We exclude users who are not giving birth for the first time. Additionally, we trim the top 5% of our data in terms of the number of posts and responses to exclude outliers, which is a standard practice with communication data (Turkiewicz 2018). We end up with 24,705 users who are either single-default-peer-group users or two-peer-group users additionally joining one of the adjacent peer groups in the final sample for the analysis.

enrollment. By design, all users have a default peer group, which we denote as PG_T. There are 2,692 users (10.9%) who additionally join the peer group of the month just before their EDD, which we denote as PG_{T-1}; there are 1,485 users (6%) who additionally join the peer group of the month just after their EDD, which we denote as PG_{T+1}.

Next we look at the information exchange – the number of posts and responses – in the peer groups. About 2.08 posts are generated per user in their default peer group, 0.29 posts in the previous month peer group (if joined), and 0.27 posts in the next month peer group (if joined). Users are more active in generating responses than posts. About 3.68 responses are generated per user in their default peer group, 1.1 responses in the previous month peer group (if joined), and 0.88 responses in the next month peer group (if joined).

Finally, besides the peer groups users can also join non-peer groups that are open to all users in the online community.[15] On average, users have 9.9 non-peer groups and generate about 3 posts and response in total in non-peer groups.

3 Identification

The empirical challenge lies in the endogenous choice of whether joining an additional peer group. In principle, users can manually join any other peer groups that are open in the community.[16] Therefore, comparing users who have two peer groups with users who have only one peer group is subject to the self-selection bias. Users who are more active in exchanging information might be more likely to join an additional peer group. If this is the case, the naive comparison will simply display a positive correlation between having additional peer groups and users' overall activeness.

We first demonstrate our identification strategy and subsequently, present the placebo checks.

3.1 Identification strategy

To overcome the identification challenge, we use the day of EDD, from 1st to 30th/31st, as an instrument for the number of peer groups. It is a valid instrument as it satisfies the following two requirements: (i) it directly affects a user's decision of joining an additional peer group, and (ii) it is quasi-random and thus uncorrelated with her characteristics that could affect her activities

[15]The non-peer groups have particular themes such as "Marriage","Sentiment", "Cooking" etc. They are not in the format of *Month Year* and the enrollment is completely unrelated to the EDD. Therefore, users in such groups are much more heterogeneous in their maternal status. In our sample, on average a user spends 74.3% of her activities in the default peer group, indicating that the default peer group is the principal group for the users in this online community.

[16]The range of such peer groups featured on the homepage is one year before and ten months after the current month. However, this asymmetry does not affect our identification.

(unless through the enrollment of an additional peer group). Therefore, our identification assumption is that the day of EDD has an effect on users' information exchange only through its effects on the enrollment in an additional peer group.[17]

Figure 2 shows the fraction of the users joining an additional peer group of the previous month and next month for each day of the month. The day of EDD directly affects the propensity of joining an adjacent peer group. The fraction of the users additionally joining the previous-month peer group is above 30% for users whose EDD is on the 1st day of the month, and it gradually decreases to about 5% for the users whose EDD is on the last day of the month. Contrarily, the fraction of the users additionally joining the next-month peer group is barely 1% for the users whose EDD is on the 1st day of the month, and it gradually increases to above 25% for users whose EDD is on the last day of the month. Therefore, the day of EDD satisfies the first requirement of a valid instrument.

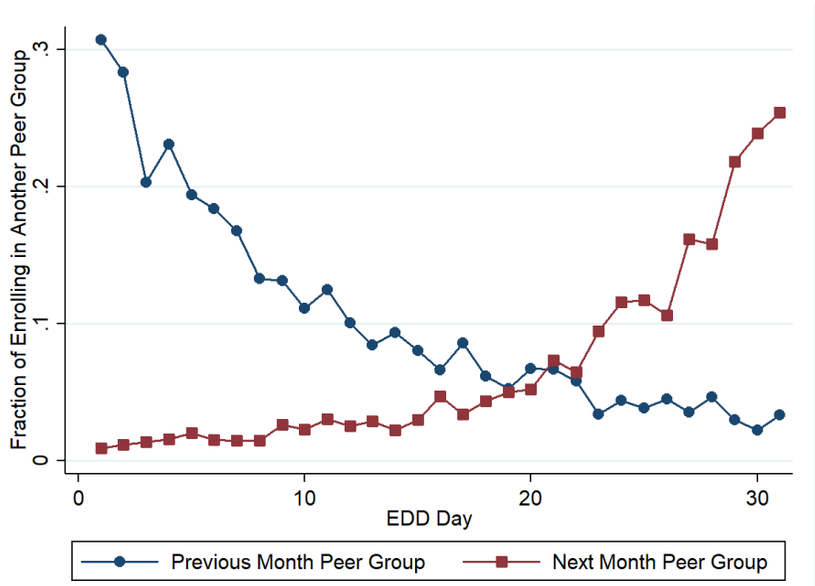

Figure 2: Fractions of users joining the adjacent peer groups, by EDD day

The day of EDD is a quasi-random variable as it is uncertain due to both natural and measurement errors. First of all, while it is feasible to plan ex ante for the year and the month to give birth, it is almost impossible to pin down the exact day of the delivery (with the exception of predetermined cesarean cut). Second, even after the EDD is determined by the doctor, the standard deviation is about $\pm$ 15 days. Thus, one can think of the day of EDD as a natural lottery ticket: each pregnant user quasi-randomly draws a number which ranges from 1 to 30/31. Therefore, the day of EDD satisfies the

[17]Notice that using the instrumental variable estimator we are identifying the local average treatment effect (LATE), which is the average treatment effect (ATE) among the compliers (Imbens & Angrist 1994). Namely, those who draw a lottery ticket number from 1 to 15/16 join the previous-month peer group and those who draw a lottery ticket number from 17 to 30/31 join the next-month peer group.

second requirement – the exclusion restriction – of a valid instrument.[18]

3.2 Placebo checks

A standard balance check would be checking some predetermined characteristics by each day of the month. However, we do not observe other demographic characteristics apart from that the users are women at fertility age. Nevertheless, we can do a placebo test using two placebo outcomes, the number of non peer groups joined and the amount of information exchange in non-peer groups. Unlike peer groups that are exclusively designed for pregnant users, non-peer groups are open to all users (not yet pregnant, pregnant, and mothers of young kids) in this community. These two outcomes can be seen as indicators of users general activeness. However, the enrollment of non-peer groups is completely unrelated to the EDD. Therefore, we should not observe any correlation between the day of EDD and the enrollment of non-peer groups or the activities in those groups. This is confirmed by the null results in Table 2.

Table 2: The Effects of the day of EDD on placebo outcomes

	(1)	(2)
VARIABLES	Number of NPGs	log(information) in NPGs
Day of EDD (1st-30th/31st)	-0.00191	0.000772
	(0.00248)	(0.000670)
Constant	9.938***	0.785***
	(0.0526)	(0.0142)
Peer group/month fixed effects	Yes	Yes
Observations	24,705	24,705
R-squared	0.003	0.001

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

4 Behavioural predictions

In this section we make behavioural predictions based on two sets of theories. On the one hand, the limited attention literature (e.g. Kahneman 1973, Camerer 2005, Falkinger 2008) emphasises that human beings have limited attention in processing information. Having an additional peer group increases the information load and switching costs and decreases attention to new information and time to encode information (O'leary et al. 2011). For example, if

[18]We do not find evidence that users manipulate their EDD in a systematic way that could threaten the identification, i.e. users who want to join an additional peer group report their day of their EDD either at the beginning of the month or end of the month. The day of EDD follows approximately a uniform distribution as shown in Figure A1 in the Appendix.

a user has already acquired desired information in the other peer group, there is no need to seek for the same information in the default peer group. Even users switch between peer groups randomly, it would cause the reduction in the information exchange in either peer group. Consequently, users in multiple peer groups will reduce the time and effort in generating information in each of the peer groups.[19] We refer to this as substitution effects. On the other hand, the social network literature (e.g. Jackson 2010, Bakshy et al. 2012)) predicts that the exposure to a wider social network promotes information diffusion and knowledge spillovers. Therefore, users having multiple peer groups may generate more information exchange due to the information spillovers between peer groups. We refer to this as spillover effects.

Table 3 summarises our behavioural predictions for the effects of joining two peer groups, taking the single default peer group as baseline. Substitution effects suggest that the volume of information in the default group would decrease but the sum of them in the default peer group and in the additional peer group should be the same as the one from those who only join the default peer group, under the assumption of a linear production function. The spillover effects suggest that the volume of information would increase in the default group and as a result, the sum of them in both peer groups will be higher than the one from those who only join the default peer group. The net effect of these two effects is an empirical question that we investigate in Section 5.

Table 3: Behavioural predictions for the effects of two PGs

Information in	Substitution Effects	Spillover Effects
Default Peer Group	$\downarrow$	$\uparrow$
Both Peer Groups	$\equiv$	$\uparrow$

5 Econometric analysis

In this section we empirically analyse the effects of having two peer groups, i.e. joining an additional peer group besides the default peer group, on the information exchange in i) the default peer group, and ii) both peer groups. The baseline compared with is the amount of information exchange generated by the users who only has the default peer group.[20] We then generalise the analysis to multiple peer groups. The results are robust and qualitatively similar (see Table A4 in the Appendix).

We estimate the econometric model in Equation (1). The outcome variable Y_i is the log-transformed information exchange generated by user i.[21] The

[19] Although users can extend their time spent on the website to some extent, we assume that the total time is fixed as in the literature of multiple team memberships.

[20] We also investigate the potential heterogeneity effects of the additional peer groups of the previous month and the next month. The results are in Section A.2 in the Appendix.

[21] As a further robustness check, we also estimate in a Poisson IV model and the results

Table 4: Two peer groups and information exchange

	(1)	(2)	(3)	(4)	(5)
	1^{st} stage	2^{nd} stages			
VARIABLES	Two PGs	log(info) in Default PG OLS	log(info) in Default PG IV	log(info) in Total PGs OLS	log(info) in Total PGs IV
$\widehat{\text{Two PGs}}$		0.232***	-0.719***	0.391***	-0.410***
		(0.0146)	(0.104)	(0.0147)	(0.102)
|Day of EDD-16|	0.0124***				
	(0.000554)				
Constant	0.0969***	1.474***	1.653***	1.471***	1.621***
	(0.00554)	(0.00946)	(0.0219)	(0.00946)	(0.0215)
Peer group fixed effects	Yes	Yes	Yes	Yes	Yes
F-tests of instrument	-	-	504.19	-	504.19
Observations	24,705	24,705	24,705	24,705	24,705

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

indicator $\mathbb{1}[\text{Two PGs}_i]$ is one if user i has two peer groups instead of one peer group by default.[22] We also include fixed effects of the default peer group – equivalent to month fixed effects – which are captured by the term α_{pg}.

$$\text{Y}_i = \gamma_1 \mathbb{1}[\text{Two PGs}_i] + \alpha_{pg} + \epsilon_i \tag{1}$$

Column (1) in Table 4 reports the first stage regressions in a linear probability model. We use |Day of EDD-16| to predict $\widehat{\text{Two PGs}}$, as the farther the day of EDD is away from the middle of the month, the more likely that a user joins in an additional (adjacent) peer group. The coefficient is positive and significantly different from zero at 1% level. A one-day deviation from the middle of the month increases the probability of having two peer groups by 1.24%.

Columns (2)-(5) in Table 4 show the estimated effects of having two peer groups on users' information exchange. Columns (2)-(3) show the information exchange in their default peer group, and Columns (4)-(5) in total peer groups (=1 if having default PG only; =2 if having an additional PG). We estimate each effect in OLS and IV regressions, respectively. The coefficients of having two peer groups are significantly positive for both default peer group and total peer groups in the OLS specification. In Columns (2) and (4), having two peer groups is associated with 26% more information in the default peer group and 48% more information in total peer groups. However, once we instrument the additional peer group, the coefficients become significantly negative (p< 0.01). In Columns (3) and (5), having two peer groups leads to about 51% less

are qualitatively unchanged as shown in Table A2 in the Appendix.

[22]The result is robust when adding controls for the number of non peer groups (NPGs) joined by user i and the log-transformed information exchange generated by user i in the non peer groups (see Table A1).

information in the default peer group and 33% less information in total peer groups.[23] Note that the null hypothesis of a weak instrument has been rejected with the F-test in Columns (3) and (5), which is consistent with the first stage results in Column (1). The direction of the biases in the OLS estimates for the variable of Two PGs is upward, which can be explained by self-selection, i.e. the unobserved nature of activeness of users is positively correlated with both information exchange and the tendency of joining an additional peer group. Therefore, after the IV correction, the causal effect of joining an additional peer group is an decrease in users' information exchange both in the default peer group as well as in the total peer groups.

6 Survey analysis

We now turn to the perceived social support in the peer groups and investigate how it relates to users' mental health and newborn's birth outcomes. To do so, we conduct a survey. We first describe the design of the survey and present the descriptive statistics of the survey data. Subsequently, we explain how perceived social support and mental health are measured in the questionnaire. Finally, we perform the empirical analysis. The full survey questionnaire can be found in Section A.3 in the Appendix.

6.1 The survey

The survey was conducted in 2020, two years after the time period of the observational data. It includes 500 participants who were or have been users of the peer groups in this online community at the time of the survey. Notice that the observational data set and the survey data set are not linked. Since they are two years apart from each other, linking the survey users to the observational users would not allow us to measure the well-being during pregnancy as the same users we observe were no longer pregnant at the time of the survey.

Among the 500 participants, 300 were pregnant and 200 gave birth within one year at the time of the survey.[24] This allows us to measure users' well-being both during and after pregnancy, and newborns' birth outcomes of users who gave birth within one year.

In the questionnaire we measure, apart from individual characteristics, i) users' perceived social support in the peer groups using a modified *Multidimensional Scale of Perceived Social Support* (MSPSS) test (Zimet et al. 1988), ii) their mental health using the *Edinburgh Depression Scale* (EDS) (Murray & Cox 1990), and iii) the newborn's birth outcomes (weight, whether planned natural birth, and whether breastfeeding) among those who gave birth within one year. We did not ask about users' activities in the peer groups or non peer groups for

[23] Since the dependent variable is log-transformed, the estimated effects $= (e^{\beta} - 1) \times 100\%$.

[24] We use a filtering question at the beginning of the survey to separate these two type of participants (See the full survey questionnaire in Section A.3 in the Appendix.)

two main reasons. First, the full cycle of the peer group is ten months. While we collected the observational data right at the end of the cycle, we surveyed the users either who are still in the cycle or left the peer group a while ago. Second, due to incomplete memory, it is unlikely that users precisely recall how many pieces of posts and responses they have generated over a period of several months.

Table 5 reports the descriptive statistics of the survey data. Column (1) is the full sample with 500 users, Column (2) is the sample with 300 pregnant users, and Column (3) is the sample with 200 users who gave birth within one year. The average day of EDD is around 16. On a scale from 1 to 7 for the statement "This community is my main source of knowledge and information about pregnancy.", users rate an average of 6.1 in both samples. Among the pregnant users, 72% of them report that they joined at least one more peer group additional to their default peer group. Among the users who gave birth within one year, 74.5% of them recall that they joined at least one or more peer groups in addition to their default peer group when they were pregnant. Among those who joined multiple peer groups, we additionally ask them about the motives of joining additional peer groups. Over 63% of the pregnant users and about 60% of the users who gave birth within one year reported that the uncertainty of the EDD is one of the reasons to join additional peer groups. About 63% of the users report that they would like to share information with more users, 94% of the users report that would like to access more information, and less than 5% of the users report other reasons. The fractions do not differ in the two samples. The majority of the users have or had 1-3 non peer groups, and about one third have or had 4-10 non peer groups.The average age of users is slightly above 28 year old in both samples. Pregnant users are, on average, in their 6th month of the pregnancy. Finally, we also include three health indicators: a binary variable of whether smoke or drink, and two categorical variables for hours of sleep and hours of physical exercise. Only 0.6% of the users smoke or drink during pregnancy and none after childbirth. Most users have adequate sleep and moderate exercise both during pregnancy and after childbirth.

It is worth noting that a much higher fraction (73.0%) of users in the survey data set report that they have or had multiple peer groups during pregnancy, comparing to 17.6% in the observational data set. This may be due to the fact that the survey participants are on average more engaged in online communities (e.g. joining additional online groups and participating in online surveys) than those in the observational data.

6.2 Measuring perceived social support in peer groups

There has been a surge in the evaluation of social support as a well-being indicator. The Multidimensional Perceived Social Support Scale (MSPSS)(Zimet et al. 1988) has evolved as one of the most extensively translated and validated

Table 5: Survey descriptive statistics

VARIABLES	(1) Full sample	(2) Pregnant	(3) After childbirth
Day of EDD (1st-30th/31st)	16.418	16.747	15.925
	(7.443)	(7.313)	(7.627)
Main source of info†	6.128	6.127	6.130
	(0.785)	(0.774)	(0.804)
Multiple PGs§	0.730	0.720	0.745
	(0.444)	(0.450)	(0.437)
Motive of joining multiple PGs⋆			
- due to EDD uncertainty	0.616	0.634	0.591
- to share info	0.627	0.625	0.631
- to access info	0.940	0.940	0.940
- other reasons	0.025	0.014	0.040
Number of NPGs			
- 0 NPG	0.138	0.153	0.115
- 1-3 NPGs	0.568	0.553	0.590
- 4-10 NPGs	0.290	0.287	0.295
- 10+ NPGs	0.004	0.007	0.000
Age	28.244	28.107	28.450
	(2.080)	(2.006)	(2.175)
Xth-Month since pregnancy‡	10.570	6.647	16.455
	(5.241)	(1.539)	(2.703)
Health indicators††			
-Whether smoke or drink	0.004	0.006	0
	(0.063)	(0.082)	(0)
-Hours of sleep (categories 1-4)	3.498	3.600	3.345
	(0.668)	(0.617)	(0.713)
-Physical exercise (categories 1-4)	1.940	2.013	1.830
	(0.643)	(0.649)	(0.619)
Observations	500	300	200

Notes: Means reported with standard deviations in parentheses.
† This a scale (1-7) of using this platform as the main source of pregnancy-related information.
§ Multiple PGs equals 1 if the user joins more than one peer groups.
⋆ sample conditional on having multiple peer groups.
‡ For all users (pregnant or after childbirth), it indicates the number of months since they were pregnant.
†† Hours of sleep is a categorical variable ranging from below 5, 5-6, 6-7, to above 7 hours; Physical exercise is a categorical variable ranging from below 30, 30-60, 60-120, to above 120 minutes.

measures for social support outcome. The original test was designed to assess the perceptions of social support from three specific sources: family, friends, and significant other.

We selected five out of the twelve items and made minimal modification to accommodate the peer groups in this online community exclusively. For example, we made the following modifications: from "*My family* is willing to help me make decisions."(original version) to "*People in the peer group(s)* are willing to help me make decisions."; from "I can talk about my problems with *my friends*." (original version) to "I can talk about my problems *in the peer group(s)*."; from "There is a *special person* who is around when I am in need." (original version) to "There are some *people in the peer group(s)* who are around when I am in need.".

Besides the five modified items, we also add one closely related item "I have learned useful information from others in the peer group(s)." Therefore, the perceived social support we measure using the modified MSPSS items exclusively refer to the *peer groups* in our context. The amount of social support is rated on a seven-point Likert scale, with responses ranging from very strongly disagree (=1) to very strongly agree (=7). The higher the score, the greater the amount of available social support. There are six items in our survey and the total score is 42. The survey results show that online social support is well received by the participants. The average MSPSS score for users during pregnancy (after childbirth) in our survey is 35.86 (36.07) with a standard deviation of 2.70 (3.53). The complete set of items in the modified MSPSS and the descriptive statistics can be found in Section A.3.1 in the Appendix.

6.3 Measuring mental health

We use the *Edinburgh Depression Scale* (EDS) (Murray and Cox 1990) to measure users' mental health. Initially, this 10-item self-report measure was developed for postnatal period (Cox et al. 1987) and was called *Edinburgh Postnatal Depression Scale*. Later, Murray & Cox (1990) validated it also for prenatal period and renamed to *Edinburgh Depression Scale*. Since then, it is used to screen for symptoms of emotional distress both during pregnancy and in the postnatal period.[25] The test is referred to the feeling in the past seven days. Therefore, in our survey, it screens for *prenatal depression* in the sample of users during pregnancy and *postnatal depression* in the sample of users after childbirth, respectively. Each item has a range of score from 0-3 and the total score is 30. The average EDS score for users during pregnancy (after childbirth) in our survey is 6.69 (5.78) with a standard deviation of 4.26 (3.74).[26] The complete items of the EDS and descriptive statistics in our

[25]The EDS was translated and validated in a wide range of languages (Government of Western Australia 2006). We use the Chinese version that was validated by Lee et al. (1998) in our survey.

[26]A score between 10-12 depression indicates presence of symptoms of distress that may be discomforting; A score above 12 requires further evaluation and possible referral to a

survey sample can be found in Section A.3.2 in the Appendix.

6.4 Empirical analysis of the survey data

In this section we present three sets of empirical analysis of the survey data. We link perceived social support to i) whether having multiple peer groups during pregnancy, ii) mental health of both pregnant users and users after childbirth, and iii) newborns' birth outcomes in the sample of users after childbirth. Notice that the perceived social support measured here is exclusively from the peer groups during pregnancy.

6.4.1 Multiple peer groups and perceived social support

We first link multiple peer groups to perceived social support from the peer groups measured by the MSPSS score.[27]

Table 6 reports the results. Column (1) is the full sample, Columns (2)-(3) are the sample of users during pregnant with and without the month-of-pregnancy fixed effects, and Column (4) is the sample of users after childbirth. The coefficients of having multiple peer groups are statistically insignificant and the sizes are negligible (average MSPSS score is about 36) in all the samples. Thus, the perceived social support in the peer groups does not seem to be affected by whether having single or multiple peer groups.

Table 6: Perceived social support measured by MSPSS

	(1)	(2)	(3)	(4)
VARIABLES	Full sample	Pregnant	Pregnant	After childbirth
Multiple PGs	0.586	0.538	0.466	0.647
	(0.387)	(0.395)	(0.411)	(0.793)
Constant	35.52***	35.48***	32.33***	35.59***
	(0.364)	(0.359)	(1.205)	(0.765)
X-th month since pregnancy FE	No	No	Yes	No
Observations	500	300	300	200
R-squared	0.007	0.008	0.042	0.006

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

6.4.2 Perceived social support and mental health

Next we link the perceived social support from the peer groups to mental health as in Equation (2). As explained in Section 6.3, the EDS test screens

perinatal mental health specialist.

[27]In principle, one could use the day of EDD as IV to predict multiple peer groups as in the analysis of the observational data. However, we could not detect any correlation between the day of EDD and multiple peer groups in the survey data. Therefore, the IV estimator (F-stats for IV <1) would not be meaningful. One plausible reason is that the survey participants seem to have clear motives to join multiple peer groups, and therefore, the EDD is no longer a good predictor for it.

for *prenatal depression* in the sample of users during pregnancy and *postnatal depression* in the sample of users after childbirth, respectively. In other words, there are two versions of Equation (2): the first version associates the contemporaneous role of perceived social support from the peer groups with prenatal mental health when the peer groups are still active (i.e. during pregnancy); the second version associates its lasting role with postnatal mental health when the peer groups are no longer active (i.e. after childbirth).

The main explanatory variable is perceived social support from the peer groups during pregnancy, measured by the MSPSS score. We control for whether having multiple peer groups, three health indicators (whether smoke or drink alcohol, hours of sleep, and physical exercise), age, as well as fixed effects of the X-th month since pregnancy in one of the specification.

$$\text{EDS}_i = \beta_1 \text{MSPSS}_i + \beta_2 \mathbb{1}[\text{Multiple PGs}_i] + \beta_3' \text{Health indicators}_i + \beta_4 \text{Age}_i + \alpha_m + \nu_i \quad (2)$$

Table 7 reports the results. Column (1) is the full sample, Columns (2)-(3) are the sample of users during pregnancy without and with fixed effects of the X-month since pregnancy , and Column (4) is the sample of users after childbirth. In the sample of users during pregnancy, a one-point increase in the MSPSS score is associated with a 0.21-point decrease in the EDS score. That is, higher perceived social support from the peer groups is associated with lower prenatal depression. This relationship is significant at 5% level without the fixed effect of the X-th month since pregnancy and at 10% level with the fixed effect. In the sample of users after childbirth, a one-point increase in the MSPSS score is associated with a 0.26-point decrease in the EDS score, and it is significant at 1% level. That is, higher perceived social support from the peer groups during pregnancy is also associated with lower postnatal depression.

It is important to control for multiple peer groups in this analysis. In the sample of users during pregnancy, higher prenatal depression is associated with having multiple peer groups ($p<0.05$). One possible explanation for this positive correlation is that users who have potentially more concerns may join multiple peer groups to seek more social support. However, ultimately, it is the actual support received from the peer groups rather than the number of peer groups that helps to relieve the concerns. While one can easily join additional peer groups, whether one successfully receives additional support from additional peer groups is not straightforward. This can be seen in Table 6, where perceived social support is not correlated with having multiple peer groups. In the sample of users after childbirth, postnatal depression is no longer uncorrelated with having multiple peer groups as the groups no longer exist. Therefore, the effects have to come from the support in the peer groups rather than having the peer groups *per se*.

Finally, users who smoke or drink alcohol score over 6 points higher in the prenatal depression, and users who have more hours of sleep have lower scores in both prenatal and postnatal depression.[28] Physical exercise and age does not seem to matter for either depression.

Table 7: Mental health measured by EDS

	(1)	(2)	(3)	(4)
VARIABLES	Full sample	Pregnant	Pregnant	After childbirth
MSPSS	-0.258***	-0.213**	-0.217*	-0.258***
	(0.0656)	(0.105)	(0.111)	(0.0838)
Multiple PGs	0.895**	1.320***	1.359***	-0.0328
	(0.399)	(0.505)	(0.524)	(0.631)
Whether smoke or drink	7.039***	6.919***	6.210***	-
	(0.789)	(0.957)	(1.375)	-
Hours of sleep†	-1.131***	-1.460***	-1.509***	-1.122***
	(0.263)	(0.374)	(0.376)	(0.356)
Physical exercise‡	0.247	0.542	0.607	-0.682
	(0.288)	(0.367)	(0.369)	(0.464)
Age	-0.0266	-0.119	-0.0729	0.193
	(0.0952)	(0.135)	(0.130)	(0.133)
Constant	19.16***	20.86***	22.44***	14.62**
	(3.992)	(5.419)	(5.451)	(5.884)
X-th month since pregnancy FE	No	No	Yes	No
Observations	500	300	300	200
R-squared	0.097	0.122	0.152	0.127

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
† Hours of sleep is a categorical variable ranging from below 5, 5-6, 6-7, to above 7 hours.
‡ Physical exercise is a categorical variable ranging from below 30, 30-60, 60-120, to above 120 minutes.

6.4.3 Perceived social support and newborn's birth outcomes

Finally, we regress newborns' birth outcomes on perceived social support from the peer groups during pregnancy, using the sample of users after childbirth as in Equation (3).

$$\begin{aligned}\text{Newborns' birth outcomes}_i = \delta_1 \text{MSPSS}_i + \delta_2 \mathbb{1}[\text{Multiple PGs}_i] \\ + \delta_3' \text{Health indicators}_i + \delta_4 \text{Age}_i + \alpha + u_i \quad (3)\end{aligned}$$

Newborns' birth outcomes include three variables: birth weight, whether being breastfed, and whether planned natural birth. We select these three variables according to the WHO technical consultation on newborn health indicators (WHO 2015). The variable of interest is perceived social support from peer

[28]The variable "whether smoke or drink" is omitted in Column (4) as no users report smoking or drinking.

groups during pregnancy (MSPSS). We also control for whether having multiple peer groups, two health indicators of the mothers (hours of sleep and physical exercise), and age.[29]

Table 8 reports the results. Babies born to mothers with higher perceived social support during pregnancy are heavier. A one-point increase in the MSPSS score is associated with a 15.31-gram increase in the newborn's weight ($p<0.01$). Mothers doing more physical exercises have lighter newborns.[30] Perceived social support from the peer groups during pregnancy and other control variables do not seem to affect the decisions of breastfeeding and natural birth.

Table 8: Newborn's birth outcomes

	(1)	(2)	(3)
VARIABLES	Newborn's weight (gram)	Breastfeeding	Planned natural birth
MSPSS	15.31***	-0.0401	0.0345
	(4.678)	(0.0342)	(0.0294)
Multiple PGs	-72.74	-0.0599	0.237
	(66.84)	(0.227)	(0.267)
Hours of sleep†	-7.178	0.150	0.171
	(37.51)	(0.138)	(0.155)
Physical exercise‡	-147.1***	-0.284*	0.198
	(41.21)	(0.163)	(0.193)
Age	-12.24	0.0233	0.0493
	(11.27)	(0.0499)	(0.0592)
Constant	3,479***	1.607	-2.542
	(401.6)	(2.074)	(2.118)
Observations	200	200	200
R-squared	0.089	0.028	0.041

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
† Hours of sleep is a categorical variable ranging from below 5, 5-6, 6-7, to above 7 hours.
‡ Physical exercise is a categorical variable ranging from below 30, 30-60, 60-120, to above 120 minutes.

7 Conclusion

Social support is one of the key social determinants of health as emphasised by the WHO. In this paper, we investigate both observable social support in terms of information exchange and perceived social support in online peer groups triggered by a significant life event: childbearing. We first identify the causal effects of having multiple peer groups (vs single peer group) – using an IV approach – on the information exchange in the peer groups. We then

[29]The variable of whether smoke or drink alcohol is again omitted as all the users after childbirth answered "no" to this question.

[30]All the newborns' weight in our sample is within the normal range of birth weight reported in previous studies (Janssen et al. 2007, Dai et al. 2014).

design a survey questionnaire to measure users' perceived social support in the peer groups, their mental health, and newborns' birth outcomes. We find that while having multiple peer groups reduces users' information exchange, it does not seem to undermine their perceived social support in the peer groups. Higher perceived social support from the peer groups is positively associated with better mental health and newborns' birth weight.

Our study contributes to several strands of literature, and in particular, the literature on prenatal environments and the well-being of mothers as well as child development. Despite that the peer groups studied here are only active during the pregnancy, the social support received from the peer groups are positively associated with both prenatal and postnatal well-being of the mothers. This is consistent with what Doyle et al. (2009) point out that the fetal environment and maternal behaviour during pregnancy have significant long-term consequences. Unlike social support from third-party interventions, peer group is a self-contained source of social support. Users are not only the recipients but also the providers of the support. This format involves minimal costs and is proven to be highly effective in our study.

Finally, this paper sets out a research agenda to further examine the relations between social support and well-being in online settings. Traditional events and activities of social support have been increasingly transformed into digital versions. Whereas both virtual and real communities could function as platforms for social support, they differ in important ways. For example, online social communities enable people to overcome physical and time constraints embedded within their offline counterparts. Moreover, the option of maintaining anonymity in online settings provides users more convenience when discussing sensitive and controversial topics. However, while obtaining a membership of online communities is often of minimal costs, users' time and attention are limited. The effectiveness of social support will therefore depend on the actual design of online communities. The peer groups studied here is one of the examples. In order to maintain both the quantity and quality of social support in the peer groups and maximise its effects on individual well-being, it is important to further identify the boundary of individual attention and investigate the driving forces behind the perceived social support in online settings.

References

Aizer, A., Stroud, L. & Buka, S. (2016), 'Maternal stress and child outcomes: Evidence from siblings', *Journal of Human Resources* **51**(3), 523–555.

Angrist, J. D. & Lavy, V. (1999), 'Using maimonides' rule to estimate the effect of class size on scholastic achievement', *Quarterly Journal of Economics* **114**(2), 533–575.

Bakshy, E., Rosenn, I., Marlow, C. & Adamic, L. (2012), The role of social networks in information diffusion, *in* 'Proceedings of the 21st international conference on World Wide Web', ACM, pp. 519–528.

Bandiera, O., Barankay, I. & Rasul, I. (2009), 'Social connections and incentives in the workplace: Evidence from personnel data', *Econometrica* **77**(4), 1047–1094.

Brown, K. M. & Laschever, R. A. (2012), 'When they're sixty-four: Peer effects and the timing of retirement', *American Economic Journal: Applied Economics* **4**(3), 90–115.

Bruhin, A., Goette, L., Haenni, S. & Jiang, L. (2020), 'Spillovers of prosocial motivation: Evidence from an intervention study on blood donors', *Journal of Health Economics* **70**, 102244.

Camerer, C. F. (2005), Behavioral economics. California Institute of Technology.

Carlson, K. (2015), 'Fear itself: The effects of distressing economic news on birth outcomes', *Journal of Health Economics* **41**, 117–132.

Carrell, S. E., Fullerton, R. L. & West, J. E. (2009), 'Does your cohort matter? measuring peer effects in college achievement', *Journal of Labor Economics* **27**(3), 439–464.

Charness, G., Cobo-Reyes, R. & Jiménez, N. (2014), 'Identities, selection, and contributions in a public-goods game', *Games and Economic Behavior* **87**, 322–338.

Charness, G., Rigotti, L. & Rustichini, A. (2007), 'Individual behavior and group membership', *American Economic Review* **97**(4), 1340–1352.

Chen, Y. & Li, S. X. (2009), 'Group identity and social preferences', *American Economic Review* **99**(1), 431–57.

Cohen, S. & Syme, S. L. (1985), 'Issues in the study and application of social support', *Social Support and Health* **3**, 3–22.

Collins, N. L., Dunkel-Schetter, C., Lobel, M. & Scrimshaw, S. C. (1993), 'Social support in pregnancy: Psychosocial correlates of birth outcomes and postpartum depression.', *Journal of Personality and Social Psychology* **65**(6), 1243.

Cornelissen, T., Dustmann, C. & Schönberg, U. (2017), 'Peer effects in the workplace', *American Economic Review* **107**(2), 425–56.

Cox, J., Holden, J. & Sagovsky, R. (1987), 'Edinburgh postnatal depression scale (epds)', *Br J psychiatry* **150**, 782–786.

Cummings, J. N. & Haas, M. R. (2012), 'So many teams, so little time: Time allocation matters in geographically dispersed teams', *Journal of Organizational Behavior* **33**(3), 316–341.

Cunha, F., Heckman, J. J., Lochner, L. & Masterov, D. V. (2006), 'Interpreting the evidence on life cycle skill formation', *Handbook of the Economics of Education* **1**, 697–812.

Dai, L., Deng, C., Li, Y., Zhu, J., Mu, Y., Deng, Y., Mao, M., Wang, Y., Li, Q., Ma, S. et al. (2014), 'Birth weight reference percentiles for chinese', *PloS one* **9**(8), e104779.

DellaVigna, S. & La Ferrara, E. (2015), Economic and social impacts of the media, *in* 'Handbook of media economics', Vol. 1, Elsevier, pp. 723–768.

Doyle, O., Fitzpatrick, N., Lovett, J. & Rawdon, C. (2015), 'Early intervention and child physical health: Evidence from a dublin-based randomized controlled trial', *Economics & Human Biology* **19**, 224–245.

Doyle, O., Harmon, C. P., Heckman, J. J. & Tremblay, R. E. (2009), 'Investing in early human development: timing and economic efficiency', *Economics & Human Biology* **7**(1), 1–6.

Duflo, E. & Saez, E. (2002), 'Participation and investment decisions in a retirement plan: The influence of colleagues' choices', *Journal of Public Economics* **85**(1), 121–148.

Dunbar, R. I. (1998), 'The social brain hypothesis', *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews* **6**(5), 178–190.

Dunbar, R. I. & Shultz, S. (2007), 'Evolution in the social brain', *Science* **317**(5843), 1344–1347.

Elsenbruch, S., Benson, S., Rücke, M., Rose, M., Dudenhausen, J., Pincus-Knackstedt, M. K., Klapp, B. F. & Arck, P. C. (2006), 'Social support during pregnancy: effects on maternal depressive symptoms, smoking and pregnancy outcome', *Human Reproduction* **22**(3), 869–877.

Falkinger, J. (2008), 'Limited attention as a scarce resource in information-rich economies', *Economic Journal* **118**(532), 1596–1620.

Francesconi, M. & Heckman, J. J. (2016), 'Child development and parental investment: Introduction', *Economic Journal* **126**(596), F1–F27.

Gächter, S., Nosenzo, D. & Sefton, M. (2013), 'Peer effects in pro-social behav-

ior: Social norms or social preferences?', *Journal of the European Economic Association* **11**(3), 548–573.

Goette, L., Huffman, D. & Meier, S. (2006), 'The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups', *American Economic Review* **96**(2), 212–216.

Government of Western Australia, D. o. H. (2006), 'Edinburgh postnatal depression scale (epds): Translated versions – validated', *State Perinatal Mental Health Reference Group* .

Heckman, J. J. & Mosso, S. (2014), 'The economics of human development and social mobility', *Annu. Rev. Econ.* **6**(1), 689–733.

Hoffman, C. S., Messer, L. C., Mendola, P., Savitz, D. A., Herring, A. H. & Hartmann, K. E. (2008), 'Comparison of gestational age at birth based on last menstrual period and ultrasound during the first trimester', *Paediatric and Perinatal Epidemiology* **22**(6), 587–596.

Imbens, G. W. & Angrist, J. D. (1994), 'Identification and estimation of local average treatment effects', *Econometrica* **62**(2), 467–475.

Jackson, M. O. (2010), *Social and economic networks*, Princeton university press.

Janssen, P. A., Thiessen, P., Klein, M. C., Whitfield, M. F., MacNab, Y. C. & Cullis-Kuhl, S. C. (2007), 'Standards for the measurement of birth weight, length and head circumference at term in neonates of european, chinese and south asian ancestry', *Open Medicine* **1**(2), e74.

Jiang, L. (2020), 'Splash with a teammate: Peer effects in high-stakes tournaments', *Journal of Economic Behavior & Organization* **171**, 165–188.

Kahneman, D. (1973), *Attention and Effort*, Vol. 1063, Citeseer.

Kail, R. V. & Cavanaugh, J. C. (2018), *Human development: A life-span view*, Cengage Learning.

Lee, D. T., Yip, S., Chiu, H. F., Leung, T. Y., Chan, K. P., Chau, I. O., Leung, H. C. & Chung, T. K. (1998), 'Detecting postnatal depression in chinese women', *The British Journal of Psychiatry* **172**(5), 433–437.

Manski, C. F. (1993), 'Identification of endogenous social effects: The reflection problem', *Review of Economics Studies* **60**(3), 531–542.

Mas, A. & Moretti, E. (2009), 'Peers at work', *American Economic Review* **99**(1), 112–45.

Murray, D. & Cox, J. L. (1990), 'Screening for depression during pregnancy with the edinburgh depression scale (edds)', *Journal of Reproductive and Infant Psychology* **8**(2), 99–107.

Olds, D. L., Kitzman, H., Anson, E., Smith, J. A., Knudtson, M. D., Miller, T., Cole, R., Hopfer, C. & Conti, G. (2019), 'Prenatal and infancy nurse home visiting effects on mothers: 18-year follow-up of a randomized trial', *Pediatrics* **144**(6).

O'leary, M. B., Mortensen, M. & Woolley, A. W. (2011), 'Multiple team membership: A theoretical model of its effects on productivity and learning for individuals and teams', *Academy of Management Review* **36**(3), 461–478.

Persson, P. & Rossin-Slater, M. (2018), 'Family ruptures, stress, and the mental health of the next generation', *American Economic Review* **108**(4-5), 1214–52.

Pluut, H., Flestea, A. M. & Curşeu, P. L. (2014), 'Multiple team membership: A demand or resource for employees?', *Group Dynamics: Theory, Research, and Practice* **18**(4), 333.

Robling, M., Bekkers, M.-J., Bell, K., Butler, C. C., Cannings-John, R., Channon, S., Martin, B. C., Gregory, J. W., Hood, K., Kemp, A. et al. (2016), 'Effectiveness of a nurse-led intensive home-visitation programme for first-time teenage mothers (building blocks): a pragmatic randomised controlled trial', *The Lancet* **387**(10014), 146–155.

Sacerdote, B. (2001), 'Peer effects with random assignment: Results for dartmouth roommates', *Quarterly Journal of Economics* **116**(2), 681–704.

Sandner, M., Cornelissen, T., Jungmann, T. & Herrmann, P. (2018), 'Evaluating the effects of a targeted home visiting program on maternal and child health outcomes', *Journal of health economics* **58**, 269–283.

Simon, H. A. (1957), 'Models of man; social and rational'.

Turkiewicz, K. L. (2018), Data trimming, *in* M. Allen, ed., 'The SAGE Encyclopedia of Communication Research Methods', SAGE Publications, Inc, Thousand Oaks, pp. 347–348.

WHO (2015), 'Who technical consultation on newborn health indicators: every newborn action plan metrics'.
URL: *https://www.who.int/maternal*$_c$*hild*$_a$*dolescent/documents/newborn−health − indicators/en/*

Yamane, S. & Hayashi, R. (2015), 'Peer effects among swimmers', *The Scandinavian Journal of Economics* **117**(4), 1230–1255.

Zhang, X. M. & Zhu, F. (2011), 'Group size and incentives to contribute: A natural experiment at chinese wikipedia', *American Economic Review* **101**(4), 1601–15.

Zimet, G. D., Dahlem, N. W., Zimet, S. G. & Farley, G. K. (1988), 'The multidimensional scale of perceived social support', *Journal of Personality Assessment* **52**(1), 30–41.

A Appendix

A.1 Robustness checks

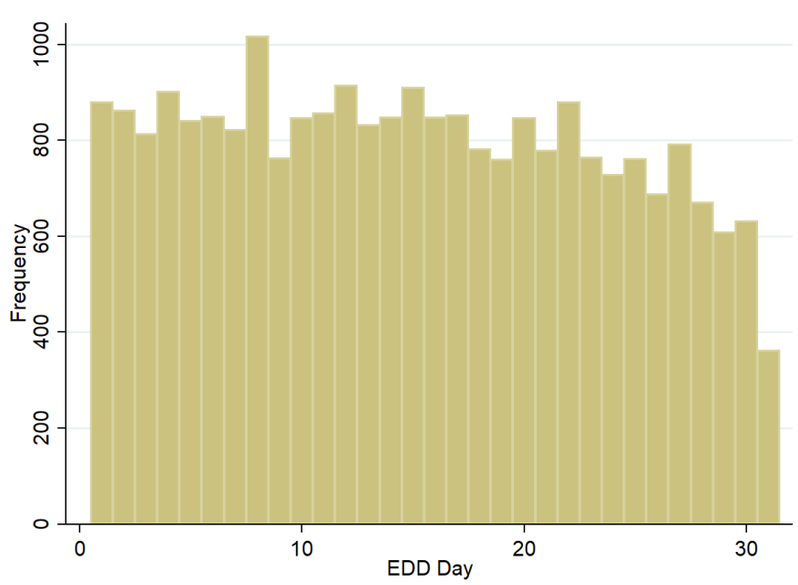

Figure A1: Number of users by the day of EDD

Note: It shows the number of users on each day of the month in the three core peer groups (March, April, and May 2018). The numbers are approximately equally spread over the month despite a slight downward trend towards the end of the month and that one of our peer groups (the month of April) does not have 31 days. There is also a spike on the 8th which is a lucky number in China and might be related to planned cesarean cut. We perform the same empirical analysis excluding users whose EDD is on the 8th of the month, the results remain robust.

Table A1: Two peer groups and information exchange

	(1)	(2)	(3)	(4)	(5)
	1^{st} stage	2^{nd} stages			
VARIABLES	Two PGs	log(info) in Default PG OLS	log(info) in Default PG IV	log(info) in Total PGs OLS	log(info) in Total PGs IV
$\widehat{\text{Two PGs}}$		0.175***	-0.671***	0.333***	-0.361***
		(0.0138)	(0.0949)	(0.0137)	(0.0931)
|Day of EDD-16|	0.0125***				
	(0.000551)				
Number of NPGs	0.00522***	-0.00312**	0.00131	-0.00354**	9.88e-05
	(0.000716)	(0.00151)	(0.00170)	(0.00151)	(0.00166)
log(info) in NPGs	0.0270***	0.339***	0.361***	0.344***	0.362***
	(0.0272)	(0.00571)	(0.00662)	(0.00570)	(0.00646)
Constant	0.0233***	1.246***	1.343***	1.243***	1.322***
	(0.00871)	(0.0170)	(0.0214)	(0.0170)	(0.0209)
Peer group fixed effects	Yes	Yes	Yes	Yes	Yes
F-tests of instrument	-	-	512.85	-	512.85
Observations	24,705	24,705	24,705	24,705	24,705

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A2: Reduced forms

	(1)	(2)	(3)	(4)
VARIABLES	log(info) in Default PG	log(info) in Total PGs	log(info) in Default PG	log(info) Total PGs
|Day of EDD-16|	-0.00894***	-0.00510***	-0.00837***	-0.00451***
	(0.00118)	(0.00121)	(0.00109)	(0.00111)
Number of NPGs			-0.00218	-0.00179
			(0.00151)	(0.00153)
log(info) in NPGs			0.343***	0.352***
			(0.00570)	(0.00575)
Constant	1.583***	1.581***	1.327***	1.314***
	(0.0127)	(0.0128)	(0.0189)	(0.0191)
Peer group fixed effects	Yes	Yes	Yes	Yes
Observations	24,705	24,705	24,705	24,705
R-squared	0.003	0.001	0.143	0.146

Notes: The reduced forms in Columns (1) and (2) are for the specifications in Table 4; The reduced forms in Columns (3) and (4) are for the specifications in Table A1.
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A3: Two peer groups and information exchange in a Poisson IV model

VARIABLES	(1)	(2)	(3)	(4)
Information in	Default PG	Total PGs	Default PG	Total PGs
Two PGs	-3.187	-0.862***	-2.513*	-0.805***
	(2.954)	(0.328)	(1.432)	(0.293)
Number of NPGs			0.00917***	0.00722**
			(0.00312)	(0.00300)
Information in NPGs			0.0404***	0.0403***
			(0.00144)	(0.00135)
Constant	1.969***	1.920***	1.735***	1.707***
	(0.0322)	(0.0329)	(0.0383)	(0.0377)
Peer group fixed effects	Yes	Yes	Yes	Yes
Observations	24,705	24,705	24,705	24,705

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A4: Multiple peer groups and information exchange

VARIABLES	(1)	(2)	(3)	(4)
log(information) in	Default PG	Default PG	Total PGs	Total PGs
	OLS	IV	OLS	IV
Multiple PGs	0.228***	-0.916***	0.309***	-0.384***
	(0.00910)	(0.135)	(0.00918)	(0.124)
Constant	1.522***	2.064***	1.519***	1.847***
	(0.00887)	(0.0649)	(0.00890)	(0.0593)
Peer group fixed effects	Yes	Yes	Yes	Yes
Observations	38,919	38,919	38,919	38,919
R-squared	0.016	0.669	0.029	0.738
F-tests of instrument	-	246.73	-	246.73

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

A.2 Heterogeneity analysis

In this section we look at potential heterogeneous effects of having an additional peer group, namely, the previous-month peer group, PG_{T-1}, versus the next-month peer group, PG_{T+1}. It is worth noting that the majority of the information generated in the peer groups is time-sensitive in our setup.[31] By construction, at any given point of time, users in PG_{T-1} are on average two months ahead of PG_{T+1} in the pregnancy status. Let $i_{T,T-1}$ and $j_{T,T+1}$ denote two users whose default peer group is PG_T, the former additionally joined PG_{T-1} while the latter additionally joined PG_{T+1}, respectively. Given the time difference, user $i_{T,T-1}$ is exposed to peers who are more advanced in the pregnancy status and has access to time-sensitive information in advance comparing to user $j_{T,T+1}$. This could have two potential consequences. On the one hand, user $i_{T,T-1}$ may be able to share more information and answer more questions in her default group comparing to user $j_{T,T+1}$, resulting in more information generated in the default peer group. On the other hand, user $i_{T,T-1}$ may have more incentives to shift her activities to PG_{T-1}, resulting in less information in the default peer group.

To examine the potential heterogeneous effects of the additional peer groups, we estimate the econometric model in Equation (4). Now we have two endogenous variables in the same regression, the enrollment in the PG_{T-1} and the enrollment in the PG_{T+1}. Thus, we use $f_1(EDD) = EDD$ and $f_2(EDD) = EDD^2$ as the instruments for the PG_{T-1} enrollment and PG_{T+1} enrollment, respectively.[32] As in the analysis in Section 5, the baseline compared with is the amount of information exchange generated by the users who have default peer group only.

$$Y_i = \kappa_1 \mathbb{1}[\text{PG}_{T-1,i}] + \kappa_2 \mathbb{1}[\text{PG}_{T+1,i}] + \alpha_{pg} + \mu_i \tag{4}$$

Table A5 shows the results. Columns (1) and (3) are the OLS regressions for the information in default peer group and total peer groups, respectively. Having either PG_{T-1} or PG_{T+1} is positively correlated with the information exchange in users' default peer group as well as in total peer groups. Columns (2a) and (2b) are IV regressions for the information in default peer group. Having PG_{T-1} leads to 43% less information ($p<0.01$) while having PG_{T+1} leads to almost 64% less information ($p<0.01$) in the default peer group. These results are consistent with the main results in Table 4 where we pooled the PG_{T-1} and PG_{T+1} together. Moreover, the reduction is significantly smaller when users additionally joining PG_{T-1} than in PG_{T+1} (χ^2 Test $\kappa_1=\kappa_2$, $p<0.001$).

[31] For example, some typical posts and follow-up responses would discuss how to deal with severe vomiting that usually starts between 5 and 10 weeks, how to read the report of Down's screening test which takes place during the 11 to 14 weeks, what is the normal size of the belly in the X-th month since pregnancy etc.

[32] We also use an alternative functional form of the EDD as the instruments: $g_1(EDD) = |EDD - 16|, \text{if } EDD \leq 16; g_1(EDD) = 0, \text{if } EDD > 16$; and $g_2(EDD) = 0, \text{if } EDD \leq 16; g_2(EDD) = |EDD - 16|, \text{if } EDD > 16$. We thank Bo Honore for this suggestion.

Columns (4a) and (4b) are the IV regressions for the information in total peer groups (=1 if default PG only; =2 if two PGs). Having PG_{T-1} leads to almost 20% less information (p<0.05) while having PG_{T+1}leads to almost 50% less information (p<0.01) in total peer groups.[33] Again, the reduction is significantly smaller when users additionally joining PG_{T-1} than in PG_{T+1} (χ^2 Test κ_1=κ_2, p<0.001).

Table A5: Heterogeneous peer groups and information exchange

VARIABLES	(1)	(2a)	(2b)	(3)	(4a)	(4b)
log(info) in	Default PG	Default PG	Default PG	Total PGs	Total PGs	Total PGs
	OLS	IV	alt IV	OLS	IV	alt IV
PG_{T-1} (κ_1)	0.311***	-0.558***	-0.560***	0.478***	-0.236**	-0.254**
	(0.0183)	(0.0990)	(0.103)	(0.0180)	(0.0973)	(0.102)
PG_{T+1} (κ_2)	0.0902***	-1.027***	-1.030***	0.232***	-0.695***	-0.718***
	(0.0214)	(0.126)	(0.132)	(0.0222)	(0.124)	(0.130)
Constant	1.472***	1.648***	1.649***	1.468***	1.613***	1.617***
	(0.00946)	(0.0209)	(0.0217)	(0.00945)	(0.0204)	(0.0213)
Peer group FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	24,705	24,705	24,705	24,705	24,705	24,705
R-squared	0.014	0.724	0.724	0.034	0.746	0.745
F-tests of instrument	-	254.52	252.17	-	254.52	252.17
χ^2 Test κ_1=κ_2	-	p<0.001	p<0.001	-	p<0.001	p<0.001

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
IV_1 : We use $f_1(EDD) = EDD$ and $f_2(EDD) = EDD^2$ for the PG_{T-1} enrollment and PG_{T+1} enrollment, respectively.
IV_2 : We use $g_1(EDD) = |EDD - 16|, \text{if } EDD \leq 16; g_1(EDD) = 0, \text{if } EDD > 16$; and $g_2(EDD) = 0, \text{if } EDD \leq 16; g_2(EDD) = |EDD - 16|, \text{if } EDD > 16$ for the PG_{T-1} enrollment and PG_{T+1} enrollment, respectively.

Table A6: First stages and reduced forms

	(1)	(2)	(3)	(4)
	First stages		Reduced forms	
VARIABLES	Additional PG PG_{T-1}	Additional PG PG_{T+1}	log(info) in Default PG	log(info) in Total PGs
Day of EDD	-0.0206***	-0.00665***	0.0183***	0.00949***
	(0.00103)	(0.000659)	(0.00241)	(0.00247)
(Day of EDD)^2	0.000411***	0.000413***	-0.000653***	-0.000384***
	(2.90e-05)	(2.49e-05)	(7.54e-05)	(7.73e-05)
Constant	0.321***	0.0323***	1.436***	1.515***
	(0.00881)	(0.00378)	(0.0181)	(0.0185)
Peer group FE	Yes	Yes	Yes	Yes
Observations	24,705	24,705	24,705	24,705

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[33]Since the dependent variable is log-transformed, the estimated effects = $(e^{\beta} - 1) \times 100\%$.

A.3 Survey Questionnaire

We first present the questionnaire of MSPSS in Section A.3.1, then the questionnaire of EDS in Section A.3.2, and finally, the full questionnaire in Section A.3.3.[34]

A.3.1 Modified Questionnaire of Multidimensional Perceived Social Support Scale (Zimet et al. 1988)

Indicate how you feel about each statement: very strongly disagree (=1) to very strongly agree (=7).

1 I have learned useful information from others in the peer group(s).

2 There are some people in the peer group(s) who are a real source of comfort to me.

3 There are some people in the peer group(s) who are around when I am in need.

4 I get the emotional help and support I need from others in the peer group(s).

5 I can talk about my problems in the peer group(s).

6 People in the peer group(s) help me make some decisions.

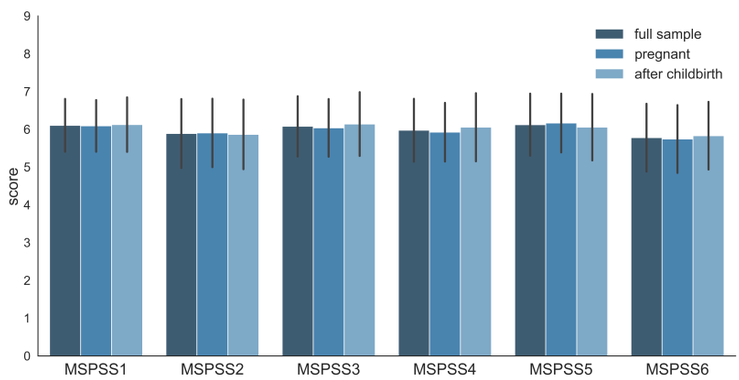

Figure A2: MSPSS by Question

Note: Each bar shows the mean score as well as the range of one standard deviation in the full survey sample, pregnant users, and users who gave birth within one year, respectively.

A.3.2 Questionnaire of Edinburgh Depression Scale (Murray & Cox 1990)

In the *last seven* days

[34]The survey was conducted in June and July in 2020 by Zhongyan Technology, a major online survey and research company in China. The online questionnaire was coded and implemented in their development web domain www.wenjuan.com. The participants were recruited from their recruitment web domain www.idiaoyan.com, where participants can earn points and convert them to cash payments which is equivalent to 50 pennies in GBP.

1. I have been able to laugh and see the funny side of things:
   - As much as I always could [0]
   - Not quite as much now [1]
   - Definitely not so much now [2]
   - Not at all [3]
2. I have looked forward with enjoyment to things:
   - As much as I ever did [0]
   - Rather less than I used to [1]
   - Definitely less than I used to [2]
   - Hardly at all [3]
3. I have blamed myself unnecessarily when things went wrong:
   - Yes, most of the time [3]
   - Yes, some of the time [2]
   - Not very often [1]
   - No, never [0]
4. I have been anxious or worried for no good reason:
   - No, not at all [0]
   - Hardly ever [1]
   - Yes, sometimes [2]
   - Yes, very often [3]
5. I have felt scared or panicky for no very good reason:
   - Yes, quite a lot [3]
   - Yes, sometimes [2]
   - No, not much [1]
   - No, not at all [0]
6. Things have been getting on top of me:
   - Yes, most of the time I haven't been able to cope at all [3]
   - Yes, sometimes I haven't been coping as well as usual [2]
   - No, most of the time I have coped quite well [1]
   - No, I have been coping as well as ever [0]
7. I have been so unhappy that I have had difficulty sleeping:
   - Yes, most of the time [3]
   - Yes, sometimes [2]
   - Not very often [1]
   - No, not at all [0]
8. I have felt sad or miserable:

- Yes, most of the time [3]
- Yes, quite often [2]
- Not very often [1]
- No, not at all [0]

9. I have been so unhappy that I have been crying:
   - Yes, most of the time [3]
   - Yes, quite often [2]
   - Only occasionally [1]
   - No, never [0]
10. The thought of harming myself has occurred to me:
   - Yes, quite often [3]
   - Sometimes [2]
   - Hardly ever [1]
   - Never [0]

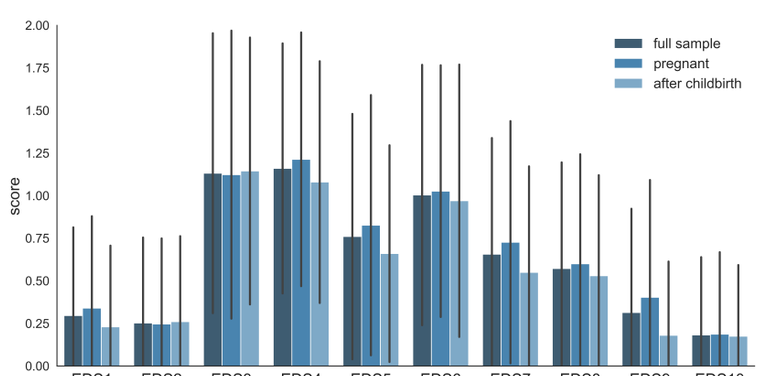

Figure A3: EDS by Question

Note: Each bar shows the mean score in the full survey sample, pregnant users, and users who gave birth within one year, respectively. The range of score is from 0 to 3 and the vertical bars indicate the range of one standard deviation.

A.3.3 Full Questionnaire

Start of the questionnaire

1. Which one of the following fits your current situation
   - You are pregnant with your first baby, and you are now a user of this community.
   - Your first baby has been born within the last year, and you are a user of this community during pregnancy.
   - None of the above. [end of survey]

2. Participation info sheet and consent form[35]
   - Agree
   - Disagree [end of survey]
3. In addition to the default peer group based on your expected due date, have you joined other peer groups?
   - Yes [to questions 3a & 3b]
   - No [to questions 4]

3a. The other peer groups you joined include (multiple choices)
   - The peer group of the previous month (e.g if your default peer group is September 2020, "previous month" refers to August 2020)
   - The peer group of the next month (e.g if your default peer group is September 2020, "next month" refers to October 2020)
   - Exactly two months apart from your default peer group
   - More than two months apart from your default peer group
   - More than three months apart from your default peer group

3b. The reasons why you joined other peer groups include (multiple choices)
   - The expected due date may not be accurate (for example: my due date is at the beginning/end of the month, so I also joined the peer group(s) of the previous/next one month)
   - I want to get more information by joining more peer groups.
   - I want to provide information to more people by joining more peer groups
   - Other reasons (e.g. miscarriage, revised due date, trembling hand, etc)

4. How many non peer groups in this community besides the peer group(s) have you joined?
   - 0
   - 1-3
   - 4-10
   - $>$ 10
5. Perceived social support in the peer groups (see Section A.3.1)
6. Rate from 1 to 7 (strongly disagree to strongly agree) the statement "This community is my main source of knowledge and information about pregnancy."
7. Mental health (see Section A.3.2)
8. In the past 7 days, how much time did you spend on physical exercise every day (e.g. walking, swimming, yoga, badminton, etc)?
   - Less than 30 minutes
   - 30-60 minutes

[35]The participation info sheet and consent form were approved by the ethical committee at the University of Essex.

- 60-120 minutes
- More than 120 minutes

9. In the past 7 days, how many hours of sleep per day did you have on average?
   - Less than 5 hours
   - 5-6 hours
   - 6-7 hours
   - More than 7 hours
10. Have you ever smoked or drank alcohol during pregnancy?
   - Yes
   - No
11. What is your expected due date (If the baby has been born, please refer to the expected due date rather than the actual date of birth)?
   - Scrolling choices
12. Your birth year is 19–?
   - Scrolling choices
13. If the baby has been born, what is the actual date of birth?
   - Scrolling choices
14. Which of the following is true?
   - My delivery was planned natural birth.
   - My delivery was unplanned Cesarean.
   - My delivery was planned Cesarean.
15. What is the birth weight of your baby? (If multiple babies, please fill for each baby)
   - Scrolling choices
16. Are/were you breastfeeding in the first 6 months?
   - Yes
   - No

End of the questionnaire