

Supplemental information for
*A dynamic model for the mutual constitution
of individuals and events*

Jürgen Lerner
University of Konstanz, Germany
juergen.lerner@uni-konstanz.de

Alessandro Lomi
University of the Italian Switzerland, Lugano, CH
alessandro.lomi@usi.ch

A Visualizations of all events in the DGG data

Figure 1 displays the distribution of statistics *individual.activity*, *dyadic.shared.activity*, and *closure* as well as the logarithm of the relative event rate of Event 2 through 14 that is predicted by the best-fitting model (Model 2 from Table 2 in the main text). Histograms show the distribution of the statistic (or log rate) over the whole size-constrained risk set (space of alternative participant lists). The dashed vertical bar displays the mean over the whole risk set. The solid red vertical bar displays the respective value of the observed event. The values for the first event are not shown, since all four values are constant over the whole risk set (note that the first event has no prior event, which implies that there is no variation in any statistic and – in turn – no variation in the event rate). Note that the event rate in the CoxPH model is only defined up to a multiplicative constant, so that the log event rate is only defined up to an additive constant. This constant has been chosen to make the mean logarithmized event rate equal to zero (corresponding to a mean relative event rate equal to one).

Several insights can be obtained from these distributions. First, most of the observed events have a higher predicted event rate than the average hyperedge from the risk set and most observed events take higher values in all three statistics than the respective average over all alternative hyperedges. That is, groups experiencing the observed events are (1) composed of previously more active actors, (2) composed of pairs of actors with higher prior shared activity, and (3) composed of pairs of actors with more common co-participants than the average group of actors of the same size.

Few events have less than average predicted event rate. Event 4 and 12 are the only two events that are less likely than the average hyperedge from the size-constrained risk set at the event time (Event 12 only by a very small margin). A manual inspection shows that both are uncommonly large events, spanning the two dense groups. However, both of these “unlikely” events (Event 4 and 12) take higher values than average in all three model statistics. A reason for their unlikeliness seems to be that *dyadic.shared.activity* (the prior dyadwise shared activity statistic, which has a positive effect on the event rate) is only slightly larger than average, while most of the “likely” events have *dyadic.shared.activity* much higher than average. From another point of view *dyadic.shared.activity* of Event 4 and 12 is not high enough to compensate the (above average) values of the two statistics individual prior activity and closure which both have a negative effect on the event rate.

Few events have statistics below the average over the risk set. Events 2, 5, and 6 have values below average in all three statistics and Event 9 takes a relatively low value in the prior individual activity (but above average in the two others). Interestingly all of these events are considered as more likely than average

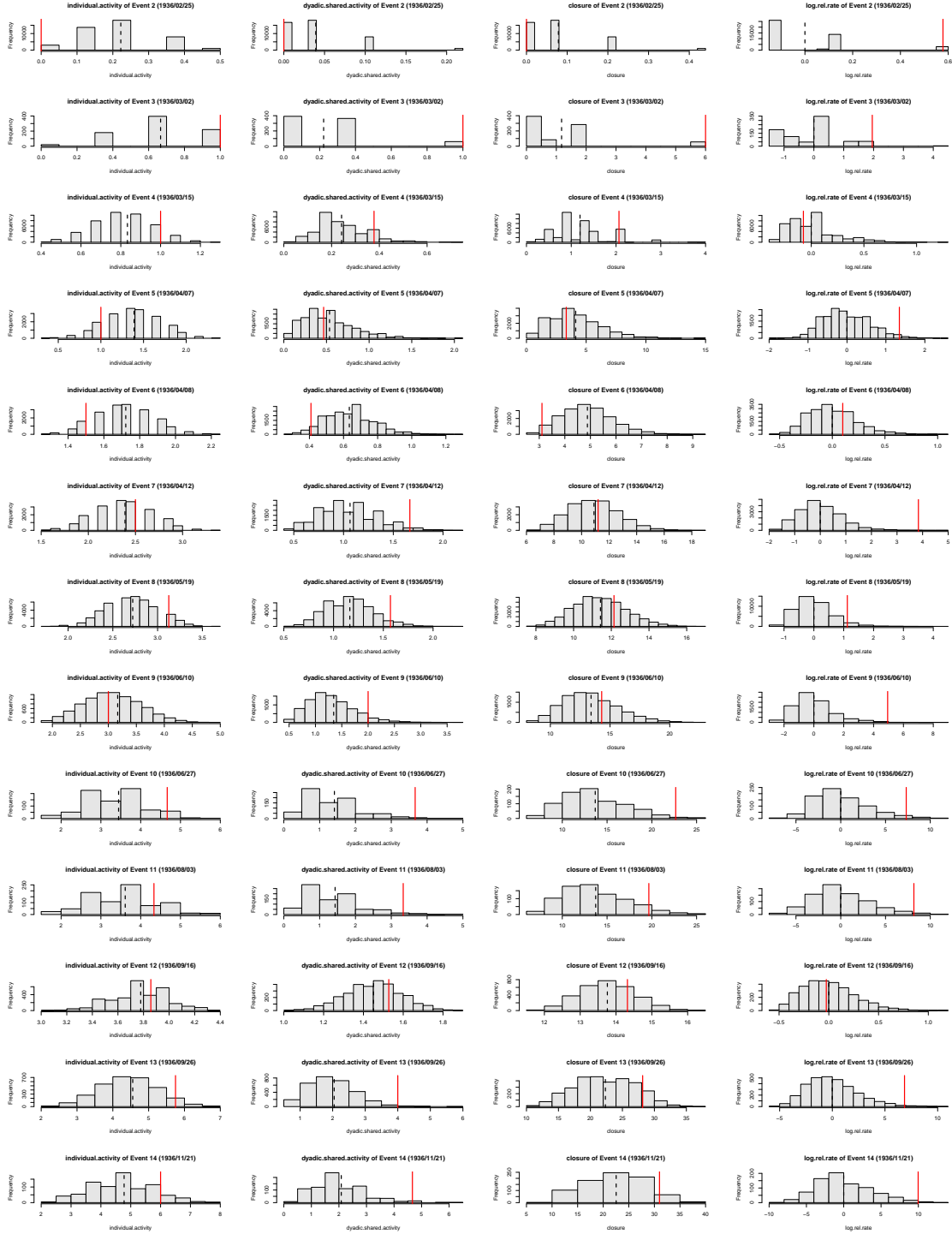


Figure 1: Statistics *individual.activity* (prior individual activity), *dyadic.shared.activity* (prior dyadwise shared activity), and *closure* as well as the logarithm of risk explained by the model of Event 2 through 14. Histograms shows the distribution of the statistic (or log risk) over the whole size-constrained risk set (space of alternative events). The dashed vertical bar is the mean over the whole risk set. The solid red vertical bar is the statistic of the observed event.

by the model (thus, they are well-explained by the model). Event 2 involves a disjoint set of participants from that of Event 1. Thus, none of the participants of Event 2 has any previous event. Event 5 involves two previously inactive actors (*Myra* and *Katherine*). Therefore, it has unusually low statistics – but is still well-explained by the model. Event 6 is a rather large “bridging event” between the two clusters.

The network visualizations of all events are shown in Fig. 2 and – in larger scale – in Figs. 3 through 16. The image at the time of Event i displays all previous events, that is, Event 1 to Event $i - 1$. Additionally, the participants of Event i are highlighted in red. Insights are discussed in the figure captions.

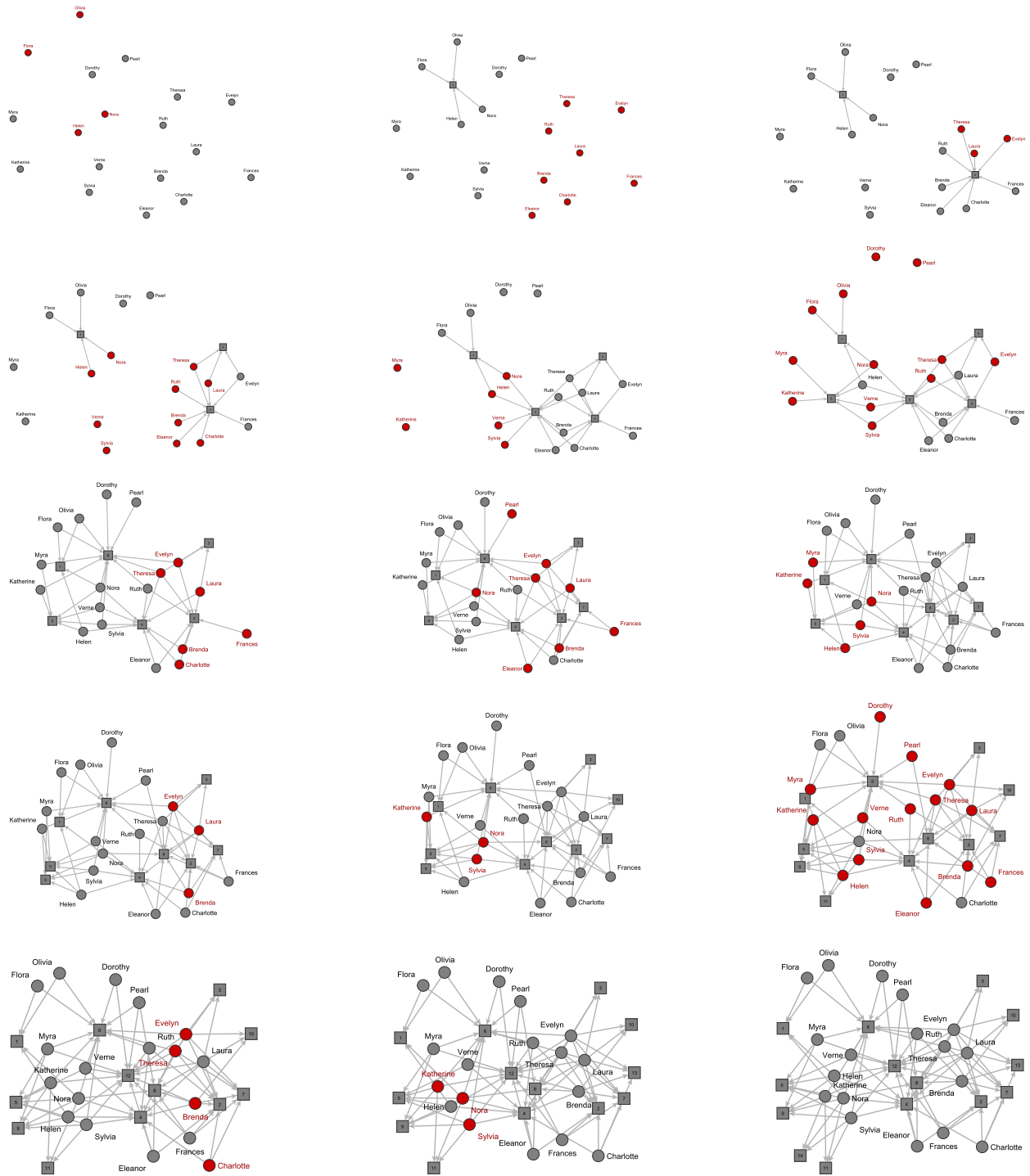


Figure 2: Network diagrams showing for each event $e_i = e_1, \dots, e_{14}$ in the DGG data how participants of Event e_i (displayed in red) are embedded in the network of previous events e_1, \dots, e_{i-1} . The last diagram (in the bottom, right) shows the co-attendance network resulting from all 14 events.



Figure 3: DGG network shortly before Event 1 whose participants are displayed in red. Since this is the first event, there are no prior events, the three statistics are constant over the whole risk set, and any group of actors has the same probability to experience a joint event as any other group. (We do not display histograms of statistics for Event 1, since statistics are constant for that event.)

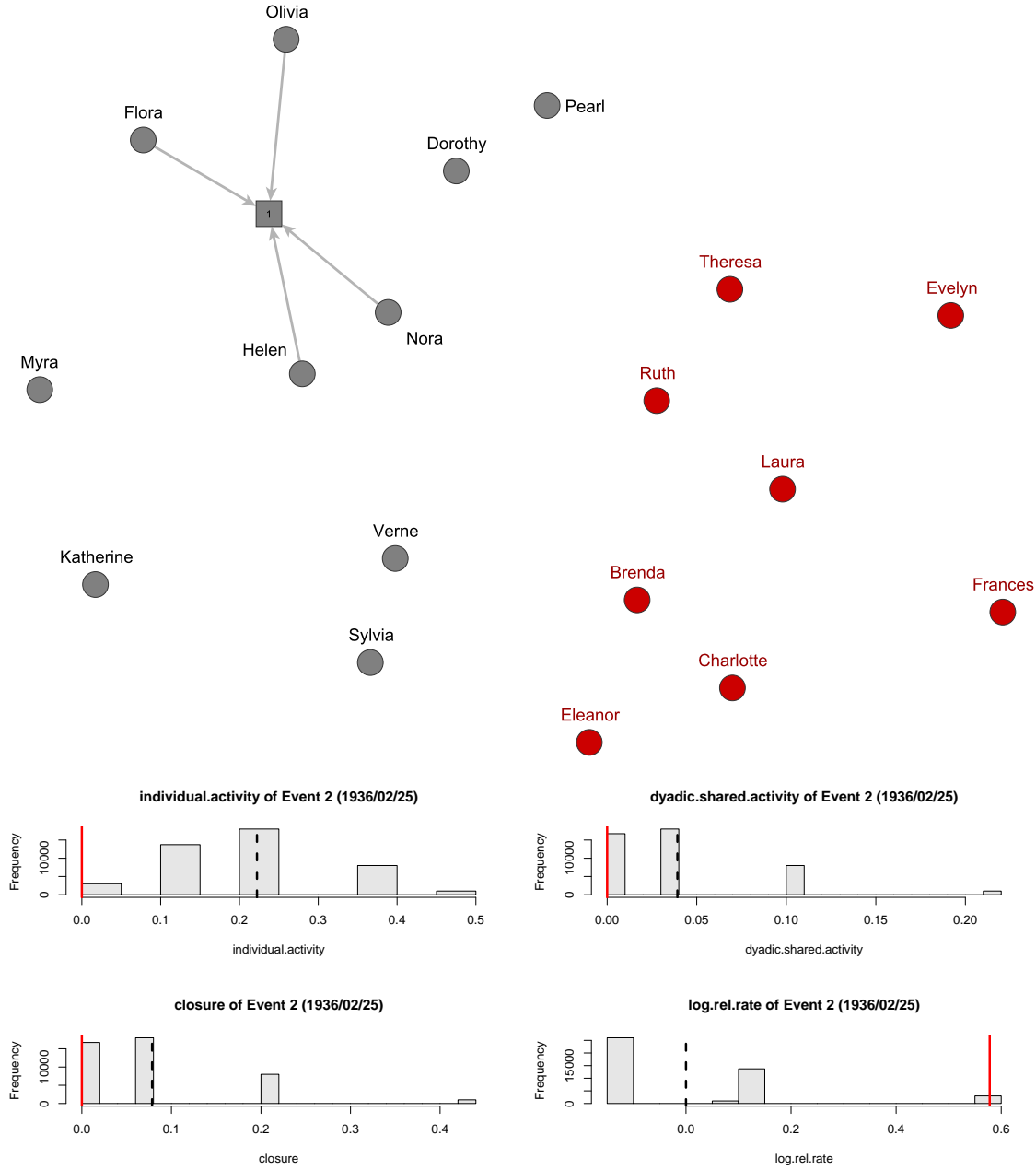


Figure 4: DGG network shortly before Event 2 whose participants are displayed in red. Participants of Event 2 have never experienced any event before and values of all statistics are below average – yet the model estimates the probability of this group of participants above average, compare Fig 1. Indeed, prior individual activity and closure (which both have a negative effect) are rather low.

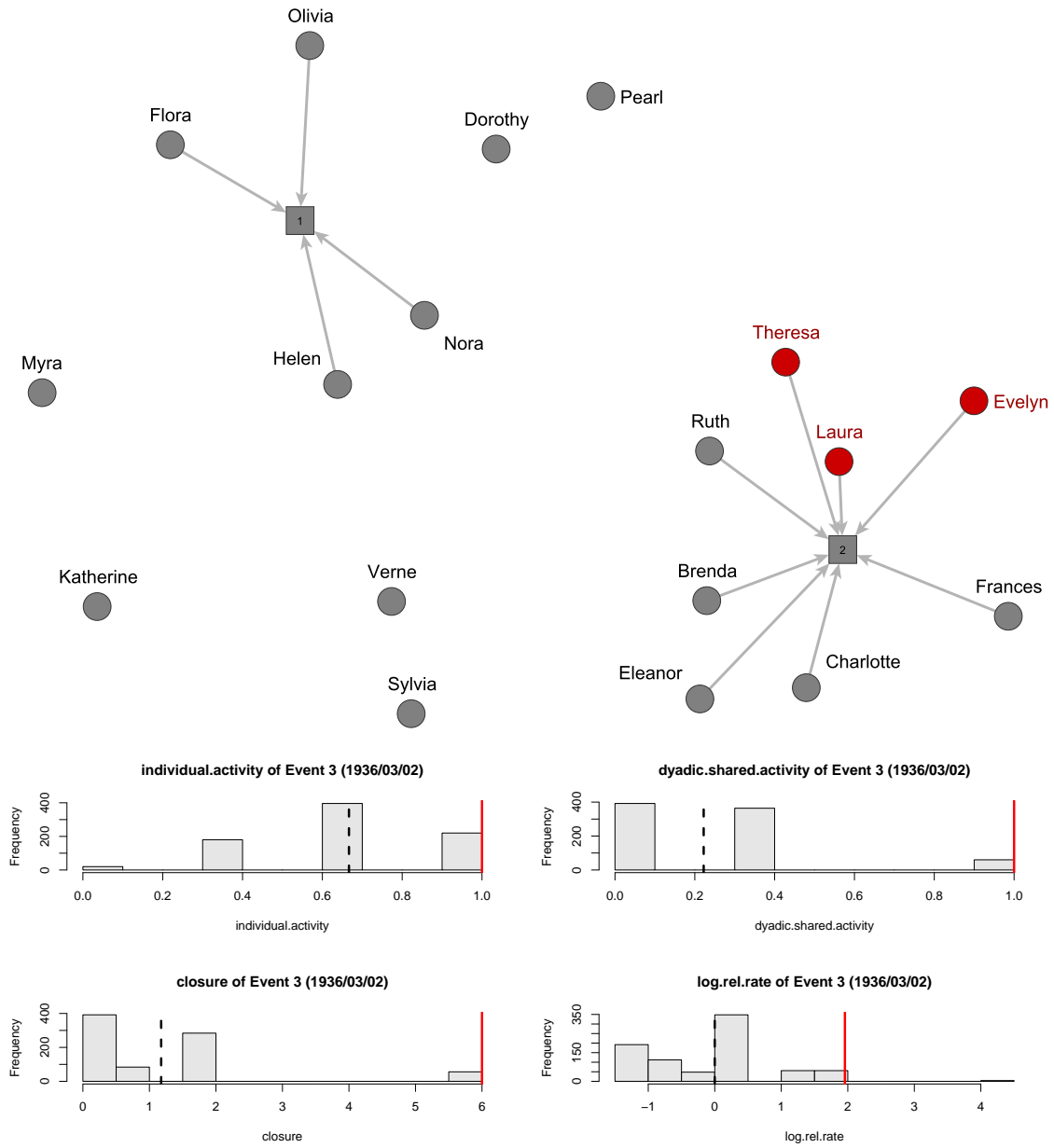


Figure 5: DGG network shortly before Event 3 whose participants are displayed in red. Participants of Event 3 are a subset of those of Event 2. Event 3 is well-explained by previous shared activity. Closure is also above average – but this is compensated by the positive effect of previous shared activity.

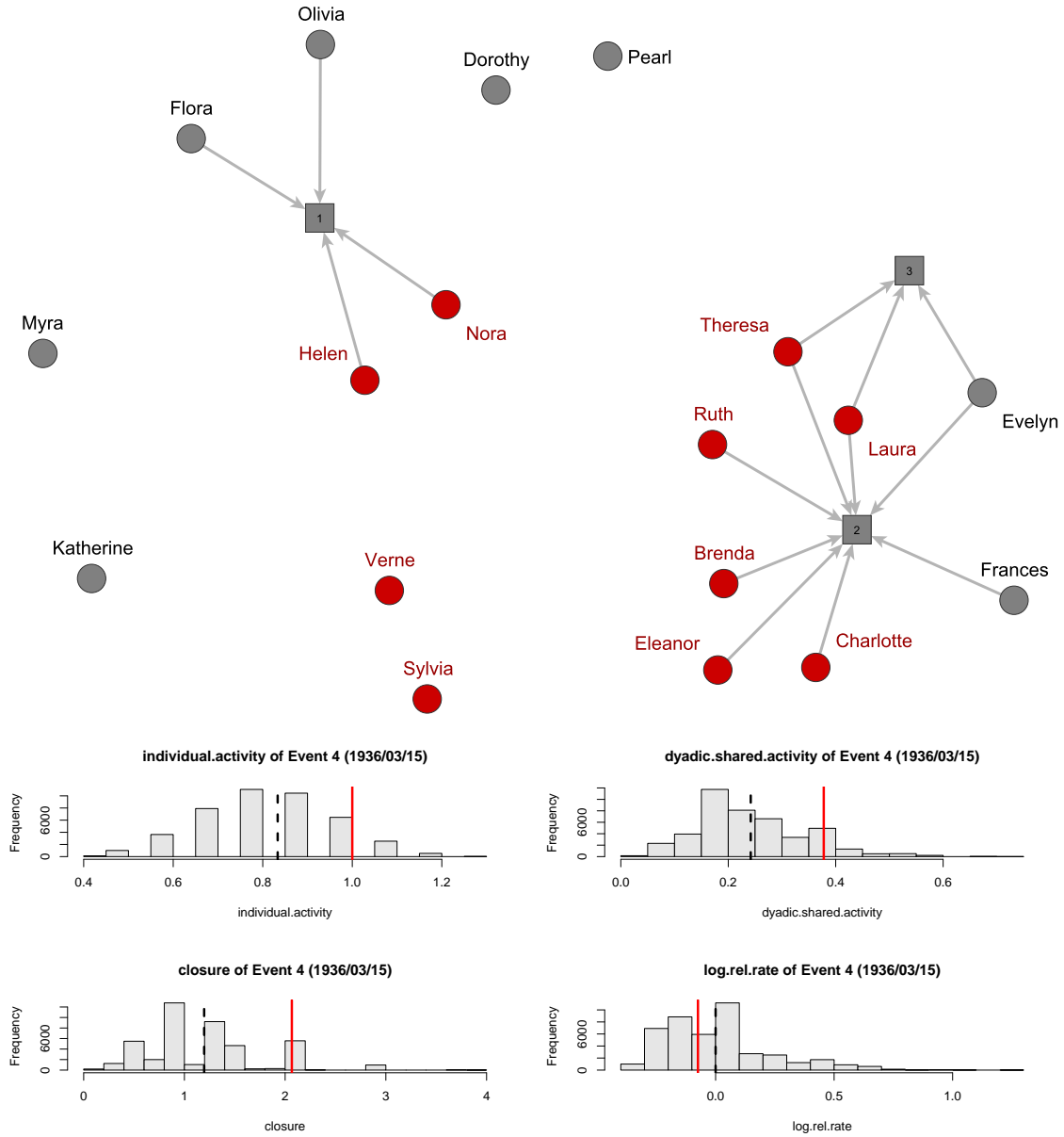


Figure 6: DGG network shortly before Event 4 whose participants are displayed in red. Event 4 is the most unlikely in the whole sequence (according to the model). It bridges between the two emerging clusters. Previous shared activity is too low to compensate for the negative effects of prior individual activity and closure.

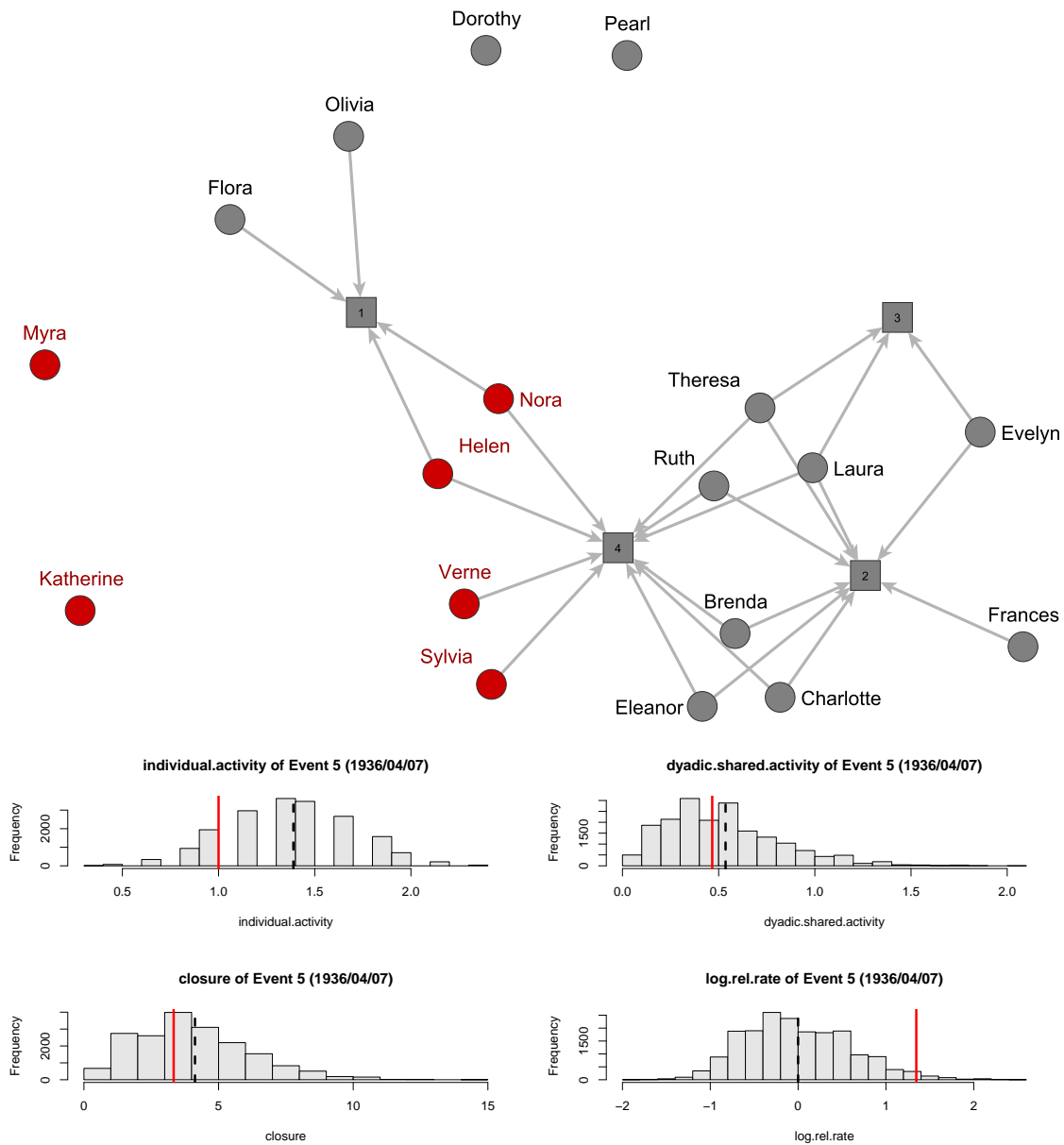


Figure 7: DGG network shortly before Event 5 whose participants are displayed in red. Event 5 has below-average values in all three statistics. Its probability is nevertheless above average, according to the model.

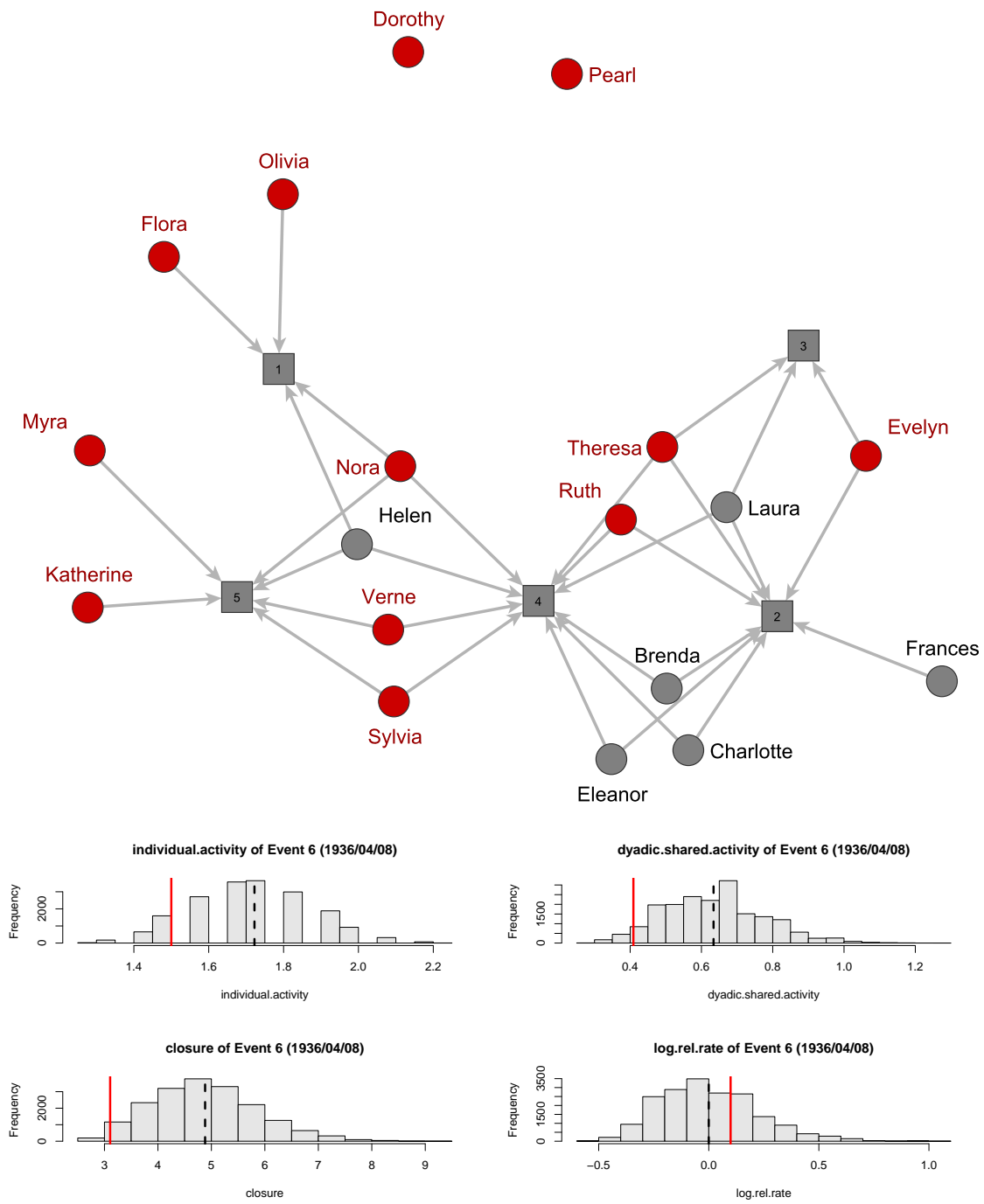


Figure 8: DGG network shortly before Event 6 whose participants are displayed in red. Event 6 is the third one that has statistics below average (it is huge and involves disparate actors). Its probability is above average, according to the model.

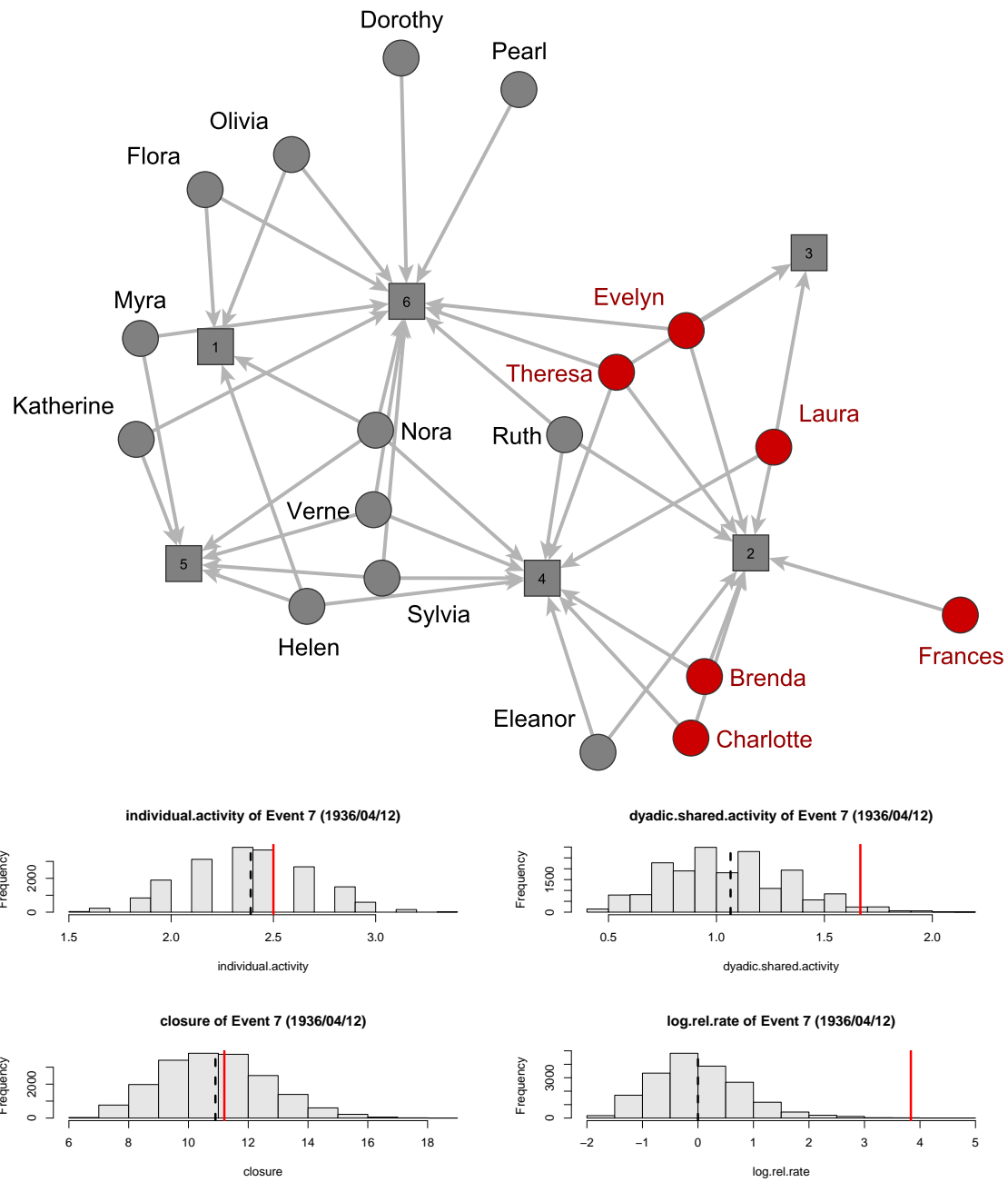


Figure 9: DGG network shortly before Event 7 whose participants are displayed in red. Event 7 is one of the most likely in the sequence of events. Its participants are a subset of those of Event 2.

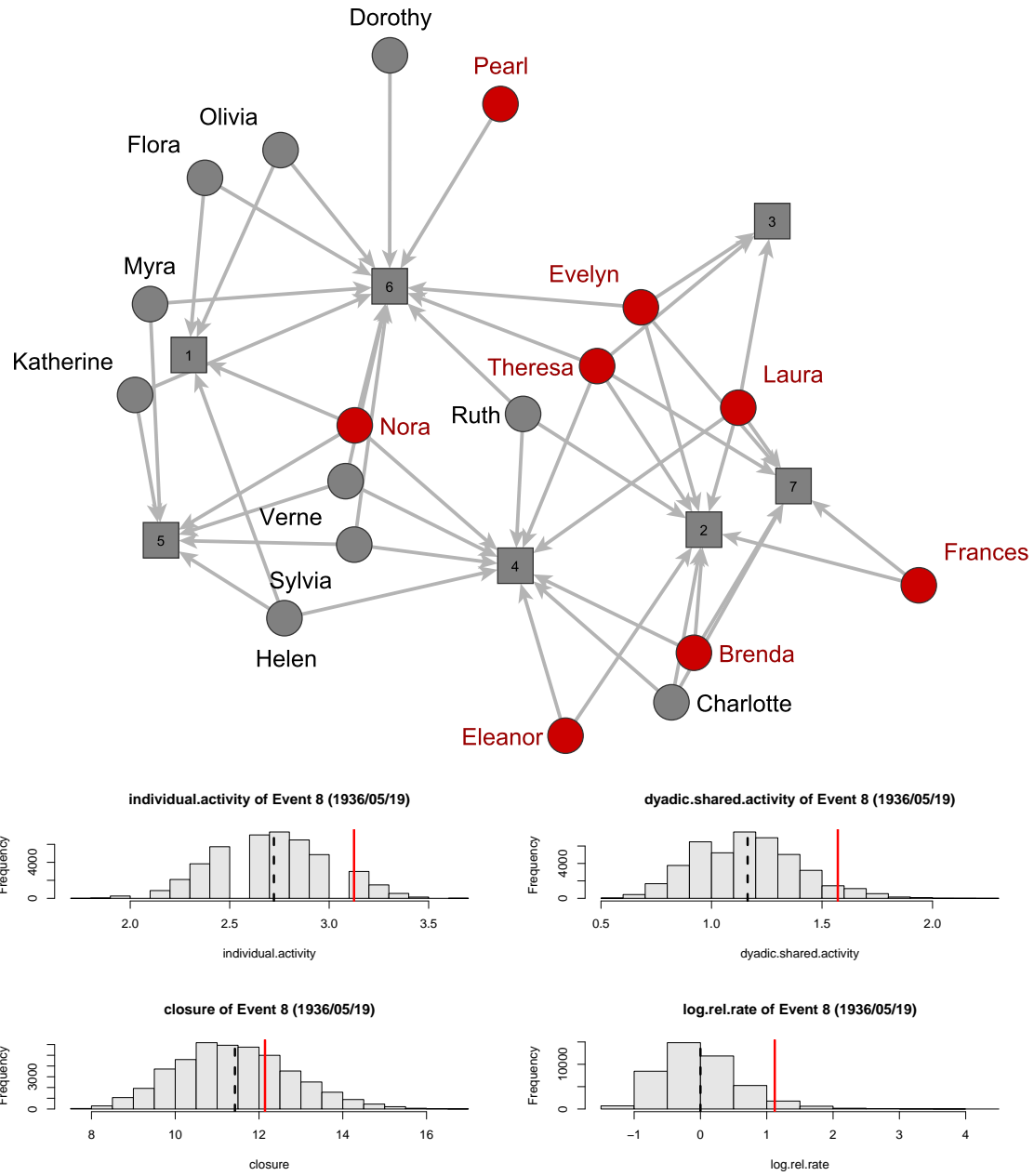


Figure 10: DGG network shortly before Event 8 whose participants are displayed in red. Event 8 connects a subset of the participants of Event 2 with two “outsiders” (*Nora* and *Pearl*). Thus, it is not entirely – but mostly – within a given cluster.

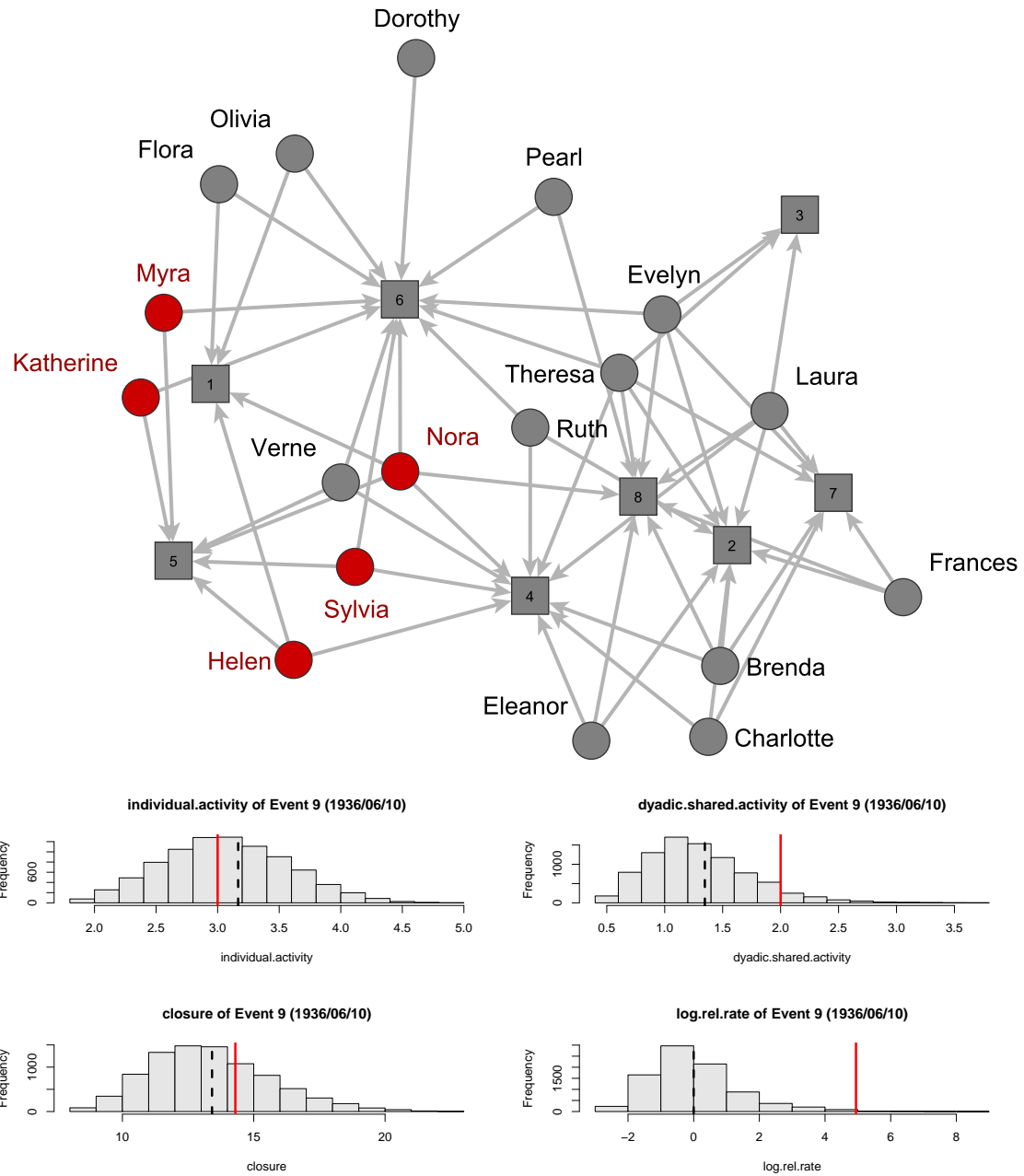


Figure 11: DGG network shortly before Event 9 whose participants, displayed in red, are participants of prior Event 5 and, thus, belong to one of the two emergent clusters.

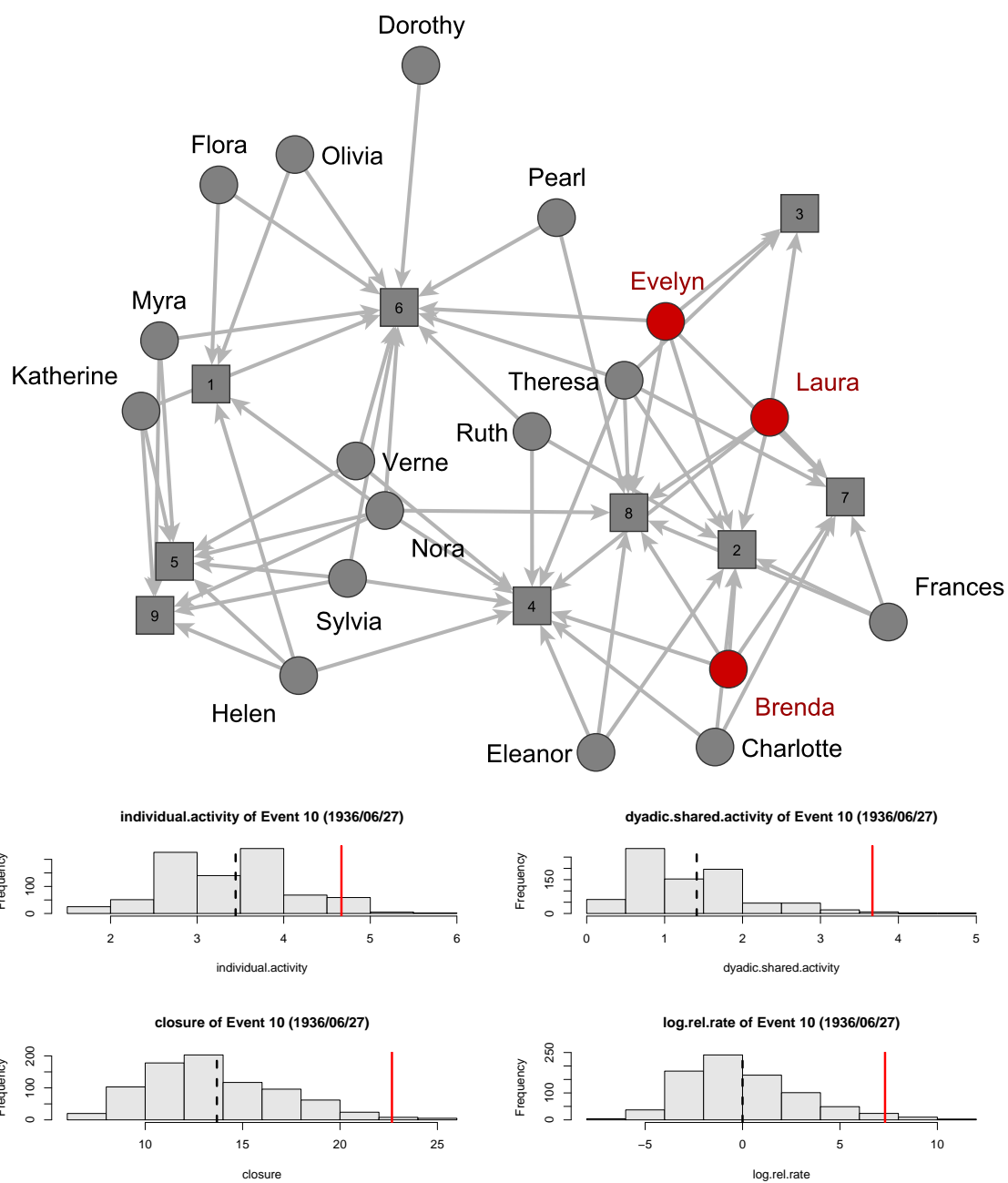


Figure 12: DGG network shortly before Event 10 whose participants are displayed in red. Event 10 occurs within the cluster centered around prior Event 2.

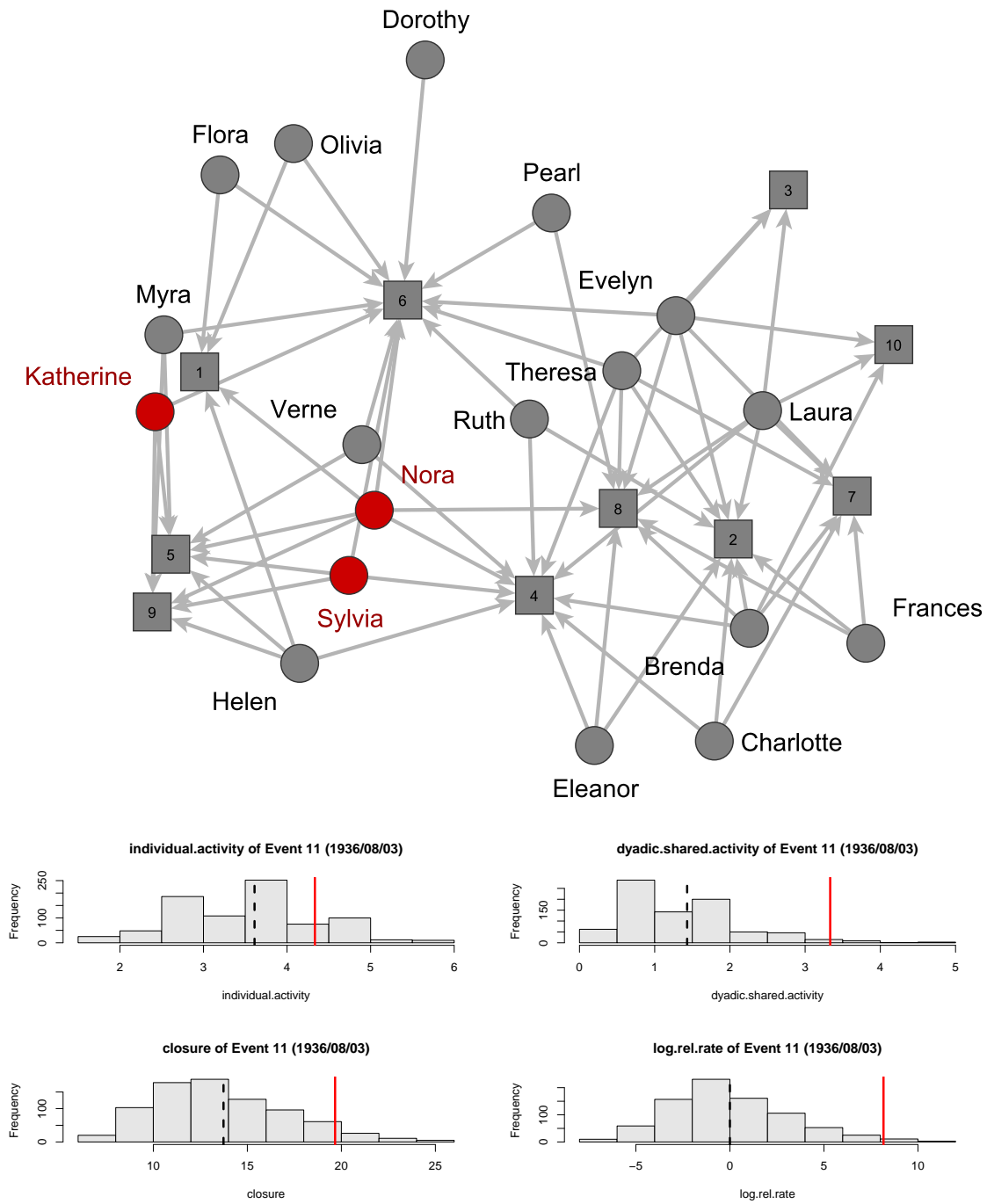


Figure 13: DGG network shortly before Event 11 whose participants, displayed in red, participated in prior Events 5 and 9.

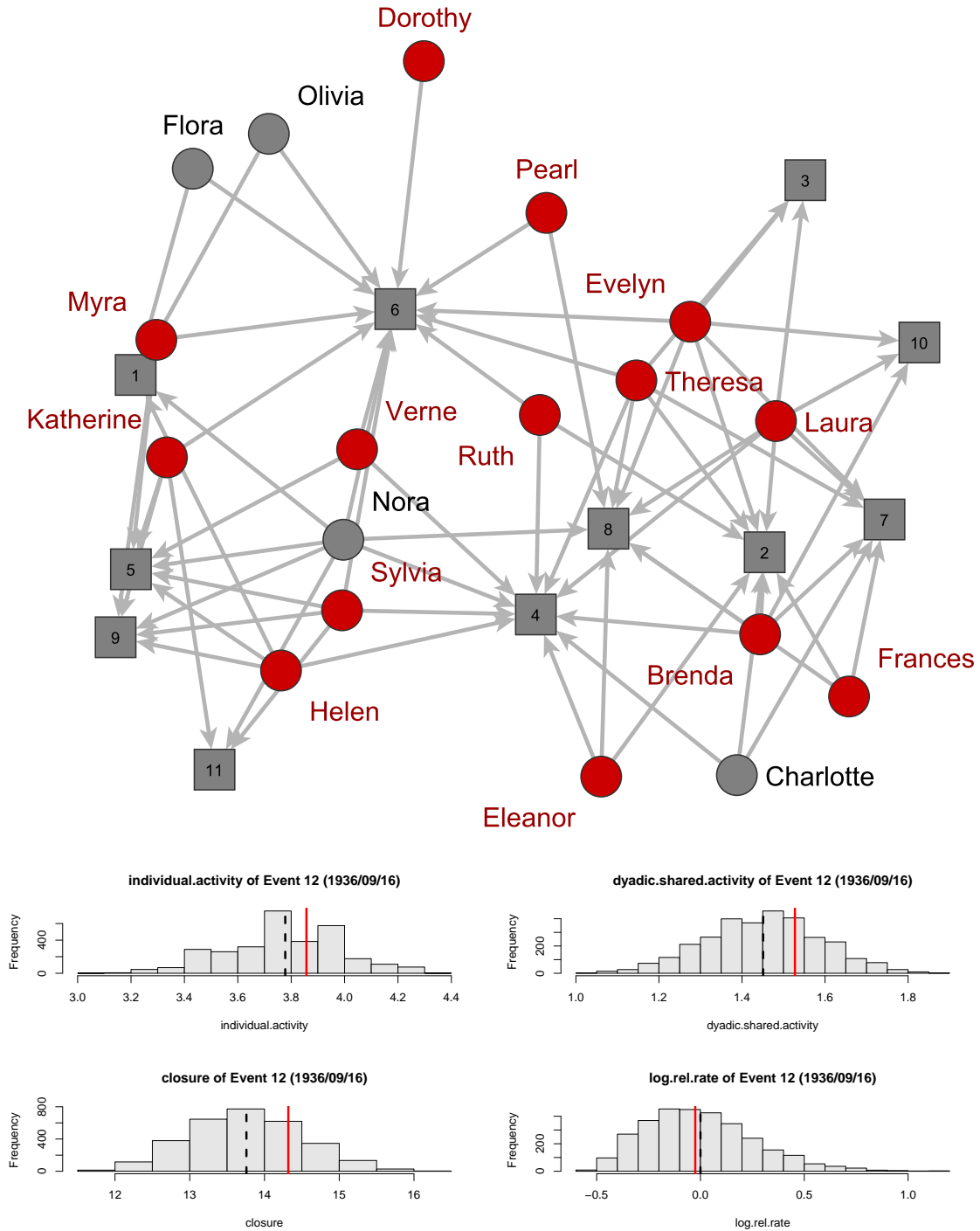


Figure 14: DGG network shortly before Event 12 whose participants are displayed in red. Event 12 is the second event (besides Event 4) whose probability is slightly below average, according to the model. It is an event spanning disparate parts of the population. However, in contrast to Event 6 – the other bridging event – (see Fig. 8), Event 12 actually closes a relatively large number of two-paths which are not well explained by prior shared activity.

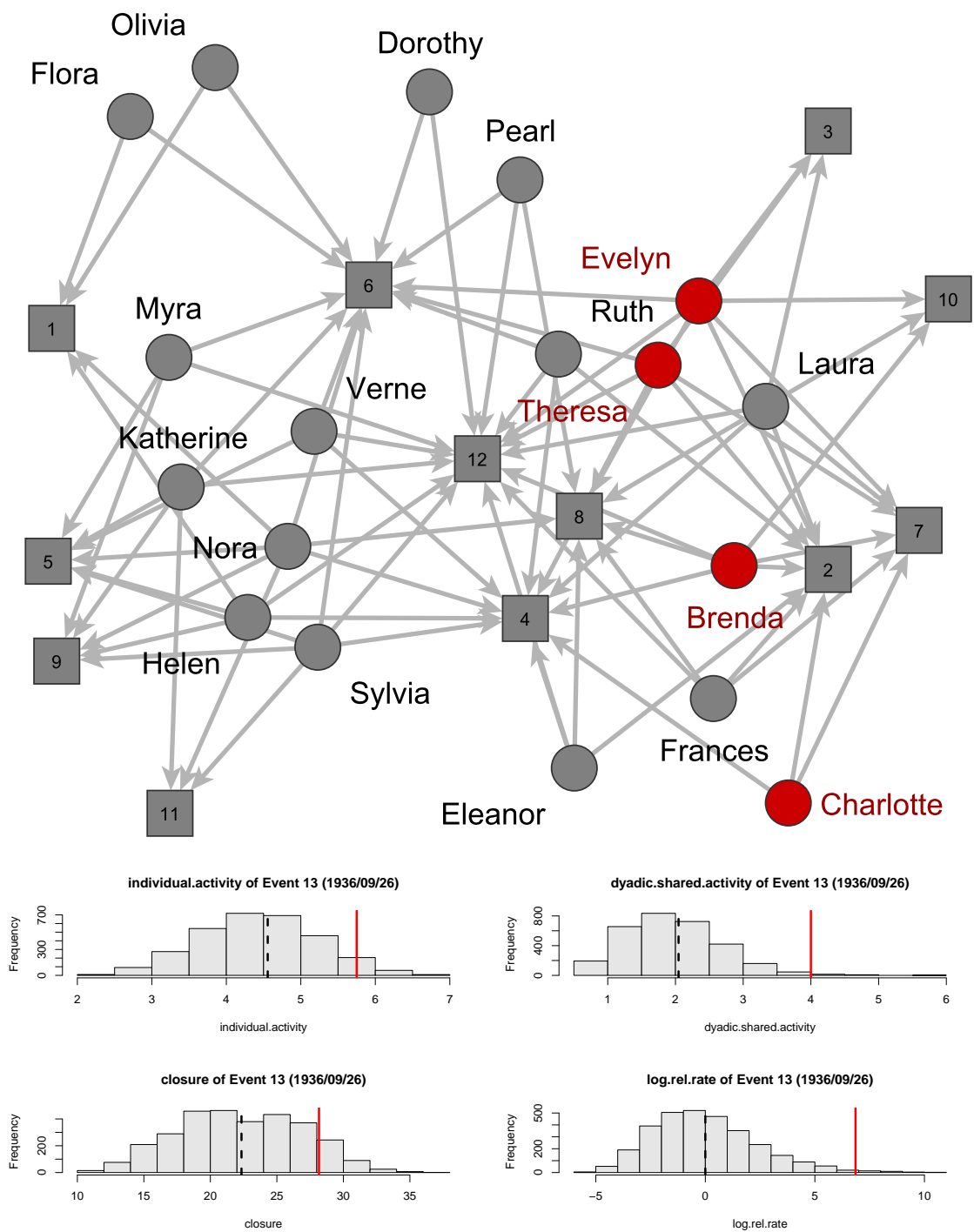


Figure 15: DGG network shortly before Event 13 whose participants, displayed in red, participated in prior Event 2.

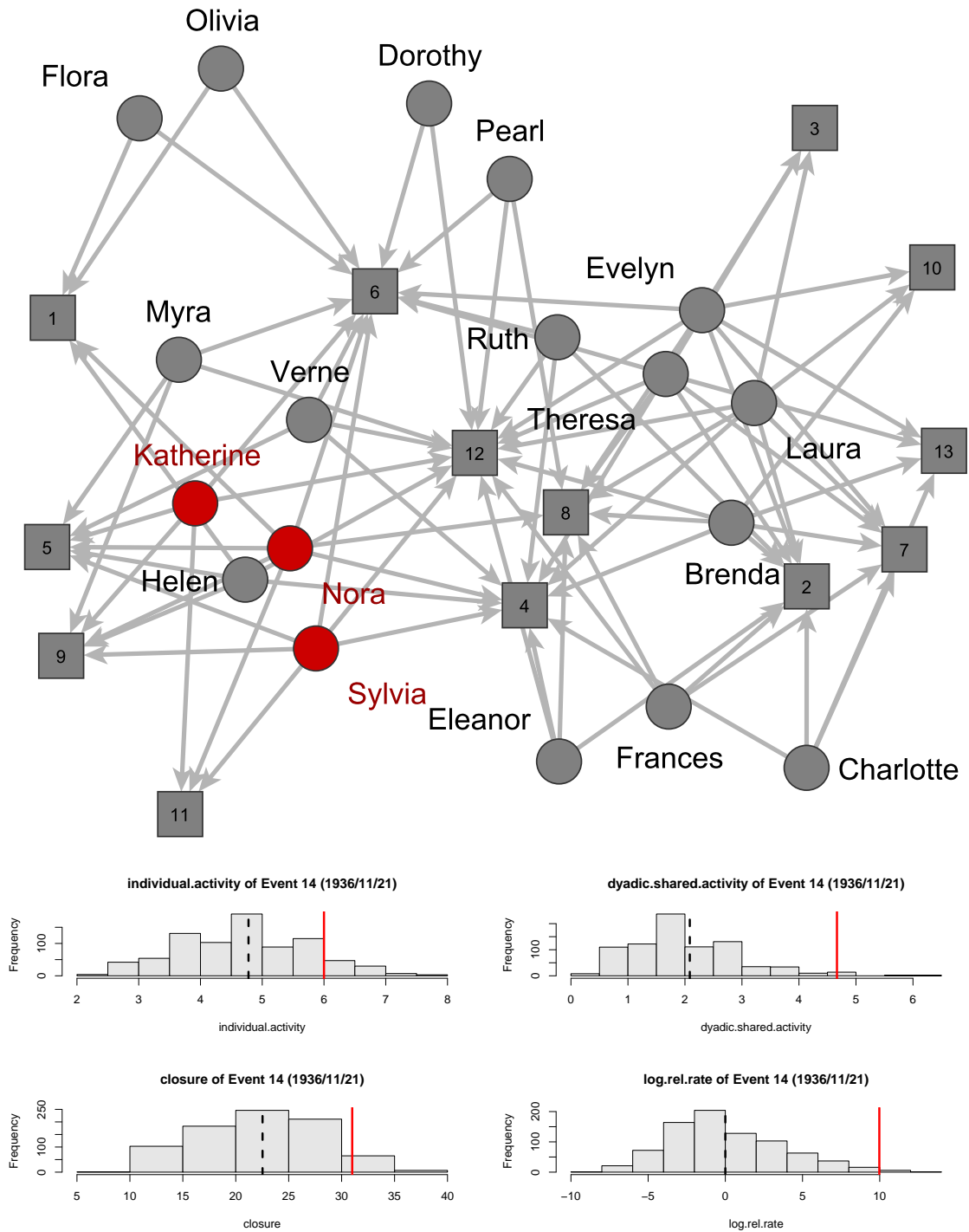


Figure 16: DGG network shortly before Event 14 whose participants, displayed in red, are exactly those of Event 11.

