Asset Pricing Implications of Parameter Uncertainty and Heterogeneous Beliefs

P. Collin-Dufresne Michael Johannes Lars Lochstoer

(SFI@EPFL) (Columbia University)[2]

November 2014

Macro risks and asset pricing

Classic tension:

- Macro risks small – consumption 'too' smooth
- Asset prices imply large risks (or very high risk aversion)

$\Rightarrow$ Equity Premium Puzzle

$\Rightarrow$ and Risk-free rate puzzle (Mehra and Prescott (1985))

Typical model: agents have full information about aggregate endowment dynamics

- **Full** Rational Expectation dominant framework (Lucas (1978)).
- In practice, estimate consumption dynamics and use point estimates of parameters as fixed quantities input to GE models
- But is it plausible that agents know the model or the parameters of the model if the econometrician does not?

$\Rightarrow$ The billiard player metaphor!

Why do prices move much more than fundamentals?

- Excess Volatility (Shiller (1981)): p^* is fundamental price, p is actual price (S&P500)

Answers:

- Rational time-varying discount rates (e.g., due to risk-aversion), or
- Mispricing (Behavioral Economics).

Motivation: What about Model Risk?

Typical Asset Pricing Model:

- Assume agents know and agree on the true model of the world (full information)
- Answer to Shiller-puzzle: Lots of variation in the equity risk premium (discount rates)

Full-information assumption counter-factual as:

1. We do not know the true model (limited information; model uncertainty)
2. We do not agree on what the 'best' model is, given available information (heterogeneous beliefs)
3. We may not use all available data consistently ('This time is different' bias).

Mark Twain: "*It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so.*"

This presentation

Relax these assumptions:

- Agents are uncertain about the macroeconomic 'model of the world'
- Agents disagree about dynamics of fundamentals
  (e.g., future economic growth, likelihood and severity of future recessions)
- Agents put less weight on old data than on recent (personal) experience.

Argue that:

- Model/parameter uncertainty is a major source of market risk premiums and price volatility
- Model/parameter uncertainty can lead to prolonged periods of significant market mispricing

  (especially when learning is slow)

$\Rightarrow$ *Long-horizon investors are well-positioned to take on this 'risk-premium'*

Why don't researchers incorporate parameter uncertainty?

Parameter uncertainty is realistic, reasonable, and theoretically sound... But

- Conventional wisdom (at least in asset pricing) is that parameter learning is not quantitatively important for (at least) two reasons:
  - Learning is fast (i.e., posterior variance decays quickly).
  - Standard (time separable) expected utility implies parameter risk is less relevant because agents do not care about long-run risks.
- Computationally complex, non-stationary. Lucas and Sargent (**1979**):

  "it has been only a matter of analytical convenience and not of necessity that equilibrium models have used the assumption of stochastically stationary "shocks" and the assumption that agents have already learned the probability distributions that they face. Both of these assumptions can be abandoned, albeit at a cost in terms of the simplicity of the model.... Equilibrium theorizing in this context thus readily leads to a model of how process nonstationarity and Bayesian learning applied by agents to the exogenous variables leads to time-dependent coefficients in agents' decision rules....While models incorporating Bayesian learning and stochastic nonstationarity are both technically feasible and consistent with the equilibrium modeling strategy, almost no successful applied work along these lines has come to light. One reason is probably that nonstationary time series models are cumbersome..."

The Call to Arms

The importance of understanding and quantifying estimation risk (Hansen, **2007**):

> *"In actual decision making, we may be required to learn about moving targets, to make parametric inferences, to compare model performance, or to gauge the importance of long-run components of uncertainty. As the statistical problem that agents confront in our model is made complex, rational expectations' presumed confidence in their knowledge of the probability specification becomes more tenuous. This leads me to ask: (a) how can we burden the investors with some of the specification problems that challenge the econometrician, and (b) when would doing so have important quantitative implications."*

Paper I: priced parameter uncertainty

Ingredients:

- Bayesian learning about (i) unknown parameters, and/or (ii) unknown models.
- Epstein-Zin recursive utility with preference for early resolution of uncertainty

Mechanism 1:

- Parameter learning induces **subjective** long run consumption risks
- Consider learning about structural fixed parameters, θ, such as consumption growth
  - Posteriors are martingales (Doob): $M_t = E_t[\theta] = E_t\left[M_{t+j}\right]$
    $\Longrightarrow$ 'shocks' to beliefs are **permanent**
    $\Longrightarrow$ *long-run* macro uncertainty due to 'uncertainty aversion'
- When learning about consumption dynamics, shocks to beliefs impact conditional distribution of consumption growth in *all* future periods

  $\Longrightarrow$ big impact on continuation utility, thus parameter uncertainty is priced and has a major impact on asset prices

Paper II: parameter uncertainty and 'this time is different' biased beliefs

Consider the Gordon growth formula:

$$\frac{P}{D} = \frac{1}{r - g}$$

Mechanism 1: Your current 'model' is $r = 9\%$ and $g = 5\%$; $P/D = 25$.

- You observe, say, low growth and revise your model to $r = 9\%$ and $g = 4\%$, $P/D = 20$ (-20%)
- If revise upward to $g = 6\%$, $P/D = 33.33$ (+30%)

Q? Why did a +/- 1% move in growth expectations have such a large price impact?

A! Because it's a **permanent** effect

Mechanism 2: Two agents: optimist ($g = 6\%$) and pessimist ($g = 4\%$)

- Since optimists holds more stocks, wealth shifts to pessimists in bear markets.
- Since wealth-weighted average investor is more pessimistic, prices drop (more)
- Since 'Young' never fully learn from the 'Old' risk-premia and volatility due to parameter uncertainty never disappear.

Outline and results

Consider different dimensions of parameter uncertainty

1. i.i.d. lognormal: Learning about mean growth rate and shock variance
2. Persistent rare events: Great Depression calibration
3. Model uncertainty: Bansal and Yaron vs. i.i.d. lognormal

Parameter learning can tremendously magnify macroeconomics risks. Helps with standard 'puzzles': risk premium, risk-free rate, excess volatility

- Up to 6 times risk premium relative to known parameter case

Endogenous time-variation in sensitivity of marginal utility to belief updates

- Long-lasting asset pricing implications, though learning is efficient (Bayesian)

'Large' wedge between subjective and objective consumption dynamics

- Interesting dynamics in asset price moments arise from homoskedastic fundamentals and isoelastic preferences
- Ex-post may see little macro risk (consumption looks i.i.d.) even though prices reflect a lot of risk.

Related Literature

Parameter learning

- Separation 'theorem': Detemple (1986), Dothan and Feldman (1986), Gennotte (1986)
- Timmermann (1993, 2001), Barberis (2000), Xia (2001), Lewellen and Shanken (2002), Pastor and Veronesi (2003, 2006), Jobert, Platania, and Rogers (2006), Johnson (2007), Weitzman (2007), Hansen (2007), Cogley and Sargent (2008), Bakshi and Skoulakis (2010), Johannes, Lochstoer, and Mou (2011), Kumar and Gvozdeva (2012), Lu and Siemer (2012)

Robustness and ambiguity

- Hansen (2007), Hansen and Sargent (2009, 2010), Leippold, Trojani, and Vanini (2008), Epstein and Schneider (2008), Collard, Mukerji, Sheppard, and Tallon (2011), Ju and Miao (2012)

Long-run risks

- Bansal and Yaron (2004), Beeler and Campbell (2012)

Related Literature

Most related papers in (large) literature on Heterogeneous beliefs:

- Ehling, Graniero, Heyerdahl-Larsen (2014), Choi and Mertens (2013), Garleanu and Panageas (2014), Borovicka (2013), Chen, Joslin, and Tran (2012), Bansal and Shaliastovich (2010), Baker, Hollifield, and Osambela (2014), Hirshleifer and Yu (2013), Collin-Dufresne, Johannes, and Lochstoer (2013), Barberis, Greenwood, Jin, and Shleifer (2014)

Preferences

- Epstein-Zin recursive utility, with stochastic discount factor:

$$M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma} \left(\frac{VC_{t+1}}{E_t \left[VC_{t+1}^{1-\gamma} \left(C_{t+1}/C_t \right)^{1-\gamma} \right]^{1/(1-\gamma)}} \right)^{\frac{1}{\psi}-\gamma}$$

where $VC_t \equiv \frac{V_t}{C_t}$, $\gamma = RRA$, $\psi = IES$, β time-discounting and C_t is consumption.

$\Rightarrow$ If $\gamma > 1/\psi$ then agents have preference for early resolution of uncertainty (and 'long-run risks' are priced).

- If $\gamma = 1/\psi$, reduces to time separable power utility case where agents are indifferent to long-run risks.

- Parameter uncertainty has primary impact on continuation utility via long-run risks which drive the price-consumption ratio (VC).

Simple case: Learning about the mean growth rate

Objective consumption dynamics:

$$d \ln C_t = \mu dt + \sigma dz_t.$$

- Investors know σ, but not μ. Initial prior: $\mu \sim \mathcal{N}\left(\mu_0, \sigma^2 A_0\right)$; posterior: $\mu \sim \mathcal{N}\left(\mu_t, \sigma^2 A_t\right)$, $A_t \leq 1$. Learn from realized consumption growth.

Subjective consumption dynamics:

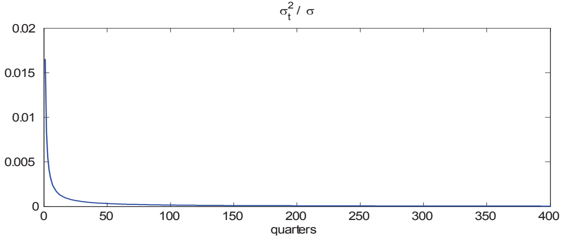
$$\begin{aligned} d \ln C_t &= \mu_t dt + \sigma d\tilde{z}_t, \\ d\mu_t &= A_t \sigma \, d\tilde{z}_t, \\ dA_t &= -A_t^2 dt. \end{aligned}$$

Thus, parameter learning yields:

1. No change in ‘short-run’ risk
2. Truly ‘long-run’ risk: μ_t is a martingale.
3. Heteroskedastic long-run risk shocks

Bayesian learning is fast

- Let $A_0 = 1$, $\sigma = 1.65\%$. Plot $A_t\sigma$ (volatility of shock to mean belief, μ_t):



- After 10 years: 0.04125%
- After 100 years: 0.004125%

Q? So small parameter learning cannot be quantitatively important for risk prices?

Analytical case: IES $= 1$ and continuous-time

RRA $= \gamma$, conts time-discounting, $\tilde{\beta} = -\ln \beta$:

$$\text{price of risk:} \qquad \gamma\sigma + \frac{\gamma - 1}{\tilde{\beta}} A_t \sigma$$

$$\text{risk premium:} \qquad \gamma\sigma^2 + \frac{\gamma - 1}{\tilde{\beta}} A_t \sigma^2$$

(A) Preference for early resolution of uncertainty required: $\gamma > 1/\psi = 1$

(B) Large multiplier on size of update in beliefs

- If, as in Bansal-Yaron, $\gamma = 10$, $\tilde{\beta} = -\ln 0.994$; $\frac{\gamma - 1}{\tilde{\beta}} = 1495$!
- Price of risk and risk premium after 100/200 years of learning relative to known mean: 1.37/1.19. Parameter uncertainty amplifies macro shocks.

Additional implications

- In model, Hall (1988) regression

$$\Delta \ln C_{t,t+1} = \alpha + \beta r_{f,t} + \varepsilon_{t,t+1}.$$

  yields (trivially) $\hat{\beta} \approx 0$ since consumption growth truly iid.

- P/D ratios do not predict consumption growth either (Beeler and Campbell, 2012), because consumption growth is actually i.i.d.

- What about IES > 1? (Need to solve numerically)
  - Excess volatility, counter-cyclical price-consumption ratio and in-sample excess return predictability.
  - Sensitivity of marginal utility increases as posterior variance of μ decreases

- Model is too simple to take (too) seriously, but highlights mechanism, provides intuition, and suggests quantitatively important effects.

$\Rightarrow$ Hansen 2007: what about cases where learning is difficult or slow?

When is learning likely to be difficult or slow?

- Rare events
  - Rietz, Barro and others suggest rare disasters can explain many asset pricing puzzles.
  - How do agents learn about events that occur once every 50 or 100 years?

  $\neq$ the billiard player metaphor.

- Learning about two models with difficult-to-discriminate components
  - Bansal and Yaron (2004) vs. an i.i.d. model.
    Bansal and Yaron: "*it is very difficult to discriminate between a purely i.i.d. process and one that incorporates a small persistent component*."
  - And yet they have very different asset pricing implications.

When is learning likely to be difficult or slow?

Barro, Nakamura, Steinsson, and Ursua (2011) use *all* available consumption data for a large a cross-section of countries and find still high parameter uncertainty

- Estimate probability of exiting a world disaster is 13.5% per year, with standard error of 2.7%
- 2 standard error bounds on bad state duration: 4.5 to 9 years.

And, these estimates are the best we can do **today**.

- Uncertainty about such states was certainly higher 100 years ago, at the start of our samples

We consider parameter uncertainty about the dynamics of a Great Depression, calibrated to U.S. data.

- Consider unbiased priors to focus on priced parameter uncertainty (see Cogley and Sargent (2008) for a case with biased priors and anticipated utility).

Rare events and model risk: summary

Learning about transition probabilities (ie, parameters governing *persistence*) has tremendous impact on asset prices

- Six-fold increase in risk premium; also high long-sample average price of risk, excess volatility
- In bad state, return volatility of over 80% p.a. with risk premium of 40% p.a. at the onset of the Depression state
- Note that volatility, in particular, cannot be matched in the corresponding known transition probability model.

Using U.S. consumption data, match risk premium with γ slightly less than 4

- Barro (2009) has $\gamma = 3 - 4$, but uses higher consumption volatility in good state and disasters where permanent consumption drop is -30% based on international consumption data (including durables)

Model does not yield interesting dynamics in normal times, unsurprisingly.
But, if add Model Risk:

- Learning is slow: takes the entire sample to modestly increase likelihood of Long-run risk model versus i.i.d. model.
- Learning generates interesting dynamics: countercyclical prices of risk and risk premia, even though individual models are homoskedastic with constant risk dynamics.

Empirical evidence on formation of macro beliefs

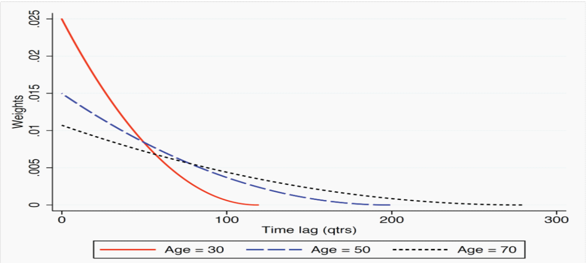
Malmendier and Nagel (2011, 2013): Experiential learning

> *"When forming macroeconomic expectations, individuals put a higher weight on realizations of macroeconomic data experienced during their life-times compared with other available historical data." - Malmendier and Nagel (2013)*

- Investors who *personally* experienced the Great Depression (the 'old') expect lower equity returns than investors that did not *personally* experience the Great Depression (the 'young') (M&N (2011))
- 'Old' macro analysts believe shocks to inflation are less persistent than 'young' macro analysts (M&N (2013))
  - i.e., the 'young' are more prone to believe in 'a new regime'
- Yet, professional analysts inflation forecasts are not biased. In fact, they beat statistical forecasting models in a horse race (Bekaert and Ang (2005)).
- Median macro beliefs from surveys very good forecasters (Ang, Bekaert, and Wei (2007))

Experiential learning evidence

- From 57 years of micro data on inflation expectations (Michigan Survey of Consumers), MN (2013) estimate following weighting scheme for experience-based learning



- **Mark Twain**:

  "My father was an amazing man. The older I got, the smarter he got.."

- My kids, traders, risk-managers, CIOs . . .

The Model: Aggregate Dynamics

- Consider an *exchange economy* with (truly) i.i.d. aggregate consumption:

$$\Delta c_t = \mu + \sigma\varepsilon_t + d_t,$$

$\varepsilon_t \overset{i.i.d.}{\sim} N(0,1)$, $d_t = \underline{d}$ with prob p, 0 otherwise

- μ **and** p **unknown.**
- We calibrate:
  - $\underline{d} = -18\%$ (U.S. Great Depression consumption drop), $p = 1.7\%$ *p.a.* (as in Barro (2006))
  - In all models: $E[\Delta c_{Annual}] = 1.8\%$, $\sigma[\Delta c_{Annual}] = 2.2\%$ as in U.S. sample from 1929-2013.

Aggregate Dynamics and Learning

Agent i born with prior $\mu \sim N\left(g_{i,0}, A_{i,0}\sigma^2\right)$ and $p \sim \beta\left(a_{i,0}, A_{i,0}^{-1} - a_{i,0}\right)$.

- Bayesian within-life learning implies subjective consumption dynamics:

$$\begin{aligned}
\Delta c_{t+1} &= g_{i,t} + E_t^i[p]\,\underline{d} + \sqrt{A_{i,t}+1}\,\sigma\tilde{\varepsilon}_{t+1} + \left(d_{t+1} - E_t^i[p]\,\underline{d}\right) \\
g_{i,t+1} &= g_{i,t} + \frac{A_{i,t}}{\sqrt{A_{i,t}+1}}\sigma\tilde{\varepsilon}_{t+1} \\
A_{i,t+1}^{-1} &= A_{i,t}^{-1} + 1 \\
a_{i,t+1} &= a_{i,t} + \mathbf{1}_{d_{t+1}=\underline{d}}
\end{aligned}$$

$E_t^i[p] = a_{i,t}A_{i,t}$, $var_t^i(p) = \frac{A_{t,t}}{1+A_{t,t}} E_t^i[p]\left(1 - E_t^i[p]\right)$.

Thus, if i lives forever, $A_{i,\infty} = 0$ and $g_{i,\infty} = \mu$, $E_\infty^i[p] = p$; $var_\infty^i[p] = 0$

The Experiential Learning Bias

Each agent lives for T years

Inherit 'parents' mean beliefs, but increase dispersion of beliefs:

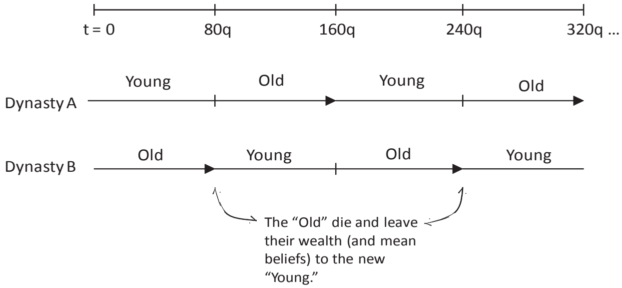
$$\begin{aligned} g_{Young,0} &= g_{Old,T} \\ E_0^{Young}[p] &= E_T^{Old}[p] \\ A_{Young,0} &= \mathbf{k} A_{Old,T}. \end{aligned}$$

where $k > 1$

- Thus, the Young update 'more' than the Old
- Sensitivity of update in expectations to shock $\approx A_{i,t}$

If $k = 1$, economy converges to one-agent, rational expectations case

The Model: OLG setup



The model: Calibration

- Learning channel: choosing A_0 (or equivalently k)
- Use numbers estimated by Malmendier and Nagel (2009):
  - The sensitivity of Young's (about 30y) updates in beliefs to macro shock: 0.025
  - The sensitivity of Old's (about 70y) updates in beliefs to macro shock: 0.01
- In our case, amounts to calibration of $A_0 = 0.025$.
  - Implies sensitivity of old of 0.005, a little on the conservative side
- *How 'wrong' are investors' expectations wrt macroeconomic quantities in model?*
  - Takes on average 175 years of time-series data to reject the model of consumption growth used by a Dynasty
  - Bias is, however, immediately identified in the cross-section of beliefs (as in MN)

Equilibrium (E-Z utility)

- The market price-dividend ratio is (i denotes agent A or B):

$$\frac{P_t}{D_t} = \sum_{j=1}^{\infty} \frac{\left(1 + E_t^i\left[g_{t+j}\right]\right)^j}{\left(1 + E_t^i\left[r_{t+j}\right]\right)^j}$$

- Thus, as argued earlier, prices can be too high or too low due to investors' growth expectations changing as they update their model belief
- Further, since optimist holds more stock, pessimists' beliefs dominate in bad times (added volatility)
- The average risk premium is high as economy features model-risk:

$$E_t^i\left[r_{t+1} - r_{f,t}\right] \approx \gamma cov_t\left(r_{t+1}, \Delta c_{t+1}^i\right) + (\theta - 1)\, cov_t\left(r_{t+1}, r_{W,t+1}^i\right)$$

where $\gamma > 0$, $\theta - 1 > 0$.

  - Here r_W^i is return on wealth of agent i, which is volatile as beliefs shift when models are updated.

Standard moments – 'Uncertain Mean'

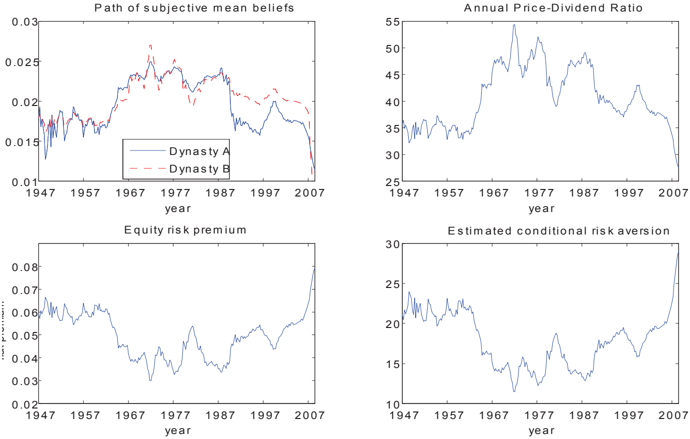
- Unknown mean, no disasters

	Data	'This Time is Different'	Known mean
	1929 – 2011	$EZ: \gamma = 10$, $\psi = 1.5$, $\beta = 0.994$	$EZ: \gamma = 10$, $\psi = 1.5$, $\beta = 0.994$
$E_T[r_m - r_f]$	5.1	5.2	1.5
$\sigma_T[r_m - r_f]$	20.2	16.6	12.9
$SR_T[r_m - r_f]$	0.25	0.31	0.12
$\sigma(pd_t)$	0.29	0.14	0
$\rho(pd_t)$	0.81	0.86	n/a

- Model risk yields: excess volatility, high Sharpe ratios, persistent P/D ratio as in the data

Feed the model consumption data from 1947-2009

Figure 6 - "Uncertain Mean": Simulated Paths



- P/D ratio ranges from about 27 to 54
  - Benchmark case (no model risk): P/D ratio is constant
  - Model risk yields prolonged ($>$ 10 year) periods of under- and over-valuation (+/- 30%)

Standard moments – 'Uncertain Probability'

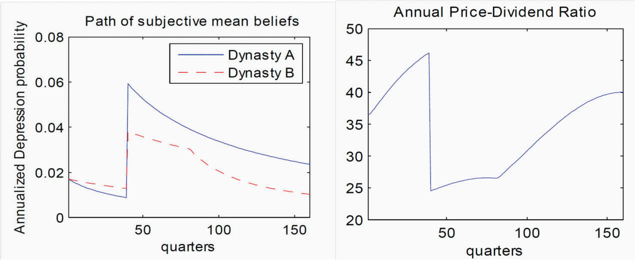
- Known mean, unknown disaster probability

	Data	**'This Time is Different'**	**Known probability**
	$1929 - 2011$	$EZ: \ \gamma = 5$, $\psi = 1.5$, $\beta = 0.994$	$EZ: \ \gamma = 5$, $\psi = 1.5$, $\beta = 0.994$
$E_T\left[r_m - r_f\right]$	5.1	4.9	1.7
$\sigma_T\left[r_m - r_f\right]$	20.2	16.7	13.2
$SR_T\left[r_m - r_f\right]$	0.25	0.30	0.13

- Again, TTiD model performs well
  - Excess volatility, high Sharpe ratios, persistent P/D ratio as in the data (next slide)

Depressions and valuation levels

'Uncertain Probability'-case



- P/D ratio ranges from about 25 to 45
  - It takes 30+ years to 'undo' pessimistic beliefs due to Depression event

Implications for long-term investment

- A very long-horizon investor (that does not suffer from the "This Time is Different"-bias, or simply believes this bias is important for asset prices) can take advantage of the mispricing

Strategic allocation:

- High average exposure to risky assets (e.g., equity)
- Why? (Many) Investors in the market perceive that shocks to the price-dividend ratio are permanent, as they update model beliefs
  - Thus, they perceive asset prices to be very volatile, also in the long-run
  - Given this, they require a high risk premium
  - In reality, asset prices mean-revert so long-run volatility is not so high

Tactical allocation:

- 'Lean against the wind'
- Decrease (increase) equity share when valuations are high (low)
  - Easier said than done: model suggests mean-reversion due to the experiential learning-bias is very slow