Empirical Asset Pricing

Part 4: Cross-sectional Asset Pricing

UCLA | Fall 2016

by Mikhail Chernov

This version: November 12, 2016

14. Moving from aggregate to specific

Focus on cash-flows

- The discussion so far focused on
  - properties of the aggregate endowment
  - preferences of economic agents
- All of this leads to a specification of the pricing kernel
- In order to value different assets, one needs to focus on differences in cash flows
- Consider a cash-flow process D_t with growth rate $G_{t,t+n} = D_{t+n}/D_t$
  - Hypothetical "zero-coupon" claims to $G_{t,t+n}$ with a price denoted by $\widehat{P}_t^n$
    - In the special case of a claim to the cash flow of one US dollar, its price is denoted by P_t^n
  - A yield on such an asset as: $\widehat{y}_t^n = -n^{-1}\widehat{p}_t^n$
    - Reserve a special notation $y_t^n \equiv -n^{-1} p_t^n$ for a yield on a US nominal bond
- Examples: foreign bonds if D_t is an exchange rate; inflation-linked bonds if D_t is price level; and equities if D_t is a dividend

Excess returns and yields

- Returns are connected to yields. Consider a hold-to-maturity $n-$period log return:

$$r_{t,t+n} = \log(G_{t,t+n}/\widehat{P}_t^n) = g_{t,t+n} + n\widehat{y}_t^n$$

- Therefore, we can express the term spread between average returns as:

$$n^{-1}Er_{t,t+n} - Er_{t,t+1} = E(\widehat{y}_t^n - \widehat{y}_t^1)$$

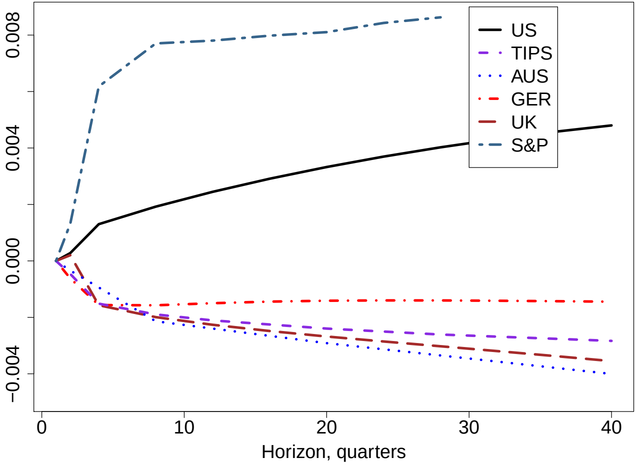
- Define excess holding return per period as

$$rx_{t,t+n} = n^{-1}(r_{t,t+n} - r_{t,t+n}^n)$$

- Therefore, the average difference between one- and n-period excess returns is equal to difference between average term spreads:

$$E(rx_{t,t+n} - rx_{t,t+1}) = E(\widehat{y}_t^n - \widehat{y}_t^1) - E(y_t^n - y_t^1)$$

Evidence: spread in excess returns at multiple horizons



Source: Backus, Boyarchenko, and Chernov (2015)

Implications

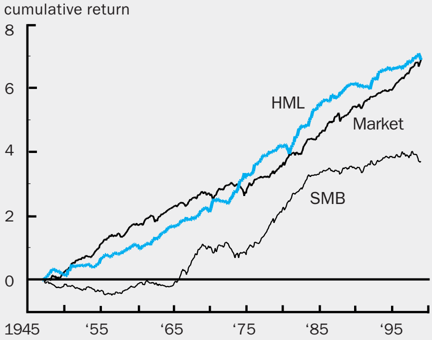
- The (US nominal) yield curve reflects the properties of the (US nominal) pricing kernel
- Departures from this curve reflect the properties of individual cash flows
- The specific forms of departures suggest different
  - persistence
  - conditional and unconditional volatility

  of the cash flows
- All of this to be clarified further

15. Cross-sectional asset pricing

Puzzles: Value and Size Premiums

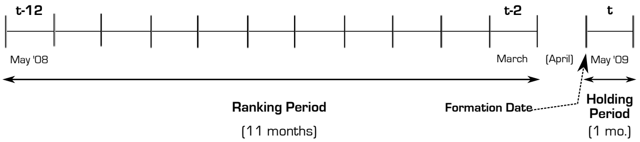
Cumulative returns on market portfolios

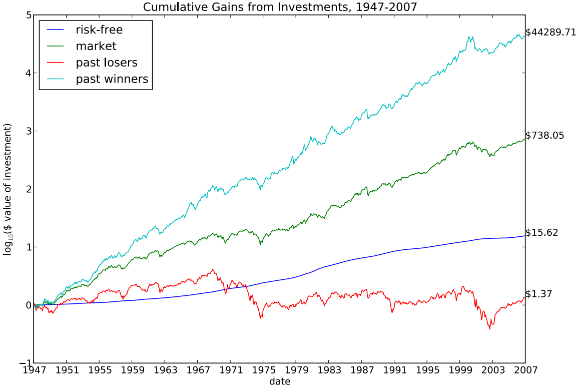


Notes: Cumulative returns on the market RMRF, SMB, and HML portfolios. The SMB return is formed by $R_t^{TB} + aSMB_t$; $a = \sigma(RMRF)/\sigma(SMB)$. In this way it is a return that can be cumulated rather than a zero-cost portfolio, and its standard deviation is equal to that of the market return. HML is adjusted similarly. The vertical axis is the log base 2 of the cumulative return or value of $1 invested at the beginning of the sample period. Thus, each time a line increases by 1 unit, the value doubles.

Source: Cochrane (1999)

Puzzles: Momentum





Source: Daniel and Moskowitz (2013)

16. Linear factor models

Zoology of consumption-based models

- Recursive:

$$m_{t+1} = \log\beta \cdot \alpha/\rho + (\rho - 1)\alpha/\rho \cdot g_{t+1} + (\alpha - \rho)/\rho \cdot r_{w,t+1}$$

- Consumption CAPM ($\alpha = \rho$):

$$m_{t+1} = \log\beta - (1-\alpha) \cdot g_{t+1} \Leftrightarrow m_{t+1} = (\alpha - 1) \cdot r_{w,t+1}$$

  (when iid)
- Conditional consumption CAPM:

$$m_{t+1} = \phi^0(X_t) - \phi^c(X_t) g_{t+1},$$

  e.g., in the Campbell-Cochrane case (with $X_t = s_t$):

$$\begin{aligned} \phi^0(X_t) &= \log\beta - (1-\alpha)[(\varphi_s - 1)(s_t - s) + g\lambda(s_t)], \\ \phi^c(X_t) &= (1-\alpha)[1 + \lambda(s_t)]. \end{aligned}$$

- Conditional CAPM:

$$m_{t+1} = \phi^0(X_t) - \phi^w(X_t) r_{w,t+1},$$

  but theoretical justification is not clear

Beta representation

- Assume a log-linear factor model:

$$m_{t+1} = \phi^0(X_t) - \phi^{f\top}(X_t) f_{t+1}$$

where f_{t+1} has covariance matrix Ω_t with a vector of diagonal elements (variances) ω_t^2, $E_t(f_{t+1}) = \mu_t^f$

- Assume that f_{t+1} and r_{it+1} (return on asset i with variance σ_{it}^2) are conditionally jointly log-normal

- Then

$$0 = \log E_t e^{\phi^0(X_t) - \phi^{f\top}(X_t) f_{t+1} + r_{it+1}}$$

and

$$\begin{aligned} E_t(r_{it+1}) + \sigma_{it}^2/2 &= -\phi_t^0 + \phi_t^{f\top}\mu_t^f - \phi_t^{f\top}\Omega_t\phi_t^f/2 + cov_t(r_{i,t+1}, f_{t+1})^\top \phi_t^f \\ &\equiv \mu_t^0 + \beta_{it}^\top \lambda_t, \end{aligned}$$

where $\lambda_t \equiv \Omega_t \phi_t^f$ are the factor risk premiums.

Time-series implications

- Suppose risk-free asset exists and f_t are log-returns on traded assets (or mimicking portfolios), then (if ι is a vector of ones)

$$\begin{aligned} E_t(r_{ft}) = \mu_t^0 &\quad\Rightarrow\quad \phi_t^0 = -r_{ft} + \phi_t^{f\top}\mu_t^f - \phi_t^{f\top}\Omega_t\phi_t^f/2, \\ \mu_t^f + \omega_t^2/2 = \iota\mu_t^0 + \Omega_t\phi_t^f &\quad\Rightarrow\quad \phi_t^f = \Omega_t^{-1}(\mu_t^f - \iota r_{ft} + \omega_t^2/2) \end{aligned}$$

- Back to the beta representation:

$$E_t(\tilde{rx}_{it+1}) \equiv E_t(r_{it+1}) - r_{ft} + \sigma_{it}^2/2 = \beta_{it}^\top(\mu_t^f - \iota r_{ft} + \omega_t^2/2)$$

- Important:
  - ϕ_t^0 and ϕ_t^f cannot be used freely to fit the cross-section: they control r_{ft} and λ_t
  - ϕ_t^0 and ϕ_t^f have time-series implications.

What if factors are not returns?

- This is particularly relevant for links to the macro economy
- We can no longer characterize ϕ_t^f, intercept of excess returns is not known
- Must use cross-sectional tests: Fama-MacBeth, or two-pass regression

Two-pass regression

- First Pass:
  - Regress (time-series)

$$r_{it+1} + \sigma_{it}^2/2 = a_i + \beta_i f_{t+1} + e_{i,t+1}$$

  - Save $\hat{\beta}_i$'s, $\bar{\bar{rx}}_{it+1}$

- Second Pass:
  - Regress (cross-section)

$$\bar{\bar{rx}}_{it+1} = \lambda\hat{\beta}_i + \alpha_i$$

  - Should impose zero intercept
  - If estimate the intercept and then test if it is zero $\Leftrightarrow$ average error is zero, weaker than all errors are zero

Fama-MacBeth

- Constant betas: the first pass is the same
- Run T cross-sectional regressions

$$\tilde{rx}_{it} = \lambda_{0t} + \hat{\beta}_i \lambda_{1t} + u_{it},\ \hat{\lambda} = (\hat{\lambda}_0, \hat{\lambda}_1)^\top$$

- Estimate

$$\hat{\lambda} = T^{-1} \sum_{t=1}^{T} \hat{\lambda}_t$$

- Time-varying betas: use rolling regressions to construct time-varying betas $\hat{\beta}_{it}$

GMM-based estimation

- Shows up under different names: GMM, SDF method, Euler equation errors
  - Based on the FTAP: $E_t(M_{t+1}R_{t+1}) = 1$

- FTAP implies two types of moment conditions $E_t(h_{t+1}) = 0$
  1. Returns: $h^i_{t+1} = M_{t+1}R^i_{t+1} - 1$
  2. Excess returns: $h^i_{t+1} = M_{t+1}(R^i_{t+1} - R_{ft})$

- Important: the second set must be complemented by $h^0_{t+1} = M_{t+1}R_{ft} - 1$

- Hansen and Singleton (1982) estimated an expected-utility-based model and did not use h^0

- After that the moment conditions can be conditioned down to $E(h_{t+1}) = 0$, or use conditioning information $E(z_t \otimes h_{t+1}) = 0$

Which one? Beta method or SDF method

- Should be equivalent (Jagannathan and Wang, 2002)
- The trouble is that it is (psychologically) harder to keep track of cross-equation restrictions implied by theory
  - Our earlier discussion of time-series implications
  - Additional restrictions implied by macro-based theory
- Example 1: Hansen and Singleton (1982) study $M_{t+1} = \beta G_{t+1}^{\alpha-1}$, so have restrictions on the magnitudes of β and α
- Example 2: Yogo (2006) (simplified here) studies

$$M_{t+1} = \beta^{\alpha/\rho} G_{t+1}^{(\rho-1)\alpha/\rho} R_{w,t+1}^{(\alpha-\rho)/\rho}$$

  - $R_{w,t+1}$ is proxied by the market return
  - Recall that

$$r_{w,t+1} = -\log\beta + g_{t+1} + \rho(\log u_{t+1} - \log\mu_t(G_{t+1}u_{t+1}))$$

  - $cov_t(R_{w,t+1}, G_{t+1})$ depends on the assumed behaviour of consumption; using the market return does not test the model

17. Rejection of CAPM and beyond

The Fama-French Portfolios

- Pre-FF evidence: a wide spread in average returns and betas of portfolios sorted on size
  - Size (ME) and betas of size are highly correlated (-0.988)
  - No power to separate ME from beta effect in average returns
- FF use double sort: 10 deciles of ME, then each into 10 pretest beta deciles
  - Within a given ME range, pre- and post-test betas can vary widely - reduces correlation between ME and beta
- Portfolio rebalanced yearly, 1963–1989. Beginning date determined by availability of Compustat data on book value of equity, earnings
- Table of 100 mean portfolio returns is main intuitive evidence against CAPM

Evidence against CAPM

	All	Low-β	β-2	β-3	β-4	β-5	β-6	β-7	β-8	β-9	High-β
Panel A: Average Monthly Returns (in Percent)											
All	1.25	1.34	1.29	1.36	1.31	1.33	1.28	1.24	1.21	1.25	1.14
Small-ME	1.52	1.71	1.57	1.79	1.61	1.50	1.50	1.37	1.63	1.50	1.42
ME-2	1.29	1.25	1.42	1.36	1.39	1.65	1.61	1.37	1.31	1.34	1.11
ME-3	1.24	1.12	1.31	1.17	1.70	1.29	1.10	1.31	1.36	1.26	0.76
ME-4	1.25	1.27	1.13	1.54	1.06	1.34	1.06	1.41	1.17	1.35	0.98
ME-5	1.29	1.34	1.42	1.39	1.48	1.42	1.18	1.13	1.27	1.18	1.08
ME-6	1.17	1.08	1.53	1.27	1.15	1.20	1.21	1.18	1.04	1.07	1.02
ME-7	1.07	0.95	1.21	1.26	1.09	1.18	1.11	1.24	0.62	1.32	0.76
ME-8	1.10	1.09	1.05	1.37	1.20	1.27	0.98	1.18	1.02	1.01	0.94
ME-9	0.95	0.98	0.88	1.02	1.14	1.07	1.23	0.94	0.82	0.88	0.59
Large-ME	0.89	1.01	0.93	1.10	0.94	0.93	0.89	1.03	0.71	0.74	0.56

Source: Fama and French (1992)

- Holding beta constant, average returns are higher for firms with smaller market caps
- But holding market cap constant, average returns show no pattern across firm betas

Evidence in favor of size and value

	Book-to-Market Portfolios										
	All	Low	2	3	4	5	6	7	8	9	High
All	1.23	0.64	0.98	1.06	1.17	1.24	1.26	1.39	1.40	1.50	1.63
Small-ME	1.47	0.70	1.14	1.20	1.43	1.56	1.51	1.70	1.71	1.82	1.92
ME-2	1.22	0.43	1.05	0.96	1.19	1.33	1.19	1.58	1.28	1.43	1.79
ME-3	1.22	0.56	0.88	1.23	0.95	1.36	1.30	1.30	1.40	1.54	1.60
ME-4	1.19	0.39	0.72	1.06	1.36	1.13	1.21	1.34	1.59	1.51	1.47
ME-5	1.24	0.88	0.65	1.08	1.47	1.13	1.43	1.44	1.26	1.52	1.49
ME-6	1.15	0.70	0.98	1.14	1.23	0.94	1.27	1.19	1.19	1.24	1.50
ME-7	1.07	0.95	1.00	0.99	0.83	0.99	1.13	0.99	1.16	1.10	1.47
ME-8	1.08	0.66	1.13	0.91	0.95	0.99	1.01	1.15	1.05	1.29	1.55
ME-9	0.95	0.44	0.89	0.92	1.00	1.05	0.93	0.82	1.11	1.04	1.22
Large-ME	0.89	0.93	0.88	0.84	0.71	0.79	0.83	0.81	0.96	0.97	1.18

Source: Fama and French (1992)

- Holding market cap constant, average returns increase with growth, BE/ME
- Holding BE/ME constant, average returns decline with ME

Fama-MacBeth on Fama-French portfolios

	7/63–12/90 (330 Mos.)			7/63–12/76 (162 Mos.)			1/77–12/90 (168 Mos.)		
Variable	Mean	Std	t(Mn)	Mean	Std	t(Mn)	Mean	Std	t(Mn)
NYSE Value-Weighted (VW) and Equal-Weighted (EW) Portfolio Returns									
VW	0.81	4.47	3.27	0.56	4.26	1.67	1.04	4.66	2.89
EW	0.97	5.49	3.19	0.77	5.70	1.72	1.15	5.28	2.82
$R_{it} = \mathrm{a} + \mathrm{b}_{2t}\ln(\mathrm{ME}_{it}) + \mathrm{b}_{3t}\ln(\mathrm{BE/ME}_{it}) + e_{it}$									
a	1.77	8.51	3.77	1.86	10.10	2.33	1.69	6.67	3.27
b_2	−0.11	1.02	−1.99	−0.16	1.25	−1.62	−0.07	0.73	−1.16
b_3	0.35	1.45	4.43	0.36	1.53	2.96	0.35	1.37	3.30
$R_{it} = \mathrm{a} + \mathrm{b}_{1t}\beta_{it} + \mathrm{b}_{2t}\ln(\mathrm{ME}_{it}) + \mathrm{b}_{3t}\ln(\mathrm{BE/ME}_{it}) + e_{it}$									
a	2.07	5.75	6.55	1.73	6.22	3.54	2.40	5.25	5.92
b_1	−0.17	5.12	−0.62	0.10	5.33	0.25	−0.44	4.91	−1.17
b_2	−0.12	0.89	−2.52	−0.15	1.03	−1.91	−0.09	0.74	−1.64
b_3	0.33	1.24	4.80	0.34	1.36	3.17	0.31	1.10	3.67

- FF do not test if b_1 equals excess market return

Ad-hoc factor models

- Pick f's and replace them with factor-mimicking portfolios
- Main goal of FF93 is to identify ME and BE/ME with stock market factors
- Idea behind identification: If low ME stocks share exposure to common risk, proxy for risk is return to portfolio of low-ME stocks less return to portfolio of high-ME stocks; same story for BE/ME
- Construction procedure:
    1. Pick a year
    2. Classify all NYSE/Amex/Nasdaq stocks into two size groups (median NYSE) as of beginning of year
    3. Classify all stocks into three BE/ME groups
    4. Use groups to construct value-weighted returns to six portfolios
    5. SMB is equal-weighted return to three small-stock portfolios less equal weighted return to three large-stock portfolios
    6. HML is equal-weighted return to two highest BE/ME portfolios less equal weighted return to two lowest BE/ME portfolios

Tautology?

- Not quite. SMB and HML explain time-series variation of the portfolios
- FF's claim that SMB and HML mimic macro factors.
- If so, they should explain other XS features, not just size and value (B/M)
- FF96 argue that the model explains variation of expected returns related to E/P, CF/P, sales growth, long-term reversals
- But all of these are related to B/M
- Momentum cannot be explained by FF96
- Since then: idiosyncratic vol, net equity issuance, gross profitability, ...

Conditional factor models

- Early studies considered factor models with constant loadings. Maybe we can explain the puzzles with time-varying risk premiums?
- More recent studies specify linear dependence of loadings on the state:

$$m_{t+1} = \phi^0(X_t) - \phi^f(X_t) f_{t+1} = \phi^0 + \phi^{0x} X_t - \phi^f f_{t+1} - \phi^{fx} X_t f_{t+1}.$$

- A one-factor conditional model becomes an unconditional multi-factor model
- Unfortunately, Lewellen and Nagel (2006) conclude that conditional models do not help

How may conditional models help?

- Suppose $m_{t+1} = \phi_t^0 - \phi_t^f f_{t+1}$ is correct
- A researcher uses $m_{t+1} = \phi^0 - \phi^f f_{t+1}$ instead
- The conditional model implies:

$$E_t(r^e_{it+1}) = cov_t(r_{it+1}, f_{t+1})\phi_t^f$$
$$E(r^e_{it+1}) = E(cov_t(r_{it+1}, f_{t+1})) \cdot E(\phi_t^f) + cov(cov_t(r_{it+1}, f_{t+1}), \phi_t^f)$$
$$E(cov_t(r_{it+1}, f_{t+1})) = cov(r_{it+1}, f_{t+1}) - cov(E_t(r_{it+1}), E_t(f_{t+1}))$$

- If f_{t+1} has a conditional mean of zero (a common normalization in these models), then $cov(cov_t(r_{it+1}, f_{t+1}), \phi_t^f)$ is the term that distinguishes implications of a conditional model
- How large can it be?

Evaluating the covariance term

- Suppose that f_t is the market index log-return
- Risk premium is $\lambda_t = \phi_t^f \omega_t^2$. Assume $\omega_t = \omega$ to maximize time-variation due to ϕ_t^f. Then $\beta_{it} = cov_t(r_{it+1}, f_{t+1})/\omega^2$ and

$$cov(cov_t(r_{it+1}, f_{t+1}), \phi_t^f) = cov(\beta_{it}, \lambda_t)$$

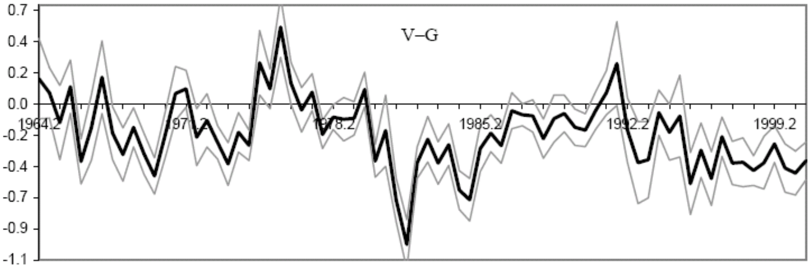
- Obtain an upper bound on the covariance if $corr(\beta_{it}, \lambda_t) = 1$
- $\sigma(\beta_{it}) \approx 0.30$ (next page) and $\sigma(\lambda_t) \leq 0.5\%$ per month. Thus,

$$max(cov(\beta_{it}, \lambda_t)) = 0.15\% \leq 0.50\% \text{ (value spread in the data)}$$

Computing vol of beta

Conditional betas, 1964 – 2001

The figure plots conditional betas for size, B/M, and momentum portfolios. The dark line is the point estimate and the light lines indicate a two-standard-deviation confidence interval. Betas are estimated semiannually (non-overlapping windows) using daily returns. The portfolios are formed from all NYSE and Amex stocks on CRSP / Compustat. We begin with 25 size-B/M portfolios (5×5 sort, breakpoints determined by NYSE quintiles) and 10 return-sorted portfolios, all value weighted. 'S–B' is the average return on the five low-market-cap portfolios (Small) minus the average return on the five high-market-cap portfolios (Big). 'V–G' is the average return on the five high-B/M portfolios (Value) minus the average return on the five low-B/M portfolios (Growth). Return-sorted portfolios are formed based on past 6-month returns. 'W–L' is the return on the top decile (Winners) minus the return on the bottom decile (Losers).



Source: Lewellen and Nagel (2006)

A puzzle?

- Jagannathan and Wang (JF, 1996) and Lettau and Ludvigson (JPE, 2001) are very famous papers making the case for conditional pricing kernels
- In cross-sectional tests one usually regresses (cross-sectional) average returns on the factor betas
- This ignores the time-series implications of these models, e.g., the loading on beta should be $\mu_t^f - \iota r_{ft} + \omega_t^2/2$
- Nagel and Singleton (2011) focus on the gap between implied and empirical dynamics of conditional moments: the implied can be more than an order of magnitude more volatile
- See Nagel (2013) for further details

Low-Dimensional Factor Structures

- There is another mechanical reason why conditional models seem to work (Lewellen, Nagel, and Shanken, 2010)
- Consider FF regressions:

$$\underset{N\times 1}{r_t^e} = \underset{N\times K}{B}\,\underset{K\times 1}{p_t^e} + e_t,$$

  where p_t^e are excess returns on the factors, $K = 3$ in FF
- Write the unconditional version of the conditional model as

$$m_{t+1} = \phi^0 - \phi^\top \tilde{f}_{t+1}, \tilde{f}_t \perp e_t, J = dim(\tilde{f}) = 3, var(\tilde{f}_{t+1}) = \Omega$$

- Then the $N \times J$ $\beta = Bcov(p_{t+1}^e, \tilde{f}_{t+1}^\top)\Omega^{-1}$
- The betas are $K = J$ linear combinations of B, so betas "explain" the cross-section as well as B regardless of the choice of p_t^e
- High R^2 carry no economic information

Addressing the low-dimensional issue

- Add more portfolios (increase N), e.g., industry portfolios

- If factors are traded assets, use additional moment conditions $E(r^e_{t+1}) = \beta\lambda$ and $E(f^e_{t+1}) = \lambda$
  - If factors are non-traded assets, then factor-mimicking portfolios will not help

- Take time-series implications seriously
  - If factors are non-traded assets, use instruments (e.g., optimal instruments as in Nagel and Singleton, 2011)

18. Recursive preferences

Overview

- Recursive preferences yield a pricing kernel that is a function of shocks to consumption
  - Consumption itself
  - Expected consumption growth
  - Consumption variance
  - ...

- If one believes that the quality of the consumption data is poor, one may want to replace consumption, or its shocks, with shocks to observables, such as the market return

- How? Using the “budget constraint”: $W_{t+1} = R_{w,t+1} W_t - C_{t+1}$

- How does one measure shocks to observables? VAR

Substituting consumption out

- We've seen earlier that

$$r_{w,t+1} = -\log\beta + g_{t+1} + \rho(\log u_{t+1} - \log\mu_t(G_{t+1}u_{t+1}))$$

- The pricing kernel is

$$m_{t+1} = \log\beta + (\alpha-1)g_{t+1} + (\alpha-\rho)(\log u_{t+1} - \log\mu_t(G_{t+1}u_{t+1}))$$

- Want to get rid of consumption growth, so substitute it from the first equation into the second equation:

$$m_{t+1} = \alpha\log\beta + (\alpha-1)r_{w,t+1} + (1-\rho)\alpha(\log u_{t+1} - \log\mu_t)$$

[$\log\mu_t$ is a shorthand for $\log\mu_t(G_{t+1}u_{t+1})$ in this section]

- Risk premiums are determined by the covariance of m with cash flows, so need to know shocks to m

Shocks to the pricing kernel (I)

- By the Campbell (1991) representation, the shock to return is

$$\begin{aligned} r_{w,t+1} - E_t r_{w,t+1} &= (E_{t+1} - E_t) \sum_{j=1}^{\infty} b_1^j g^d_{w,t+1+j} - (E_{t+1} - E_t) \sum_{j=1}^{\infty} b_1^j r_{w,t+1+j} \\ &\equiv N_{CF,t+1} - N_{DR,t+1} \end{aligned}$$

- The shock to the value function is

$$\log u_{t+1} - E_t \log u_{t+1} = b_1 (\log \mu_{t+1} - E_t \log \mu_{t+1})$$

  - Recall that $u_{t+1} \approx b_0 + b_1 \log \mu_{t+1}$

- Use the expression for m_{t+1} to characterize $\log \mu_t$

Shocks to the pricing kernel (II)

- Under log-normality

$$0 = \log E_t e^{m_{t+1}+r_{i,t+1}} = E_t(m_{t+1}+r_{i,t+1}) + 1/2\, Var_t(m_{t+1}+r_{i,t+1})$$

- Applying this to $i = w$:

$$\begin{aligned}\log \mu_t &= \frac{1}{1-\rho}\log\beta + \frac{1}{1-\rho}E_t r_{w,t+1} + E_t \log u_{t+1} \\ &+ \frac{1}{2\alpha(1-\rho)} Var_t(m_{t+1}+r_{w,t+1})\end{aligned}$$

- Therefore, the shock to $\log u$ is

$$(E_{t+1}-E_t)b_1\left(\frac{1}{1-\rho}r_{w,t+2} + \log u_{t+2} + \frac{1}{2\alpha(1-\rho)}Var_{t+1}(m_{t+2}+r_{w,t+2})\right)$$

Shocks to the pricing kernel (III)

- Solve the shock to log u forward:

$$
\begin{aligned}
\log u_{t+1} - E_t \log u_{t+1} &= \frac{1}{1-\rho}(E_{t+1} - E_t) \sum_{j=1}^{\infty} b_1^j r_{w,t+1+j} \\
+ \quad & \frac{1}{2\alpha(1-\rho)}(E_{t+1} - E_t) \sum_{j=1}^{\infty} b_1^j Var_{t+j}(m_{t+1+j} + r_{w,t+1+j}) \\
\equiv \quad & \frac{1}{1-\rho} N_{DR,t+1} + \frac{1}{2\alpha(1-\rho)} N_{Var,t+1}
\end{aligned}
$$

- Thus, innovation to the pricing kernel is

$$
\begin{aligned}
& m_{t+1} - E_t m_{t+1} = \alpha N_{DR,t+1} + 1/2 N_{Var,t+1} + (\alpha - 1)(r_{w,t+1} - E_t r_{w,t+1}) \\
\equiv \quad & N_{DR,t+1} + 1/2 N_{Var,t+1} + (\alpha - 1) N_{CF,t+1}
\end{aligned}
$$

- See, e.g., Bansal, Kiku, Shaliastovich, and Yaron (2014); Campbell, Giglio, Polk, and Turley (2013)

Measuring shocks

- Specify a VAR

$$x_{t+1} = \mu + \Phi(x_t - \mu) + \sigma_t u_{t+1}, \; u_{t+1} \sim N(0, \Sigma)$$

- Select $x_{1,t} = r_{w,t+1}$, $x_{2,t} = \sigma_t^2$, $x_{3,t} = \ldots$

- Then one can show that

$$\begin{aligned}
N_{DR,t+1} &= e_1^\top b_1 \Phi (I - b_1 \Phi)^{-1} \sigma_t u_{t+1} \\
N_{Var,t+1} &= \omega b_1 e_2^\top (I - b_1 \Phi)^{-1} \sigma_t u_{t+1}, \; \omega \equiv Var_t(m_{t+1} + r_{w,t+1})/\sigma_t^2 \\
N_{CF,t+1} &= e_1^\top (I + b_1 \Phi (I - b_1 \Phi)^{-1}) \sigma_t u_{t+1}
\end{aligned}$$

- Is this VAR consistent with the model?

- The argument is that one can always find a process for consumption growth that is consistent with this structure
    - Can one?
    - Would it be consistent with observed consumption growth?

Cross-sectional asset pricing

- The model implies

$$\begin{aligned} E_t \tilde{rx}_{i,t+1} &= -cov_t(N_{DR,t+1}, r_{i,t+1}) - 1/2 cov_t(N_{Var,t+1}, r_{i,t+1}) \\ &- (\alpha - 1) cov_t(N_{CF,t+1}, r_{i,t+1}) \end{aligned}$$

- Switch to w=M and to simple returns, condition down, multiply/divide by $\sigma_M^2 = Var(r_{M,t} - E_{t-1} r_{M,t})$

$$E(R_i - R_f) = \sigma_M^2 \beta_{i,DR_M} - 1/2 \omega \sigma_M^2 \beta_{i,V_M} - (\alpha - 1) \sigma_M^2 \beta_{i,CF_M}$$

- The betas can be constructed from the VAR estimates of shocks, so take them for a ride in the cross-section

Consumption-based asset pricing

- Specify a VAR

$$x_{t+1} = \mu + \Phi(x_t - \mu) + \sigma_t u_{t+1}, \; u_{t+1} \sim N(0, \Sigma)$$

- Select $x_t = (g_t; \pi_t; y_t^1, \sigma_t^2)^\top$ (Zviadadze, 2015)
- Then Bellman-based techniques imply

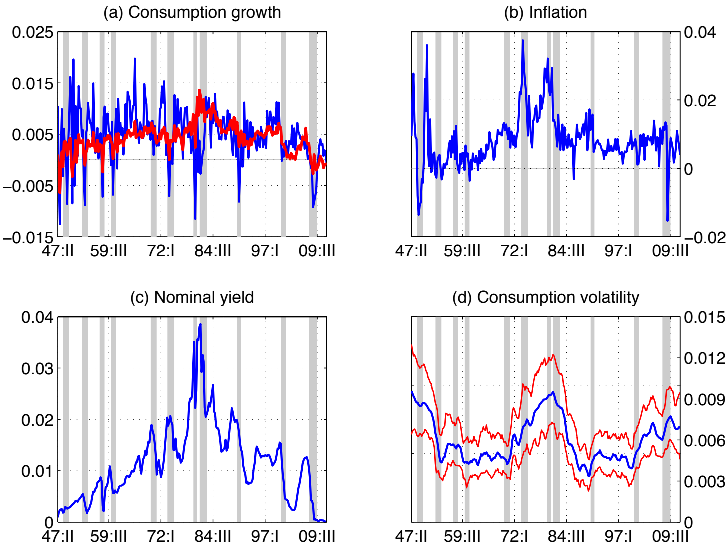
$$m_{t+1} = m + \eta^\top x_t + q^\top \sigma_t u_{t+1}$$

- The equilibrium (nominal) short interest rate is

$$y_t^1 = A g_t + B \pi_t + C y_t^1 + D \sigma_t^2 + E$$

  implies $A = B = D = E = 0$ and $C = 1$.
- $D = 0$ imposes non-linear restriction on parameters of the VAR and coefficients in the solution of $\log u_t$ (drops out in a model considered by Hansen, Heaton, and Li, 2008 as σ_t is constant)
- The rest impose linear restrictions on μ and Φ

The state



Source: Zviadadze (2015)

Cash flows

- Model cash-flows as a function of the state:

$$g^d_{i,t+1} = g^d_i + \eta_i^\top x_t + q_i^\top \sigma_t u_{t+1}$$

- Hansen, Heaton, and Li (2008) consider the case of constant σ_t

$$g^d_{i,t+1} = g^d_i + \eta_i^\top x_t + q_i^\top \sigma u_{t+1} = g^d_i + \gamma_i(L)\sigma u_{t+1}$$

- They highlight the permanent (martingale) component of dividends Beveridge-Nelson decomposition

$$\begin{aligned}
g^d_{i,t+1} &= g^d_i - \eta_i^{*\top} x_{t+1} + \eta_i^{*\top} x_t + \gamma_i(1)^\top \sigma u_{t+1}, \\
\eta_i^{*\top} &= \eta_i^\top (I - \Phi)^{-1}, \\
\gamma_i(1)^\top &= q_i^\top + \eta_i^\top (I - \Phi)^{-1}
\end{aligned}$$

- HHL extract both short- and long-term dynamics from VARs containing consumption, aggregate earnings and i's dividends

Difficulties with modelling cash flows

- Model cash-flows as a function of the state:

$$g^d_{i,t+1} = \mu^d_{i,t} + \sigma_{it}^\top u_{t+1}$$

- If the rep agent holds the market, and dividends are not storable then $C_t = \sum_i D_{i,t}$
- Log-consumption growth:

$$\Delta c_{t+1} = \log \sum_i \frac{D_{i,t+1}}{C_t} = \log \sum_i S_{i,t} e^{\mu^d_{i,t} + \sigma_{it}^\top u_{t+1}},$$

where $S_{i,t} = D_{i,t}/C_t$ – a consumption share

- Not a pretty picture ... Cochrane, Longstaff, and Santa-Clara (2008); Martin (2013); Santos and Veronesi (2010) struggle with this in different ways

Reduced-form modelling of cash flows

- Lettau and Wachter (2011) construct an affine model to capture the value-growth spread, among other things
- They use duration logic: value stocks are more front-loaded with dividends; growth stocks are back-loaded
- Introduce deterministically moving dividend $N = 200$ shares: $S_{i,t} = D_{i,t} / \sum_i D_{i,t}$

$$
\begin{aligned}
S_{1,1} &= S, \\
S_{1,t} &= (1+g_s) S_{1,t-1} = (1+g_s)^t S, \quad t \leq N/2+1 \\
S_{1,t} &= (1+g_s)^{-1} S_{1,t-1}, \quad N/2+1 < t \leq N+1 \\
S_{N/2,1} &= (1+g_s)^{N/2-1} S, \\
S_{N,1} &= (1+g_s) S
\end{aligned}
$$

- If $\widehat{P}_t^n = E_t(M_{t,t+n} D_{t+n})$, then $\widehat{P}_{it} = \sum_{n=1}^{\infty} S_{i,t+n} \widehat{P}_t^n$

Direct evidence about cash flows

- Equities are generating uncertain cash flows every period, so we observe only

$$S_t = \sum_{j=1}^{\infty} \widehat{P}_t^j, \ \widehat{P}_t^n = E_t(M_{t,t+n} D_{t+n})$$

- Our theories are suited nicely for a single cash flow, so it would be nice to measure $\widehat{P}_t^n$ (like we do with bonds)

- van Binsbergen, Brandt, and Koijen (2012) propose to use put-call parity, or futures:

$$\text{Call}_t^n + Ke^{-ny_t^n} = \text{Put}_t^n + S_t - \sum_{j=1}^{n} \widehat{P}_t^j, \text{or}$$

$$F_t^n = e^{ny_t^n}(S_t - \sum_{j=1}^{n} \widehat{P}_t^j)$$

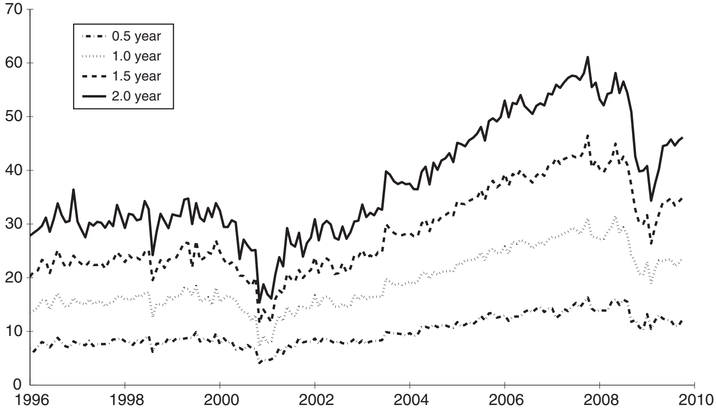
The dividend strips

- These relationships imply:

$$
\begin{aligned}
\widehat{P}_t^{n,m} &\equiv \sum_{j=n}^{m} \widehat{P}_t^j \\
&= \mathrm{Put}_t^m - \mathrm{Put}_t^n - \mathrm{Call}_t^m + \mathrm{Call}_t^n - K\left(e^{-my_t^m} - e^{-ny_t^n}\right), \text{or} \\
\widehat{P}_t^{n,m} &= e^{-ny_t^n} F_t^n - e^{-my_t^m} F_t^m
\end{aligned}
$$

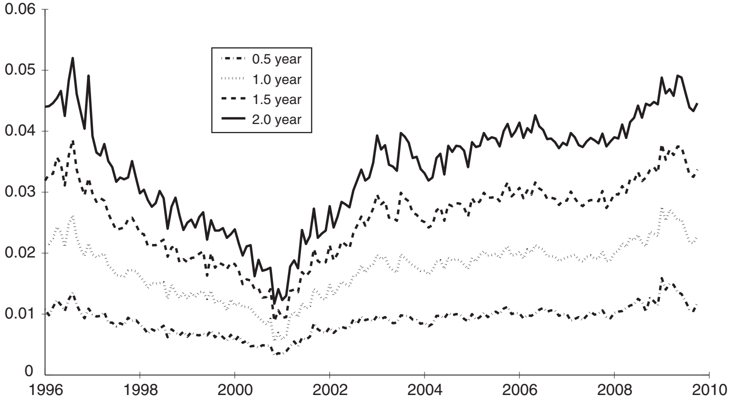
- van Binsbergen, Brandt, and Koijen (2012) consider maturities of 0.5, 1, 1.5, 2 years, and "the rest"

Prices of dividends



Source: van Binsbergen, Brandt, and Koijen (2012)

Relative prices of dividends



Source: van Binsbergen, Brandt, and Koijen (2012)

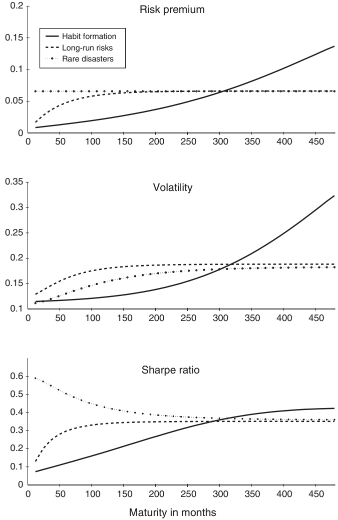
Properties of returns

- Long short-term asset $R_{1,t+1} = (\sum_{j=1}^{n-1} \widehat{P}_{t+1}^{j} + D_{t+1}) / \sum_{j=1}^{n} \widehat{P}_{t}^{j} - 1$
- Dividend steepener $R_{2,t+1} = \widehat{P}_{t+1}^{n-1,m-1} / \widehat{P}_{t}^{n,m} - 1$

	$R_{1,t}$	$R_{1,t} - R_{f,t}$	$R_{2,t}$	$R_{2,t} - R_{f,t}$	$R_{SP500,t}$	$R_{SP500,t} - R_{f,t}$
Mean	0.0116	0.0088	0.0112	0.0084	0.0056	0.0027
	(0.0044)	(0.0044)	(0.0044)	(0.0045)	(0.0047)	(0.0047)
Standard deviation	0.0780	0.0781	0.0965	0.0966	0.0469	0.0468
	(0.0136)	(0.0136)	(0.0171)	(0.0171)	(0.0050)	(0.0050)
Sharpe ratio	0.1124	—	0.0872	—	0.0586	—
	(0.0520)	—	(0.0494)	—	(0.1058)	—
Observations	165	165	165	165	165	165

Source: van Binsbergen, Brandt, and Koijen (2012)

Rejection of all models?



Source: van Binsbergen, Brandt, and Koijen (2012)

Alternative dividend dynamics

- Belo, Collin-Dufresne, and Goldstein (2015) argue that it is easy to fix the issue in both recursive and habit models
- Consider the BY2 setup:

$$
\begin{aligned}
g_t^c &= g^c + x_{t-1} + v_{t-1}^{1/2} w_t \\
d_t - d_{t-1} = g_t^d &= g^d + \eta_d x_{t-1} + q_c v_{t-1}^{1/2} w_t + q_d v_{t-1}^{1/2} u_t
\end{aligned}
$$

- This setup ignores that dividends are leveraged cash flows. Is this distinction important?
- Suppose g^d represent changes in unlevered cash-flows, than claims to levered cashflows will be more volatile and have higher expected returns than those to g^d
- Moreover, levered and unlevered cash-flows are co-integrated, so they have the same volatility over long horizons $\Rightarrow$ volatility of claims to levered cashflows declines with horizon Beveridge-Nelson decomposition

Leverage and co-integration

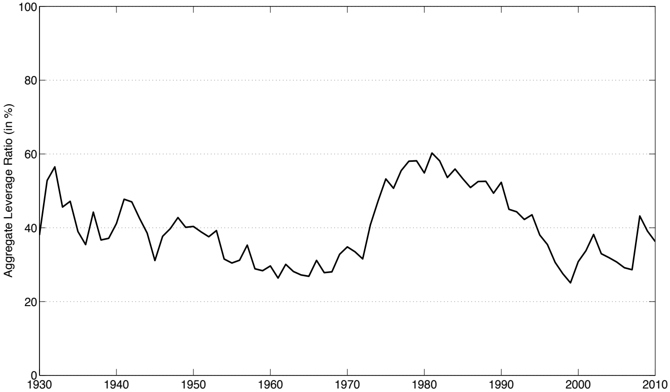
- Consider very basic type of debt: every period a firm issues riskless debt that matures next period with a present value of
$$B(s_t, d_t, x_t, v_t) = e^{s_t} V(d_t, x_t, v_t),$$
where V is the value of the claim on unlevered cashflows, and e^{s_t} is leverage
- Consistent with empirical evidence, s_t is modelled as stationary:
$$s_{t+1} = s + (s_t, x_t, v_t)\eta + \text{shocks}$$
- Dividends are equal to earnings (unlevered cashflows) + change in the bond position:
$$D^s_{t+1} = B(s_{t+1}, d_{t+1}, x_{t+1}, v_{t+1}) - e^{y^1_t(x_t, v_t)} B(s_t, d_t, x_t, v_t) + e^{d_{t+1}}$$
- s_{t+1} is stationary, so is d^s_{t+1}/d_{t+1} (similar to cointegration)
- Because of MM (capital structure irrelevance), implications for returns on perpetuities are the same as for the unlevered case, so the BY2's ability to match certain moments is not affected

Evidence on vol of dividends

	Horizon (years)								Diff	Diff
	1	2	4	6	8	10	15	20	$1-10$	$1-20$
Dividend definition 1: Cash dividends										
$\sigma^T_{D,1}$	15.52	14.15	9.84	7.88	8.15	7.45	7.77	7.34	8.07	8.18
$\sigma^T_{D,2}$	14.98	13.26	8.60	6.49	6.38	5.53	5.09	4.31	9.45	10.67
VR	1	0.84	0.43	0.27	0.33	0.30	0.48	0.40	–	–
p-value	–	0.42	0.09	0.09	0.17	0.19	0.37	0.36	–	–
p-value IID	–	0.16	0.01	0.01	0.04	0.06	0.29	0.28	–	–

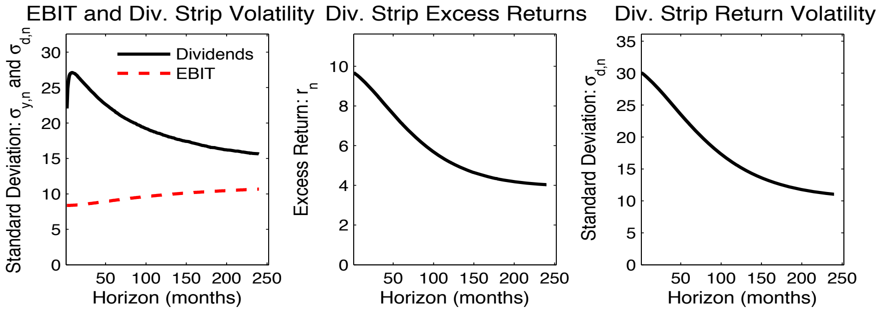
Source: Belo, Collin-Dufresne, and Goldstein (2015)

Aggregate leverage ratio is stationary



Source: Belo, Collin-Dufresne, and Goldstein (2015)

Dividend strips



Source: Belo, Collin-Dufresne, and Goldstein (2015)

Variation on the KLV model

- From Backus, Boyarchenko, and Chernov (2015), motivated by Koijen, Lustig, and Nieuwerburgh (2014)
- Add a jump distubance to Vasicek

$$
\begin{aligned}
m_{t,t+1} &= \log\beta + \theta_m x_{1t} + \lambda_0 w_{t+1} + \lambda_2 z^m_{t+1} \\
x_{1t+1} &= \varphi_1 x_{1t} + w_{t+1}
\end{aligned}
$$

- Stir in some cash flow growth (dividends)

$$
\begin{aligned}
g_{t,t+1} &= \log\gamma + \theta x_{1t} + \theta_g x_{2t} + \eta_0 w_{t+1} + \eta_2 z^g_{t+1} \\
x_{2t+1} &= \varphi_2 x_{2t} + w_{t+1}
\end{aligned}
$$

- The value is $\widehat{P}^1_t = E_t(M_{t,t+1} G_{t,t+1}) \equiv E_t(\widehat{M}_{t,t+1})$
- Transformed pricing kernel

$$
\begin{aligned}
\widehat{m}_{t,t+1} &= (\log\beta + \log\gamma) + (\theta_m + \theta) x_{1t} + \theta_g x_{2t} \\
&\quad + (\lambda_0 + \eta_0) w_{t+1} + \lambda_2 z^m_{t+1} + \eta_2 z^g_{t+1}
\end{aligned}
$$

- iid jump component has no impact on term spreads

KLV model: parameter values

- To calibrate m, consider short rate

$$y_t^1 \quad = \quad -\log E_t(M_{t,t+1}) \quad = \quad \text{const} - \theta_m x_t$$

- Choose
  - $(\varphi_1, \theta_m) = (0.95, 0.003)$ match autocorrelation and variance of y^1
  - $\lambda_0 = -0.12$ matches mean 40-quarter term spread

$$E(y_t^n - y_t^1) = -n^{-1}\left[\lambda_0 \sum_{j=0}^{n-1} b_j + 1/2 \sum_{j=0}^{n-1} b_j^2\right], \quad b_j = \theta_m \frac{1-\varphi_1{}^n}{1-\varphi_1}$$

- To calibrate $\hat{m}$, select $\theta = -\theta_m$ for convenience:

$$\hat{y}_t^1 = -\log E_t(\widehat{M}_{t,t+1}) = \text{const} - \theta_g x_{2t}$$

- Choose
  - $(\varphi_2, \theta_g) = (0.68, 0.029)$ match autocorrelation and variance of one-period dividend yield, $\widehat{y}^1$
  - $\eta_0 = -0.02$ matches mean 40-quarter term spread

Representative agent with EZ preferences

- Assume consumption growth

$$g^c_{t,t+1} = g^c + \theta_c x_{2t} + \sigma w_{t+1} + \eta z^c_{t+1}$$

- Then *real* pricing kernel is

$$\begin{aligned} \widehat{m}_{t+1} &= \widehat{m} + (\rho - 1)\theta_c x_{2t} + [(\alpha - 1)\sigma + (\alpha - \rho)u_x]w_{t+1} \\ &+ (\alpha - 1)\eta z^c_{t+1}, \end{aligned}$$

  where $u_x = b_1\theta_c(1 - b_1\varphi_2)^{-1}$

- Because of iid jumps, persistence of expected cons. growth, φ_2, can be dedicated to matching the real term spreads

- If there are no jumps, φ_2 has to be high to match high one-period risk premiums; this leads to unrealistic yield curves

Nominal pricing kernel

- Assume a process for inflation:

$$g^{\pi}_{t,t+1} = g^{\pi} - \theta_m x_{1t} + \theta_g x_{2t} + \eta_0 w_{t+1} + \eta_2 z^{g}_{t+1}$$

- *Nominal* pricing kernel is

$$\begin{aligned} m_{t+1} &= \widehat{m}_{t+1} - g^{\pi}_{t,t+1} \\ &= m + \theta_m x_{1t} + [(\rho-1)\theta_c - \theta_g] x_{2t} \\ &+ [(\alpha-1)\sigma + (\alpha-\rho) u_x - \eta_0] w_{t+1} + \lambda_2 z^{m}_{t+1} \end{aligned}$$

where $\lambda_2 = [(\alpha-1)^2\eta^2 + \eta_2^2]^{1/2}$, arrival rate ω, jump sizes $\mathcal{N}(\mu_m, 1)$, $\mu_m = [(\alpha-1)\eta\mu_c - \eta_2\mu_g]/\lambda_2$

- Compare this to the affine kernel:

$$\log m_{t,t+1} = \log\beta + \theta_m x_{1t} + 0 x_{2t} + \lambda_0 w_{t+1} + \lambda_2 z^{m}_{t+1}$$

Bibliography I

Backus, David, Nina Boyarchenko, and Mikhail Chernov, 2015, Term structures of asset prices and returns, working paper.

Bansal, Ravi, Dana Kiku, Ivan Shaliastovich, and Amir Yaron, 2014, Volatility, the macroeconomy, and asset prices, *Journal of Finance* 69, 2471–2511.

Belo, Frederico, Pierre Collin-Dufresne, and Robert Goldstein, 2015, Dividend dynamics and the term structure of dividend strips, *Journal of Finance* 70, 1115–1160.

Campbell, John, 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.

———, Stefano Giglio, Christopher Polk, and Robert Turley, 2013, An intertemporal CAPM with stochastic volatility, Working paper, Harvard.

Bibliography II

Cochrane, John, 1999, New facts in finance, *Economic Perspectives* 23, 36–58.

———, Francis Longstaff, and Pedro Santa-Clara, 2008, Two trees, *Review of Financial Studies* 21, 347.

Daniel, Kent, and Tobias Moskowitz, 2013, Momentum crashes, working paper.

Fama, Eugene, and Kenneth French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Hansen, Lars, John Heaton, and Nan Li, 2008, Consumption strikes back? Measuring long-run risk, *Journal of Political Economy* 116, 260 – 302.

Hansen, Lars Peter, and Kenneth J. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, *Econometrica* 50, 1269–1286.

Bibliography III

Jagannathan, Ravi, and Zhenyu Wang, 2002, Empirical evaluation of asset-pricing models: A comparison of the sdf and beta methods, *Journal of Finance* 57, 2337–2367.

Koijen, Ralph, Hanno Lustig, and Stijn Van Nieuwerburgh, 2014, The cross-section and time series of stock and bond returns, working paper.

Lettau, Martin, and Jessica Wachter, 2011, The term structures of equity and interest rates, *Journal of Financial Economics* 101, 90–113.

Lewellen, Jonathan, and Stefan Nagel, 2006, The conditional capm does not explain asset pricing anomalies, *Journal of Financial Economics* 79, 289–314.

———, and Jay Shanken, 2010, A skeptical appraisal of asset-pricing tests, *Journal of Financial Economics* 96, 175–194.

Bibliography IV

Martin, Ian, 2013, The lucas orchard, *Econometrica* 81, 55–111.

Nagel, Stefan, 2013, Empirical cross-sectional asset pricing, *Annual Review of Financial Economics* p. forthcoming.

———, and Kenneth Singleton, 2011, Estimation and evaluation of conditional asset pricing models, *Journal of Finance* 66, 873–909.

Santos, Tano, and Pietro Veronesi, 2010, Habit formation, the cross section of stock returns and the cash-flow risk puzzle, *Journal of Financial Economics* 98, 385–413.

van Binsbergen, Jules, Michael Brandt, and Ralph Koijen, 2012, On the timing and pricing of dividends, *American Economic Review* 102, 1596–1618.

Yogo, Motohiro, 2006, A consumption-based explanation of expected stock returns, *Journal of Finance* 61, 539580.

Bibliography V

Zviadadze, Irina, 2015, Term-structure of consumption risk premia in the cross-section of currency returns, Journal of Finance, forthcoming.

Beveridge-Nelson decomposition

- Moving average:

$$\Delta x_t = C(B)\varepsilon_t = C(1)\varepsilon_t + [C(B) - C(1)]\varepsilon_t$$

- Levels: Beveridge-Nelson decomposition

$$\begin{aligned} x_t &= x_0 + \sum_{s=1}^{t} \Delta x_s = x_0 + C(1)\xi_t + C^*(B)\varepsilon_t && \text{(BN dec)} \\ \xi_t &= \sum_{s=1}^{t} \varepsilon_s && \text{(Permanent component)} \\ C^*(B) &= (1-B)^{-1}[C(B) - C(1)] \end{aligned}$$

- Per period variance:

$$t^{-1} var_0(x_t) = C(1)C(1)^\top + t^{-1} C^*(B)C^*(B)^\top \to C(1)C(1)^\top$$

- Cointegration: $\alpha^\top x_t$ is stationary $\Leftrightarrow \alpha^\top C(1) = 0$
- Thus, $t^{-1} var_0(\alpha^\top x_t) \to 0$

Back 1 Back 2