2. In section 6.7 of that same paper, the authors say their preferred tool for assessing balance is an empirical QQ plot. What's a QQ plot? Compare and contrast the use of QQ plots and a balance table. Neither of these tools in dominate, so what are the benefits and drawbacks to each?

A QQ plot is used for comparing the distribution of two groups. You put the quantiles of one group on the x-axis and the quantiles of the other group on the y-axis. A lot of times a quantile-quantile plot compares a set of observations to a theoretical distribution (e.g., Gaussian). If the points in the QQ plot look like they're close to the x=y line then the two distributions are quite similar. Strong departures from the x=y line suggest the two groups have different distributions. (Here's a sssuuper detailed, basic explanation. You may want to watch at 1.5 speed.) Instead of using a QQ plot, you could overlay two histograms to see if they look similar.

When used in matching, the QQ plot compares the distribution of a particular variable (example: age) between the treatment group and the control group. Ideally, the quantiles of the two groups will lay right on the x=y. If the points are substantially off the x=y then we've detected a difference in the two groups.

The QQ plot looks at the entire distribution of a given covariate. Perhaps the mean of age for the treatment group and the control group are the same but the treatment group has a very small standard deviation of age, whereas the control group has quite a much larger standard deviation. This kind of imbalance may not be obvious if one is only looking at the standardized mean differences.

In practice, balance tables tend to only report the means of the covariates, which can lead to oversights due to outliers (or higher moment imbalances). A plus for the balance tables is that they report the means and standard deviations of each covariate so readers can get a sense of if the population under study may be informative to the population the reader is interested in. Balance tables can also be quite information-dense, reporting information about 50-60 variables per page.