Dear Professor Giles

I've thoroughly enjoyed your recent posts and associated links on distributed lags. I'd like to throw in a slightly different perspective.

To give you some brief background on myself: I did a PhD in econometrics 1993-1998 at Southampton University. I remain good friends with Jan Podivinsky who I suspect you may know from Monash days. My supervisor was John Driffill, a macro guy, so I actually completed most of my PhD in Florence working with Mark Salmon who remains a close friend. Mark and I have worked together since in three stints, at a research group he founded, at the Bank of England and at a hedge fund. I now manage capital and am heavily influenced by my study of econometrics and in particular exploring the historical foundations of many things that today that look new and funky.

I wanted to draw attention to the fact that many finance practitioners have long used 'models' that in my view are robust and heuristic versions of nonlinear ADL models. I'm not sure this interpretation is as widely recognised as it could be.

Probably the single most popular and resilient technical analysis indicator is formed by taking the difference between two moving averages on price levels. The most common inference traders make is that when the short-run trend is above the long-run trend then it indicates prices are more likely to continue rising.

In practice, for many years traders have used exponential moving averages to form the short-run ('fast') and long-run ('slow') averages. Usually (as for example would be standard practice via commercial financial systems like Bloomberg and Reuters etc) people eyeball price level charts with the technical indicator below it. They then make trading decisions.

This process has long since been automated and forms the basis of the trend-following industry, one that has grown in size rapidly in the last decade and continues to do so. There is at least a trillion dollars of capital (the 'industry' AUM is by most guesses around 300-500 mm, and leverage is probably 3 times) being explicitly traded by simple variations on this rule.

While traders (and probably most humans!) tend to more naturally think in terms of price levels, if we think in terms of price differences then the implicit ADL function that technical traders use is very close to the logarithmic one that J.N.M Wit discussed in his note on Irving Fisher. Specifically, as soon as we take differences of two EWMAs we create a decay function that first rises to a peak and then decays again.

For example, for a EWMA scripted in R as:

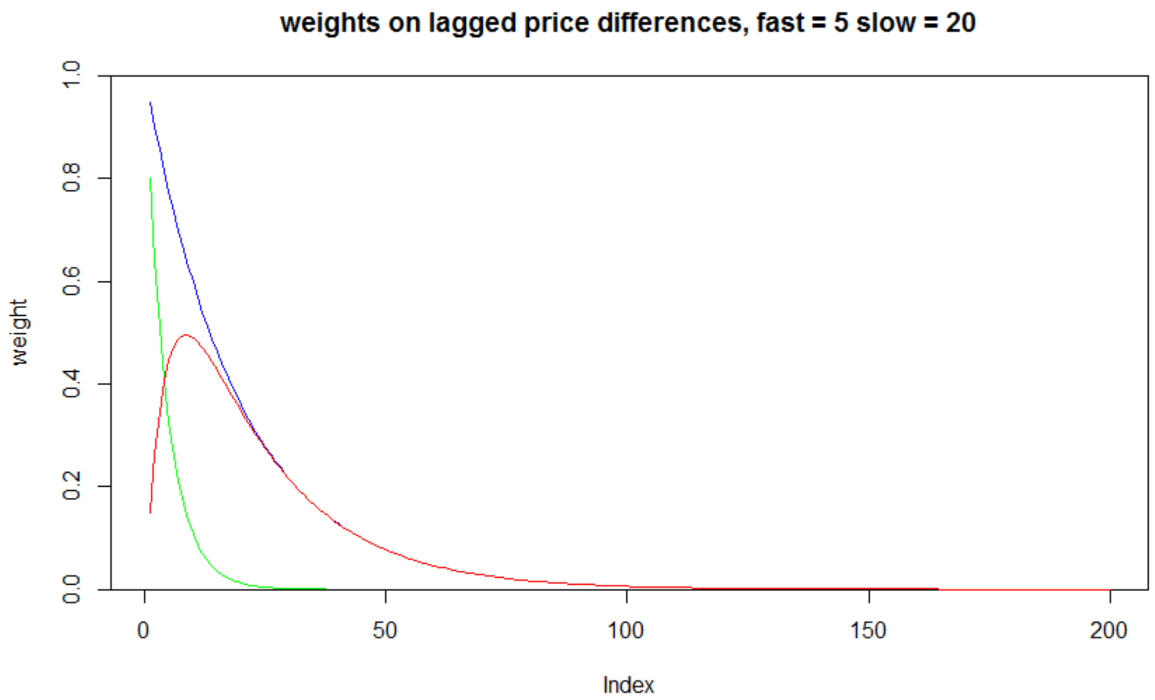
```
EWMA <- function(x,alpha,lags=0) {
  # alpha is the coef on the actual data so higher is alpha less memory ha
s the EWMA
  if(lags>0) alpha <- 1/lags
  s <- x
  for(i in 2:length(x)) s[i] <- alpha*x[i] + (1-alpha)*s[i-1]

  s

}
```

Then

fastSpeed <- 5

slowSpeed <- 20

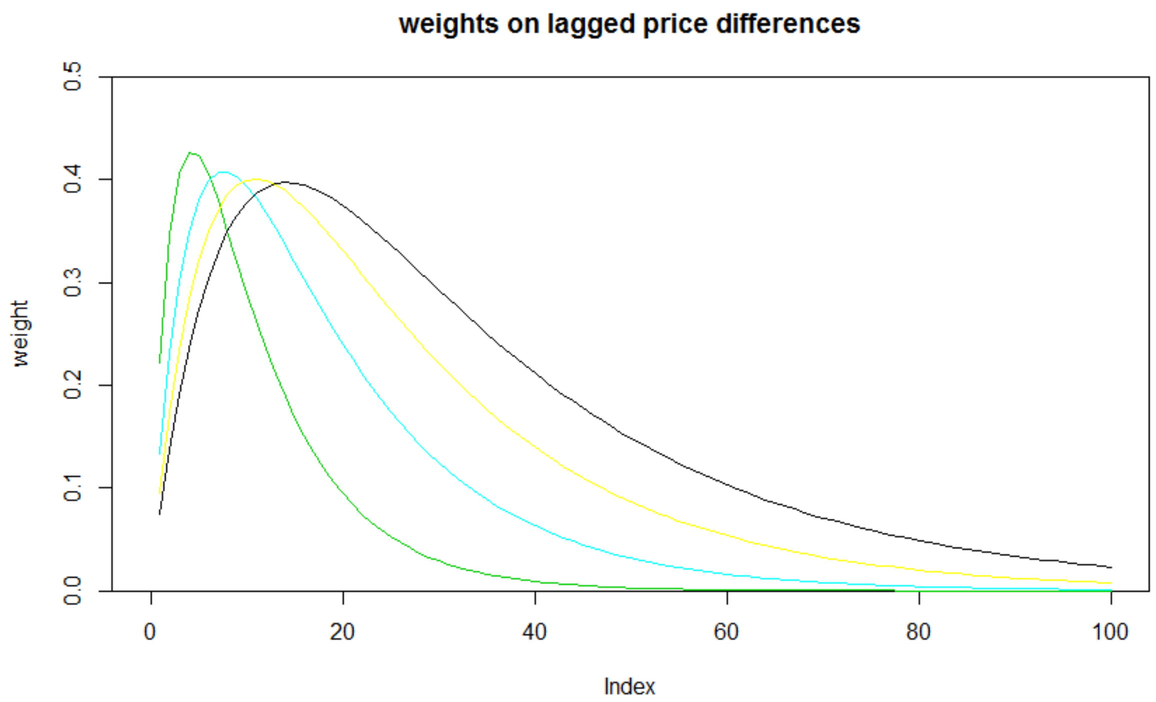
```
weightsFast <- EWMA(c(1,rep(0,200)),lags=fastSpeed)[-1]

weightsSlow <- EWMA(c(1,rep(0,200)),lags=slowSpeed)[-1]

netWeights <- weightsSlow-weightsFast

plot(weightsFast,col="green",type="l",ylim=c(0,1),main=paste("weights on
lagged price differences, fast = ",fastSpeed," slow =
",slowSpeed,sep=""),ylab="weight")

lines(weightsSlow,col="blue",type="l")

lines(netWeights,col="red",type="l")
```



Varying the fast and slow speeds is a nice way of creating different shapes, similar in spirit to the idea of modelling the ADL coefficients as following a polynomial (Almon).

```
for(fastSpeed in seq(1,10,2)) {


    slowSpeed <- 3*fastSpeed

    weightsFast <- EWMA(c(1,rep(0,100)),lags=fastSpeed)[-1]

    weightsSlow <- EWMA(c(1,rep(0,100)),lags=slowSpeed)[-1]

    netWeights <- weightsSlow-weightsFast

    if(fastSpeed==1)

      plot(weightsFast,type="l",ylim=c(0,.5),xlim=c(0,100),main="weights
      on lagged price differences",ylab="weight")
```

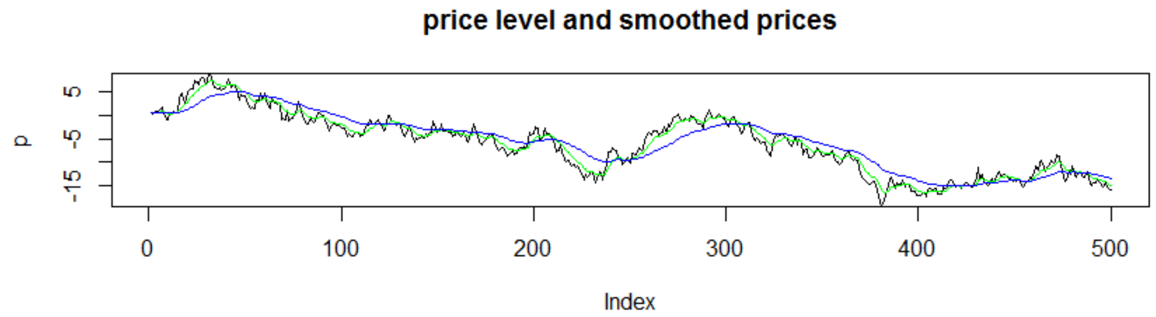
```
    else lines(netWeights,type="l",col=fastSpeed)
}
```

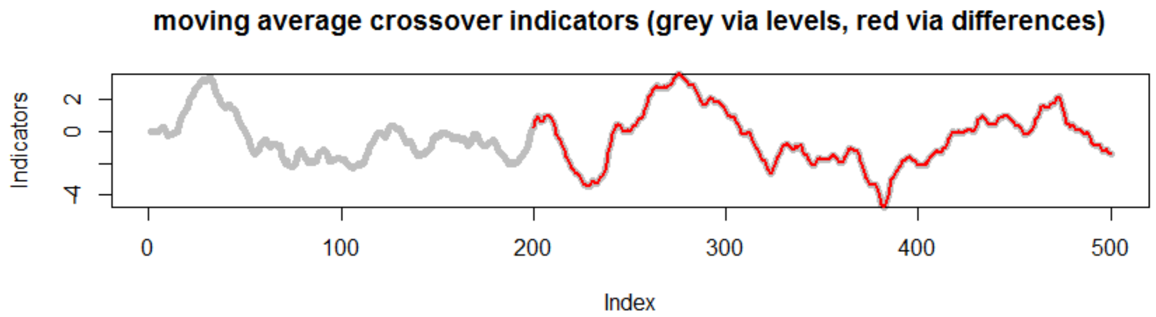


Just to demonstrate specifically this code makes the point if not obvious

```
r <- rnorm(500)
p <- cumsum(r)
par(mfrow=c(2,1))

fastSpeed <- 5
slowSpeed <- 20

# working in LEVELS
fastEWMA <- EWMA(p,lags=fastSpeed)
slowEWMA <- EWMA(p,lags=slowSpeed)

plot(p, type="l",main="price level and smoothed prices")
lines(fastEWMA,type="l",col="green")
lines(slowEWMA,type="l",col="blue")

levelsSignal <- fastEWMA-slowEWMA
```

```
plot(levelsSignal, type="l", col="grey", lwd=5, ylab="Indicators",
main="moving average crossover indicators (grey via levels, red via
differences)")

#working in DIFFERENCES
weightsFast <- EWMA(c(1,rep(0,200)),lags=fastSpeed)[-1]
weightsSlow <- EWMA(c(1,rep(0,200)),lags=slowSpeed)[-1]

netWeights <- weightsSlow-weightsFast

returnsSignal <- filter(r,netWeights,sides=1)

lines(returnsSignal,type="l",lwd=2,col="red")
```





(N.B. The filter function simply applied doesn't burn in so the red line only starts 200 days in).

A few comments

'Estimation'/ model selection

In practice traders tend to 'prefer' different combinations of fast and slow decays (or in other words different shaped decay functions) in terms of their relative ability to make money when used in conjunction with a trading rule. I see this as consistent with the literature on using economic cost functions to select and estimate models.

Also, the more automated practitioners tend to use more than one decay function at once, by weighting different shapes, e.g. they may equally weight a 5-10 day, 10-20 day etc.

Now, with the same input series this simply amounts to a single net decay function, but maybe their preference for thinking about it like this is based on a second level of function decomposition, a kind of constrained optimisation working with a small number of basis functions instead of a full nonparametric approach like a spline. This is also sympathetic to a heterogeneous participant view of the world, and MIDAS.

The delayed 'peak'

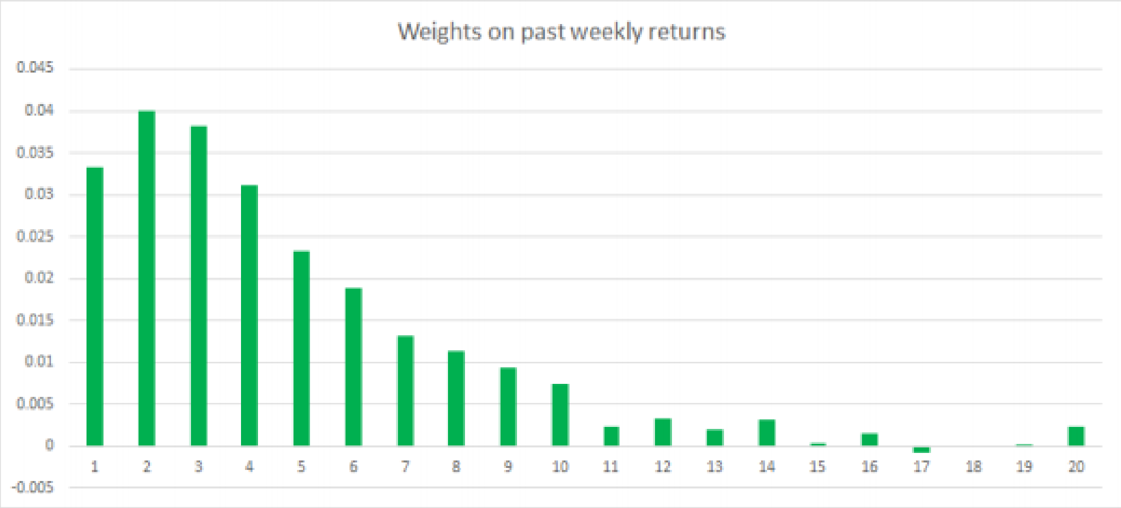
So from a trading perspective we can think in terms of the weighting function as providing an extrapolative expectation of future price movements. The desire to have a peak input sometime in the past is consistent with the idea that we need to wait for shorter-term price action to be compared to longer-term. i.e. we don't react immediately to a sudden rise in price. We need to put that rise in price in perspective. This leads to the maximum impact of past returns to be some days or weeks in the past.

I think there's another angle based on frictions. If we think like a Keynesian beauty-contest participant then when we see prices change we have to consider how others might interpret and act on it. Suppose we think that in general rising prices lead to higher expectations and that other traders will therefore end up buying which pushes prices up. It may well be the case that different participants react at different speeds and in any case even if traders wanted to immediately adjust their holdings they are constrained by frictions like market liquidity, trading costs etc. In this case we might reasonably expect the market reaction to be drawn out. I realise this is an incredibly loose argument.

What's the underlying time unit? Obviously the shape of decay could look different were we to be using daily, weekly or monthly data. So for example a daily model may well benefit from the fully nonlinear shape incorporating the rise and then fall in weights. But a monthly model might be fine with the Fisher approximation as shown in Wit Figure 2, ignoring the initial rise in weight. In the absence of stronger theory or insights about expectations formation then I guess it is largely an empirical issue. In my own work I have found that speculator positioning in futures markets can be best fitted by a nonlinear approximation that does indeed peak a few weeks in the past, see for example the chart below taken from a note I wrote here

http://neuronadvisers.com/Documents/Trend%20Follower%20Capacity%20-%20A%20Cautionary%20Observation:

To explore the influence of past prices changes on these positions we run a simple regression. On the left hand side of the equation is the dependent variable the weekly net active position, and on the right hand side the explanatory variables are lagged price changes. The position data is weekly so we use lagged weekly price changes (measured as log price differences divided by volatility).



The chart above shows the regression weights (coefficients). To be clear, the coefficient reported with index 1, is the weight on the previous week's price difference. The weight on index 2 uses the week before that and so on. The model has a good fit explaining nearly 50% of the variation in the positioning data (R^2=0.49).

Econometricians / technical analysts / heuristics etc

Finally, I would just like to repeat my thanks for your work on this subject and bringing people's attention to the older literature which is so often incredibly insightful about today's issues. My passion is all about uncovering (rediscovering) links between different areas of research and in this example I feel that traders and finance guys could learn a lot from seeing how econometricians and statisticians have developed techniques that are often more rigorously founded than theirs, but on the flip-side I think econometricians could spend more time thinking about why apparently simple (or more likely to be considered daft!) heuristic techniques might actually be a lot smarter than they first appear.

Many thanks again

Robert