Discussion of:

The Implied Equity Term Structure

by Lieven Baele, Joost Driessen and Tomas Jankauskas

Christian Julliard

London School of Economics

LTI Conference 2022

The bigger picture

Absent arbitrage opportunities, it exists an SDF (M) s.t. the price (P_t) of an asset that delivers the cashflows $\{D_{t+i}\}_{i=i}^{\infty}$ can be decomposed as

$$P_t = \underbrace{\sum_{i=1}^{T} \mathbb{E}_t \left[M_{t,t+i} D_{t+i} \right]}_{\text{price of ST asset}=:P_t^{(1:T)}} + \underbrace{\sum_{i=T+1}^{\infty} \mathbb{E}_t \left[M_{t,t+i} D_{t+i} \right]}_{\text{price of LT asset}=:P_t^{(T+1:\infty)}} = \sum_{i=1}^{\infty} \underbrace{P_t^{(i)}}_{\text{price of "bullet" CF}\equiv P_t^{(i:i)}}$$

The dividend strip prices $P_t^{(i)}$ are very salient since economic theory has stringent predictions on the shape of the term structure of

$$\mathbb{E}\left[R_{t,t+i}^{(i)}\right] := \mathbb{E}\left[\frac{D_{t+i}}{P_t^{(i)}}\right] \text{ and } \mathbb{E}\left[R_{t,t+1}^{(i)}\right] := \mathbb{E}\left[\frac{P_{t+1}^{(i-1)}}{P_t^{(i)}}\right]$$

E.g. monotonically increasing risk premia in Habit models, increasing but with horizontal asymptote in vanilla LRR, flat in disaster models.

The bigger picture

Absent arbitrage opportunities, it exists an SDF (M) s.t. the price (P_t) of an asset that delivers the cashflows $\{D_{t+i}\}_{i=i}^{\infty}$ can be decomposed as

$$P_t = \underbrace{\sum_{i=1}^{T} \mathbb{E}_t \left[M_{t,t+i} D_{t+i} \right]}_{\text{price of ST asset}=:P_t^{(1:T)}} + \underbrace{\sum_{i=T+1}^{\infty} \mathbb{E}_t \left[M_{t,t+i} D_{t+i} \right]}_{\text{price of LT asset}=:P_t^{(T+1:\infty)}} = \sum_{i=1}^{\infty} \underbrace{P_t^{(i)}}_{\text{price of "bullet" CF}\equiv P_t^{(i:i)}}$$

The dividend strip prices $P_t^{(i)}$ are very salient since economic theory has stringent predictions on the shape of the term structure of

$$\mathbb{E}\left[R_{t,t+i}^{(i)}\right] := \mathbb{E}\left[\frac{D_{t+i}}{P_t^{(i)}}\right] \quad \text{and} \quad \mathbb{E}\left[R_{t,t+1}^{(i)}\right] := \mathbb{E}\left[\frac{P_{t+1}^{(i-1)}}{P_t^{(i)}}\right]$$

E.g. monotonically increasing risk premia in Habit models, increasing but with horizontal asymptote in vanilla LRR, flat in disaster models.

The bigger picture cont'd

Problem: $P_t^{(i)}$s are NOT generally available.

Solutions:

1. Use implied $P_t^{(i)}$ from derivative markets (index options, index futures, dividend futures): e.g. van Binsbergen-Kojen 2017, van Binsbergen-Brandt-Koijen 2012, Bansal-Yaron-Miller-Song 2021
2. Affine modelling of SDF, price, and dividend processes: Kelly-Giglio-Kozak 2021
3. **This paper:** estimates directly $\mathbb{E}\left[R_{t,t+i}^{(i)}\right]$ using the fact that

$$P_t = \sum_{i=1}^{\infty} \mathbb{E}_t\left[M_{t,t+i} D_{t+i}\right] \equiv \sum_{i=1}^{\infty} \frac{\mathbb{E}_t\left[D_{t+i}\right]}{\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]}$$

$\Rightarrow$ clever and fresh idea - I like it!

But needs:

i. forecasting model/method for $\mathbb{E}_t\left[D_{t+i}\right]$ (analysts' forecasts + AR(1))
ii. Enough restrictions for identification of $\left\{\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]\right\}_{i=1}^{\infty}$ (truncation at T+ parametric form)

The bigger picture cont'd

Problem: $P_t^{(i)}$s are NOT generally available.

Solutions:

1. Use implied $P_t^{(i)}$ from derivative markets (index options, index futures, dividend futures): e.g. van Binsbergen-Kojen 2017, van Binsbergen-Brandt-Koijen 2012, Bansal-Yaron-Miller-Song 2021
2. Affine modelling of SDF, price, and dividend processes: Kelly-Giglio-Kozak 2021
3. **This paper:** estimates directly $\mathbb{E}\left[R_{t,t+i}^{(i)}\right]$ using the fact that

$$P_t = \sum_{i=1}^{\infty} \mathbb{E}_t\left[M_{t,t+i} D_{t+i}\right] \equiv \sum_{i=1}^{\infty} \frac{\mathbb{E}_t\left[D_{t+i}\right]}{\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]}$$

$\Rightarrow$ clever and fresh idea - I like it!

But needs:

i. forecasting model/method for $\mathbb{E}_t\left[D_{t+i}\right]$ (analysts' forecasts + AR(1))
ii. Enough restrictions for identification of $\left\{\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]\right\}_{i=1}^{\infty}$ (truncation at T+ parametric form)

The bigger picture cont'd

Problem: $P_t^{(i)}$s are NOT generally available.

Solutions:

1. Use implied $P_t^{(i)}$ from derivative markets (index options, index futures, dividend futures): e.g. van Binsbergen-Kojen 2017, van Binsbergen-Brandt-Koijen 2012, Bansal-Yaron-Miller-Song 2021
2. Affine modelling of SDF, price, and dividend processes: Kelly-Giglio-Kozak 2021
3. **This paper:** estimates directly $\mathbb{E}\left[R_{t,t+i}^{(i)}\right]$ using the fact that

$$P_t = \sum_{i=1}^{\infty} \mathbb{E}_t\left[M_{t,t+i} D_{t+i}\right] = \sum_{i=1}^{\infty} \frac{\mathbb{E}_t\left[D_{t+i}\right]}{\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]}$$

$\Rightarrow$ clever and fresh idea - I like it!

But needs:

i. forecasting model/method for $\mathbb{E}_t\left[D_{t+i}\right]$ (analysts' forecasts + AR(1))

ii. Enough restrictions for identification of $\left\{\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]\right\}_{i=1}^{\infty}$ (truncation at T+ parametric form)

The bigger picture cont'd

Problem: $P_t^{(i)}$s are NOT generally available.

Solutions:

1. Use implied $P_t^{(i)}$ from derivative markets (index options, index futures, dividend futures): e.g. van Binsbergen-Kojen 2017, van Binsbergen-Brandt-Koijen 2012, Bansal-Yaron-Miller-Song 2021
2. Affine modelling of SDF, price, and dividend processes: Kelly-Giglio-Kozak 2021
3. **This paper:** estimates directly $\mathbb{E}\left[R_{t,t+i}^{(i)}\right]$ using the fact that

$$P_t = \sum_{i=1}^{\infty} \mathbb{E}_t\left[M_{t,t+i} D_{t+i}\right] \equiv \sum_{i=1}^{\infty} \frac{\mathbb{E}_t\left[D_{t+i}\right]}{\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]}$$

$\Rightarrow$ clever and fresh idea - I like it!

But needs:

i. forecasting model/method for $\mathbb{E}_t\left[D_{t+i}\right]$ (analysts' forecasts + AR(1))
ii. Enough restrictions for identification of $\left\{\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]\right\}_{i=1}^{\infty}$ (truncation at T+ parametric form)

The bigger picture cont'd

Problem: $P_t^{(i)}$s are NOT generally available.

Solutions:

1. Use implied $P_t^{(i)}$ from derivative markets (index options, index futures, dividend futures): e.g. van Binsbergen-Kojen 2017, van Binsbergen-Brandt-Koijen 2012, Bansal-Yaron-Miller-Song 2021
2. Affine modelling of SDF, price, and dividend processes: Kelly-Giglio-Kozak 2021
3. **This paper:** estimates directly $\mathbb{E}\left[R_{t,t+i}^{(i)}\right]$ using the fact that

$$P_t = \sum_{i=1}^{\infty} \mathbb{E}_t\left[M_{t,t+i}D_{t+i}\right] \equiv \sum_{i=1}^{\infty} \frac{\mathbb{E}_t\left[D_{t+i}\right]}{\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]}$$

$\Rightarrow$ clever and fresh idea - I like it!

But needs:

i. forecasting model/method for $\mathbb{E}_t\left[D_{t+i}\right]$ (analysts' forecasts + AR(1))
ii. Enough restrictions for identification of $\left\{\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]\right\}_{i=1}^{\infty}$ (truncation at T+ parametric form)

Comments

I. Restrictions on $\mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]$: time & cross-sectional variation + monotonicity

- Define (as in the paper): $r_{t,i} \equiv \mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]^{\frac{1}{i}} - 1$ (the annualized discount rate)
- The authors assume the parametric form ("kind of" Nelson-Siegel):

$$r_{t,i} - r^f_{t,i} = \beta_1 - \beta_2\left[\frac{1-\exp(-i/\lambda)}{i/\lambda}\right] \quad \forall i \le T, \quad r_{t,i} = r_{t,T} \quad \forall i > T; \quad \forall t$$

and estimate β's but calibrate λ (☹)

1. Conditional risk premia are constant (cross-sectionally and in time)

... but they don't need to be for identification!

- If parameters are fixed in time, there are enough d.o.f. to allow them to vary across assets (3 periods are enough).
- If parameters are fixed cross-sectionally, there are enough d.o.f. to allow them to vary across time (3 assets are enough)

Note: the authors show that they vary both in time (BC) and across characteristics

⇒ I'd model variation in β's and λ formally: e.g. i) linear functions of characteristics (to capture cross-sectional variation) and aggregated state variables (to capture time variation) or ii) (Bayesian) latent TVPs

- more than 500 assets and 40 years of data make it highly feasible!

I. Restrictions on $\mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]$: time & cross-sectional variation + monotonicity

- Define (as in the paper): $r_{t,i} \equiv \mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]^{\frac{1}{i}} - 1$ (the annualized discount rate)
- The authors assume the parametric form ("kind of" Nelson-Siegel):

$$r_{t,i} - r^f_{t,i} = \beta_1 - \beta_2\left[\frac{1-\exp(-i/\lambda)}{i/\lambda}\right] \quad \forall i \le T, \quad r_{t,i} = r_{t,T} \quad \forall i > T; \quad \forall t$$

and estimate β's but calibrate λ (☹)

1. Conditional risk premia are constant (cross-sectionally and in time)

... but they don't need to be for identification!

- If parameters are fixed in time, there are enough d.o.f. to allow them to vary across assets (3 periods are enough).
- If parameters are fixed cross-sectionally, there are enough d.o.f. to allow them to vary across time (3 assets are enough)

Note: the authors show that they vary both in time (BC) and across characteristics

⇒ I'd model variation in β's and λ formally: e.g. i) linear functions of characteristics (to capture cross-sectional variation) and aggregated state variables (to capture time variation) or ii) (Bayesian) latent TVPs

- more than 500 assets and 40 years of data make it highly feasible!

I. Restrictions on $\mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]$: time & cross-sectional variation + monotonicity

- Define (as in the paper): $r_{t,i} \equiv \mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]^{\frac{1}{i}} - 1$ (the annualized discount rate)
- The authors assume the parametric form ("kind of" Nelson-Siegel):

$$r_{t,i} - r^f_{t,i} = \beta_1 - \beta_2\left[\frac{1-\exp(-i/\lambda)}{i/\lambda}\right] \quad \forall i \le T, \quad r_{t,i} = r_{t,T} \quad \forall i > T; \quad \forall t$$

and estimate β's but calibrate λ (☹)

1. Conditional risk premia are constant (cross-sectionally and in time)

... but they don't need to be for identification!

- If parameters are fixed in time, there are enough d.o.f. to allow them to vary across assets (3 periods are enough).
- If parameters are fixed cross-sectionally, there are enough d.o.f. to allow them to vary across time (3 assets are enough)

Note: the authors show that they vary both in time (BC) and across characteristics

⇒ I'd model variation in β's and λ formally: e.g. i) linear functions of characteristics (to capture cross-sectional variation) and aggregated state variables (to capture time variation) or ii) (Bayesian) latent TVPs

- more than 500 assets and 40 years of data make it highly feasible!

I. Restrictions on $\mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]$: time & cross-sectional variation + monotonicity

- Define (as in the paper): $r_{t,i} \equiv \mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]^{\frac{1}{i}} - 1$ (the annualized discount rate)
- The authors assume the parametric form ("kind of" Nelson-Siegel):

$$r_{t,i} - r^f_{t,i} = \beta_1 - \beta_2\left[\frac{1-\exp(-i/\lambda)}{i/\lambda}\right] \quad \forall i \leq T, \quad r_{t,i} = r_{t,T} \quad \forall i > T; \quad \forall t$$

and estimate β's but calibrate λ (☹)

1. Conditional risk premia are constant (cross-sectionally and in time)

... but they don't need to be for identification!

- If parameters are fixed in time, there are enough d.o.f. to allow them to vary across assets (3 periods are enough).
- If parameters are fixed cross-sectionally, there are enough d.o.f. to allow them to vary across time (3 assets are enough)

Note: the authors show that they vary both in time (BC) and across characteristics

⇒ I'd model variation in β's and λ formally: e.g. i) linear functions of characteristics (to capture cross-sectional variation) and aggregated state variables (to capture time variation) or ii) (Bayesian) latent TVPs

- more than 500 assets and 40 years of data make it highly feasible!

I. Restrictions on $\mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]$: time & cross-sectional variation + monotonicity

- Define (as in the paper): $r_{t,i} \equiv \mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]^{\frac{1}{i}} - 1$ (the annualized discount rate)
- The authors assume the parametric form ("kind of" Nelson-Siegel):

$$r_{t,i} - r^f_{t,i} = \beta_1 - \beta_2 \left[\frac{1 - \exp(-i/\lambda)}{i/\lambda}\right] \quad \forall i \leq T, \quad r_{t,i} = r_{t,T} \quad \forall i > T; \ \ \forall t$$

and estimate β's but calibrate λ (☹)

1. Conditional risk premia are constant (cross-sectionally and in time)

... but they don't need to be for identification!

- If parameters are fixed in time, there are enough d.o.f. to allow them to vary across assets (3 periods are enough).
- If parameters are fixed cross-sectionally, there are enough d.o.f. to allow them to vary across time (3 assets are enough)

Note: the authors show that they vary both in time (BC) and across characteristics

⇒ I'd model variation in β's and λ formally: e.g. i) linear functions of characteristics (to capture cross-sectional variation) and aggregated state variables (to capture time variation) or ii) (Bayesian) latent TVPs

- more than 500 assets and 40 years of data make it highly feasible!

I. Restrictions on $\mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]$ cont'd

- Define (as in the paper): $r_{t,i} \equiv \mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]^{\frac{1}{i}} - 1$ (the annualized discount rate)
- The authors assume the parametric form ("kind of" Nelson-Siegel):

$$r_{t,i} - r^f_{t,i} = \beta_1 - \beta_2\left[\frac{1-\exp(-i/\lambda)}{i/\lambda}\right] \quad \forall i \le T, \quad r_{t,i} = r_{t,T} \quad \forall i > T; \quad \forall t$$

and estimate β's but calibrate λ

2. Conditional risk premia are strictly monotonic

- Why not full Nelson-Siegel?
- Only one more parameter, and curvature likely to depend on characteristics too.

⇒ I'd like formal estimation and selection of the β's and λ.

- hard to take the slope findings at face value without formal testing of e.g. $H_0 : \beta_2 \neq 0$

Note: the estimation problem can be written as a NLS (just divide P_t by D_t and add an error)

⇒ can use standard methods (GMM, MLE) for inference. Can also test formally which characteristics drive the heterogeneity.

I. Restrictions on $\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]$ cont'd

- Define (as in the paper): $r_{t,i} \equiv \mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]^{\frac{1}{i}} - 1$ (the annualized discount rate)
- The authors assume the parametric form ("kind of" Nelson-Siegel):

$$r_{t,i} - r_{t,i}^f = \beta_1 - \beta_2\left[\frac{1-\exp(-i/\lambda)}{i/\lambda}\right] \quad \forall i \leq T, \quad r_{t,i} = r_{t,T} \quad \forall i > T; \;\; \forall t$$

and estimate β's but calibrate λ

2. Conditional risk premia are strictly monotonic

- Why not full Nelson-Siegel?
- Only one more parameter, and curvature likely to depend on characteristics too.

⇒ I'd like formal estimation and selection of the β's and λ.

- hard to take the slope findings at face value without formal testing of e.g. $H_0 : \beta_2 \neq 0$

Note: the estimation problem can be written as a NLS (just divide P_t by D_t and add an error)

⇒ can use standard methods (GMM, MLE) for inference. Can also test formally which characteristics drive the heterogeneity.

I. Restrictions on $\mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]$ cont'd

- Define (as in the paper): $r_{t,i} \equiv \mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]^{\frac{1}{i}} - 1$ (the annualized discount rate)
- The authors assume the parametric form ("kind of" Nelson-Siegel):

$$r_{t,i} - r^f_{t,i} = \beta_1 - \beta_2\left[\frac{1-\exp(-i/\lambda)}{i/\lambda}\right] \quad \forall i \leq T, \quad r_{t,i} = r_{t,T} \quad \forall i > T; \quad \forall t$$

  and estimate β's but calibrate λ

2. Conditional risk premia are strictly monotonic

- Why not full Nelson-Siegel?
- Only one more parameter, and curvature likely to depend on characteristics too.

⇒ I'd like formal estimation and selection of the β's and λ.

- hard to take the slope findings at face value without formal testing of e.g. $H_0 : \beta_2 \neq 0$

Note: the estimation problem can be written as a NLS (just divide P_t by D_t and add an error)

⇒ can use standard methods (GMM, MLE) for inference. Can also test formally which characteristics drive the heterogeneity.

I. Restrictions on $\mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]$ cont'd

- Define (as in the paper): $r_{t,i} \equiv \mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]^{\frac{1}{i}} - 1$ (the annualized discount rate)
- The authors assume the parametric form ("kind of" Nelson-Siegel):

$$r_{t,i} - r^f_{t,i} = \beta_1 - \beta_2\left[\frac{1-\exp(-i/\lambda)}{i/\lambda}\right] \quad \forall i \leq T, \quad r_{t,i} = r_{t,T} \quad \forall i > T; \quad \forall t$$

and estimate β's but calibrate λ

2. Conditional risk premia are strictly monotonic

- Why not full Nelson-Siegel?
- Only one more parameter, and curvature likely to depend on characteristics too.

$\Rightarrow$ I'd like formal estimation and selection of the β's and λ.

- hard to take the slope findings at face value without formal testing of e.g. $H_0 : \beta_2 \neq 0$

Note: the estimation problem can be written as a NLS (just divide P_t by D_t and add an error)

$\Rightarrow$ can use standard methods (GMM, MLE) for inference. Can also test formally which characteristics drive the heterogeneity.

II. Restrictions on $\mathbb{E}_t\left[D_{t+i}\right]$

Cash-flows are modelled observing that:

$$D_t = \underbrace{BE_{t-1}}_{\text{book equity}} \times \left(ROE_t - \underbrace{\%BE_t}_{\text{BE growth}} \right)$$

And assuming:

1. $\mathbb{E}_t\left[D_{t+i}\right] =$ analysts' forecast $\forall i < 6$ years
2. AR(1) for both ROE and BE (stationarity? maybe typo: $\%BE$ instead?) $\forall i > 6$ years with coefficients from a pooled regression.

⇒ Same identical expected growth from 6 years onward... but don't we expect e.g. "growth" and "value" stocks to have different growth rates?

Issue: If growth rates are different, the above will generate spurious heterogeneity in discount rates.

⇒ I'd estimate firm/portfolio specific ARMA processes (again, tons of dof)

Note: $\%BE_t$ proxied by sales growth – why? BE is "observable"

II. Restrictions on $\mathbb{E}_t\left[D_{t+i}\right]$

Cash-flows are modelled observing that:

$$D_t = \underbrace{BE_{t-1}}_{\text{book equity}} \times \left(ROE_t - \underbrace{\%BE_t}_{\text{BE growth}} \right)$$

And assuming:

1. $\mathbb{E}_t\left[D_{t+i}\right]$ = analysts' forecast $\forall i < 6$ years
2. AR(1) for both ROE and BE (stationarity? maybe typo: $\%BE$ instead?) $\forall i > 6$ years with coefficients from a pooled regression.

$\Rightarrow$ Same identical expected growth from 6 years onward... but don't we expect e.g. "growth" and "value" stocks to have different growth rates?

Issue: If growth rates are different, the above will generate spurious heterogeneity in discount rates.

$\Rightarrow$ I'd estimate firm/portfolio specific ARMA processes (again, tons of dof)

Note: $\%BE_t$ proxied by sales growth – why? BE is "observable"

II. Restrictions on $\mathbb{E}_t [D_{t+i}]$

Cash-flows are modelled observing that:

$$D_t = \underbrace{BE_{t-1}}_{\text{book equity}} \times \left(ROE_t - \underbrace{\%BE_t}_{\text{BE growth}} \right)$$

And assuming:

1. $\mathbb{E}_t [D_{t+i}] =$ analysts' forecast $\forall i < 6$ years
2. AR(1) for both ROE and BE (stationarity? maybe typo: $\%BE$ instead?) $\forall i > 6$ years with coefficients from a pooled regression.

⇒ Same identical expected growth from 6 years onward... but don't we expect e.g. "growth" and "value" stocks to have different growth rates?

Issue: If growth rates are different, the above will generate spurious heterogeneity in discount rates.

⇒ I'd estimate firm/portfolio specific ARMA processes (again, tons of dof)

Note: $\%BE_t$ proxied by sales growth – why? BE is "observable"

III. Inference and interpretation

1. Many of the term structure estimates have discount rates in the $[-8\%, \ +8\%]$ range, and rates tend to be negative for maturities below $\approx$ 6 years...

$\Rightarrow$ needs proper confidence bands to trust the conclusions on the slope (again, cast the estimation as a big GMM – including the ARs – and use e.g. the Delta method)

Note: the negative discount rates corresponding to analyst forecasts suggest the analysts' medium run forecast are too low relative / observed prices too high – very different story...

$\Rightarrow$ Large literature on biases in analyst forecasts' (cf. Bradshaw-Ertimur-O'Brien 2016)

2. The estimated $\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]^{\frac{1}{i}} \forall i$ is nothing but a (filtered) asset pricing model

$\Rightarrow$ needs to show that it is a good pricing model if we have to believe the estimates. Report the pricing errors and standard fit metrics (MAPE, R^2, J-stat etc.)

Note: might be hard to to do well without heterogeneity – all assets have a $\beta^{(i)} = 1$

III. Inference and interpretation

1. Many of the term structure estimates have discount rates in the $[-8\%, \ +8\%]$ range, and rates tend to be negative for maturities below $\approx$ 6 years...

$\Rightarrow$ needs proper confidence bands to trust the conclusions on the slope (again, cast the estimation as a big GMM – including the ARs – and use e.g. the Delta method)

Note: the negative discount rates corresponding to analyst forecasts suggest the analysts' medium run forecast are too low relative / observed prices too high – very different story...

$\Rightarrow$ Large literature on biases in analyst forecasts' (cf. Bradshaw-Ertimur-O'Brien 2016)

2. The estimated $\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]^{\frac{1}{i}}$ $\forall i$ is nothing but a (filtered) asset pricing model

$\Rightarrow$ needs to show that it is a good pricing model if we have to believe the estimates. Report the pricing errors and standard fit metrics (MAPE, R^2, J-stat etc.)

Note: might be hard to to do well without heterogeneity – all assets have a $\beta^{(i)} = 1$

III. Inference and interpretation

1. Many of the term structure estimates have discount rates in the $[-8\%,\ +8\%]$ range, and rates tend to be negative for maturities below $\approx$ 6 years...

$\Rightarrow$ needs proper confidence bands to trust the conclusions on the slope (again, cast the estimation as a big GMM – including the ARs – and use e.g. the Delta method)

Note: the negative discount rates corresponding to analyst forecasts suggest the analysts' medium run forecast are too low relative / observed prices too high – very different story...

$\Rightarrow$ Large literature on biases in analyst forecasts' (cf. Bradshaw-Ertimur-O'Brien 2016)

2. The estimated $\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]^{\frac{1}{i}}$ $\forall i$ is nothing but a (filtered) asset pricing model

$\Rightarrow$ needs to show that it is a good pricing model if we have to believe the estimates. Report the pricing errors and standard fit metrics (MAPE, R^2, J-stat etc.)

Note: might be hard to to do well without heterogeneity – all assets have a $\beta^{(i)} = 1$

III. Inference and interpretation cont'd

3. This paper estimates $\mathbb{E}_t\left[R^{(i)}_{t,t+i}\right]^{\frac{1}{i}}$ but literature focuses on $\mathbb{E}_t\left[R^{(i)}_{t,t+1}\right]$ or $\mathbb{E}_t\left[\frac{1}{i}\ln R^{(i)}_{t,t+i}\right]$

$\Rightarrow$ it's ok, but the direct comparison should be careful (e.g. Kelly-Giglio-Kozak (2021))

Note: from the literature we know what to expect for the latter quantitites in canonical models but not the former $\Rightarrow$ simulate LRR and Habit models to give us a benchmark.

4. In the absence of arbitrage opportunities:

$$\underbrace{\mathbb{E}\left[\frac{P^{(i)}_t}{P_t}\right]}_{\text{values share of strip}} \equiv \underbrace{Cov\left(M_{t,t+i}; \frac{D_{t+i}}{P_t}\right)}_{\text{risk adjustment}} + \underbrace{\mathbb{E}\left[M_{t,t+i}\right]\mathbb{E}\left[\frac{D_{t+i}}{P_t}\right]}_{\text{risk neutral valuation}}$$

$\Rightarrow$ Use a sound SDF (e.g. Kozak-Nagel-Santosh 2020, Huang-Bryzgalova-Julliard 2022) to verify your findings.

III. Inference and interpretation cont'd

3. This paper estimates $\mathbb{E}_t\left[R_{t,t+i}^{(i)}\right]^{\frac{1}{i}}$ but literature focuses on $\mathbb{E}_t\left[R_{t,t+1}^{(i)}\right]$ or $\mathbb{E}_t\left[\frac{1}{i}\ln R_{t,t+i}^{(i)}\right]$

$\Rightarrow$ it's ok, but the direct comparison should be careful (e.g. Kelly-Giglio-Kozak (2021))

Note: from the literature we know what to expect for the latter quantitites in canonical models but not the former $\Rightarrow$ simulate LRR and Habit models to give us a benchmark.

4. In the absence of arbitrage opportunities:

$$\underbrace{\mathbb{E}\left[\frac{P_t^{(i)}}{P_t}\right]}_{\text{values share of strip}} \equiv \underbrace{Cov\left(M_{t,t+i};\frac{D_{t+i}}{P_t}\right)}_{\text{risk adjustment}} + \underbrace{\mathbb{E}\left[M_{t,t+i}\right]\mathbb{E}\left[\frac{D_{t+i}}{P_t}\right]}_{\text{risk neutral valuation}}$$

$\Rightarrow$ Use a sound SDF (e.g. Kozak-Nagel-Santosh 2020, Huang-Bryzgalova-Julliard 2022) to verify your findings.

Conclusion

Conclusion

(+) A very good and novel idea, and I really enjoyed reading the paper.

(+) Results are (potentially) very salient and interesting.

(-) But I would like the analysis to be much more econometrically formal to fully trust the findings (relatively “easy” to fix).