Why Value Moves

Discount rates, the stochastic discount factor, and ESG

Jae Yung Kim

2026

Summer ESG Module | Valuation | Theory

One number, swinging wildly

The price/dividend ratio of the U.S. market (or the CAPE) is not stable. It has ranged from single digits to the high tens, over and over, for a century.

The question that organises everything

When a market's price is high relative to its dividends, *why*?

Hold the thought. The answer reorganises how you think about valuation — and where ESG enters it.

PREDICT · Two stories about a high price

A high price-to-dividend ratio. Which story is right?

Story A — cash flows
Prices are high because investors expect high future growth. Value moves with the *numerator*.

Story B — discount rates
Prices are high because investors will accept low future returns. Value moves with the *denominator*.

Pick one before the next slides. *Almost everyone starts with Story A.*

The fundamental equation:

$$P = E[m\,x]$$

Value is expected, discounted cash flow

Everyone agrees on this much:

$$P_t \;=\; E_t\left[\sum_{j=1}^{\infty} M_{t,t+j}\, D_{t+j}\right].$$

- D_{t+j} — the cash flows (dividends).
- $M_{t,t+j}$ — the discount factor: how a dollar in period $t+j$, in each possible state of the world, is worth today.

The whole subject is: what is that discount factor, and does it move?

The central equation of asset pricing

Strip it to one period. For *any* asset with payoff x_{t+1} and price p_t:

$$\boxed{p_t \;=\; E_t[\, m_{t+1}\, x_{t+1} \,]}$$

m_{t+1} is the stochastic discount factor (SDF), a.k.a. the pricing kernel.

What m_{t+1} *means*, in words

It is how much you value an extra dollar next period — state by state.

- High in bad states (recession, you are poor and hungry — a dollar is precious).
- Low in good states (boom, you are rich — a dollar is just another dollar).

Where risk premia come from: four lines

Work in gross returns $R^i = x/p$, so the price equation $p_t = E[m_{t+1}x_{t+1}]$ is just $E[m_{t+1}R^i] = 1$ for *every* asset.

$$\textbf{risk-free:}\quad E[m_{t+1}R^f] = R^f\,E[m_{t+1}] = 1 \;\Rightarrow\; E[m_{t+1}] = \tfrac{1}{R^f}.$$

$$\textbf{any asset:}\quad 1 = E[m_{t+1}R^i] = E[m_{t+1}]\,E[R^i] + \operatorname{cov}(m_{t+1}, R^i).$$

$$\textbf{solve:}\quad E[R^i] = \tfrac{1}{E[m_{t+1}]} - \tfrac{\operatorname{cov}(m_{t+1},R^i)}{E[m_{t+1}]} = R^f - R^f\operatorname{cov}(m_{t+1}, R^i).$$

$$E[R^i] - R^f = -R^f\operatorname{cov}(m_{t+1}, R^i)$$

Gross or net returns give the same excess-return statement.

Where risk premia come from

From the four-line derivation, the expected *excess* return on any asset is:

$$E_t[R^i] - R^f \;=\; -\,R^f \operatorname{cov}_t\!\big(m_{t+1},\, R^i\big).$$

Pays off in bad states
(high when m_{t+1} is high)

$\Rightarrow$ it is insurance $\Rightarrow$ you accept a low expected return.

Pays off in good states
(high when m_{t+1} is low)

$\Rightarrow$ it is risky $\Rightarrow$ you demand a risk premium.

The discount rate on an asset $= R^f +$ its risk premium $= R^f - R^f \operatorname{cov}(m_{t+1}, R^i)$.

Make it concrete: a three-state world

Numbers make the SDF tangible. Let next period have three states, with m_{t+1} high in bad states:

state	prob. π	SDF m_{t+1}	
Bad (recession)	0.20	1.50	a dollar is precious
Normal	0.60	0.90	
Good (boom)	0.20	0.60	a dollar is cheap

The risk-free rate drops straight out, since $R^f = 1/E[m_{t+1}]$ and $E[m_{t+1}] = 0.2(1.5) + 0.6(0.9) + 0.2(0.6) = 0.96$:

$$R^f \;=\; \tfrac{1}{0.96} - 1 \;=\; 4.2\%.$$

PREDICT · Two assets, *identical* expected payoffs

	Bad	Normal	Good	$E[x]$
Asset A — pays in booms	50	100	150	**100**
Asset B — pays in busts	150	100	50	**100**

Both pay $100 on average. A is procyclical (more when times are good); B is countercyclical — it pays you precisely when you are poor. Which is worth more today?

Commit before the next slide.

EVIDENCE · Price $= E[m_{t+1}\, x]$ — and they are *not* equal

Weight each payoff by $\pi\, m_{t+1}$ and sum across states:

	Bad (πm_{t+1}=0.30)	Normal (0.54)	Good (0.12)	**Price**
Asset A	$0.30{\times}50 = 15$	$0.54{\times}100 = 54$	$0.12{\times}150 = 18$	**87**
Asset B	$0.30{\times}150 = 45$	$0.54{\times}100 = 54$	$0.12{\times}50 = 6$	**105**

Same cash flows, different prices

$E[x]$ is identical (\$100), yet $P_A = 87$ and $P_B = 105$. The recession-hedge B is worth more — purely because of *when* it pays.

VERDICT · Read the returns: it is *all* discount rate

Expected gross return $= E[x]/P$, so the expected net return is $E[x]/P - 1$:

	price	$E[R] = 100/P - 1$	risk premium $= E[R] - R^f$
Asset A (risky)	87	$+14.9\%$	$+10.8\%$
Asset B (insurance)	105	-4.8%	-8.9%

A pays when m_{t+1} is *low* (good states) $\Rightarrow \mathrm{cov}(m_{t+1}, R_A) < 0 \Rightarrow$ a risk premium. B pays when m_{t+1} is *high* $\Rightarrow$ insurance $\Rightarrow$ a negative premium ($-R^f \mathrm{cov}$ from two slides ago). Identical \$100 cash flows, opposite returns — the discount rate did all the work.

This is Cochrane's thesis, in arithmetic. Keep this table; ESG will reuse it.

The upgrade: the SDF *moves over time*

The one idea the rest of the lecture rests on

Risk aversion — the market's willingness to bear risk — is not constant. In bad times people are scared and demand more to hold risk; in good times they are calm.

m_{t+1} time-varying $\implies$ discount rates move $\implies$ expected returns are predictable.

This is Cochrane's *Discount Rates*. Now we test it. *(Cochrane, "Discount Rates," Journal of Finance 2011.)*

Campbell–Shiller: a high price must forecast *something*

Start from the definition of a return

A return is just next period's price-plus-dividend over this period's price:

$$R_{t+1} \;=\; \frac{P_{t+1} + D_{t+1}}{P_t}.$$

Take logs, linearise around the mean (Campbell–Shiller 1988), and iterate forward.

With $dp_t \equiv \log(D_t/P_t)$ and a constant $\rho \approx 0.96$:

$$dp_t \;\approx\; \underbrace{\sum_{j=1}^{k} \rho^{j-1} r_{t+j}}_{\text{future returns}} - \underbrace{\sum_{j=1}^{k} \rho^{j-1} \Delta d_{t+j}}_{\text{future div. growth}} + \rho^k dp_{t+k}.$$

PREDICT · It is an identity — so something must give

This is not a theory; it is algebra. A high price today (low dp_t) is only possible if:

- future returns r_{t+j} are low *(Story B — discount rates)*, *or*
- future dividend growth Δd_{t+j} is high *(Story A — cash flows)*, *or*
- the price/dividend ratio keeps rising forever *(a "rational bubble")*.

Which one does the data pick? Predict, then look.

The data: it is discount rates, not growth

EVIDENCE · Dividend yields forecast returns

Regress future returns on today's dividend yield: $r_{t\to t+k} = a + b\,(D_t/P_t) + \varepsilon$.

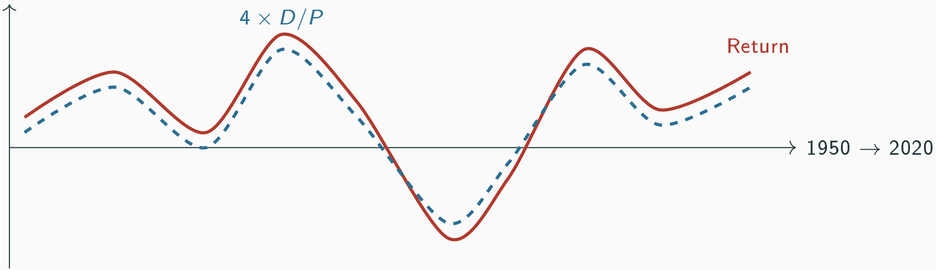
Horizon	b	$t(b)$	R^2	strength
1 year	3.4	2.5	0.07	modest
5 year	18.0	3.5	0.22	strong

When the dividend yield is high (prices low), subsequent returns are high — and the effect is much stronger over long horizons.

Read it the other way

A *high price* today reliably precedes *years of low returns*.

Source: Cochrane, "Discount Rates," JF 2011, Table I (sample to 2010).



The dividend yield (read it as "prices, upside down") and the following 7-year return move together: high prices in 2000 preceded poor returns; low prices in 1980 preceded high returns.

Source: Original redraw of Cochrane (JF 2011), Fig. 1. Shape only.

VERDICT · What moves the price/dividend ratio?

Decompose the historical variation of P/D into the two channels of the identity:

Channel	share of P/D variation	coefficient
Future returns (discount rates)	$\approx 100\%$	1.01
Future dividend growth (cash flows)	$\approx 0\%$	-0.11

Story B wins, decisively

Essentially all of the movement in market valuations is discount rates, and almost none is changing growth expectations.

Source: Cochrane (JF 2011); long-run regression / VAR. Implied-by-VAR ($k \to \infty$): returns 1.35, growth 0.35.

Why discount rates move

The same pattern is *everywhere*

"High valuation $\Rightarrow$ low subsequent return" is not special to stocks. The same predictive pattern shows up across asset classes:

- Stocks — price/dividend
- Government bonds — the yield curve
- Corporate bonds — credit spreads
- Currencies — interest-rate carry
- Sovereign debt
- Real estate — price/rent

A single force — time-varying discount rates — runs through all of them.

Example: housing (Cochrane Fig. 2)

Old view: prices are high because *rents will grow.*

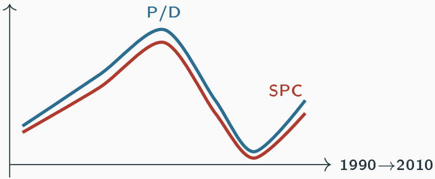
Data: a high price/rent ratio predicts lower future housing returns — exactly the stock-market pattern.

Source: Original redraw of Cochrane (JF 2011), Fig. 2.

The economics: risk premia are *countercyclical*

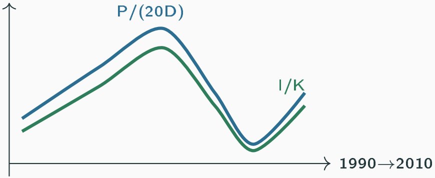
Why do discount rates move? Because the willingness to bear risk moves with the economy.

Habit (consumption) and prices



Good times → calm → high P/D → low future returns.

Investment and prices



Heavy investment → high prices → low expected returns.

In bad times: fear → high discount rates → low prices → high subsequent returns.

Source: Original redraws of Cochrane (JF 2011), Figs. 11–12 (habit; investment / I/K, ME/BE).

Now watch the SDF *move*: the same asset, two dates

Take one risky asset — payoffs $(50, 100, 150)$, the *same* $E[x] = \$100$ at both dates. Only risk aversion differs: calm vs. fearful (a more *volatile* m_{t+1}).

state	prob. π	m_{t+1} (calm)	m_{t+1} (fear)
Bad	0.20	1.20	1.90
Normal	0.60	0.95	0.82
Good	0.20	0.80	0.50

Same machine $p = E[m_{t+1}x]$, two states of the world (πm_{t+1} weights, then sum):

$$P_{\text{calm}} = 0.24(50) + 0.57(100) + 0.16(150) = 93$$

$$P_{\text{fear}} = 0.38(50) + 0.49(100) + 0.10(150) = 83$$

VERDICT · The valuation moved — on the *discount rate*

Identical cash flows, yet the price fell $93 \to 83$. Read the returns ($E[R] = 100/P - 1$):

	price	$E[R]$	risk premium
Calm (low risk aversion)	93	7.5%	+4.4%
Fear (high risk aversion)	83	20.2%	+17.3%

The whole lecture, in one asset

Nothing about the cash flows changed — and the risk-free rate barely moved ($3.1\% \to 2.9\%$), so this is a pure risk-premium move, not a rate move. Fear made m_{t+1} more volatile $\Rightarrow$ a higher risk premium $\Rightarrow$ a higher discount rate $\Rightarrow$ a lower price, and a higher expected return going forward. A low price *predicts* high returns.

Cross-section (the A/B example) and time-series (this one): one machine, m_{t+1}, doing both.

ESG: the same discount rate, in the cross-section

So where does ESG enter value?

$$P_t = E_t\left[\sum_j M_{t,t+j} D_{t+j}\right] \implies \text{ESG can move either piece.}$$

- **Cash flows** (D): ESG risks/opportunities change future profits. *(The obvious channel.)*
- **The discount factor** (M): ESG changes the required return — the cost of capital. *(The deeper, and today's, channel.)*

We focus on the discount-rate channel — it is where the asset-pricing machine bites.
There are three ways ESG enters M.

Channel 1 — tastes (Pástor–Stambaugh–Taylor, JFE 2021)

Suppose investors get utility from holding green assets, beyond the money. In equilibrium that preference shows up as a wedge in expected returns. Reduced to one line:

$$\underbrace{E[R_i] - R^f}_{\text{expected excess return}} = \beta_i \lambda - \delta g_i, \qquad \delta > 0, \quad g_i = \text{greenness.}$$

Consequence

The greener the firm (high g_i), the lower its expected return: a lower cost of capital and a higher valuation — *for the same cash flows.* δ is the green expected-return wedge (the "greenium"). Brown firms get the mirror image.

Investors literally pay — in foregone return — for the non-pecuniary benefit of green.

Source: Pástor, Stambaugh & Taylor, "Sustainable investing in equilibrium," JFE 2021.

Channel 1, *in numbers*: add a green asset

Return to our three-state example. Asset A (pays in booms) had price 87 and expected return 14.9%. Now build asset **G** with the *same payoffs*, but investors have a green taste worth $\delta = 3\%$ of return:

$$E[R_G] = E[R_A] - \delta = 14.9\% - 3\% = 11.9\% \;\Rightarrow\; P_G = \frac{E[x]}{1 + E[R_G]} \approx 89.3.$$

This *is* PST

Same \$100 cash flows as A, but G is worth 89.3 vs 87: a higher price, a lower required return, a lower cost of capital. The kernel m_{t+1} is *unchanged* — it still prices A, B and G — yet green's discount rate falls below what risk alone dictates, because investors accept less for the green benefit. *One m_{t+1}; the taste moves the asset's required return, not the kernel.*

VERDICT · The greenium puzzle *is* Cochrane

A real tension: green stocks outperformed through the 2010s — yet PST says green should earn *less*. How can both be true?

Resolution (Pástor–Stambaugh–Taylor, JFE 2022)

The outperformance came from an unexpected rise in climate concern, which lifted green valuations — i.e. lowered the green discount rate. Going forward, expected green returns are now *lower*.

$$\text{taste} \uparrow \Rightarrow \text{green discount rate} \downarrow \Rightarrow \underbrace{\text{price} \uparrow \text{ now}}_{\text{realised gain}} + \underbrace{\text{expected return} \downarrow \text{ later}}_{\text{Cochrane}}.$$

This is exactly Parts 1–3 — high price ⇒ low future return — now driven by tastes, in the green factor.

Source: Pástor, Stambaugh & Taylor, "Dissecting green returns," JFE 2022.

Channel 2 — risk (Bolton–Kacperczyk, JFE 2021)

A different engine, same direction for brown. Suppose carbon is a *risk* investors must be paid to bear — transition risk, stranded assets, policy. Test it as a cross-sectional regression:

$$R_{i,t} - R^f_t \;=\; a \;+\; \lambda\,\mathrm{Emissions}_{i,t-1} \;+\; \gamma'\mathrm{Controls}_{i,t-1} \;+\; \varepsilon_{i,t}.$$

Finding: $\lambda > 0$, a carbon premium

High-carbon firms earn higher average returns — compensation for carbon risk. It is priced on both the level of total emissions (not intensity) and their growth, and survives the usual factor controls. Brown $\Rightarrow$ higher risk $\Rightarrow$ higher discount rate.

Source: Bolton & Kacperczyk, "Do investors care about carbon risk?," JFE 2021.

Two engines, one direction

	Channel 1 — PST	**Channel 2 — BK**
mechanism	preference (tastes)	risk (compensation)
who moves	green: required return ↓	brown: risk premium ↑
why	investors pay for green	investors paid to bear carbon
cross-section	brown earns *more* (expected)	brown earns *more* (expected)

Opposite *stories* — preferences vs. risk — but the same cross-sectional prediction: brown has a higher expected return, green a lower one. And both are discount-rate channels: neither touches the cash flows. That is the link to Parts 1–4.

Channel 3 is a *different question* (Pedersen et al., JFE 2021)

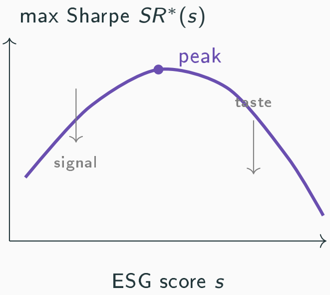
Channels 1–2 ask "why do valuations differ?" (asset pricing). Pedersen, Fitzgibbons & Pomorski ask "how should I *invest*?" (portfolio choice). Three investor types:

- **ESG-unaware** — plain mean–variance, ignores ESG.
- **ESG-aware** — uses ESG as a signal of risk / future returns.
- **ESG-motivated** — has a taste, will sacrifice return for ESG.

Their tool is the ESG-efficient frontier: the best Sharpe ratio attainable at each ESG level.

Source: Pedersen, Fitzgibbons & Pomorski, "Responsible investing: the ESG-efficient frontier," JFE 2021.

The frontier, decomposed — and your question answered



Rising part — ESG as a signal: it carries information about the cross-section of expected returns, so using it *raises* Sharpe (this is Channels 1–2, from the investor's seat).

Falling part — ESG as a taste: past the peak you *pay* Sharpe for ESG (this is PST).

Is the frontier a discount-rate story or a portfolio story?

A portfolio story — but its *shape* is set by the discount-rate channels. The hump is ESG's *information value* fading against its *taste cost*. It *packages* Channels 1–2; it does not replace them.

VERDICT · ESG *is* time-varying discount rates — in the cross-section

	green / high-ESG	**brown / low-ESG**
taste channel (PST)	required return ↓	—
risk channel (BK)	—	risk premium ↑
discount rate	lower	higher
valuation (same CFs)	higher	lower

The whole lecture in one line

Parts 1–4: the discount rate moves over time (Cochrane). Part 5: it also varies across firms by greenness — and that green premium is itself time-varying. Same machine, new axis.

Subtlety to keep: "lower discount rate" = lower expected return, not necessarily lower realised return while the market re-prices.

Bridge to practice

Where ESG meets the DCF

A discounted-cash-flow valuation has exactly one place for all of this to land:

$$V_0 = \sum_t \frac{\mathrm{FCF}_t}{(1+\mathrm{WACC})^t}, \qquad \mathrm{WACC} = \frac{E}{V} R_e + \frac{D}{V} R_d (1 - T_c).$$

The link

The discount rate (the WACC, via the cost of equity R_e) is where time-varying risk premia *and* ESG enter the valuation — disciplined, not hand-waved.

In the valuation practice: we build that DCF for a real, just-listed company — SpaceX.

Five things to carry away

1. Market valuations move on discount rates, not changing growth expectations ($\approx$*100% vs* $\approx$*0%*).
2. The discount rate is the stochastic discount factor m_{t+1} — and it is time-varying, so expected returns are predictable.
3. Discount rates move with risk appetite and the economy — countercyclically, across every asset class.
4. ESG is priced into the discount rate: tastes lower green's cost of capital; carbon risk raises brown's.
5. So ESG valuation = adjusting the discount rate (and the cash flows) *with discipline* — which is exactly the DCF we turn to next.

Value moves because discount rates move.

ESG is one of the forces that moves them.