Project 2

Physics 250 Econophysics

In this project, the goal is to write a computer code that will implement the Stauffer-Penna trading model. The output will be a series of plots similar to Figure 1a of the Stauffer-Penna paper.

1. You will use equation 2b of SP to select the size of our clusters of traders. In other words, we will compute our cluster size from the equation:

$$n = \frac{n_0}{s^\tau} \qquad (*)$$

where s is the size of the trader-cluster, n_0 is the number of trials, $\tau = 5/2$, and n is a random number.

More specifically, we will assume that $p \equiv \frac{n}{n_0}$ is a uniformly distributed random number between [0,1]. So to pick a cluster size of traders at random, you will generate a random number p, and then solve equation (*) for s.

2. You will assume 1 trade (1 cluster) on each time step, and you will use 10,000,000 time steps, or the maximum you can use for your dev environment (Excel may have a limit on cell sizes). Note that there are about volume=100,000,000 or more trades of the SPY per trading day.

3. You will also assume that on each time step, the cluster of traders always either buys or sells (i.e., $\Phi_i = \pm 1$ in equation (1) of Stauffer-Penna, thus $\Phi_i \neq 0$). Assume that a trader buys or sells with equal probability (i.e., a 50-50 chance of either buying or selling). Thus you will again generate a random number to determine whether the traders buy or sell.

4. You will then generate a time series of 10,000,000 price changes using equation (1) of Stauffer-Penna. Note that in the case of interest, where we have only one cluster per time step, that $\Delta P(t_i) = P(t_i) - P(t_{i-1})$ and $P(t_i) = s_i \Phi_i$ (i.e., no sum, since there is only 1 cluster on each time step).

5. From this time series, you will generate a histogram $H(\Delta P)$ of the 10,000,000 price changes (if you don't know what a histogram is, see http://en.wikipedia.org/wiki/Histogram). Note that all dev environments such as MatLab, Excel, and IDL have built-in routines to compute histograms from a time series, so you shouldn't have to do any work.

6. This histogram corresponds to a PDF, a probability density function (which is just a histogram normalized so that the area under the curve is 1). This PDF has a mean μ (or

expectation) and a standard deviation σ, which is the square root of the variance, that can be computed by standard methods (see http://en.wikipedia.org/wiki/Mean and http://en.wikipedia.org/wiki/Variance). Again note that dev environments such as MatLab, Excel, and IDL have built-in routines to compute these, so again you shouldn't have to do any work.

7. You will then compute the Gaussian histogram $H_G(\Delta P')$ corresponding to the same number of trades, and the same μ, σ as for $H(\Delta P)$:

$$H_G(\Delta P') = \frac{10{,}000{,}000 * \exp\left[-\left(\frac{1}{2}\right)\left\{\frac{\Delta P' - \mu}{\sigma}\right\}^2\right]}{\sigma\sqrt{2\pi}} \qquad (**)$$

Note that to compute $H_G(\Delta P')$, you just compute a vector of equidistantly tabulated price changes $\Delta P'$ from $["-\infty","\infty"]$ and numerically evaluate the expression (**). Note that by $"\infty"$ we just mean a large number, say 1000.

8. Now take the Log_{10} of both $H(\Delta P)$ and $H_G(\Delta P')$ and plot. The result should be a plot that looks like the below figure. In the figure, the blue symbols are for $H(\Delta P)$, and the red symbols are for the Gaussian $H_G(\Delta P')$.

9. Now repeat this exercise using $\tau = 2,\ 2.222,\ 2.857$, and plot the results. The blue histogram $H(\Delta P)$ is similar to real distributions of tick-by-tick trading statistics, displaying *Leptokurtosis.*