Personal Finance and Forecasting

AGEC 317: Economic Analysis for Agribusiness Management
Instructor: Michael Black

TEXAS A&M UNIVERSITY
Agricultural
Economics

An amuse-bouche

Moving beyond supply and demand analysis, we can do some cool things with econometrics that can benefit us personally. This lecture is about the application of econometrics to building a portfolio of investments, and forecasting future behavior. But we are looking at the tip of the iceberg! For more, take a class dedicated to financial analysis.

Time Series

Time series data had many observations over time for a single entity.

Year	Inflation	Unemployment
1948	8.1	3.8
1949	−1.2	5.9
1950	1.3	5.3
1951	7.9	3.3
.	.	.
.	.	.
.	.	.
1998	1.6	4.5
1999	2.2	4.2
2000	3.4	4.0
2001	2.8	4.7
2002	1.6	5.8
2003	2.3	6.0

Time Series

- Stock prices
- Crude oil
- Cattle prices
- Crime rates
- Housing expenditures
- etc

Investing

Let's say you want to invest in a or some stocks. You may be interested in:

- **Expected return**: what you expect an investment will return
- **Risk**: the likelihood that the real return will differ from the expected return

Investing

Return for a time-series of stock prices:

$$Return_t = \frac{(Price_t - Price_{t-1})}{Price_{t-1}}$$

Investing

Expected return for a time-series of stock prices:

$$ER = \frac{\sum_{t-T}^{t} Return_t}{T}$$

...or the average of returns over a time period with T increments.

Investing

Risk is the likelihood that the real return is different than expected return. Two broad types of risk:

- Systematic risk: risk that affects the entire system
- Unsystematic risk: risk that affects an individual stock (eg company)

Investing

It is hard to protect yourself from systematic risk: COVID-19 is affecting everyone, so how could you have adjusted your investment strategy to protect against this?
We will focus on protecting ourselves against unsystematic risk, through *diversification*.

Investing

Two broad ways of measuring the risk of investing in a stock:

- **Standard deviation**: measures the volatility of the stock by looking at the stocks historical returns
- **Beta**: measures the volatility of the stock by comparing returns of a stock to a standard market benchmark, like the S&P 500.

Standard deviation is a formula we have already seen:

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

If the standard deviation is high, the stock is volatile. This is considered high risk, which is attractive for aggressive-growth investment strategies, and unattractive for conservative-growth strategies.

Investing

Beta is the measurement of risk of a stock compared to a market benchmark:

$$\beta = \frac{COV(R_s, R_m)}{VAR(R_m)}$$

where R_s is the return of an individual stock, and R_m is the return of the market benchmark.

Investing

Hold the phone! Does that definition of β look familiar? That's the *exact* same definition we derived for the estimate of a slope coefficient in a linear regression! Clearly, we'll be able to use a linear regression here...

Investing

Importantly, we can diversify risk by creating a *portfolio* of investments. The portfolio is a collection of stocks we invest in simultaneously, and that portfolio can be designed to minimize the amount of unsystematic risk we are exposed to.

Investing

The standard deviation of a portfolio:

$$\sigma_p = \sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1 w_2 \sigma_1 \sigma_2 \rho_{1,2}}$$

where w_i is the weight (in percentage terms) of a stock in the portfolio, σ_i^2 is the variance of returns of stock i, and $\rho_{1,2}$ is the correlation between stock 1 and 2. As you get more stocks in the portfolio, this definition becomes *long*. We won't do that here.
The beta of a portfolio:

$$\beta = \frac{COV(R_p, R_m)}{VAR(R_m)}$$

where R_p is the return of the portfolio, and R_m is the return of the market.

Investing

Let's pretend to invest some initial principle in a single stock and a portfolio of two stocks, and decide how to diversify...

Forecasting

So we can look at the past to understand current conditions. What about trying to predict the future? This is known as **forecasting**.

Forecasting

Some preliminaries:

- We aren't going to incorporate uncertainty, so these forecasting models aren't as good as they could be.
- "A model is always wrong, but a good model is useful"

Forecasting

We will focus on:

- Static models
- Finite discrete lagged models (FDL)

Static model

A static model is:

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

where we say that β_1 is the *contemporaneous* effect of x on y.

Static model: examples

$$\begin{aligned} inflation_t &= \beta_0 + \beta_1 unemp_t + u_t \\ murder_t &= \beta_0 + \beta_1 conviction_t + \beta_2 unemp_t + u_t \end{aligned}$$

Finite discrete lag models

In FDL models, we allow the independent variables to be lagged. The x of today *and* yesterday are allowed to affect y today. In fact, we can lag x any number of times:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q}$$

Time trend

Huge potential problem: is y changing because of x, or is it just naturally changing over time? We can get a better answer with a **time trend**.

Forecasting

With time series data, we can make predictions about some outcome in the future

- Linear trend models
- Moving average (MA) models
- Autoregressive (AR) models

Before any model, you should plot the data as a line graph: y_t on the y-axis, and time on the x-axis.

Linear time trend

$$y_t = \alpha + \beta_1 t + \varepsilon_t$$

where $t = (1, \cdots, T)$. Then, β_1 measures the contemporaneous change in y given a one unit passage of time. We can add other explanatory variables:

$$y_t = \alpha + \beta_1 t + \delta_0 x_t + \delta_1 x_{t-1} + \varepsilon_t$$

and β_1 measures the same thing, but now δ_0 measures the contemporaneous effect of x on y, *holding time constant*.

Nonlinear time trend

Exponential time trend:

$$\ln(y)_t = \alpha + \beta_1 t + \varepsilon_t$$

Quadratic time trend:

$$y_t = \alpha + \beta_1 t + \beta_2 t^2 + \varepsilon_t$$

Example

Consider these models on investment as a function of the housing price index. Without a time trend we get a **spurious relationship**.

$$\log(\widehat{invpc_t}) = \underset{(.043)}{-.550} + \underset{(.382)}{1.241}\log(\widehat{price_t})$$

$$\log(\widehat{invpc_t}) = \underset{(.043)}{-9.13} - \underset{(.382)}{.381}\log(\widehat{price_t}) + \underset{(.0035)}{.0098}t$$

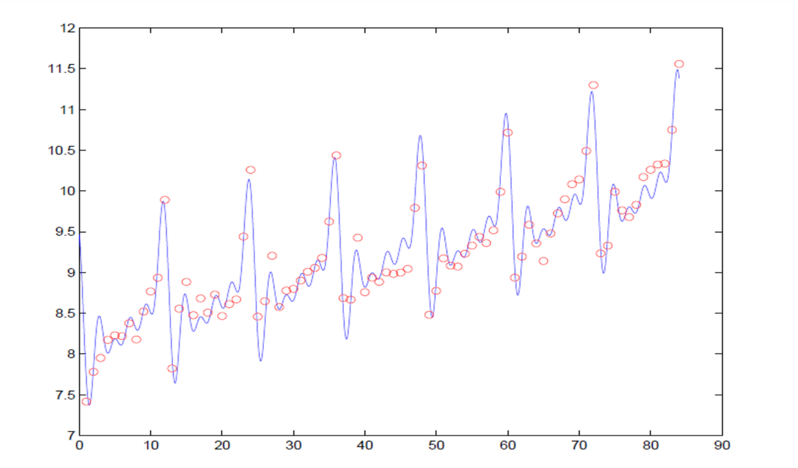
Seasonality

What if we think y depends on the season? Maybe demand for consumer goods spikes in November and December. We can capture these seasonal effects with dummy variables:

$$y_t = \alpha + \beta_1 feb_t + \beta_2 mar_t + \cdots + \beta_{11} dec_t + \cdots$$

where we exclude one base case.

Seasonality



Time series regression assumptions

- Linear in parameters
- No perfect colinearity
- Zero conditional mean of error
- Homoskedasticity
- No serial correlation (error between time is uncorrelated)
- Normality of errors

Time series regression assumptions

Under the first three assumptions, OLS is unbiased. Under the first five assumptions, OLS estimators are BLUE. Under all six assumptions, we can make the same type of inferences from cross-sectional data on time series data.

Moving average

With MA models, the outcome of interest in the future is predicted using average of the last p observations:

$$\hat{y}_{t+1} = \frac{y_t + y_{t-1} + \cdots + y_{t-p+1}}{p}$$

Autoregressive models

With AR models, the outcome of interest in the future is predicted using a linear combination of the last p observations:

$$\hat{y}_{t+1} = \alpha + \beta_0 y_t + \beta_1 y_{t-1} + \cdots + \beta_{t-p} y_{t-p+1}$$

With p lags, the model above is known as an AR(p) model.

How to forecast

Linear time trend

1. Recover coefficients on the time trend model
2. Use coefficients to estimate any number of periods ahead

How to forecast

Simple moving average:

1. Decide on how many periods back to go
2. Use the average of the previous periods to predict one step ahead
3. Use the predicted step ahead plus other previous periods to predict the next step, and so on...

How to forecast

AR(p) models

1. Choose number of lags (p)
2. Estimate model to recover $\hat{\beta}_0, \hat{\beta}_1, \cdots, \hat{\beta}_p$
3. Use the estimated coefficients to predict one step ahead, repeat

How to forecast

What is a good forecast? Lots of measures, but we'll focus on **MSE**: mean-squared error. The error is:

$$\hat{\varepsilon}_{t+h} = y_{t+h} - \hat{y}_{t+h}$$

And the MSE is:

$$\frac{1}{N} \sum_{j=t}^{t+N-1} \hat{\varepsilon}^2_{j+h}$$

Practice

Let's go back to our stock prices and do some forecasting, and decide which model works best for our data.