Applying Machine Learning to Identify Regimes for Asset Allocation and ALM

Professor John M. Mulvey

Operations Research and Financial Engineering

Bendheim Center for Finance

Princeton University

October 5, 2016

Agenda

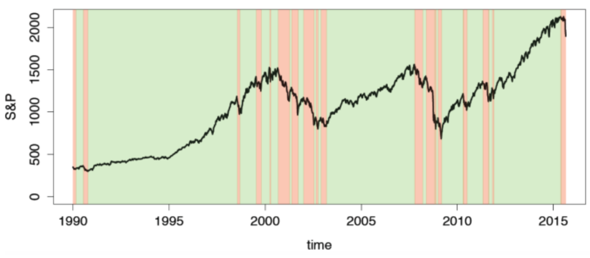
1. Factor investing for asset allocation (feature selection)

2. Motive regimes and apply 2-regime simulation to a university endowment

3. Factor investing for ALM and a large pension plan study

4. Future steps

Two-Regimes for the S&P 500 Index (1990-2015)

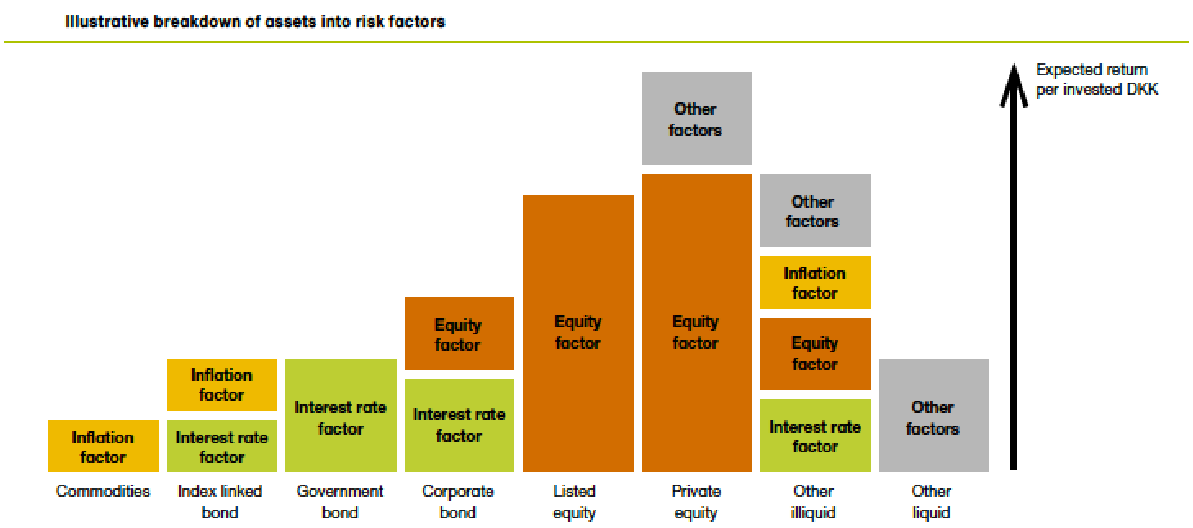


Factor Investing

- Improve diversification -- get paid to take on risks
- Help explain newer securities and asset categories
  - Examples:
    - Junk bonds
    - FTSE Dynamic Commodity Index (long short commodity via futures)
- Assist during crash periods when contagion is present
- Intensive search for higher returns to achieve goals and meet liabilities

Example: Factors as Building Blocks (critical ingredients)

Danish Pension System ATP – Factors for Assets (Ang)



Liability-related factors – Real economic growth, inflation (hard to link to asset returns)

Why Factor Investing?

- Many institutional investors have made the shift to "alternative" asset categories

  - 57% on average for U.S. endowments above $1 B
  - Main categories: private equity, real assets, hedge funds
  - There is great diversity in this domain
  - The newer asset categories capture many types of risks

Average Allocation Across University Endowments

Asset Allocation for U.S. Colleges and Universities 2014 (NACUBO 2015)

	Survey Average	Endowments over $1b
Equities	36%	31%
Domestic Equities	*17%*	*13%*
International Equities	*19%*	*18%*
Fixed Income	9%	8%
Alternatives	51%	57%
Short-term securities/cash/other	4%	4%

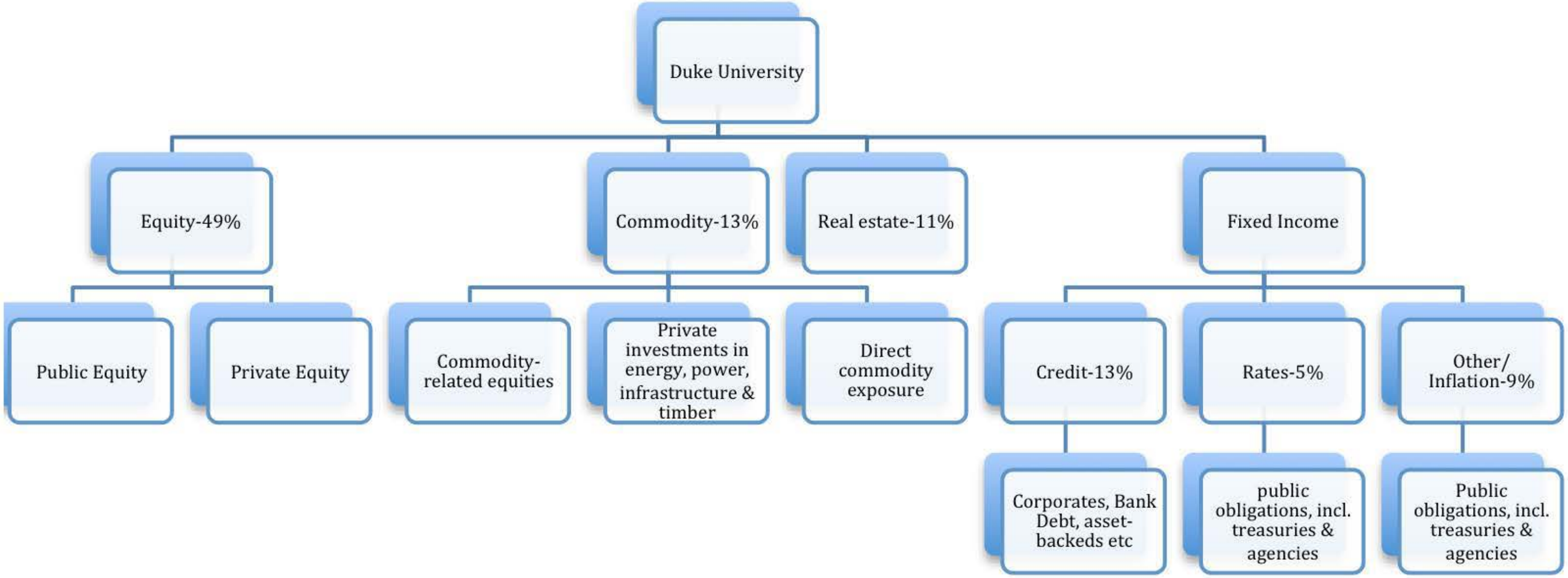
Examples of Subcategories in Hedge Fund Land

Institutions Subdividing Absolute Return/Hedge Fund Categories

Institution	#	Equity-related Hedge Fund Category	Other Categories
Georgia Institute of Technology	2	L/S equity hedge funds	Multi-strategy hedge funds
Williams College	2	Global long/short equity	Absolute return
U. Illinois Foundation	2	Marketable strategies: Hedged equity	Marketable strategies: Credit/absolute return/distressed
U. North Carolina at Chapel Hill	2	L/S equity	Diversifying strategies
University of California	3	Opportunistic equity	Absolute return Strategies, Cross-Asset Class Strategy
University of Texas System	6	Developed country equity, Emerging markets	Credit-related fixed income, Investment grade fixed income, Real estate, Natural resources
University of Washington	2	Capital appreciation: Opportunistic	Capital preservation: Absolute return
Pennsylvania State University	3	Hedged strategies: Equity related	Hedged strategies: Credit relate, Hedged strategies: other
The Ohio State University	2	Long/short equities	Relative value/macro, Credit funds
University of Virginia	2	Long/short equity	Marketable alternatives & Credit

An Example of the Diversity in Defining Asset Categories

Duke's Target Asset Allocation & Asset Category Descriptions (June 30 2014)
Source: Duke (2014)



Alternative Factor Approaches *(Feature selection in machine learning)*

- Purely statistical factors
  - Factor analysis, principle component analysis
- Fundamental macro economic
  - Chen, Roll, and Ross
    - Maturity premium (long- short government bonds), expected inflation, unexpected inflation, industrial production growth, and default premium (corporate high versus low grade bonds)
- Micro factors
  - Fama and French
    - Equity markets risks, small minus large stock returns, value minus growth
    - Profitability (high-low profit), and investment (conservative-aggressive investment)
    - Momentum, and low volatility
- Number of factors determined by sensitivity analysis
  - Most studies have shown that about 5 factors are best (little improvement above 5 for equities)

Factor Loadings for Harvard Endowment

Exhibit 4. Example of an asset class mapping matrix

	World Equities	U.S. Treasuries	High Yield	Inflation Protection	Currency Protection
U.S. Equities	1.0			0.1	0.5
Foreign Equities	1.0			0.1	-0.5
Private Equity	1.3		0.2	0.1	0.3
Real Assets	0.3		0.8	0.3	
Commodities				2.0	-0.5
U.S. Treasuries		1.0			
TIPS		1.0		1.0	
Corporate Bonds		0.8	0.2		
Foreign Bonds		0.8			-1.0
Absolute Return	0.2		0.2		

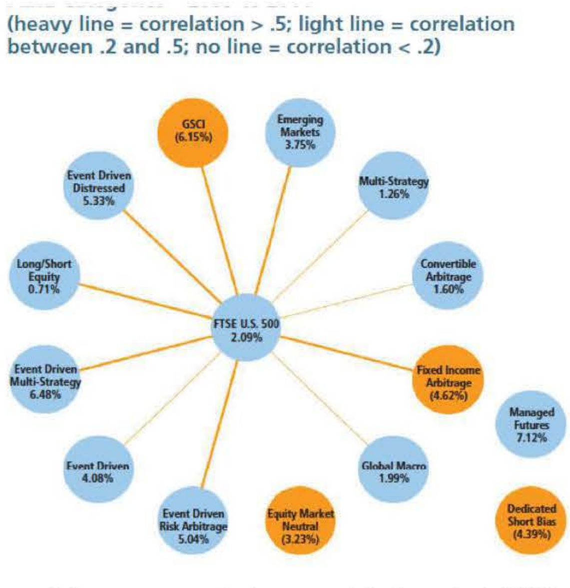
Apply cross validation from machine learning

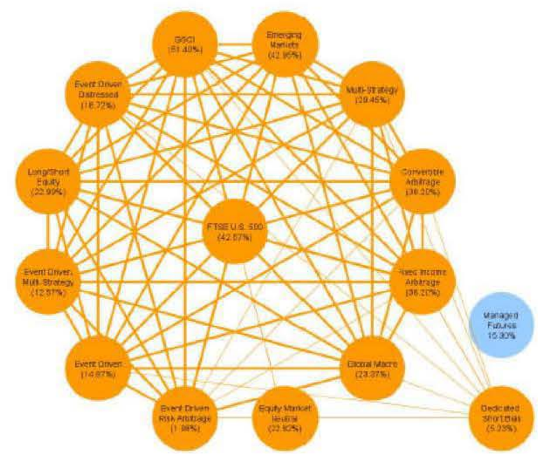
Motivate Regimes with a Study of a University Endowment

Most Hedge Funds Experienced Contagion during 2008 Crash –> massive change in covariance matrix

Most Hedge Funds Experienced the Classic Pattern of Contagion during the 2008 Crash

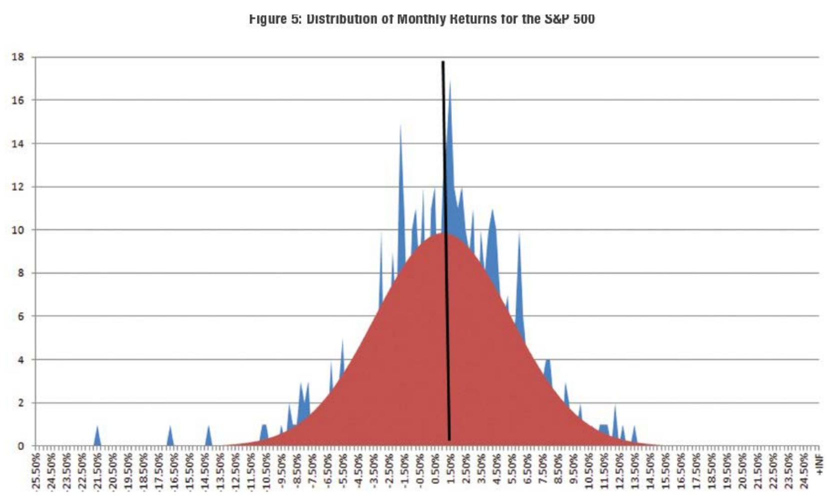
Note: stark differences between normal 2001-2007 period (left side) and crash 2008 period (right side)





Princeton University ORFE Report

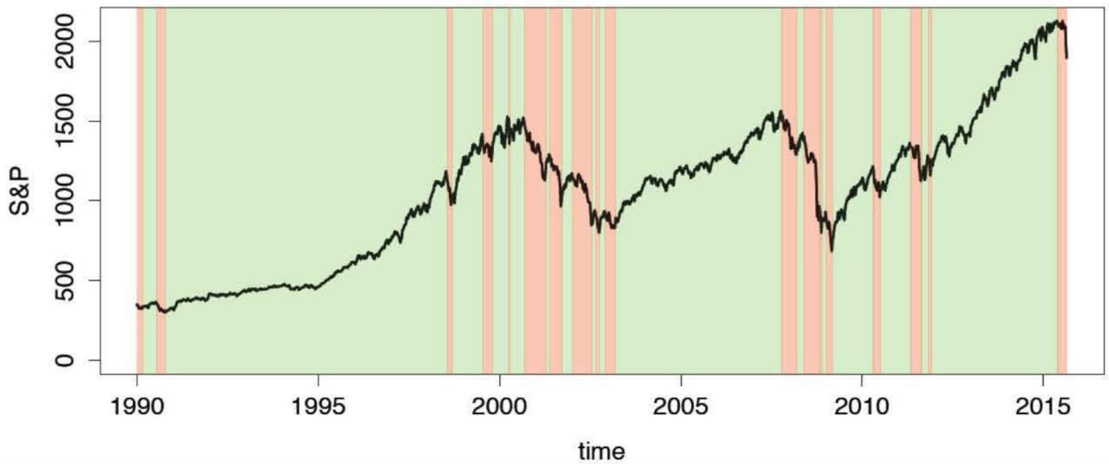
Current Assumption – Multi-Normal Distribution Underestimates Downside Risks



Multi-Regime Models can provide more accurate downside risk estimates in long run simulation studies

Identify Two Regimes via Trend Filtering Algorithm*

Two-Regimes for the S&P 500 Index (1990-2015)



*See appendix

Historical Performance

Annualized Rate	Private Equity	Real Estate	Hedge Fund	Real Assets	U.S. Equities	International Equity - Developed	International Equity - Emerging	U.S. Government Bond
Single Regime Returns	6.50%	5.50%	5.00%	4.00%	4.50%	4.20%	4.80%	1.00%
Annual Volatility	14.212%	9.827%	7.681%	6.045%	17.988%	19.503%	26.344%	5.546%

Panel C: Historical Returns for Assets under Growth Regime (Inflation adjusted)

Annualized Rate	Private Equity	Real Estate	Hedge Fund	Real Assets	U.S. Equities	International Equity - Developed	International Equity - Emerging	U.S. Government Bond
Return under growth	15.00%	9.00%	8.50%	4.50%	17.00%	16.50%	18.00%	0.30%
Annual Volatility	12.089%	7.881%	5.887%	6.157%	13.556%	15.731%	23.762%	4.236%

Panel D: Historical Returns for Assets under Contraction Regime (Inflation adjusted)

Annualized Rate	Private Equity	Real Estate	Hedge Fund	Real Assets	U.S. Equities	International Equity - Developed	International Equity - Emerging	U.S. Government Bond
Return under crash	-21.66%	-7.41%	-7.91%	2.02%	-33.50%	-33.31%	-34.80%	3.85%
Annual Volatility	9.455%	12.345%	8.678%	5.561%	12.898%	13.654%	18.716%	7.601%

Transition Matrix

Equilibrium Transition Matrix
(Probability, period *t* to *t+1*)

	Growth Regime at Time *t+1*	Contraction Regime at Time *t+1*
Growth Regime at Time *t*	0.9	0.1
Contraction Regime at Time *t*	0.4	0.6

Employ these probabilities in the two-regime simulation

Multi-Period Simulation

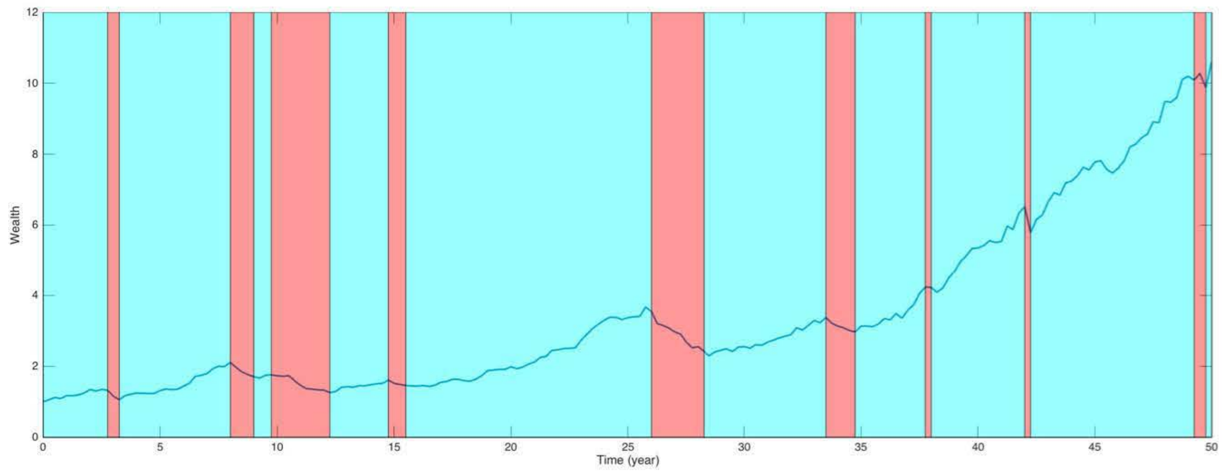
- Single regime: select asset returns each period via a single multi-normal distribution; pay operating budget at 4% of capital (averaged over 4 years)

- Two regime: start with non-crash distribution; switch between non-crash and crash each quarter depending upon transition probability; pay operating budget the same as above

- ***The two regime model more accurately projects the worst events (left tails) than the single regime approach***

Forward Looking Simulation – Sample Path

A Representative Scenario Path over the 50-Year Planning Horizon (S&P 500 Index)



Compare Single and Two-Regime Models

Summary Statistics for Baseline Monte Carlo Simulations (4% Target Spending Target)

	Without spending-cut rule		**Spending Cut by 20%**	
Simulation Results	1-Regime	2-Regime	1-Regime	2-Regime
Crash Prob, 5 years	10.3%	18.4%	10.3%	18.4%
Crash Prob, 10 years	20.5%	31.8%	20.0%	31.5%
Crash Prob, 50 years	4.9%	19.9%	2.2%	13.1%
mean-5 years	1.0644	1.0780	1.0644	1.0780
mean-10 years	1.1635	1.1793	1.1692	1.1879
mean-20 years	1.3998	1.4230	1.4173	1.4514
mean-50 years	2.6630	2.6197	2.7152	2.7147
# of simulations	10000	10000	10000	10000
Average % time in "adverse"	2.88%	5.60%	2.22%	4.60%

Advantages of Reducing Spending

Panel B: Performance Statistics for 3.5% Spending Target

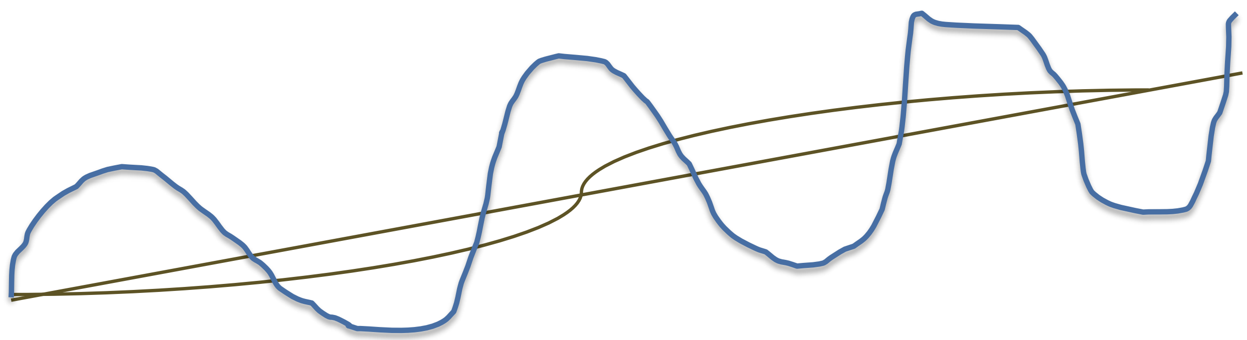
	Without spending-cut rule		**Spending Cut by 25%**	
Simulation Results	1-Regime	2-Regime	1-Regime	2-Regime
Crash Prob, 5 years	8.3%	16.3%	8.3%	16.3%
Crash Prob, 10 years	14.6%	26.6%	14.4%	26.4%
Crash Prob, 50 years	1.2%	8.5%	0.4%	5.6%
mean-5 years	1.0977	1.1112	1.0977	1.1112
mean-10 years	1.2421	1.2609	1.2462	1.2680
mean-20 years	1.5883	1.6267	1.6018	1.6499
mean-50 years	3.6169	3.6499	3.6585	3.7310
# of simulations	10000	10000	10000	10000
Average % time in "adverse"	1.39%	3.41%	1.06%	2.68%

Factor Investing for a Large Pension System via ALM

Mention attention to liabilities –
CFA level 3

Difficulties Explaining Assets and Liabilities via Common Factors

1. Assets returns are driven by short and mid-term factors (interest rates, risk premium, cash flows, and micro factors such as momentum, value and so on)
2. Liabilities are driven by mid to long term factors
3. Longevity issues



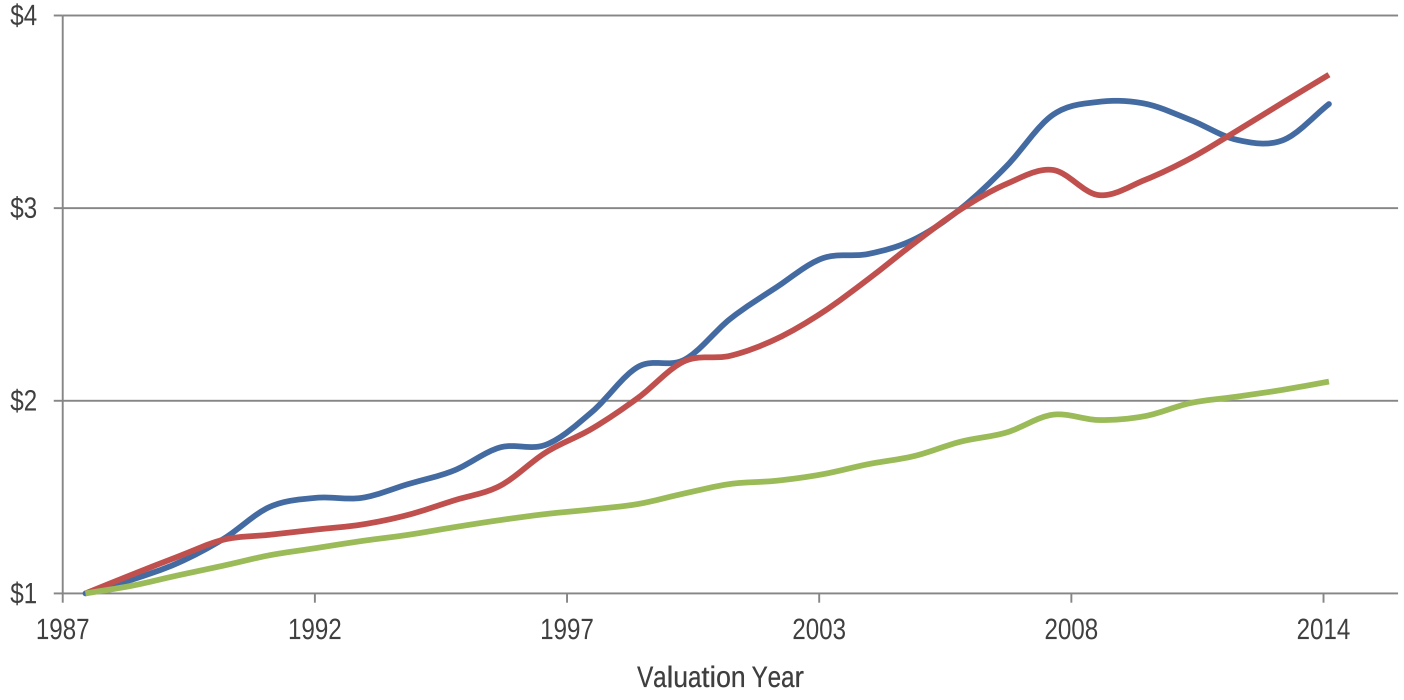
Market factors (short horizon), Macro-economic factors (intermediate horizon), demographic factors (long horizon)

Macro-Economic Factors and Pension Liabilities

- Two macro-economic factors that affect liability cash flows are:
  - Growth – Real Gross Domestic Product (GDP)
  - Inflation
- Impact on salaries and retirement benefits
- Assumption: liabilities are discounted at the blended return rate (for now)

History - Pension Payroll, GDP and US Inflation

- Pension Payroll and GDP showed a strong relationship



Potential Advantages of Regime-Aware ALM

1. Consistency between asset performance and changes in liability cash flows

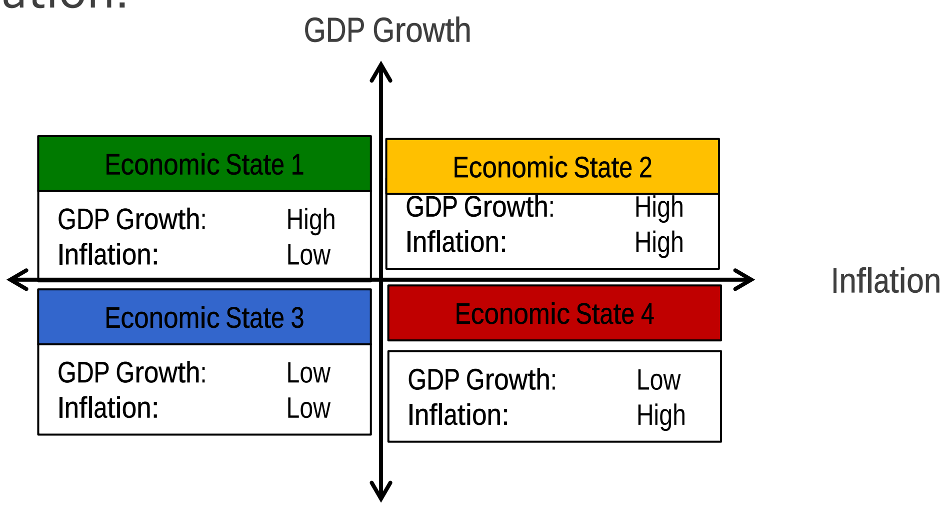
1. Improved estimates of downside risks

1. Enhanced asset performance (possibly)

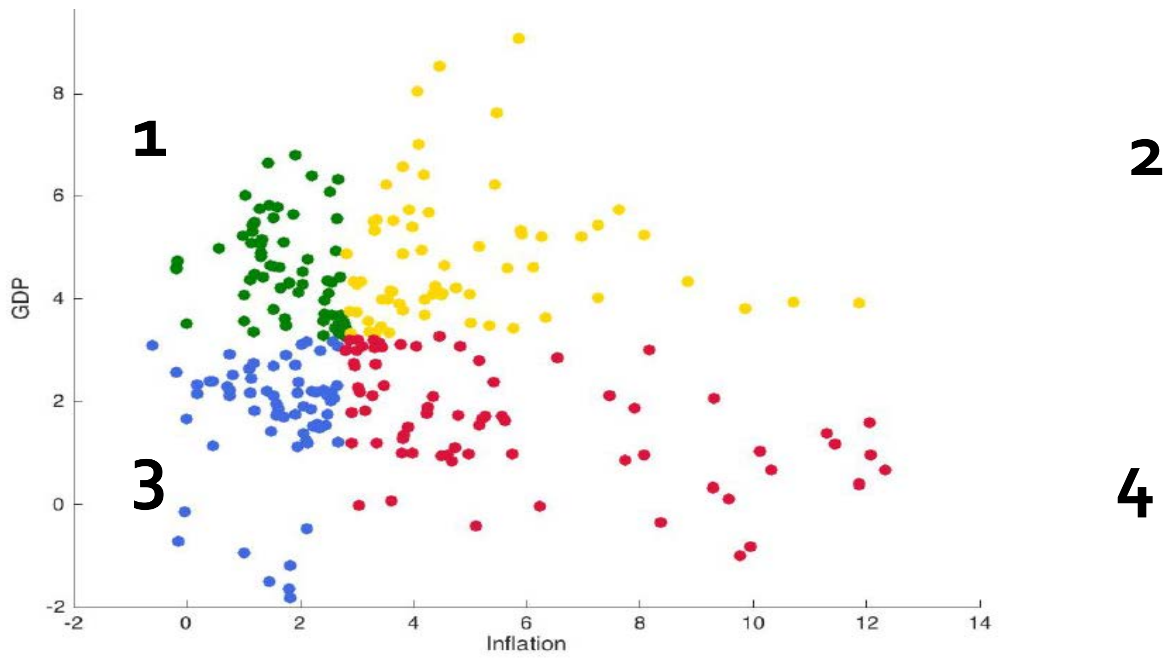
2. Streamlined estimates of liability cash flows

An Illustrative Example

- Assume that the current state of the economy is defined by Real GDP growth and inflation:

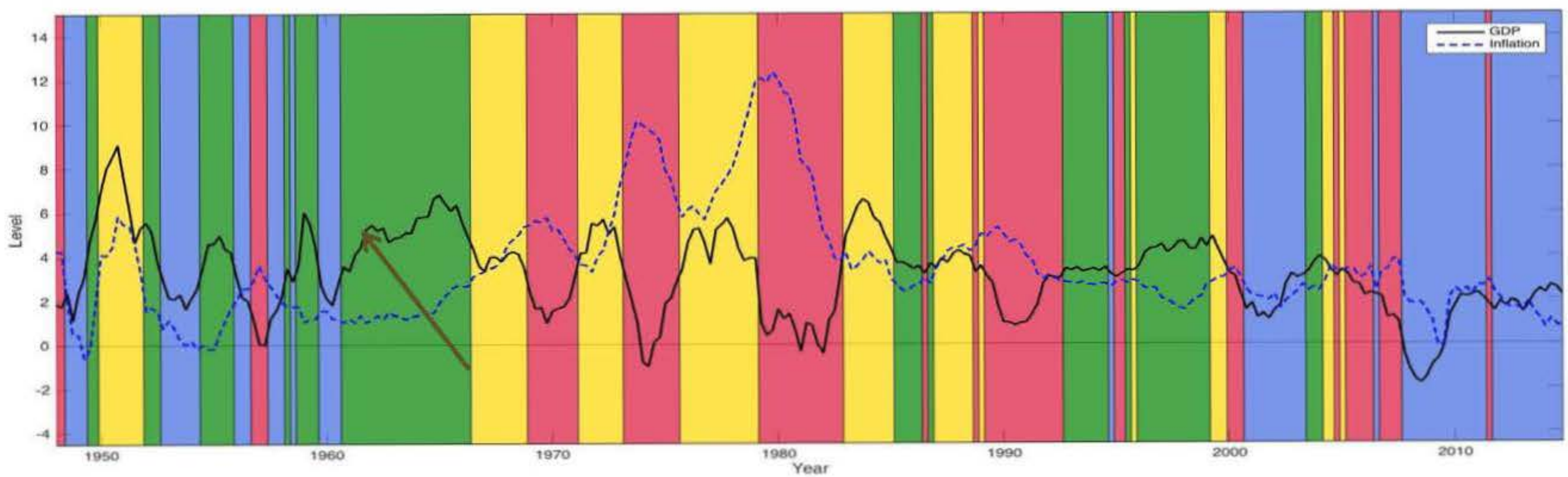


Scatter Plot of Inflation and Real GDP 1948-2015 (4-Way Split)



Princeton University ORFE Report

Historical Patterns (time series of inflation and real GDP)



Four Regimes are Stable Across Time

Time Period **1948-2014**
Frequency **Quarterly**

	Regime 1	Regime 2	Regime 3	Regime 4
Regime 1	0.83	0.09	0.07	0.01
Regime 2	0.05	0.83	0.00	0.13
Regime 3	0.10	0.00	0.84	0.06
Regime 4	0.04	0.07	0.08	0.80

Performance of Asset Categories

Real Returns of Major Asset Categories (1973-2015 monthly)1

Geometric Mean of Return (Annually)

U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
5.8354%	4.6518%	4.0300%	3.5651%	5.3212%	2.4234%	2.9045%	0.6919%

Geometric Mean of Return (Annually)

	U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
Regime 1	15.2834%	13.4923%	6.9086%	5.1891%	9.1384%	9.7553%	3.7975%	1.4420%
Regime 2	0.8939%	6.4105%	-0.1339%	1.4991%	3.9968%	4.7818%	1.0627%	0.7616%
Regime 3	11.0503%	10.6285%	5.1594%	7.4233%	13.7547%	-0.8822%	6.1820%	0.2001%
Regime 4	-2.8658%	-10.2222%	4.3173%	0.3051%	-4.6992%	-3.4542%	0.6718%	0.3682%

Regime 1 = growth+ and inflation-, Regime 2= growth+, inflation+, Regime 3 = growth-, inflation-, Regime 4 = growth-, inflation+

Equity Micro-Factor Performance

Real Return of Equity Micro-Factors over Four Regimes – 1970-2015 monthly

Geometric Mean of Return (Quarterly)	High Value	Low Value	High Vol	Low Vol	High Investment	Low Investment	High Profitability	Low Profitability	High Momentum	Low Momentum
Regime 1	5.36%	4.20%	4.73%	4.08%	4.37%	4.99%	5.42%	4.30%	6.36%	3.53%
Regime 2	3.40%	0.97%	1.91%	2.19%	1.44%	2.49%	1.96%	1.82%	2.51%	1.23%
Regime 3	2.86%	0.97%	1.01%	2.19%	0.78%	2.54%	2.39%	1.23%	1.12%	1.88%
Regime 4	2.78%	0.86%	2.14%	1.63%	1.63%	2.26%	1.82%	2.01%	2.48%	1.22%

Equity Segments are Highly Correlated

Correlation of Return of S&P 500 and Equity Micro-Factors

S&P500	0.8726655	0.95527	0.8534567	0.8816909	0.9536146	0.8337706	0.9485551	0.8847728	0.9515912	0.8821828

- ***Motivate – long/short versions to improve diversification***

Comparing Traditional MVO and Regime-Aware Approach

Inputs to a Markowitz Portfolio Model
January 1973-December 2007, Real Monthly Returns

Geometric Mean of Return (Monthly)

U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
0.4803%	0.4870%	0.2995%	0.2806%	0.4222%	0.4727%	0.2507%	0.0968%

Conditional Value at Risk (Monthly)

U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
-10.0627%	-10.6460%	-5.9763%	-4.2674%	-10.7883%	-11.5732%	-6.8351%	-0.3874%

Volatility (Monthly)

U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
0.044211	0.047802	0.030183	0.021728	0.045102	0.055778	0.028858	0.002063

Sharpe Ratio (Monthly)

U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
0.086752	0.081635	0.067182	0.084619	0.072162	0.067397	0.053333	0.000000

Correlation

	U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
U.S.Equity	1.000000	0.574599	0.211152	0.236128	0.565631	0.005373	0.114511	0.011030
Intl.Equity	0.574599	1.000000	0.134243	0.204139	0.387313	0.062845	0.070981	0.022911
U.S.Treasur	0.211152	0.134243	1.000000	0.741640	0.238839	-0.049413	0.481781	0.174982
Corp.Bond	0.236128	0.204139	0.741640	1.000000	0.288717	-0.050624	0.387406	0.197569
Real Estate	0.565631	0.387313	0.238839	0.288717	1.000000	-0.041503	0.214121	-0.034849
Commodity	0.005373	0.062845	-0.049413	-0.050624	-0.041503	1.000000	0.101095	-0.015984
TIPS	0.114511	0.070981	0.481781	0.387406	0.214121	0.101095	1.000000	0.096095
Risk Free	0.011030	0.022911	0.174982	0.197569	-0.034849	-0.015984	0.096095	1.000000

Comparing Traditional MVO and Regime-Aware Approach

Performance of Assets Under the Four Regimes
January 1973 to December 2007

Geometric Mean of Return (Monthly)

	U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
Regime 1	1.1820%	1.1216%	0.4543%	0.3829%	0.6396%	1.1698%	0.3033%	0.1574%
Regime 2	0.0742%	0.5191%	-0.0112%	0.1036%	0.3271%	0.3900%	0.0881%	0.0632%
Regime 3	0.5105%	0.6285%	0.6368%	0.5022%	0.7955%	0.0284%	0.6426%	0.0943%
Regime 4	0.2245%	-0.2829%	0.2715%	0.2300%	0.0574%	0.1740%	0.1165%	0.0759%

Geometric Mean of Return (Annually)

	U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
Regime 1	15.1435%	14.3208%	5.5896%	4.6924%	7.9512%	14.9772%	3.7006%	1.9056%
Regime 2	0.8939%	6.4105%	-0.1339%	1.2503%	3.9968%	4.7818%	1.0627%	0.7616%
Regime 3	6.3007%	7.8078%	7.9149%	6.1962%	9.9747%	0.3409%	7.9896%	1.1377%
Regime 4	2.7279%	-3.3423%	3.3074%	2.7946%	0.6906%	2.1084%	1.4074%	0.9152%

Conditional Value at Risk (Monthly)

	U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
Regime 1	-5.4773%	-7.0090%	-4.7597%	-2.6150%	-6.5103%	-8.4059%	-3.8300%	-0.1358%
Regime 2	-11.7270%	-11.3778%	-4.9681%	-3.8713%	-10.8949%	-11.0935%	-7.0140%	-0.2242%
Regime 3	-9.6752%	-11.2141%	-6.6980%	-2.4451%	-8.7136%	-11.6188%	-3.2554%	-0.1359%
Regime 4	-10.8876%	-11.9761%	-7.0830%	-6.8815%	-14.0816%	-14.5317%	-8.9285%	-0.5232%

Conditional Value at Risk (Annually)

	U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
Regime 1	-49.1331%	-58.1890%	-44.3009%	-27.2381%	-55.4172%	-65.1333%	-37.4140%	-1.6173%
Regime 2	-77.6161%	-76.5303%	-45.7459%	-37.7359%	-74.9485%	-75.6105%	-58.2161%	-2.6576%
Regime 3	-70.5094%	-76.0045%	-56.4795%	-25.6995%	-66.5133%	-77.2848%	-32.7766%	-1.6182%
Regime 4	-74.9241%	-78.3626%	-58.5865%	-57.4960%	-83.8179%	-84.8063%	-67.4469%	-6.1009%

Volatility (Monthly)

	U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
Regime 1	0.036762	0.040043	0.024361	0.014356	0.034961	0.048538	0.018276	0.001563
Regime 2	0.044392	0.045772	0.026912	0.018935	0.042762	0.057525	0.027643	0.001952
Regime 3	0.042659	0.050499	0.032279	0.014211	0.036556	0.052504	0.021081	0.001437
Regime 4	0.051203	0.054603	0.036895	0.032596	0.060046	0.062507	0.041122	0.002771

Sharpe Ratio (Monthly)

	U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free
Regime 1	0.278706	0.240775	0.121850	0.157036	0.137920	0.208577	0.079799	0.000000
Regime 2	0.002465	0.099597	-0.027649	0.021309	0.061706	0.056803	0.009002	0.000000
Regime 3	0.097557	0.105774	0.168056	0.287058	0.191804	-0.012562	0.260078	0.000000
Regime 4	0.029019	-0.065716	0.053010	0.047248	-0.003093	0.015691	0.009871	0.000000

Princeton University ORFE Report

Compare Traditional MVO and Regime-Aware Approach

Opt Portfolios	U.S.Equity	Intl.Equity	U.S.Treasury	Corp.Bond	Real Estate	Commodity	TIPS	Risk Free	Volatility
Full period	18.89%	14.70%	8.30%	26.28%	8.73%	23.10%	0.00%	0.00%	**8.15%**
Regime 3	0.00%	7.54%	26.28%	0.00%	27.39%	0.00%	38.79%	0.00%	**8.43%**

Exhibit 22
Total Return for the MVO allocation and the Regime-Aware allocation

	Full Period	Regime 3 Aware
2008	-30.25%	-10.9%
2008-2014	5.6%	43.1%

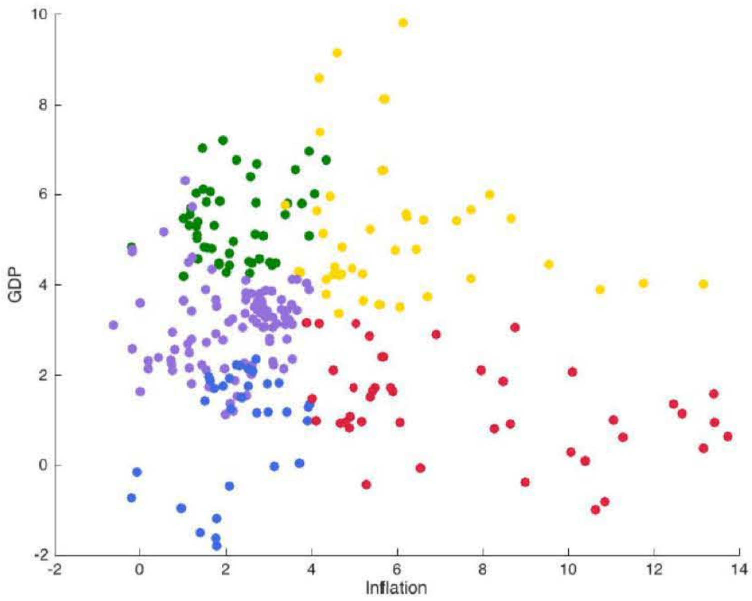
Future Directions

Current Research

- Identify regimes via alternative approaches
- Evaluate performance of assets and micro-factors over the stated regimes
  - Factors (features) over regimes
  - Apply machine learning as appropriate
- Develop robust allocations in light of current and projected regimes (over the mid-term)

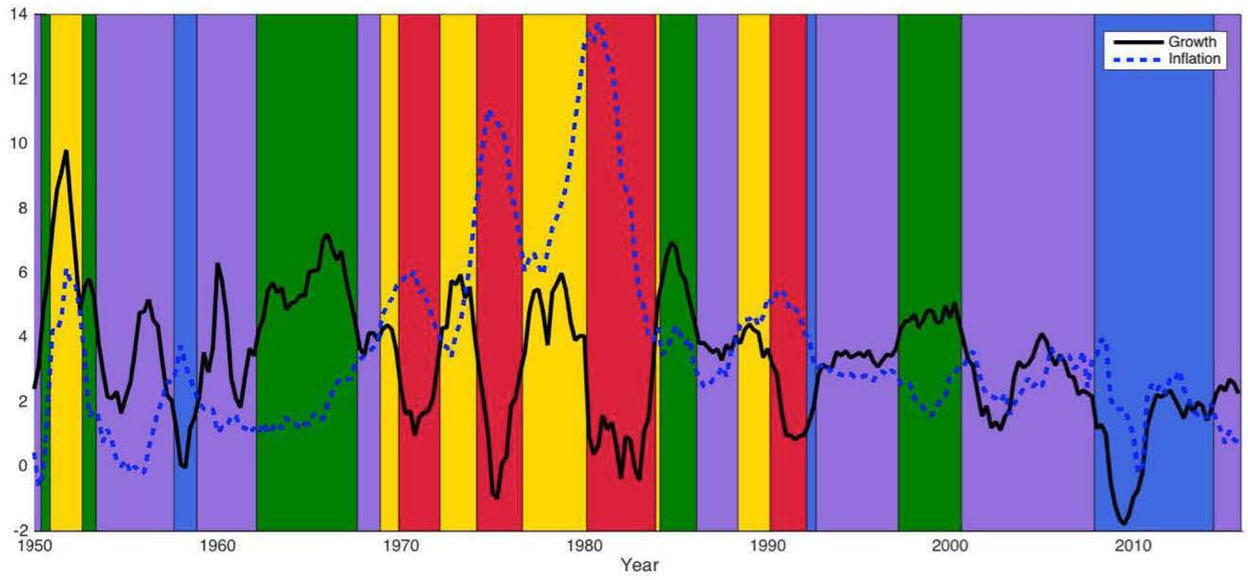
Alternative Regime Definitions

Defining Five Economic Regimes based on Machine Learning Methods



A Crash Regime Looks Sensible

Historical Record of Five Regimes 1948-2015



References

- Mulvey, J.M. and M. Holen, "What Can We Learn from University Endowments: Insights from Self-Reports," *Journal of Investment Consulting, 2017 (to appear)*.

- "Identifying Economic Regimes: Reducing Downside Risks for University Endowments and Foundations, *Journal of Portfolio Management*, Fall 2016.

- "A Nonparametric Smoothing Approach to Financial Market Regime Identification," Princeton University Report, J. Mulvey, H. Liu, and T. Zhang, August 2015.

- "Dynamic Asset Allocation for Varied Financial Markets under Regime Switching Framework," *European Journal of Operational Research* ,G. Bae., W. Kim, and J. Mulvey 2013.

- "Assisting Defined-Benefit Pension Plans," *Operations Research*, 2009 (with K. Simsek, Z. Zhang, and F. Fabozzi).

- Towers Watson, Global Pension Asset Study 2015.

- **The Elements of Machine Learning**, 2nd Edition, T. Hastie, R. Tibshirani, and J. Friedman (2009).

Appendix

Global Pension Assets Study 2015 (Towers Watson)

Australia	1,675	113.0%
Brazil[1]	268	12.0%
Canada	1,526	85.1%
France	171	5.9%
Germany[2]	520	13.6%
Hong Kong	120	41.2%
Ireland	132	53.7%
Japan[3]	2,862	60.0%
Malaysia	205	60.7%
Mexico	190	14.6%
Netherlands	1,457	165.5%
South Africa	234	68.6%
South Korea	511	35.3%
Switzerland[4]	823	121.2%
UK	3,309	116.2%
US[5]	22,117	127.0%
Total	36,119	84.4%

Trend Filtering Algorithm

A detailed formulation of the trend filtering is as follows. Let $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^T \in \mathbb{R}^n$ denote the price of a series at n evenly spaced time points. For a given integer k, the general form of a trend filtering estimator $\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_1, \ldots, \widehat{\beta}_n)^T \in \mathbb{R}^n$ is defined as the solution to the following penalized optimization problem:

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta} \in \mathbb{R}^n}{\operatorname{argmin}} \|\boldsymbol{Y} - \boldsymbol{\beta}\|_2^2 + \lambda \left\|\mathbf{D}^{(k+1)} \boldsymbol{\beta}\right\|_1, \tag{2.1}$$

where $\lambda \geq 0$ is a regularization parameter, and $\mathbf{D}^{(k+1)} \in \mathbb{R}^{(n-k-1)\times n}$ denotes the operator for computing the $(k+1)$-th order discrete derivative. For example, when $k = 0$ and $k = 1$,

$$\mathbf{D}^{(1)} = \begin{pmatrix} 1 & -1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & -1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -1 \end{pmatrix}, \quad \mathbf{D}^{(2)} = \begin{pmatrix} 1 & -2 & 1 & \ldots & 0 & 0 \\ 0 & 1 & -2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & -2 & 1 \end{pmatrix}$$

so that $\|\mathbf{D}^{(1)}\boldsymbol{\beta}\|_1 = \sum_{i=1}^{n-1} |\beta_i - \beta_{i+1}|$, and $\|\mathbf{D}^{(2)}\boldsymbol{\beta}\|_1 = \sum_{i=1}^{n-2} |\beta_i - 2\beta_{i+1} + \beta_{i+2}|$.

Trend Filtering Algorithm

With different choices of k, the solution takes on different structures. When $k = 0$, the solution to (2.1) is a piecewise step function. When $k = 1$, the solutions is piecewise linear. When $k = 2$, the solution is piecewise quadratic, and so on. To see the intuition of piecewise linear when $k = 1$: Eq. (2.1) is in a form of generalized lasso (least absolute shrinkage and selection operator) problem (Tibshirani et al., 2011, 2012):

$$\hat{\boldsymbol{\beta}} = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \|\boldsymbol{Y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \left\|\mathbf{H}\boldsymbol{\beta}\right\|_1, \tag{2.2}$$

where $\mathbf{X} \in \mathbb{R}^{n \times p}$, and $\mathbf{H} \in \mathbb{R}^{m \times p}$. When $m = p$ and $\mathbf{H}$ is the identity matrix, then (2.2) becomes the regular lasso estimator (Tibshirani, 1996). The lasso was originally proposed

Trend Filtering Algorithm

to handle high dimensional sparse regression and variable selection problems. Its solution has the remarkable property of being sparse (i.e., many entries of the solution vector are zero), while the optimization problem remains convex and efficient to solve. There is a straightforward geometric interpretation for the sparse property of lasso. By Langrange multiplier theory, the formulation of lasso is equivalent to a constraint optimization problem:

$$\widehat{\beta} = \operatorname*{argmin}_{\beta \in \mathbb{R}^d} \frac{1}{n} \sum_{i=1}^{n} (Y_i - X_i^T \beta)^2 \quad \text{s.t.} \quad \|\beta\|_1 \leq \mu,$$

for some $\mu \in \mathbb{R}$ and X_i is the i-th row of $\mathbb{X}$. Similar transformation is true for ℓ_0- and ridge regression, which uses $\|\cdot\|_0$ and $\|\cdot\|_2$ norm as regularization function.