Decomposing long bond returns

Liuren Wu

2021
Different modeling frameworks are built for different purposes:

- **Expectation hypothesis (EH):** *Predict short rate move* with yield curve slope.
  - Yield curve shape combines information from expectation, risk premium, and convexity, but expectation dominates the short term.

- **Dynamic term structure models (DTSM):** Value the whole yield curve based on assumptions on the *full* risk-neutral dynamics of the *short rate*.
  - Uses one yardstick (the short rate) to measure everything else for cross-sectional consistency.
  - Deviations from DTSM valuation can be used to construct statistical arbitrage trading on the yield curve.

- **HJM-type models:** Price interest rate options based on the current forward curve and views of *forward rate volatility*.
  - Highlight the volatility contribution for option valuation while delta-hedge the yield curve exposure.
Our objective: Analyzing returns on long-dated bonds

- EH uses long rates to predict short rate move, not the other away around.
  - How to predict long rate movements based on the yield curve shape, while accounting for risk premium and convexity?
  - More importantly, how to predict excess returns on long bonds?
- Modeling long rates with DTSM stretches the modeler’s imagination on how short rate should move in the next 30-60 years...
  - Mean reversion calibrated to time series or short end of the yield curve implies much smaller movements than observed from long rates.
  - Long rates are neither (easily) predictable, nor converging to a constant. — They move randomly, and with substantial volatility.
  - Can we say something useful about a 50-year bond without making a 50-year projection?
- Volatility can be much more accurately estimated than mean using historical data or options, how can one effectively use such information in predicting long-bond return behaviors?

The distinct behaviors of long bonds ask for a distinct modeling approach.
We propose a new modeling framework that is particularly suited for analyzing long bond returns:

- Link pricing of a long bond directly to the P&L attribution of investments in that particular bond.
  - Link pricing to short-term investment risks, rather than long-run expected cash flows.
- Price each bond based on its own behavior, not that of the short rate.
  - Localization allows one to make less ambitious but more confident statements, leveraging domain expertise.
- The model can say something useful about a 50-yr bond investment without making a 50-year projection, especially if one just wants to hold the bond for short term.
Notation and the classic *centralized* setting

- Let $B_t$ be the time-$t$ price of a default-free coupon bond (portfolio) with fixed future cash flows $\{C_j\}$ at times $\{t + \tau_j\} \geq t$ for $j = 1, 2, \cdots, N$.

- The classic valuation of this coupon bond can be represented as

$$B_t = \sum_j C_j \mathbb{E}_t^\mathbb{P} [M_{t, T}] = \sum_j C_j \mathbb{E}_t^\mathbb{P} \left[ \left( \frac{dQ}{d\mathbb{P}} \right) e^{-\int_t^{T_j} r_u du} \right] = \sum_j C_j \mathbb{E}_t^\mathbb{Q} \left[ e^{-\int_t^{T_j} r_u du} \right],$$

- $M_{t, T}$ — the pricing kernel linking value at time $t$ to value at time $T$
- $\mathbb{P}$ — the real world/physical/statistical/subjective probability measure
- $\mathbb{Q}$ — the so-called risk-neutral measure

- Bond pricing can start with the dynamics specification of the pricing kernel, the $\mathbb{P}$-dynamics for $r_t$ and market pricing, or directly the $\mathbb{Q}$-dynamics for $r_t$.

- DTSMs: Start with short rate $\mathbb{Q}$-dynamics with *common factor structure* $X$: $r_t = r(X)$, $dX_t = \mu(X_t)dt + \sqrt{\Sigma(X_t)}dW_t$, pricing is determined by the PDE,

$$r(X_t) = \mathbb{E}_t^\mathbb{Q} \left[ \frac{dB_t}{B_t dt} \right] = \frac{\partial B_t}{B_t \partial t} + \frac{\partial B_t}{B_t \partial dX_t} \mu(X_t) + \frac{1}{2} tr \left[ \frac{\partial^2 B_t}{B_t \partial XX^\top} \Sigma(X_t) \right]$$

*Integrate from today to expiry* to get pricing solutions $B(\tau, X)$.
A localized return attribution of a bond investment

- Represent the variation of a coupon bond in its own yield to maturity,
  \[ B_t(t, y_t) \equiv \sum_j C_j \exp(-y_t \tau_j). \]

- Attribute bond return wrt the variation of its own yield:
  \[
  \frac{dB_t}{B_t} = \frac{\partial B_t}{B_t \partial t} dt + \frac{\partial B_t}{B_t \partial y} dy + \frac{1}{2} \frac{\partial^2 B_t}{B_t \partial y^2} (dy)^2
  \]

- no longer short rate centric, but localized to the bond’s own rate/yield

- Bond returns come from three sources:
  1. **Carry**: \[ \frac{\partial B_t}{B_t \partial t} = y_t. \] Carrying the bond makes the yield \( y_t \) per year
  2. **Duration**: \( D \equiv -\frac{\partial B_t}{B_t \partial y} = \sum_j \frac{C_j e^{-y_t \tau_j}}{B_t} \tau_j \) (weighted average maturity)
    measures the sensitivity of the bond price to the yield change. Rate hike reduces bond return due to the inverse relation.
  3. **Convexity**: \( C \equiv \frac{\partial^2 B_t}{B_t \partial y^2} = \sum_j \frac{C_j e^{-y_t \tau_j}}{B_t} \tau_j^2 \) (weighted average maturity squared) captures the convexity of the price-yield relation. The convex relation increases the bond return with higher yield variation.
Given the *Carry-Duration-Convexity* bond decomposition

\[
\frac{dB_t}{B_t} = \frac{\partial B_t}{B_t \partial t} dt + \frac{\partial B_t}{B_t \partial dy} dy + \frac{1}{2} \frac{\partial^2 B_t}{B_t \partial y^2} (dy)^2 = y_t dt - D dy + \frac{1}{2} C(dy)^2
\]

the expected return from the bond investment is

\[
\frac{1}{dt} \mathbb{E}_t^P \left[ \frac{dB_t}{B_t} \right] = y_t - D \mu_t + \frac{1}{2} \sigma_t^2 C
\]

- \( (y_t, D, C) \) capture the bond exposures to time, directional yield move, and yield change variance.
- \( (\mu_t, \sigma_t^2) \) are the expected rates of yield change and variance:

\[
\mu_t = \mathbb{E}_t[dy_t]/dt, \quad \sigma_t^2 = \mathbb{E}_t [(dy_t)^2]/dt
\]

- Under diffusion dynamics, mean-variance forecasts capture everything.
- This is actually very old, ... much older than DTSMs
Given the decomposition \( \frac{dB_t}{B_t} = y_t dt - D dy + \frac{1}{2} C(dy)^2 \)

Take expectation under \( Q \), and set the instantaneous expected return to \( r_t \) by no dynamic arbitrage (NDA):

\[
E^Q_t \left[ \frac{dB_t}{B_t dt} \right] = r_t = y_t - D\mu_t^Q + \frac{1}{2} C\sigma_t^2
\]

with \( \mu_t^Q = E^Q_t [dy] / dt = \mu_t + \lambda_t \sigma_t \).

NDA leads to a simple \textit{pricing relation} for the long bond yield:

\[
y_t - r_t = \mu_t D + \lambda_t \sigma_t D - \frac{1}{2} \sigma_t^2 C
\]

The fair value of the yield spread \( (y_t - r_t) \) on any bond investment \( i \) is determined by the three conditional moment condition and pricing forecasts: \( \mu_t, \lambda_t, \sigma_t \)

There should be a superscript/subscript \( i \) on \((y, D, C; \mu_t, \lambda_t, \sigma_t^2)\) to indicate that both the risk exposures and moment conditions (and risk premiums) are particular to the \textit{ith} bond.
Bond pricing based on *local, near-term dynamics*

\[ y_t - r_t = \mu_t D + \lambda_t \sigma_t D - \frac{1}{2} \sigma_t^2 \tau^2 \]  \hspace{1cm} (1)

- **Local**: The fair valuation of the bond investment in (1) does not depend on short-rate dynamics, but only depend on the *behavior of its own yield*.

- **Near-term**: The pricing of the yield does not even depend on its own full dynamics, but only depends on the current estimates of first two moments.
  
  - The drift \( \mu_t \) and volatility \( \sigma_t \) can each follow some stochastic process, and/or depend on other rates/economic state variables ...
  
  - None of these dynamics specifications enter into the pricing relation

- **Views, not (much) dynamics**: One can bring in forecasts/estimates/opinions on volatility, risk premium, & directional forecasts on the yield, and examine their implications on the yield level.
  
  - The estimates can come from any (other) model assumptions, algorithms, or information sources, allowing maximum flexible cross-domain collaboration, without interference.
Three term decomposition of bond yields

\[ y_t - r_t = \mu_t D + \lambda_t \sigma_t D - \frac{1}{2} \sigma_t^2 C \]  

The new decentralized theory decomposes the yield spread into 3 terms, all related to the movement and pricing of the *yield itself*:

1. **Expectation**: The yield is higher, the faster it is expected to grow
2. **Risk premium**: The yield is higher, the higher the risk premium on the underlying bond
3. **Convexity**: The yield is lower, the more volatile its movements.

⇒ The yield curve cannot be flat if it moves (a lot) in parallel

The classic centralized theory can also be used for decomposing zero-bond yield to 3 terms, but all related to the movement and pricing of the short rate:

\[
y_t(\tau) = \frac{1}{\tau} \mathbb{E}_t^P \left[ \int_t^T r_u du \right] \quad \text{(Expectation)}
\]

\[
+ \frac{1}{\tau} \mathbb{E}_t^P \left[ \left( \frac{dQ}{dP} - 1 \right) \int_t^T r_u du \right] \quad \text{(Risk premium)}
\]

\[
- \frac{1}{\tau} \ln \mathbb{E}_t^Q \left[ \exp \left( -\left( \int_t^T r_u du - \mathbb{E}_t^Q \int_t^T r_u du \right) \right) \right] \quad \text{(Convexity)}
\]
Different frameworks serve different purposes

**DTSM**

- Starting PDE:
  $$r(X_t) = \frac{\partial B_t}{B_t \partial t} + \left[ \frac{\partial B_t}{B_t \partial X_t} \right]^\top \mu_t + \frac{1}{2} tr \left[ \frac{\partial^2 B_t}{B_t \partial XX^\top} \cdot \Sigma_t^X \right].$$

- *Full* short rate dynamics prices bond of *all* maturities.

- Maintain cross-sectional *consistency* across the whole curve.

- Hard to reconcile long rates with *actual* short rate dynamics.

- Better suited to construct smooth curves with cross-sectional consistency.

**New mean-variance analysis**

- Starting PDE:
  $$r_t = \frac{\partial B_t}{B_t \partial t} + \frac{\partial B_t}{B_t \partial y} \mu_t^Q + \frac{1}{2} \frac{\partial^2 B_t}{B_t \partial y^2} \sigma_{\tau t}^2$$

- Each yield is priced according to *its own near-term mean-variance* predictions.

- Lever *domain expertise*.

- Hard to maintain cross-sectional consistency across all bonds.

- Better suited to analyze specific bonds and connect to (expert views on) their own, current behaviors.
Different angles lead to different perceptions

**Classic centralized framework**

- Stationarity of short rate/states  
  \[ \Rightarrow \text{long rates can never fall} \]
  \[ \ldots \text{asymptotically as } \tau \to \infty. \]
- If one starts with short rate to derive its implication for super long rates, “stationarity” is a natural assumption to make for the impossibly long projections.
- The gap between observed long maturities and the asymptote remains very long.  
  *Infinity minus 60 is still infinity.*

**New local framework**

- \[ y_t = r_t + \mu^Q_t \tau - \frac{1}{2} \sigma_t^2 \tau^2 \]
  \[ \Rightarrow \text{long rates must fall} \]
  \[ \ldots \text{as long as they show substantial movements.} \]
- If one derives implications on a long rate based on its own, near-term dynamics, one just need to look at its own, recent variance.
- A must easier task, much less likely to go wrong.  
  *Aim small, miss small.*
Bond return prediction

- **EH**: Predict long yield changes/bond returns with the interest rate slope
  \[ \frac{1}{hD} ER_{t+h}(\tau) = \beta_0 + \beta_1 (y_t(\tau) - r_t) + e_{t+h}, \]

- **Cochrane-Piazzesi (2005)**: Use the shape of the whole curve
  - Use five forward rates and use regression to determine the loading
    \[ ER_{t+h} = \beta_0 + \sum_{j=1}^{5} \beta_j f_j^t + e_{t+h}, \]

- Predict bond returns with the extracted bond risk premium from the CP decentralized theory:
  \[ \lambda_t \sigma_t = \frac{y_t - r_t - \mu_t D + \frac{1}{2} \sigma_t^2 C}{D}. \]  
  (3)

- Estimate \( \sigma_t \) with a rolling-window historical volatility estimator on interest rate changes
- Predict future rate change with yield curve sliding: \( \mu_t = -s_t \)
  - Long constant-maturity rate is hard to predict

- Predict bond returns with macro variables...
Out-of-sample predicting performance

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Panel A. CP</th>
<th>Panel B. CW</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>DM</td>
</tr>
<tr>
<td>2</td>
<td>-0.35</td>
<td>-0.96</td>
</tr>
<tr>
<td>3</td>
<td>-0.43</td>
<td>-1.61</td>
</tr>
<tr>
<td>4</td>
<td>-0.41</td>
<td>-1.61</td>
</tr>
<tr>
<td>5</td>
<td>-0.39</td>
<td>-1.51</td>
</tr>
<tr>
<td>7</td>
<td>-0.31</td>
<td>-1.30</td>
</tr>
<tr>
<td>10</td>
<td>-0.21</td>
<td>-1.02</td>
</tr>
<tr>
<td>15</td>
<td>-0.13</td>
<td>-0.76</td>
</tr>
<tr>
<td>20</td>
<td>-0.08</td>
<td>-0.59</td>
</tr>
<tr>
<td>30</td>
<td>-0.06</td>
<td>-0.47</td>
</tr>
</tbody>
</table>

- CP bond risk premium has high in-sample $R^2$ but breaks down out of sample.
- CW bond risk premium can significantly outperform historical average when the maturity is 5 years and longer.
- These are on swaps; similar results on Treasury par bonds.
Source of outperformance

- CP bond risk premium is constructed with a portfolio of forward rates, with tent-shaped weight.
- CW bond risk premium is constructed with the yield spread adjusted with convexity, estimated based on historical variance.

The connection: The tent-shaped forward portfolio captures the yield curve curvature, which is highly related to the interest rate variance (Heidari & Wu, 2003). Both constructions can be related to the yield curve.

The distinction: The CP predictive regression determining the forward portfolio weights is not stable. Large estimation errors overwhelm the information in out-of-sample tests.

  - The CW bond risk premium construction does not involve a predictive regression, and does not experience much out-of-sample degeneration.

- Macro and other bond return predictors: Econometric issue is the main concern.
  - What specification generates the best risk structure is still up to debate.
Investing in *decentralized* butterflies

- Butterflies are a staple trade in fixed income (Duarte, Longstaff, Yu, 2007)
  - GA2+ is the starting point of many quant fixed income desks.
  - A 3-leg fly can hedge away the 2 factors, resulting in a stable portfolio.
  - Stat-arb: Take the model pricing errors as the alpha source, while using the fly construction to remove all (modeled) risk.
  - As long as model is good, one can use *any* 3 maturities to hedge away the 2 factors — *Model centric, but contract agnostic*.

- Instead of using any global common factor structures, our new theory inspires us to propose the concept of *local commonality*:
  - Contracts of nearby maturities naturally move together more closely,
  - ... regardless of model assumptions

- In the US, the most actively traded maturities center around Treasury bond futures, written on bonds at 6 maturity segments: 3, 5, 7, 10, 15, and 30

- We propose to construct *decentralized* butterfly bond portfolios around these reference maturities, *without* global factor structure assumptions.
  - *Contracts centric, but model assumption agnostic.*
Local commonality of swap rate movements

Cross-correlations of weekly changes of swap rates at different maturity gaps

- Swap contracts of nearby maturities *naturally* move together, independent of dynamics assumptions.
- Centralized models focus on getting the *global factor structure* correct.
- We rely more on contract specifications, to come up strategies less reliant on model assumptions.
- *Local commonality* is an approximate, but robust concept.
  - An estimated DTSM is exact, ... but can be fragile.
  - We’d rather be approximately right, than exactly wrong.
Global factor structures of swap rate movements

- The common wisdom: Three factors capture the level, slope, & curvature.
- The reality: More likely capture movements of different frequencies
- Number of frequencies/components “necessary” to capture the majority of movements depends on sample length/frequency and maturity range

Principal component analysis on swap rate changes:

<table>
<thead>
<tr>
<th>Maturity range</th>
<th>Explained percentage variation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>2-5</td>
<td>96.5</td>
</tr>
<tr>
<td>2-10</td>
<td>93.3</td>
</tr>
<tr>
<td>2-30</td>
<td>89.5</td>
</tr>
<tr>
<td>2-50</td>
<td>88.5</td>
</tr>
</tbody>
</table>

- A “global” factor structure is only valid relative to the targeted data range.
  - Earlier works focus on 2-10 yrs, and hence 2-factor structures.
  - Variations over 2-30 yrs need 3 factors, probably more with longer maturities...
- The only thing that is globally true is the local commonality.
Construct decentralized butterflies with nearby maturities

- We do not make global factor dynamics assumptions, just estimate the covariance $\Sigma_t$ of weekly changes of the three swap rate series in the butterfly, choose the weights to minimize the portfolio variance,

$$\min_{b_t} b_t^T \Sigma_t b_t, \quad \text{subject to} \quad b_t^T i_2 = 1,$$

(4)

- Regress swap rate changes in middle against the two on the wings

### Annualized vol of weekly swap rate changes and variance ratio for butterflies:

<table>
<thead>
<tr>
<th>Reference maturity</th>
<th>Swap Vol</th>
<th>Variance ratio of butterflies with maturity gap:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.89</td>
<td>0.008</td>
</tr>
<tr>
<td>5</td>
<td>0.94</td>
<td>0.008</td>
</tr>
<tr>
<td>7</td>
<td>0.95</td>
<td>0.006</td>
</tr>
<tr>
<td>10</td>
<td>0.94</td>
<td>0.014</td>
</tr>
<tr>
<td>15</td>
<td>0.90</td>
<td>0.005</td>
</tr>
<tr>
<td>30</td>
<td>0.85</td>
<td>0.007</td>
</tr>
</tbody>
</table>

- *The smaller the maturity gap, the smaller is the variance ratio.*
Mean reversion in swap rates and butterfly portfolios

Autocorrelation estimates on weekly changes of swap rates and butterflies:

<table>
<thead>
<tr>
<th>Reference maturity</th>
<th>Swap rates</th>
<th>Butterflies with maturity gap:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>-0.003</td>
<td>-0.394</td>
</tr>
<tr>
<td>5</td>
<td>-0.020</td>
<td>-0.517</td>
</tr>
<tr>
<td>7</td>
<td>-0.028</td>
<td>-0.435</td>
</tr>
<tr>
<td>10</td>
<td>-0.035</td>
<td>-0.401</td>
</tr>
<tr>
<td>15</td>
<td>-0.051</td>
<td>-0.438</td>
</tr>
<tr>
<td>30</td>
<td>-0.033</td>
<td>-0.222</td>
</tr>
</tbody>
</table>

- Changes in butterflies constructed with nearby maturities are much more predictable than changes in swap rates.
- The smaller the maturity gap, the stronger the predictability and stability.
  - Instead of predicting each swap rate separately, we can directly predict the expected change of the swap-rate butterfly portfolio, $\mu_{t,f}$, based on mean reversion.
- Super wide flies are no longer stable, and no more predictable than the single swap rate series.
Out of sample investment on decentralized butterflies

- Make investment based on predicted excess returns (from carry, convexity, and expected rate change):
  \[ EER_{t,f} = w_t^\top (y_t - r_t) + w_t^\top \hat{a}_t - D_2 \mu_{t,f}, \]
  \[ \hat{a}_{t,i} = \frac{1}{2} C_i \hat{\sigma}_{t,i}^2 \]
  and variance \( \hat{\nu}_{t,f} = D_2^2 \hat{\sigma}_{t,e}^2 \).
- Carry is observable, convexity is based on variance estimators.
- Predict the expected fly rate change \( \mu_{t,f} \) based on mean reversion.
- Allocate to each fly based on risk-return tradeoff: \( n_t = s \frac{EER_{t,f}}{\hat{\nu}_{t,f}} \).

| Allocation weights on butterflies constructed with adjacent maturities: |
|-----------------------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Maturity | 3 | 5 | 7 | 10 | 15 | 30 |
| 25 | -16.51 | -9.29 | -7.97 | -5.14 | -5.51 | -5.38 |
| 50 | -0.43 | -0.15 | -0.25 | 0.05 | -0.05 | 0.00 |
| 75 | 14.94 | 7.24 | 7.55 | 4.67 | 6.73 | 5.29 |
| 90 | 40.44 | 21.07 | 23.21 | 15.31 | 24.57 | 18.98 |
| Corr | -0.04 | 0.01 | 0.07 | 0.09 | -0.05 | -0.06 |

- Allocation weights are reasonably symmetric around 0.
- With the same coefficient \( s \), weights vary more at shorter maturities.
- **Decentralized flies**: Weights across flies do not co-vary.
Out of sample investment performance

in butterflies constructed with adjacent maturities

<table>
<thead>
<tr>
<th>Maturity</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>10</th>
<th>15</th>
<th>30</th>
<th>Aggregate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.13</td>
<td>0.12</td>
<td>0.09</td>
<td>0.09</td>
<td>0.14</td>
<td>0.20</td>
<td>0.12</td>
</tr>
<tr>
<td>Stdev</td>
<td>0.10</td>
<td>0.09</td>
<td>0.07</td>
<td>0.09</td>
<td>0.08</td>
<td>0.16</td>
<td>0.05</td>
</tr>
<tr>
<td>IR</td>
<td>1.26</td>
<td>1.44</td>
<td>1.17</td>
<td>1.03</td>
<td>1.61</td>
<td>1.24</td>
<td>2.66</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.33</td>
<td>2.87</td>
<td>4.25</td>
<td>1.37</td>
<td>3.65</td>
<td>2.06</td>
<td>0.99</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>32.07</td>
<td>27.25</td>
<td>42.38</td>
<td>56.78</td>
<td>33.86</td>
<td>36.11</td>
<td>24.32</td>
</tr>
<tr>
<td>Corr</td>
<td>-0.01</td>
<td>0.08</td>
<td>0.11</td>
<td>0.10</td>
<td>0.06</td>
<td>0.09</td>
<td>0.07</td>
</tr>
</tbody>
</table>

- All flies constructed with adjacent maturities generate good investment performance:
  - All IR > 1, returns all show positive skewness.
  - Decentralized: low correlation among the returns of the butterflies. Diversification further enhances the performance of the aggregate portfolio.
Performance deterioration with increasing maturity gaps

<table>
<thead>
<tr>
<th>Maturity</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>10</th>
<th>15</th>
<th>30</th>
<th>Aggregate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A. Maximum maturity gap = 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR</td>
<td>0.87</td>
<td>0.78</td>
<td>0.53</td>
<td>0.54</td>
<td>0.65</td>
<td>0.79</td>
<td>1.18</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.02</td>
<td>0.96</td>
<td>-1.70</td>
<td>-4.07</td>
<td>-0.30</td>
<td>0.39</td>
<td>-2.00</td>
</tr>
<tr>
<td>Corr</td>
<td>0.01</td>
<td>0.13</td>
<td>0.34</td>
<td>0.32</td>
<td>0.35</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>Panel C. Maximum maturity gap = 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR</td>
<td>0.62</td>
<td>0.40</td>
<td>0.33</td>
<td>0.45</td>
<td>0.45</td>
<td>0.72</td>
<td>0.66</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.98</td>
<td>0.86</td>
<td>-1.86</td>
<td>-4.45</td>
<td>-2.25</td>
<td>-0.14</td>
<td>-2.47</td>
</tr>
<tr>
<td>Corr</td>
<td>0.22</td>
<td>0.48</td>
<td>0.61</td>
<td>0.57</td>
<td>0.52</td>
<td>0.43</td>
<td>0.47</td>
</tr>
</tbody>
</table>

- As the gap becomes wider, the mid-maturity flies become less stable.
  - IR becomes lower; returns show negative skewness
  - *Chained together:* Mid-maturity flies become *more intertwined*. Correlations become higher. Diversification effect becomes smaller.
- Contract choice can play a big role in the success of an investment strategy.
  - Smart quants try to design an omnipotent model.
  - Wise investors try to pick & choose the appropriate contracts.
Statistical arbitrage investments based on DTSMs

Form flies to neutralize 2 DTSM factors, invest to arb pricing errors

Ref. maturity: 3 5 7 10 15 30 Aggregate

<table>
<thead>
<tr>
<th>Maturity gap</th>
<th>Panel A. two-factor model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.65 1.12 0.86 0.57 0.83 0.74 1.54</td>
</tr>
<tr>
<td>2</td>
<td>0.66 0.52 0.29 0.33 0.51 0.74 0.86</td>
</tr>
<tr>
<td>4</td>
<td>0.40 0.07 -0.02 0.24 0.34 0.66 0.35</td>
</tr>
</tbody>
</table>

- Stat-arb trading with adjacent maturities generates reasonable performance..
- Maturity choice remains important even under centralized modeling
- Dilemma between choosing an ill-specified 2-factor model and deciding which factors to hedge in a higher-dimensional better-specified model.
  - Estimate a 2-factor structure locally ...
- Stock market analogy
  - Pair trading; multiples comparison (or local weighting) within a selective universe ...
  - Careful universe/pair/group selection can reduce reliance on modeling
    - Industry dummy in valuation model removes industry bias; industry dummy in return structures induces industry neutrality in portfolio construction