NATIONAL RESEARCH UNIVERSITY

Art of Yield Curve Modelling:

Joint Consistency of Russian Government Bond Quotes

Victor Lapshin

joint work with Marat Kurbangaleev

Higher School of Economics

vlapshin@hse.ru

Yield Curve Modelling

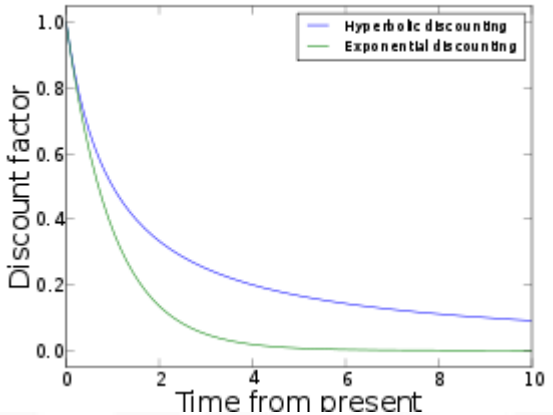
- What?
- Why?
- How?

Yield Curve Modelling

- **Why?**
- What?
- How?

What's a Yield Curve?

- Discount function: time value of money.
  - \$1 in time ***t*** is worth \$***d(t)*** now.
- Zero-coupon yield curve: $d(t) = \exp(-r(t) \cdot t)$
- Instantaneous forward rate: $d(t) = \exp\left(\int_0^t f(\tau)\,\mathrm{d}\tau\right)$
- Other representations…
- The main problem:
  it is a fiction notion,
  a pure mind construct.



What: Main Problems of Yield Curve Modelling

- Estimation.
  - Given the data, estimate the (unobservable) yield curve.
  - Heavily depends on the nature of the data.
- Forecasting.
  - Forecast the future yield curve.
- Stochastic modelling.
  - Estimate the distribution.
  - Risk-neutral / real-world.
- Scenario analysis.
  - What if…

Why: Main Applications of Yield Curve Modelling

- Pricing.
  - Discounted cash flows: investment-like instruments (e.g. bonds).
  - Using the yield curve to price financial instruments (derivatives).
- Risk management, actuarial assessment and regulation.
  - Fair value accounting (re-pricing): mark-to-market / mark-to-model.
  - Sensitivity analysis: interest rate risk capital.
  - Hedging.
  - Stress-testing.
- Investment analysis and decision making.
  - NPV, IRR, real options.
- Macroeconomic analysis.
  - Expectations analysis and forecasting.
  - Policy making.
  - All kinds of macroeconomic modelling.

How: Available Data

- Ready yield curve estimates (Bloomberg, Reuters, central banks, exchanges, reports, etc.)
  - Black box: methodologies and source data are rarely thoroughly described.
  - Must be suitable for the purpose.
- Market data: interest rate is not a traded asset.
  - Money market instruments: REPOs, Ibors, xxONIAs (FedFund, EONIA,…), T-bills and the like.
  - Bonds: government, corporate…
  - Interest rate swaps.
  - Other interest rate derivatives: FRAs, caps/floors/collars, swaptions, etc.
- Non-market data.
  - Macroeconomic statistics (e.g. Ultimate Forward Rate).
  - Expert estimates and forecasts.

#82365610

How: Market Data Quality

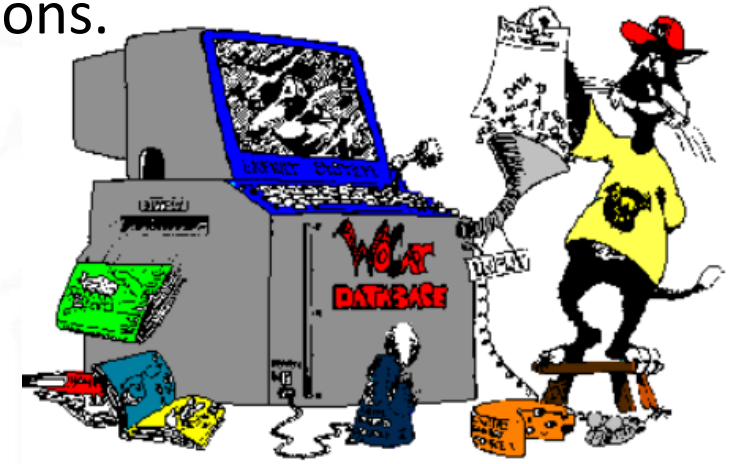
- How credible is the data?
  - Deals (marketplace): *ex post*
  - Quotes (marketplace, fixing by a 3rd party): *ex ante*
    - Commitment / no commitment
  - Valuation (data provider, internal, anybody…): *ex machina*
- Raw data / aggregated
  - Time (best, average, last, close)
  - Cross-section (best, average, mid, filtering)
  - Update rules (sync, async)

Always Start with ‘Why?’

- We are going to study the problem of estimating the Russian yield curve.
- Different applications require different data sources for the yield curve.
  - See the PWS-2015 talk.
- We are going to assess data quality: for which purposes can we use this data?

How: Models and Methods

- We are not solving the problem of term structure estimation.
  - We are studying the possible ways of solving it and the possible mishaps.
- There are various term structure fitting methods which produce a zero-coupon yield curve from coupon bond price data.
- These methods have their assumptions.
  - The whole notion of the zero-coupon yield curve is based on these assumptions.



The Usual Assumptions

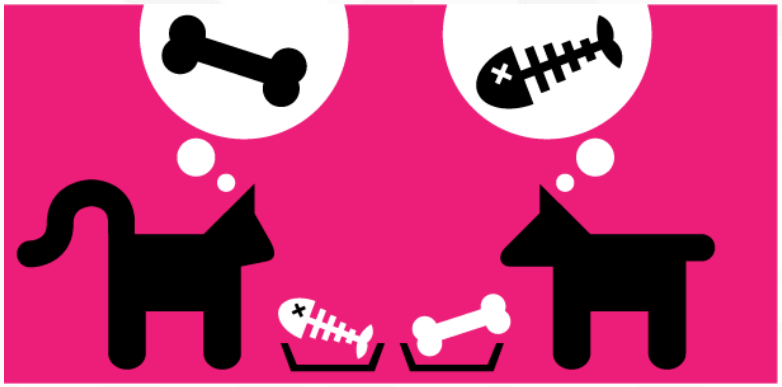
- No risk of any kind.
  - Especially, no default risk.
- Absolute liquidity.
  - No transactional costs.
  - Infinite depth (no price impact).
  - Immediate transactions.
  - Buy and sell at the same price.
- Infinitely divisible instruments.
- Unrestricted short selling.
- No taxation.
- No arbitrage.

The Data

- It would be nice if the input data didn't contradict the assumptions, at least explicitly.
  - Deals data is usually OK.
  - Quotes data includes different buy and sell prices.
- Fortunately, this can easily be incorporated into a new model.
  - However, the other assumptions remain in place.

Model vs. Data

- Our fitting model has a certain number of general financial assumptions.
  - Sometimes they are not realistic, but at least they are common and generally accepted.
- However, if the data clearly contradicts these assumptions, we might expect irregularities.
  - Some assumption violations are neglectable, some are not.



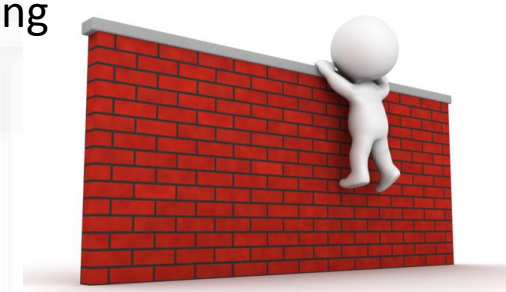
Neglectable vs. Unneglectable

A neglectable assumption violation.

- Bond 1 is risk-free,
  Bond 2 has default risk.
- The ‘no risk’ assumption is violated, but we can choose to neglect this, because:
  - Nobody really cares about this particular default risk.
  - Our other models ignore default risk.
  - This is the only way for us to gather enough data.
  - …

An unneglectable assumption violation.

- Bond 1 is risk-free,
  Bond 2 has default risk.
- We cannot neglect the violation of the ‘no-risk’ assumption, because:
- They are otherwise similar, and such large a price difference in similar bonds would yield arbitrage opportunities under the assumptions.
- This may be happening in another model, to which the results are being fed.



Our Objective

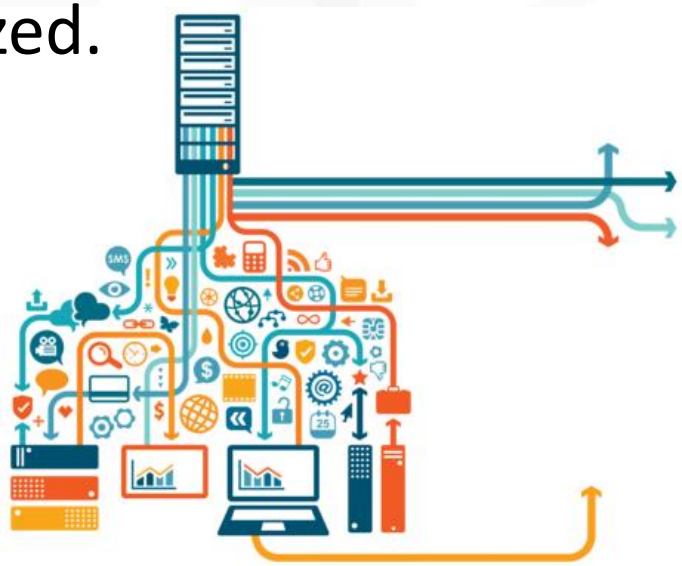
- We know that the data doesn't satisfy the assumptions. It never does.
- Can we safely ignore it?
  - Supposedly, we are aware of the precision loss.
- Or will something happen?
  - Theoretical arbitrage opportunities in the data can sometimes totally ruin a model.
  - Even if it is not a real arbitrage, a portfolio optimization algorithm might easily go crazy.

Outline

- Describe the data and its nature.
- Check the data for internal consistency given the common financial assumptions.
- Report the discrepancies and study them.
- Present an example.

The Data

- Russian government bonds.
  - Fixed-coupon, maybe amortized.
- Data Provider: Cbonds.
  - Source: Moscow Exchange.
- Daily quotes: bid, ask.
- Quote type: close.
- Dates: January 2007 – April 2015.

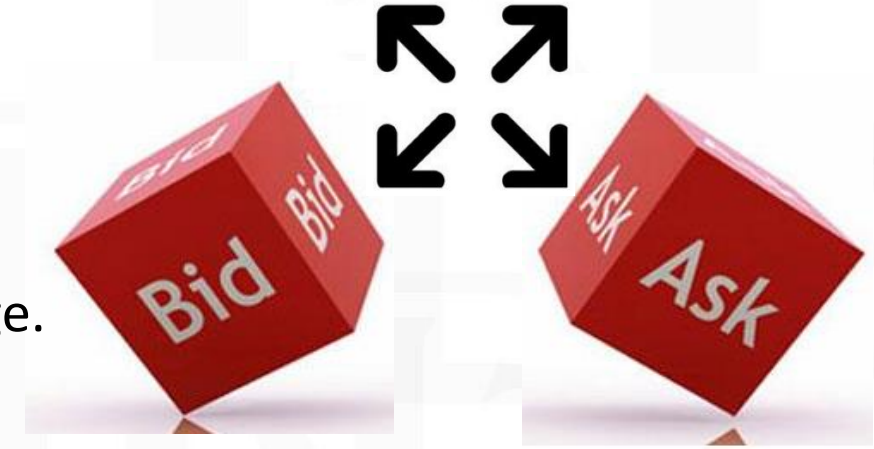


The Bid-Ask Spread

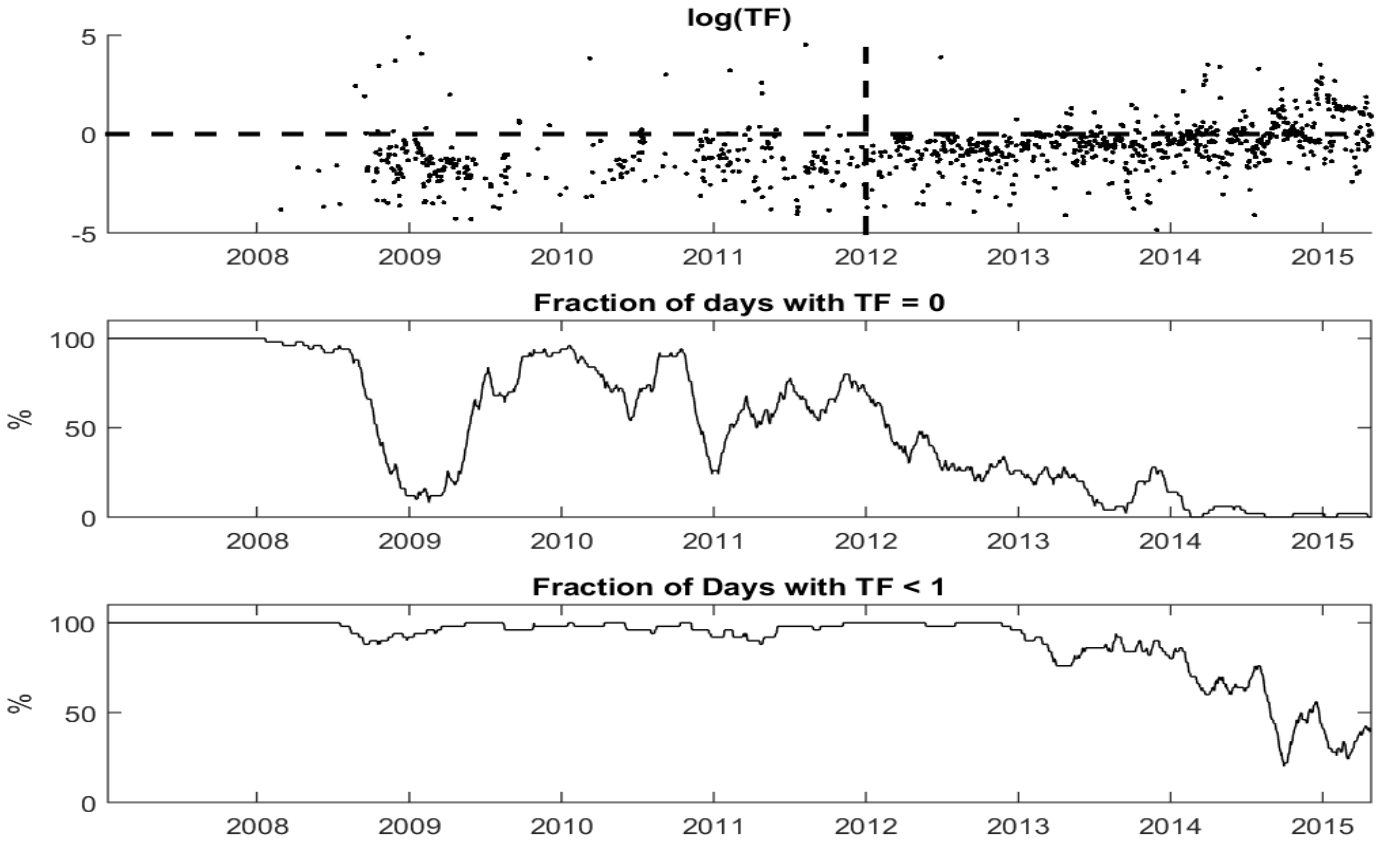
- A single price is easier to deal with: just use it.
- Bid-ask spreads supposedly give us more information.
  - The prices (we can always take the mid-points).
  - The boundaries.
  - The degree of precision.
- But… do they really?

Tightness Factor

- Suppose there are theoretical arbitrage opportunities: buy something, sell something, profit (within the assumptions).
  - If bid-ask spreads were very high, there would be no arbitrage: we would be buying high and selling low.
- TF is the least factor by which one must widen the bid-ask spreads to eliminate the theoretical arbitrage opportunities.
  - TF = 2 means the spreads must be twice as wide.
  - TF = 0.5 means the spreads could be twice as narrow and still there would be no arbitrage.
  - TF = 0 means the spreads could be absent.



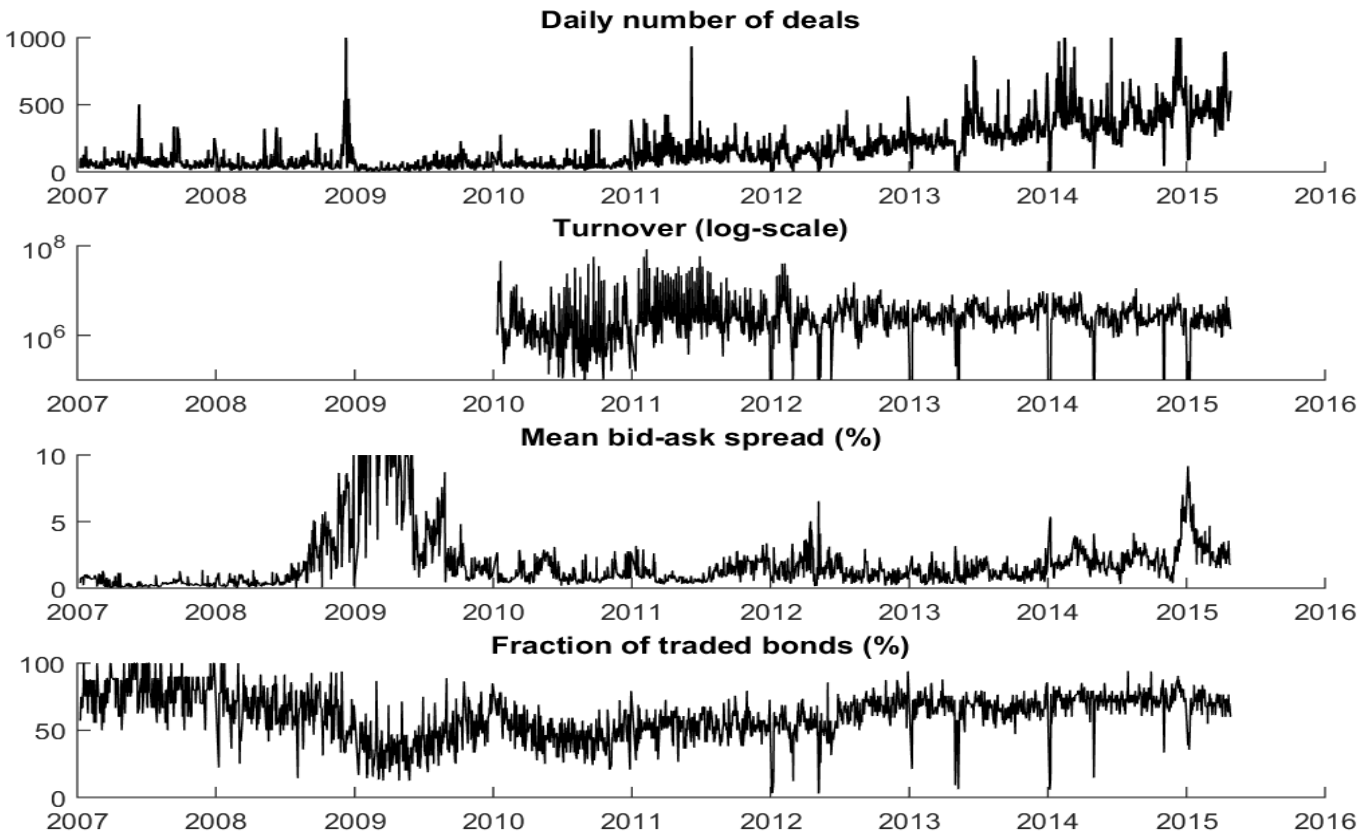
Tightness Factor



Findings

- There is a structural break in early 2012.
- Before that, the dataset is usually consistent.
- After that, its quality is deteriorating.
- Traditional liquidity measures are unable to capture this change.

Traditional Liquidity Measures



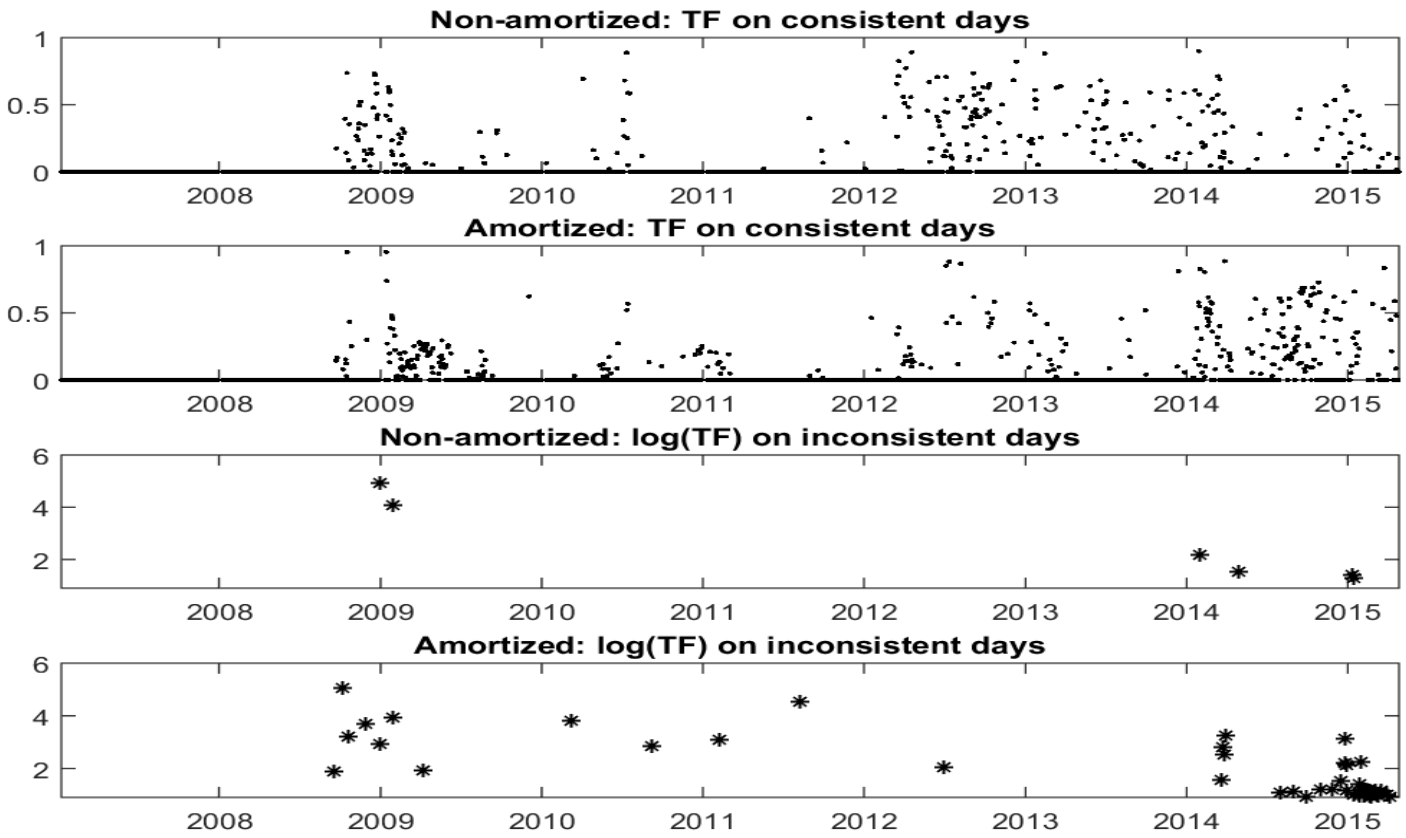
Filtering Procedure

- Which bonds introduce these arbitrage opportunities?
  - We propose an iterative algorithm to spot them.
- Filter the offending bonds out.
  - Get a consistent dataset.
  - Find something common in the arbitrage-introducing bonds.

Filtering Results

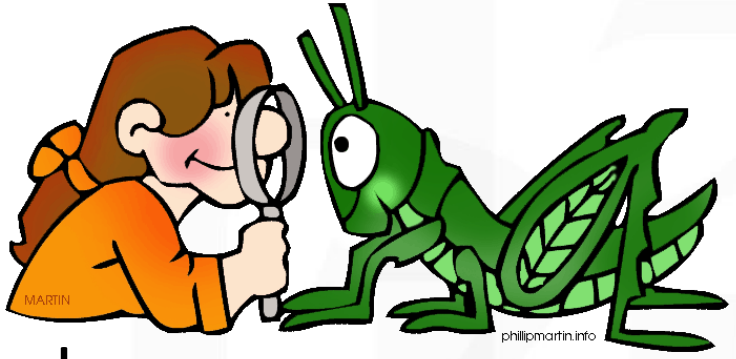
- Turns out all ‘top-offenders’ are amortized bonds.
- They also happen to be relatively illiquid.
- The logical next step is to split the dataset into two: non-amortized and amortized bonds.

Separate Estimation



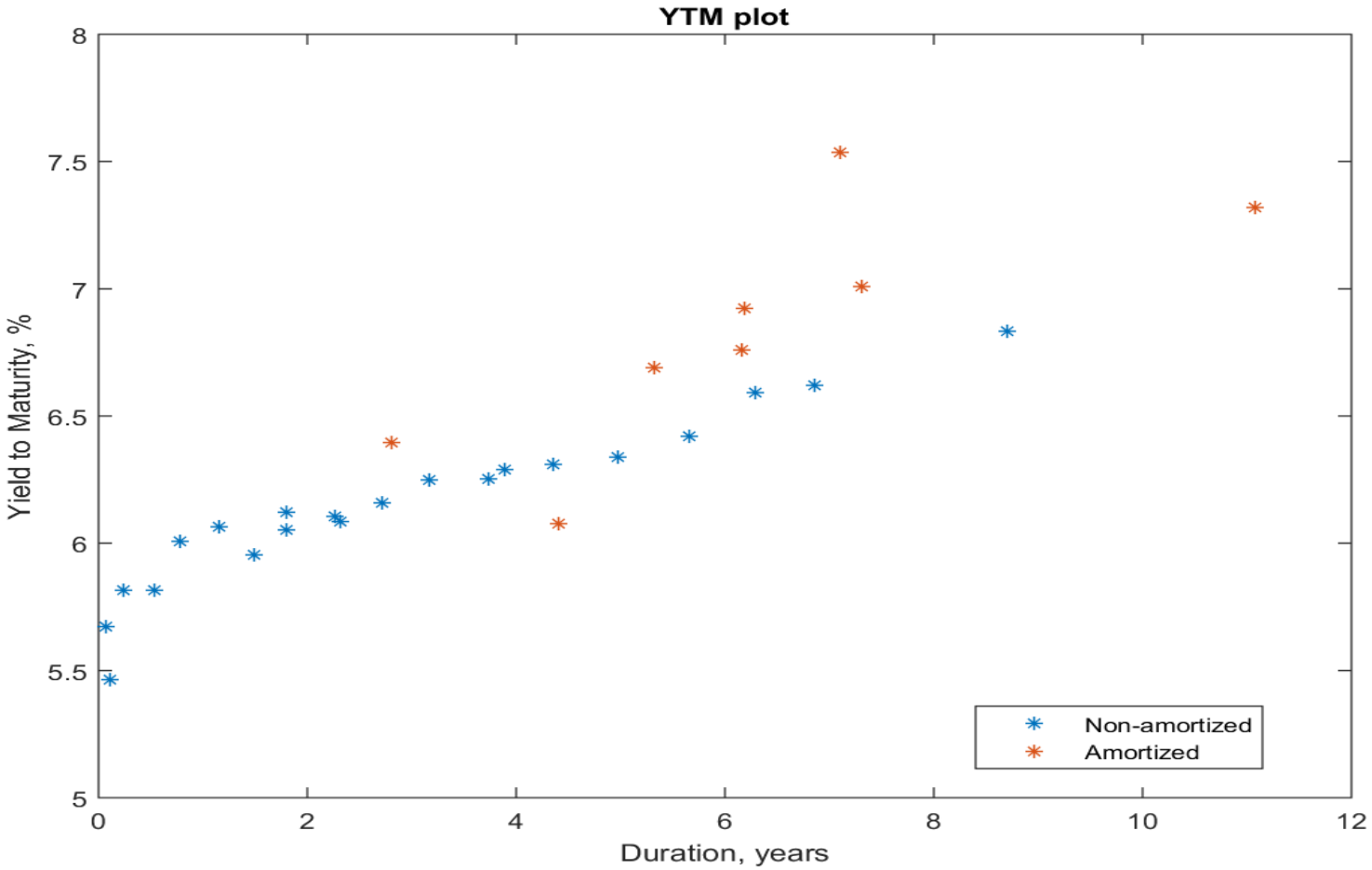
Observations

- Market segmentation: amortized vs. non-amortized.



- Amortized bonds are a less homogeneous group: they are more often inconsistent even among themselves.
- This analysis further supports the hypothesis of a structural break: prior to 2012 non-zero TFs were observed only in crises. After 2012 it has become normal.

Example


YTM plot
Yield to Maturity, %
Duration, years
8
7.5
7
6.5
6
5.5
5
0
2
4
6
8
10
12
Non-amortized
Amortized

Conclusions

- A methodologically sound approach to modelling starts with ‘why?’
- A sound approach doesn’t guarantee successful modelling, because of the data.
- The data comes in various forms. Using the wrong data might be a worse mistake than using a wrong model.
- We introduce a tool for studying the consistency of quote-type data.
- We apply this tool to study the segmentation of the Russian government bond market.

NATIONAL RESEARCH UNIVERSITY