Modelling the data

1. Initial idea: use models from my thesis to fit the data. So, have 30 min as a unit of time (as min. size of research unit is 30 min). So total units = 110\*48. And actions as whether or not to do a unit at each timestep.
2. But in the end, I would like to display my results in days, so why do this. Plus, why deal with decisions between times of the day (irrelevant).  
   So, instead, the decisions are modelled as number of units one wants to do each day (110 days). Then the transition function is specified by the binomial distribution from the repeated Bernoulli trials with prob = efficacy. Units that are not used for work are used for shirk. There is also some ambiguity about how the reward schedule should be treated.   
     
   a. Rewards at the end with min. 14 units needed to be completed and extra credits upto 22 credits. As usual, with discount=1, all the actions are equally good until some point in the end depending on the efficacy. With discount < 1, better to shirk in the beginning, and then complete as many as possible in the end (starting times depend on the efficacy). Even for low efficacies, can finish in a few timesteps.   
     
   b. Always better to choose action of highest number of units to do, even if this > 22 (even if beyond this, there is no extra reward). Why not try a small number so that effort is minimized – this requires distributing in time but due to discounting, it looks very unattractive to do some effort now. So instead of trying small number each timestep, minimize effort by trying as many as possible at the end (and the higher you try, higher you can achieve even if marginally more). What if effort is increased:   
     
   c. So that’s why if 30 units are allowed per day, one can try 30 and achieve > 22 – this is not how people actually do it, they do it sequentially (so they could stop at 22). So I keep max number per day to 22.

d. What if the number of allowed units per day is lower, to allow spreading out. The logical number is the max. number of units a person completed in the task (4hrs=8 units). So needs to start a little earlier, but again last few days enough (and again try the max possible). Only spreads out a bit.

1. But above is without softmax rule, maybe this must be implemented to get a good distribution. This does make it better, but still not that much of a difference between efficacies (even for lowest efficacies, start 4-5 days before). Too many timsteps!   
     
   a. But can’t we just adjust discount factor and rewards to get same structure? There is also the notion of efficacy and by allowing agent the action of trying all in one day vs in a week, having 10 days is like having a lot of timesteps. One solution is to have fewer units available per day but seems a bit arbitrary.
2. Probably people are allocating work over weeks rather than days. It doesn’t matter when during the week they decide to do. Aggregating data over days into weeks doesn’t eliminate much information (the patterns are preserved), so good to use this. Playing with efficacy (with some softmax noise) reproduces many/ most patterns
3. Reward schedule: when rewards are given immediately if 7 credits are completed, then there are no longer any delays and the optimal thing is to complete as soon as possible. So maybe this is actually a problem with no delays!