Marriage and children, but not childless dating or cohabitation, may be immediately competitive with close friendships

Goh Wei Zhong

20th December 2017

Abstract

Do romantic relationships displace friendships? Understanding the prognosis of friendships in early adulthood is crucial to our understanding and navigation of relationships. A regression model was performed to predict the number of close friends one has, based on the length of time one is in a romantic relationship, one's current relationship status, the number of children one has and their ages, controlled for major demographic variables and personality indicators. Married people, regardless of marriage length, are expected to have 10% fewer close friends, and people with children can expect 5% fewer close friends per child. No effects (duration-dependent or duration-independent) were found for childless cohabitation, dating and current pregnancies.

Introduction

The late twenties is a time of transition, especially in urban America, where young adults get married later in life. Friendship plays an important role in the emotional lives of Americans in their late twenties, but this is also a time when many people are pursuing committed long-term romantic relationships in various forms, such as short-term dating, long-term dating, cohabitation and marriage. These relationships may displace friendships, as romantic partners rely less on their friends and more on each other for emotional support and company, and may share domestic and childcare responsibilities as the relationship progresses.

To understand how friendship changes as one enters romantic relationships, based on a nationwide survey, I performed regression models to predict the number of close friends one has, based on the length of time one is in a romantic relationship, one's current relationship status, the number of children one has and their ages, controlled for major demographic variables and personality indicators.

Exploratory data analysis

Figure 1 shows that men, people with higher income and educational levels, and people who socialise widely in social situations have statistically significantly more close friends, while

Black and Hispanic people have statistically significantly fewer close friends. (It will be seen that these demographic effects do not exist independently.) No age effect was found in the relatively narrow age range of the study, but age will still be controlled for in subsequent sections as with sociology convention.

Results and discussion

With the above personality indicators and demographic variables controlled for, a negative binomial regression was performed to investigate the effect of current relationship status, current relationship length (interacted with current relationship status), the number of children born alive and the number of fetal deaths (pregnancies that did not result in live births, such as abortions, miscarriages, stillborns, etc.). The significant predictors are summarized in Figure 2. The only significant parameters, other than the control parameters, were the number of children born alive and whether one is married as opposed to single. The number of fetal deaths, the length of the current romantic relationship in general, or a marriage, cohabiting relationship, pregnancy-bound relationship or dating relationship in particular, as well as the presence of any kind of relationship status other than marriage does not significantly affect the expected number of close friends. Including these factors in isolation did not yield significant effects. In other words, once a person is in a marriage, the expected number of friends appears insensitive to whether it is a new marriage or an old marriage.

These results suggest that marriage and children, and not other types of romantic commitments, are strong indicators of commitment to a lifestyle that may be competitive with close friendships. In addition, the institution of marriage, and not less formal types of romantic relationships, may deter one's friends from seeking one out, contributing to the dissolution of friendships.

To investigate the effect of the child's age on the parent's expected number of friendships, a separate negative binomial regression was performed for subjects (male or female) who have had or caused at least one live birth. Only a trending negative effect was found with the youngest child's age (Holm–Sidak corrected $p \sim 0.06$), and no effect was found with the oldest child's age. In other words, as the youngest child ages, the expected number of friends shows a decreasing trend. If this trend bears out, it may be related to particular times of the child's life that imposes large stresses on the parents, and having a delayed effect on the parents' friendships.

Within the sample of subjects who have had or caused at least one live birth, marital status, the number of births and Hispanic origin are no longer significant, while the number of years in a current dating relationship becomes significant (Holm–Sidak corrected $p < 0.02$). This suggests that a long-term dating relationship may also be competitive with close friendships if children are involved.

Note that while males have statistically significantly more close friends than females (Figure 1), the effect disappears when income and social behaviour is controlled for in the regression model, suggesting that the gender dependence is not an independent effect but a result of income, social behaviour (and other factors).

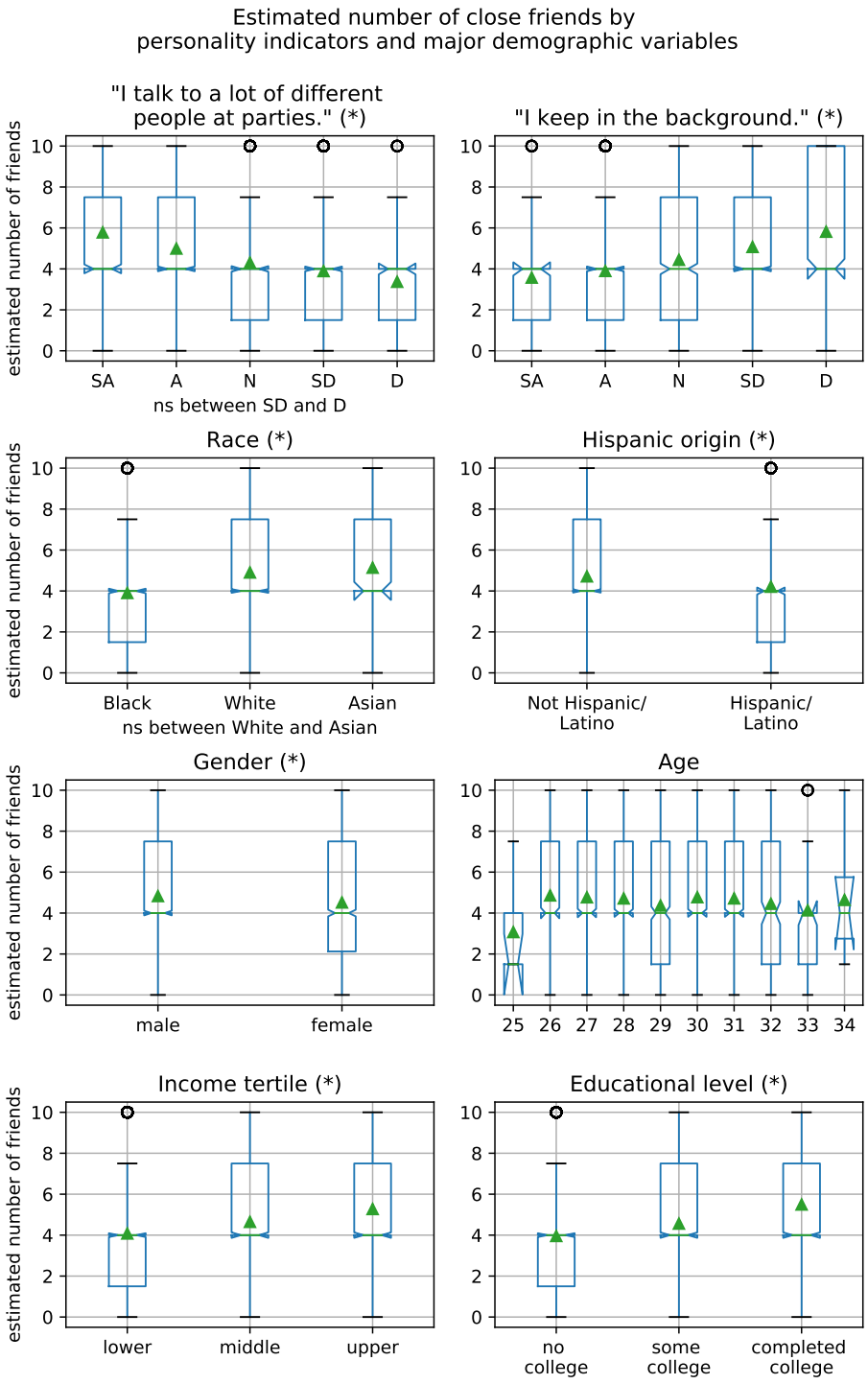

Figure 1: Race, Hispanic origin, gender, income, education and self-assessed behaviour in social situations correlate well with the estimated number of close friends. Green triangles are means. Stars in subfigure titles indicate that all pairwise Mann–Whitney U tests on the unweighted data are significant (Bonferroni-corrected for multiple comparisons within each subfigure) except where indicated below the subfigure. SA, strongly agree; A, agree; N, neutral; D, disagree; SD, strongly disagree.

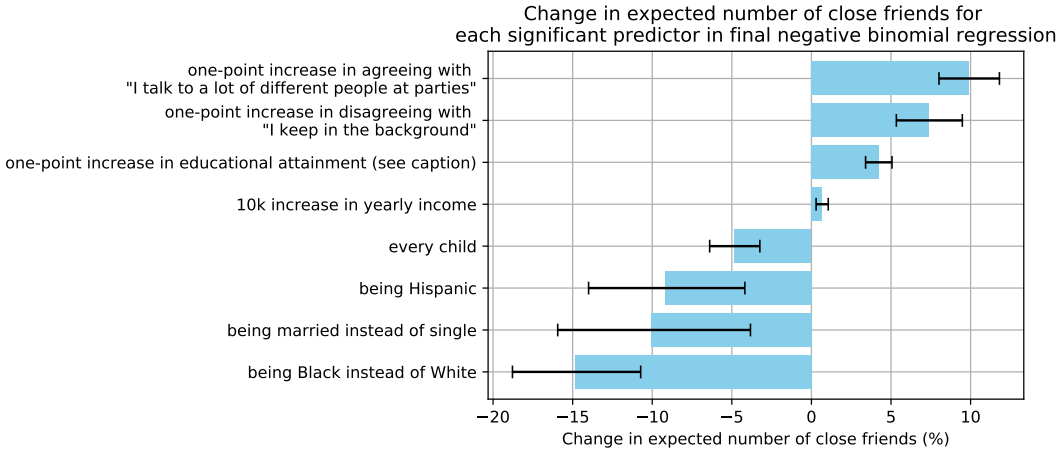

Figure 2: Only the number of children born alive and whether one is married as opposed to single are significant in a regression model of the number of close friends that one has. Age, gender and number of fetal deaths were used as controls, but were not significant. Holm–Sidak correction for multiple comparisons was applied in determining whether variables were significant, but the error bars here are 95% confidence intervals with no correction. Educational attainment is scored on a 13-point scale, with high school graduation scored at 3, some college scored at 6, college graduation scored at 7, and so on. The full scale is provided here: http://www.cpc.unc.edu/projects/addhealth/documentation/ace/tool/codebookssearch?field=varname&text=H4ED2

Conclusion

The only effects of romantic relationships on friendships that were found were that married people are expected to have 10% fewer close friends and people with children can expect 5% fewer close friends per child. With an average of 4.6 close friends in the population, an average married person with two children can expect to lose one close friend. No duration effects were found.

This number is surprisingly low, judging from anecdotal evidence, as married people may spend significantly less time with friends as they juggle their work and family commitments, taking care of newborns who refuse to be put to sleep. Perhaps people are reluctant to admit, or do not recognise, that their friendships have changed as they enter another stage of their lives. The lack of availability and participation in activities of partnered, non-married friends is also a common observation that does not seem to mesh with the present data. Thus, a future step would be to mine the American Time Use Survey and other sources for data on the time spent with friends as a proxy for the quality of friendships.

Finally, it is jarring that a Black person can expect 15% fewer close friends, and a Hispanic White person 10% fewer close friends, as compared to a non-Hispanic White person, controlled for income, education, age, sex and personality self-assessment. This puzzling observation may be the result of an omitted variable, racial discrimination, or simply differing standards in friendship judgement that are culturally entwined. Again, as a first step, data from the American Time Use Survey may be used for a more objective measure of friendship. In addition, looking at the number of close friends that a Black person has in an area where Black

people are the majority versus where they are the minority may shed some light into the role of racial discrimination in this disparity.

Methods

Data

This paper is based on public use microdata from Wave IV (2008–09) of the National Longitudinal Study of Adolescent to Adult Health (Add Health), which is an ongoing study of a nationally representative group of more than 10,000 youths dating back to 1994. It has been administered by the National Institute of Child Health and Human Development, a child institute of the National Institute of Health.

The Wave IV data, which is the latest available, was collected when the subjects were aged between 24 and 32 via 90-minute computer-assisted in-person interviews across the nation. Detailed information regarding all aspects of the subjects' lifestyle, physical and emotional health was collected, including information about subjects' romantic partners and children.

Importantly for this paper, the subjects were asked to indicate whether they had zero, 1–2, 3–5, 6–9 or 10 or more "close friends" ("can talk to about private matters, and can call on for help"). Almost half (45%, weighted) of the subjects responded that they had 3–5 close friends, while 22% of respondents responded 1–2, and 17%, 6–9. Surprisingly, perhaps reflecting American attitudes, 1 in 8 subjects (13%) responded that they had 10 or more close friends. 3% of subjects responded that they had no close friends.

For this analysis, data regarding subjects who did not respond regarding the number of close friends was discarded. Non-responses to the remaining relevant questions were replaced with the sample mean.

The data contains grand sample weights, which were taken into account where possible during the analysis, including in the regression. The weights were rescaled so that the mean weight per subject is 1.

Dealing with binned data

Perhaps the biggest hurdle to performing a regression on this data is the fact that the data on the number of close friends is binned. To make matters worse, the bin widths are unequal, and the data are top-coded at 10, creating a category into which 13% of respondents fall. This question is not the focus of the survey designers, who presumably did not want respondents to delay the interview while enumerating the number of close friends they have. The loss of information leads to less powerful statistical conclusions. On the other hand, in any case, the number of friends a subject is able to report would not be as precise as, say, the number of children one has, and expecting subjects to state an integer for the number of close friends may lead to problems such as subjects' rounding tendencies.

Since the distances between the bins are not the same (e.g., the gap between having zero friends and 1–2 friends may be greater than that between having 1–2 and 3–5 friends), the most statistically robust method to deal with this data may be to perform a type of ordinal regression, such as ordered logit. However, the statistical implementation of ordered logit is lacking in Python (the `mord` package is a machine learning implementation which allows for

prediction but does not give confidence bounds on the parameters). Moreover, such methods are unlikely to be sophisticated enough to incorporate the sample weights.

Thus, I have opted for a negative binomial regression. A subject who responded that they have 1–2 (3–5, 6–9, >10) close friends was considered to have 1.5 (4, 7.5, 10) close friends. In so doing, I am putting a naive linear prior on the distances between the responses.

This analysis could be made more robust by manually implementing ordinal regression or by using a package from advanced statistical software. Better yet, a more powerful analysis can be done using results from a survey that specifically probes the quality and quantity of a subject's friendships. However, such data may not have been collected at a level comparable in scope to Add Health.

Data exploration

To screen for candidate features for the regression, I performed univariate analyses on personality indicators, and demographic variables such as age, race and gender. I performed the non-parametric Mann–Whitney U tests pairwise to test for statistical significance, and performed the Bonferroni correction for multiple comparisons per variable, as at this stage I am not concerned about false negatives.

Regression

A negative binomial regression was chosen because the number of friends is count data that is overdispersed. This is a parametric technique. Since the regression involves multiple comparisons, the p-values from the regression are adjusted using the Holm–Sidak step-down method, which is the default in the `statsmodels` package in Python and gives a more powerful test while controlling the family-wise error rate.

Acknowledgements

This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.