

## **Appendix:** Description of the Sampling Method

The probability sampling we used is a modified form of two-stage clustering sampling based on the types of on-campus student dormitories. In all three universities, almost all the students live on campus except for very few married graduate students. There are different types of dormitory rooms with 1 to 6 students living in them. To guarantee the representativeness of the final sample, we first sampled dormitory rooms from each room type and then selected one or two students from each selected room. We selected one student from dormitory rooms with 3 or fewer students and two students from dormitory rooms with 4 or more students to complete the questionnaires. The different probabilities for a student to be selected in different room types require separate sampling for different room types. Eventually the probability of a student being selected into the final sample is the product of the probabilities at the two stages, that is, (1) the probability of each room being selected times (2) the probability of each person being selected from the room. We set this probability to be the sample size divided by the population size. In other words, in our sampling scheme, all students have the equal chance of being selected into the final sample.

Take Tsinghua University for example. There are 3811 double rooms, 2670 triple rooms, and 3528 quadruple rooms that altogether house over 29700 students. 500 questionnaires were expected to be distributed in Tsinghua University. To make sure that all the students have the same chance of being selected as survey participants, we calculated that 128 double rooms, 135 triple rooms and 118 quadruple rooms need to be selected. To select these rooms, systematic sampling was applied within each type of dormitory rooms. Then one student in each of the selected double and triple rooms and two students in each of the selected quadruple rooms were

selected randomly by interviewers to complete the questionnaire.<sup>1</sup> In this way, every student in THU has the equal probability (0.017) to be selected into the survey.<sup>2</sup> Similar sampling procedures were used in PKU and RUC.

When conducting the survey we anticipated the potential problem of self-censorship and tried our best to eliminate its impact. The survey was totally anonymous (not just confidential) and self-administered. We left the questionnaires for respondents to complete by themselves and mixed the completed questionnaires in a large envelope. There is no way to match completed questionnaires with specific individuals. We trained our survey team carefully before the survey. The survey conductors explained this complete anonymity of the survey when distributing questionnaires. In doing so, we hope to minimize the impact of self-censorship in respondents' answers.

---

<sup>1</sup> If the selected student refused to complete the questionnaire, another student in the same room would be chosen randomly. If there were not enough students willing to participate, one or two students in the nearest room would be selected.

<sup>2</sup> The probability (p) of being selected into the sample is 0.017 for all THU students. For students living in double rooms:  $p = (1/2) * (128/3811) = 0.017$ ; for students living in triple rooms:  $p = (1/3) * (135/2670) = 0.017$ ; for students living in quadruple rooms:  $p = (2/4) * (118/3528) = 0.017$ .