NORTHFIELD | modeling financial markets worldwide

Commodities, Exotic Currencies, and Crypto

T.J. Blackburn

Newport Seminar
June, 2022

Outline of Today's Presentation

- New and old asset classes
  - Commodities
  - Frontier currencies
  - Cryptocurrencies
- Approaches
  - Machine learning
  - A priori factors based on economics
  - Dealing with idiosyncratic returns
- Data issues
  - Regime changes
  - Data history and frequency

Northfield Everything Everywhere Model

- Global, multi-asset class risk model introduced in 2001
  - 90 factors in total, combination of specified exogenous and statistical factors for equities from our equity models
  - Multiple geographic regions, economic sectors
  - Observed yield curves for major markets, implied yield curves for small bond markets
  - Links fixed income credit risk to equity market risks via contingent claims model from Merton (1974)
  - "On demand" data creation for derivatives

Northfield Everything Everywhere Model

- Currently provides factor representation of more than six million individual securities, currencies and commodities
  - Annual horizon (update monthly), 10 Day horizon (updated daily)
- Extensions to non-traded asset classes including real estate, infrastructure and private equity

Factor Model Comments

- One possible approach to including commodities, currencies is simply to include each as its own additional factor (basically a full covariance matrix)
  - We reject the full covariance approach for several reasons
  - The observed correlation of commodities to each other and other asset classes is very unstable over time. A factor approach will separate persistent from transient effects
  - Adding lots of new factors to the model increases the potential for an ill-conditioned factor covariance matrix, which could impact the quality of forecast for all asset classes

NORTHFIELD | modeling financial markets worldwide

Commodities

Commodities in Global Model

- Represent commodity contracts in terms of existing Northfield factors, like any other security
- Two stage estimation procedure

- First subdivide the universe of commodities into four groups.
- Agricultural, Energy, Precious Metals, Industrial Metals
- Strikes a balance between variance explained by each of next two steps

Commodities Old Procedure – 1st Stage

- First stage: Using return history create four Principal Component Analysis (PCA) factor models
  - Orthogonal linear transformation to uncorrelated factor returns

$$r_{sec,i} = \sum_j w_{ij}\, r_{PC,j} \qquad \left(\langle r_{sec,i} \rangle = 0\right)$$

$$C_{PC,ij} = \langle r_{PC,i} r_{PC,j} \rangle = \sum_{kl} w_{jk} \langle r_{sec,k} r_{sec,l} \rangle w_{li} = \sum_{kl} w_{jk}\, C_{sec,kl} w_{li} = \sigma_{PC,i} \delta_{ij}$$

- Each commodity has exposures to the two (usually) strongest PC's for its cluster
  - Separates 'signal' from random 'noise'
- PCA updated monthly, rolling 60 month estimation

Commodities Old Procedure - 2nd Stage

- Second Stage: Restate the PCA factor exposures into factor set defined by the Global model.

$$r_{PC,i} = \sum_j l_{ij}\, r_{fac,j} + \epsilon_i$$

- Each commodity PCA factor is treated as a new security
- Factor loading are established by time series regressions of the PC returns against apparently relevant factors

Update

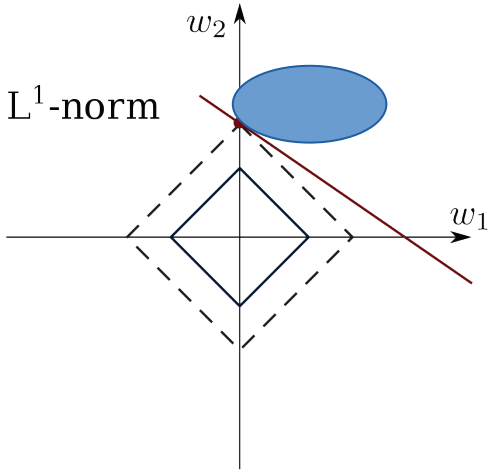
- First step objective and can easily be updated monthly
- Second step somewhat subjective and not as simple to update
- Automated model building or ‘machine learning’ techniques can update objectively

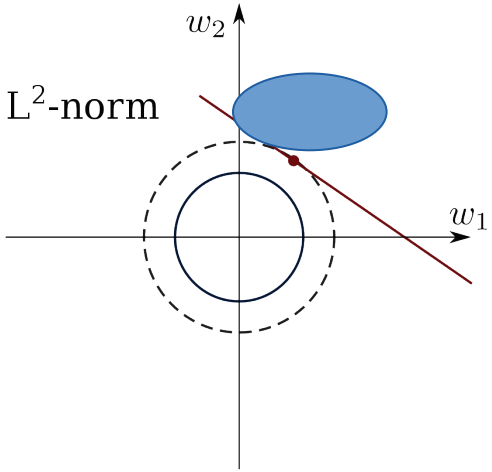
Selecting a method

- Tried a number of automated model building procedures:
  - Stepwise regression (bad) based on AIC , t stat, r^2
  - Mallow C, PRESS
  - Ridge regression, LASSO
- Simulated returns to securities with known loadings
  - Randomly selecting Northfield factors and loadings
  - Calculated returns from actual historic factor returns
  - Added random idiosyncratic returns
- Tested all methods
  - Evaluate based on prediction error, r^2, correct factors
- LASSO consistently performed best

LASSO

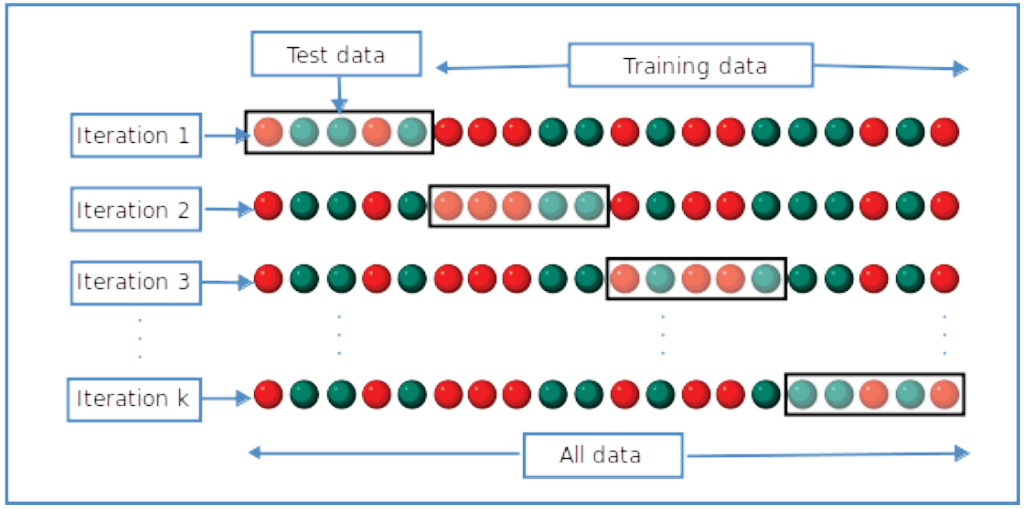
- LASSO combines factor selection with shrinkage toward zero
- Uses L1 penalty
  - Minimize $\left(y - \hat{\beta} \cdot X\right)^2 - \lambda \sum_i |\beta_i|$, or minimize $\left(y - \hat{\beta} \cdot X\right)^2$ subject to $\sum_i |\beta_i| = C$





Cross validation

- Lasso requires specifying a penalty value
- Chosen by cross validation
  - Divide historical data used for calibration into two sets
  - One set to estimate parameters using different penalty values
  - Remaining data used to evaluate which parameter performs best



Empirical Test

- Estimate the model as of December 31 of each year from 2004 to 2017
- Form thousands of equal weighted portfolios of commodity contracts
- Estimate portfolio returns for each month of the subsequent year based on known contemporaneous EE factor returns
- Calculate the correlation of forecast returns with actual realized returns out of sample for each portfolio, as well as predicted and actual risk
- Results very similar to previous method

NORTHFIELD | modeling financial markets worldwide

Frontier Currencies

Approach

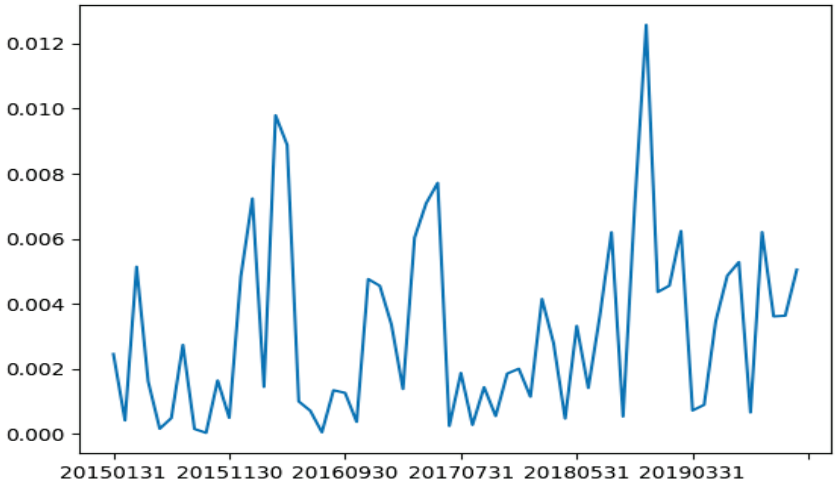
- Hoped to use same approach as commodities, possibly with some restrictions on factors used
- Just not enough explanatory power in existing factors- frontier currencies too idiosyncratic
- New approach
  - Dependence on currencies of trade partners
  - Since largest common effect is due to fluctuations of US dollar, basket of currencies to represent US dollar dependence
  - Also use oil, non-energy minerals sector, market development factors

Handling common 'idiosyncratic' risk

- Frontier currencies have significant 'idiosyncratic' risk
  - While this risk is idiosyncratic from the perspective of the currency, all securities denominated in that currency will share that common risk
  - Therefore it is important not to let this 'idiosyncratic' risk diversify away.
- The same can be said of frontier market risk more broadly: if there is, e.g., an outbreak of civil war in a country, all securities in that country will suffer
- We handle this by giving all securities denominated in a given frontier currency an identical 'issuer' code

Looking out for regime changes

- Model only good to the extent future looks like the past
- Frontier market currencies can have significant regime changes – pegged or unpegged, runs on currency, hyperinflation, etc.
- We automatically check to see if residuals are increasing with time. If so, we shorten the time period used to calibrate the model



Adjust for market pricing of risk

- Finally, to account for effects not captured by our model, we compare modeled risk to market indications of risk
- Each country has sovereign credit rating. To convert this into a risk, we use the credit spread on financial sector companies in the same region and credit rating
  - For a derivation of this relationship see Estimating an Investor's Volatility/Return Tradeoff: The Answer is Always Six (northinfo.com).
- If the market implied risk is less than our predicted risk, we scale the idiosyncratic risk up to make them match

NORTHFIELD | modeling financial markets worldwide

Cryptocurrencies

Approach

- Again using PCA-LASSO approach used for commodities
- Main challenge here is lack of history
  - Most crypto only dates back a few years
  - Northfield factor returns are monthly, so not many observations for regression
- Several work-arounds
  - 'Disaggregating' lower frequency returns using proxy data – Chow-Lin method
  - Just use Bitcoin returns for regression, map other crypto onto bitcoin – main PC is crypto market anyway
- Data quality also a challenge

Dealing with Idiosyncratic risk

- Cryptocurrencies can have very volatile and non normal return distribution
- Potentially a lot of unrealized risk with short history
- We have several approaches to deal with this

Dealing with Trends

- The first is the use of "root mean square" (RMS) rather than standard deviation as the measure of dispersion of factor returns.
  - We are treating crypto returns as if markets are very efficient so mean returns to a factor should be close to zero, rather than whatever time series mean is observed.
  - An asset that goes up 10% per month every month for two years (as with Internet stocks in the late 1990s) would have a std. dev. of zero but a significant RMS.
  - Northfield models have captured these effects for many years as described in Getting an Early Jump on Market Anomalies: Lessons from the Internet Stock Phenomenon (northinfo.com), later published in Journal of Index Investing

Open Wide

- The second technique is the idea of "range based" volatility measures, also replacing the usual definition of standard deviation of returns.
  - One way to think about the volatility of an asset is to consider the percentage distance between the highest and lowest prices observed during a particular period (e.g. day, month, year).
  - If the high and low prices are close together, the asset has low volatility. If the high and low prices are far apart, the asset is volatile.
  - Several papers starting with Parkinson (JoB, 1980) have shown that if returns are IID, there is a direct algebraic transformation between traditional return volatility and range-based measures.

Carry Trade

- The third input to currency risk estimation is the availability of a "carry trade" wherein bank deposits denominated in a particular currency offer higher interest rates than in major currencies.
- These accounts do not carry any form of government deposit insurance, so the risk of counterparty failure is substantial.
- "Bitcoin savings accounts" are available with yields over 8% annually, as compared to close to zero for retail accounts in the US

Crypto Saves the Cost of a Trip to Panama

- Our final key input is the concept of "convenience yield". The anonymity and ease of global transactions has material economic value to certain market participants.
- While this is hard to quantify directly, there is a long history of low or negative interest rates in countries with tough banking secrecy laws
- In the 1980s, Swiss banks routinely offered negative interest rates on deposit accounts while US banks were offering a rate of around 5% (the maximum allowable under Federal Reserve Regulation Q until 1986).

Implied Risk Aversion

- At the current time the combination of convenience yield and interest premium is probably around 12-13% which implies a volatility equivalent (i.e. inclusive of higher moments) of at least 72%.
  - For a derivation of this relationship see Estimating an Investor's Volatility/Return Tradeoff: The Answer is Always Six (northinfo.com).
- There is also a thinly traded Bitcoin Volatility Index (BVOL) and even less liquid linked "tokens".
  - A useful discussion is presented in Alexander and Imeraj (SSRN, 2019). Values have been in the range of 5-6% per day which is consistent with an annualized volatility in excess of 80%.

NORTHFIELD | modeling financial markets worldwide

Conclusions

Conclusions

- Three new or improved asset classes
  - Commodities in production
  - Frontier currencies basically available
  - Crypto coming soon
- Require flexibility and a range of approaches
  - Machine learning works for some cases
  - Others require some judgment
- Particular care to idiosyncratic returns of exotic classes like frontier and cryptocurrency