ECONOMIC MODELLING & MACHINE LEARNING

A PROOF OF CONCEPT

NICOLAS WOLOSZKO, OECD

TECHNOLOGY POLICY INSTITUTE – FEB 22 2017

OECD BETTER POLICIES FOR BETTER LIVES

Economic forecasting with Adaptive Trees

1 Motivation

2 Adaptive Trees

3 Proof of Concept

4 Perspectives

I. MOTIVATION

Machine learning and economic complexity

Linear models are constrained by economic complexity

Non-linearities

Multiple interactions

Multiple discontinuities

Structural change

Relationships between variables may change over time, suddenly or incrementally

Context-specific impact of policies

Depending on countries

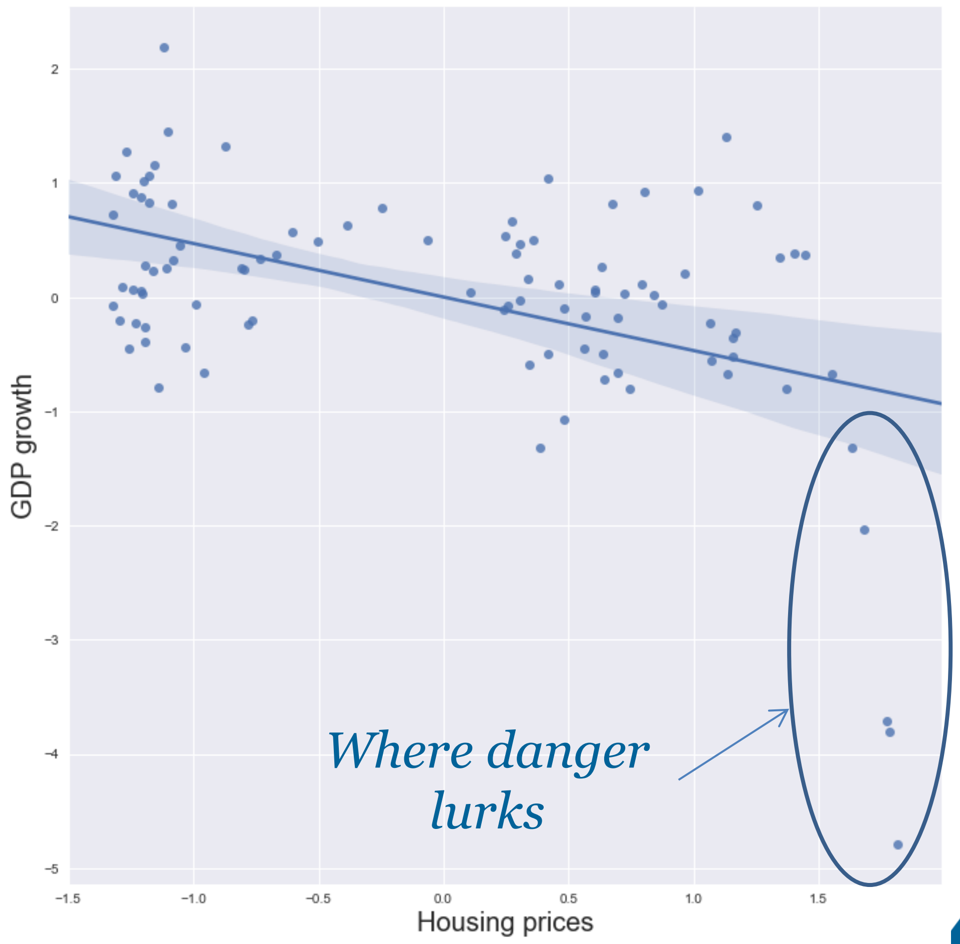
Depending on people's place in income, skills, or age distribution

Especially around **turning points**

Non-linearities

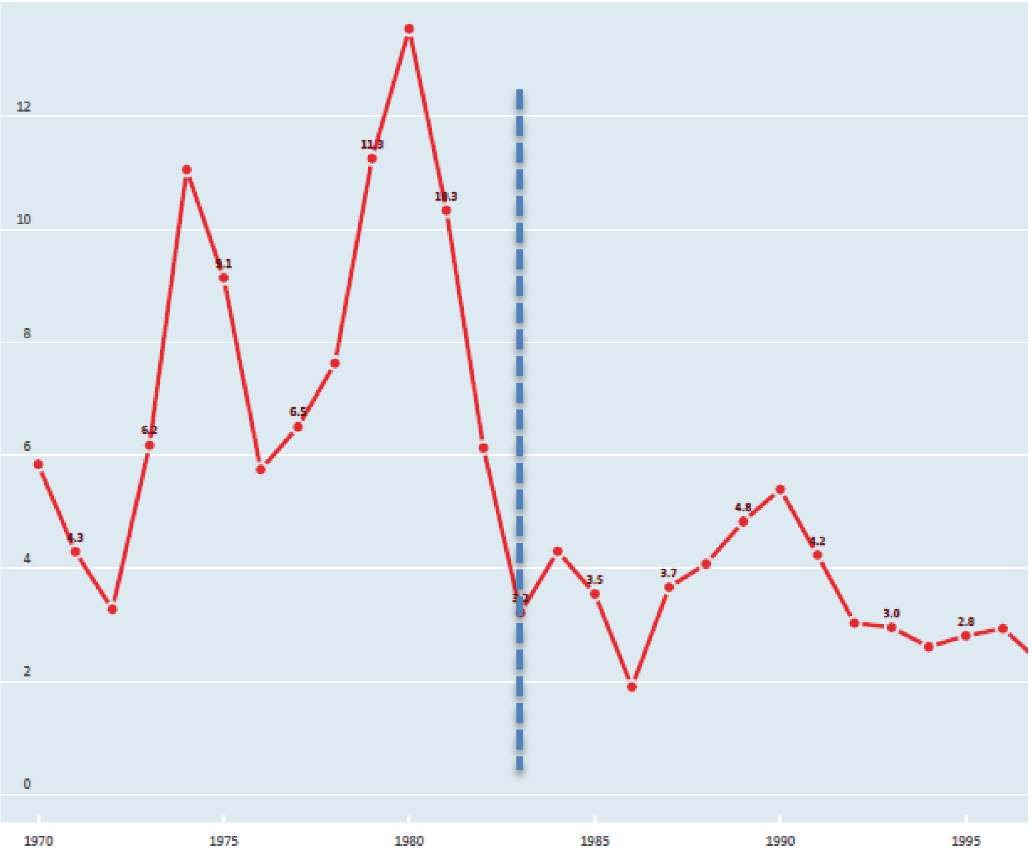
Housing prices against GDP growth, UK

Non-linear behaviour past a given threshold, at a tipping point



Structural change

Inflation in the US, 1970-2017



Monetary policy helped tame inflation and changed the nature of the Philipps Curve, by stabilising inflation expectations.

Machine learning provides tools to tackle these challenges

What is machine learning ?

- Powerful methods designed to extract information from data

How is different from econometrics ?

- Modelling without a model: no prior knowledge is required
- Relies on cross-validation to prevent overfitting and underfitting

How can it be useful ?

- Uncover complex patterns in data, even from a vast array of variables
- Data comes first, theory comes next

II. ADAPTIVE TREES

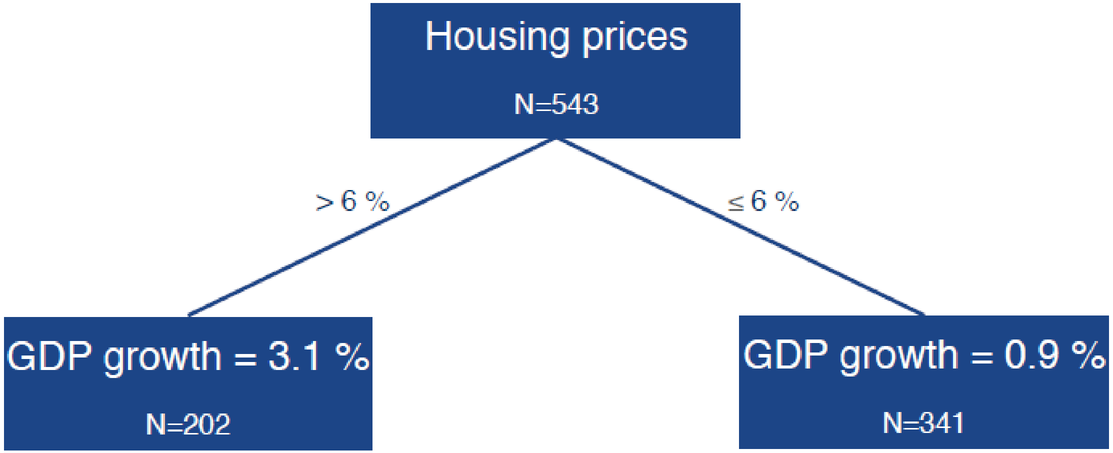
A non-linear approach to capture structural change in the economy

Adaptive Trees: two steps

1. Tackling *non-linearities* with regression trees

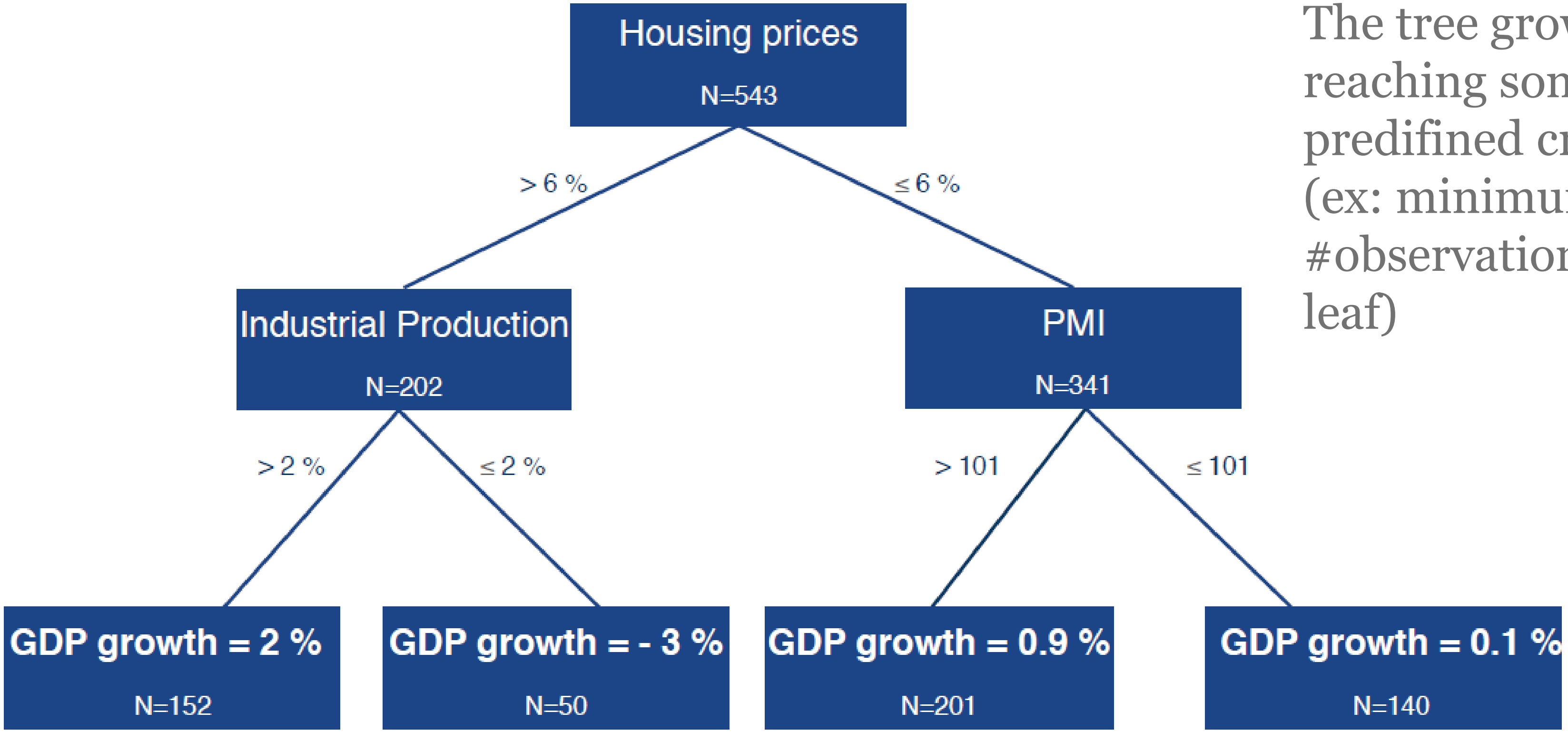
2. Adressing *structural change*: adaptive trees

Training regression trees



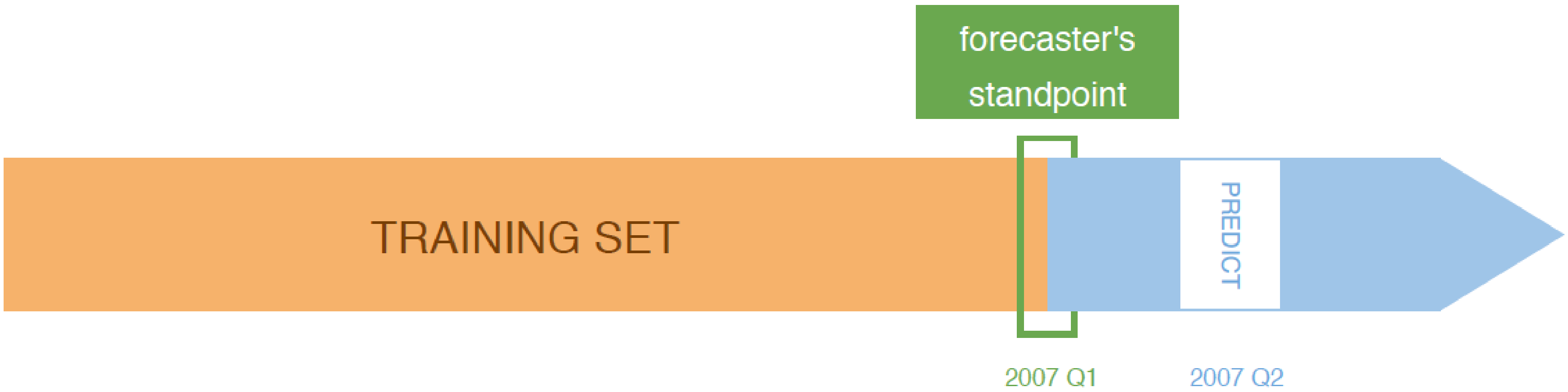
At each node, the algorithm selects the splitting variable + splitting point that minimises sub-group variance of GDP growth

Training regression trees



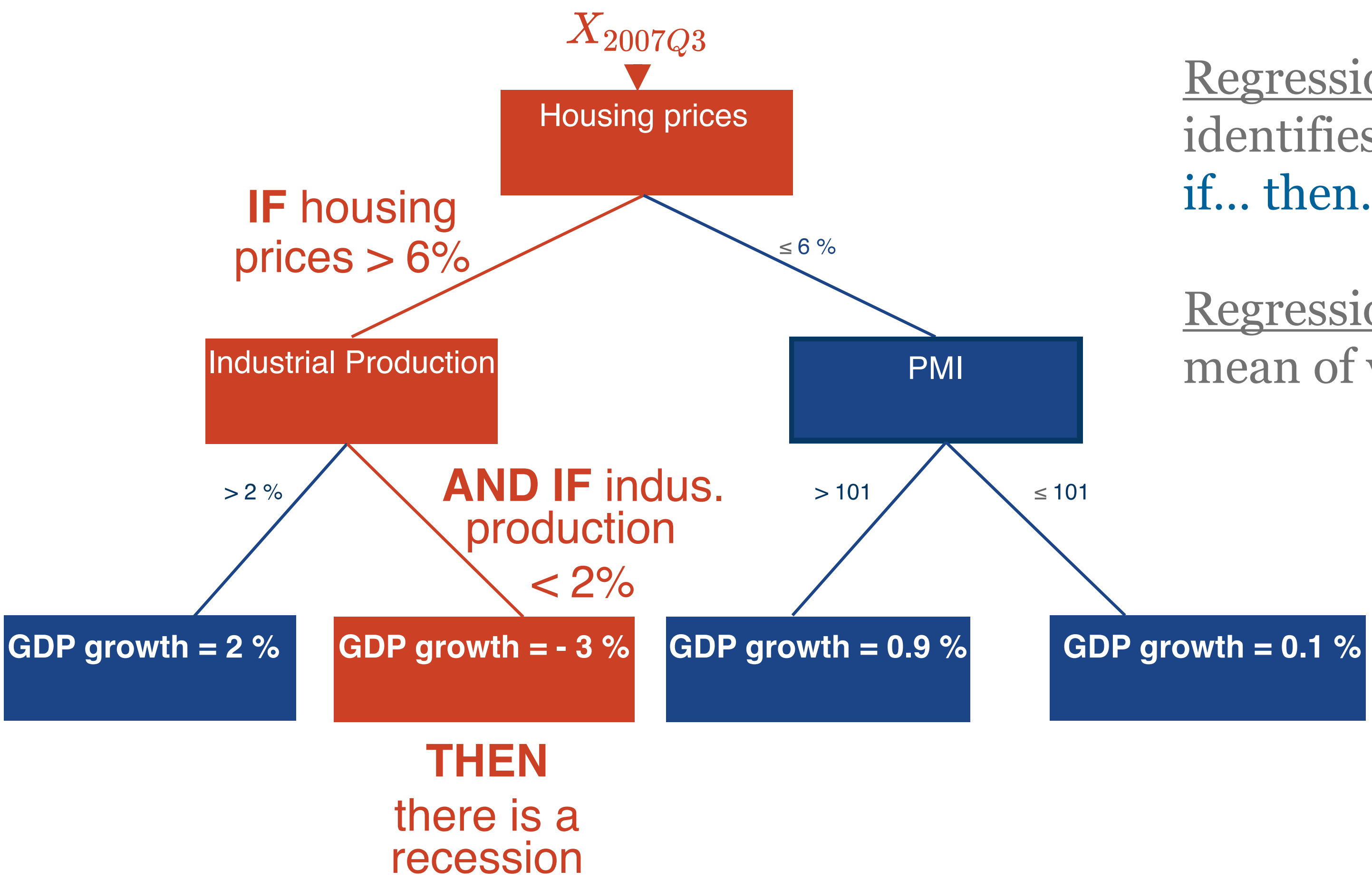
The tree grows until reaching some predifined criteria (ex: minimum #observations per leaf)

Train & Predict



The tree is grown using past data (training). Then it makes a prediction about the future (here, Q+1), using contemporaneous and past data.

Prediction



Regression trees: Prediction identifies complex structure: if... then...

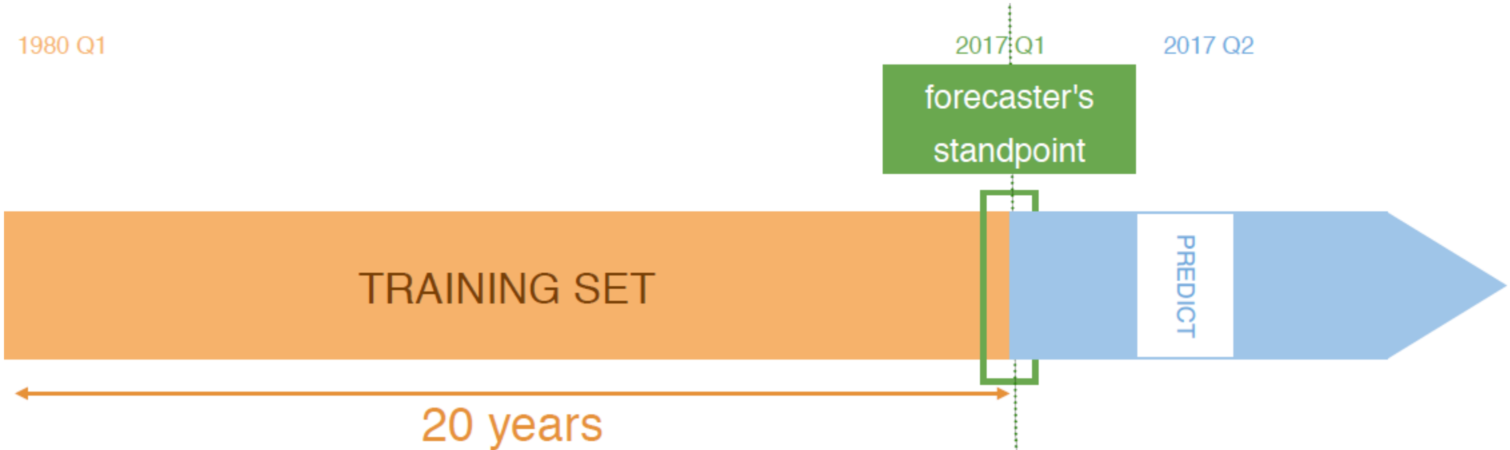
Regression: simple weighted mean of variables

Adaptive Trees: two steps

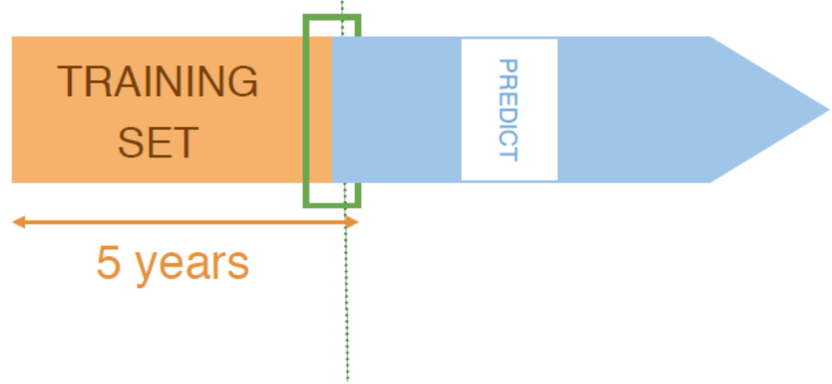
1. Tackling *non-linearities* with regression trees

2. Adressing *structural change*: adaptive trees

Trade off quantity/relevance



Without structural change, we want to use as much data as possible



But in presence of strong structural change, we need to focus on most recent data

Structural change

- The economy is ever-changing**.** That is part of « economic complexity ».
- Consequence: recent past more informative about near future than remote past
- There may be **sudden structural breaks** (during crises), or **incremental structural change**
- We tackle structural change using an original technique that we developped for the purpose of economic forecasting: « Adaptive Trees »

Adaptive Trees

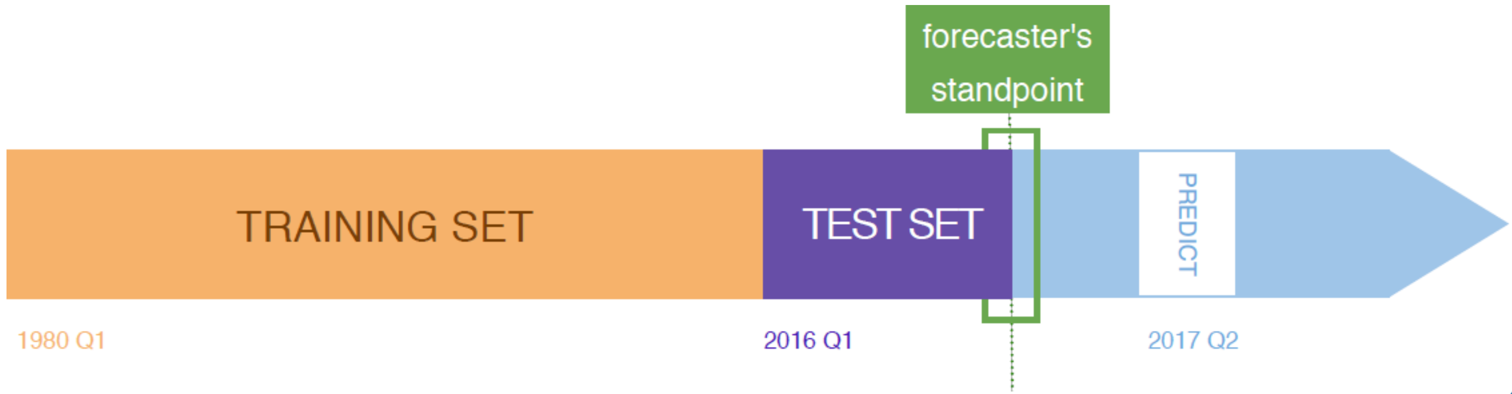
Adaptive Trees are a transformation of the Gradient Boosting algorithm

Tackling incremental structural change:

- Give more weight to the recent past

Tackling sudden structural change:

- Detect structural change: measuring how accurately the algorithm trained on the training set can predict the latest observations
- If not well: gives even more weight to the recent observations that are hard to predict



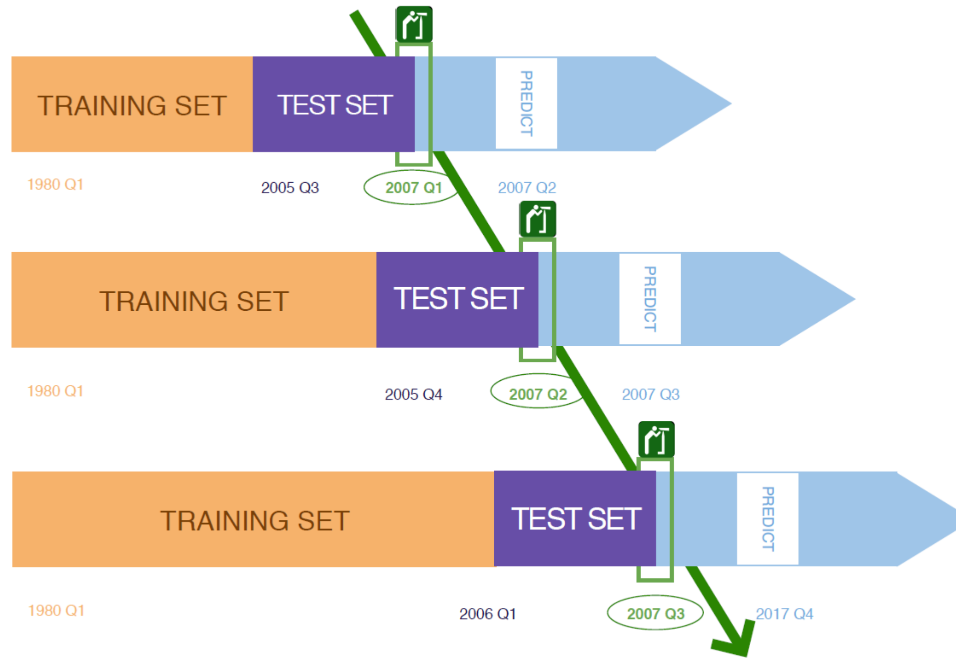
III. PROOF OF CONCEPT

FORECAST OF GDP GROWTH IN G6 COUNTRIES

Setting of forecast simulations

- Simulations in pseudo-real time of a forecast of GDP growth in G6 countries

- Using the exact same data as benchmark OECD Indicator Model (housing prices, indutrial production, PMI…) so as to provide a *methodological benchmark*



Benchmark forecasts

Compare with a benchmark forecast:			
OECD Indicator Model	M+3 & M+6	VAR	2007 – 2016, quarterly, q-o-q

Measuring performance:

- Accuracy: Root Mean Square Error (RMSE)
- Forecast Directional Accuracy (FDA): % times forecasts right direction

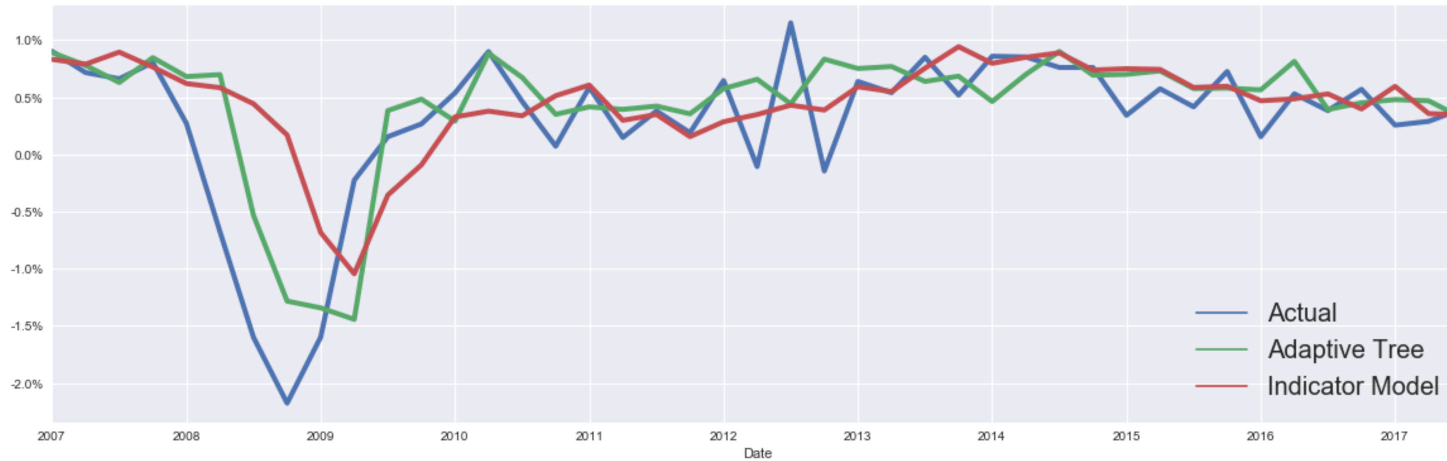
Quantitative

Qualitative

Comparison with OECD Indicator Model

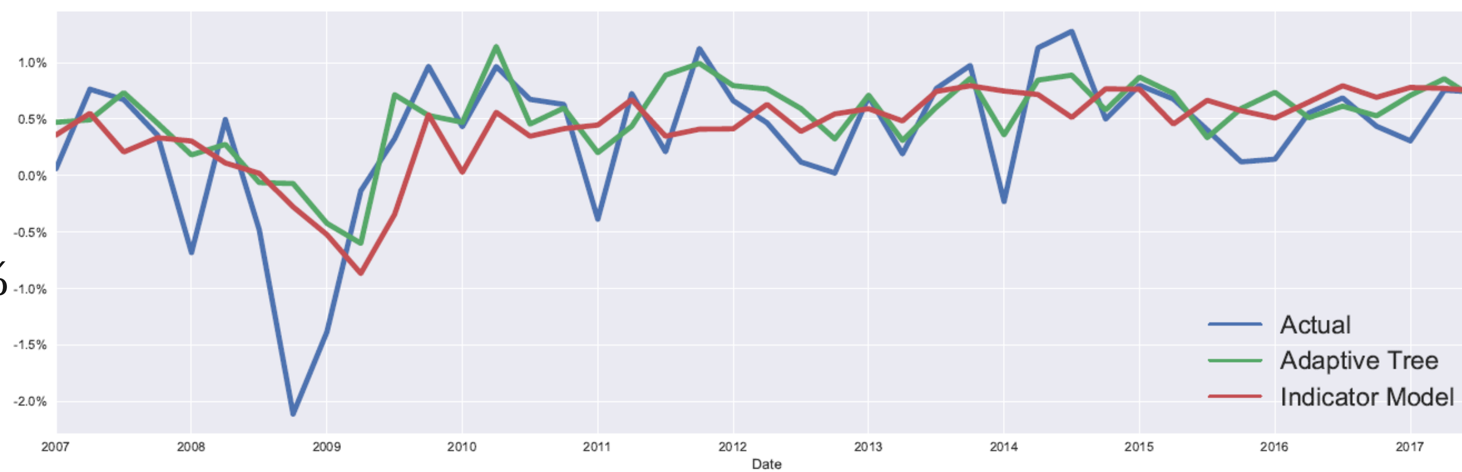
1. UK, M+3

Accuracy: +25 %
Dir. Accuracy: +4 %



2. USA, M+3

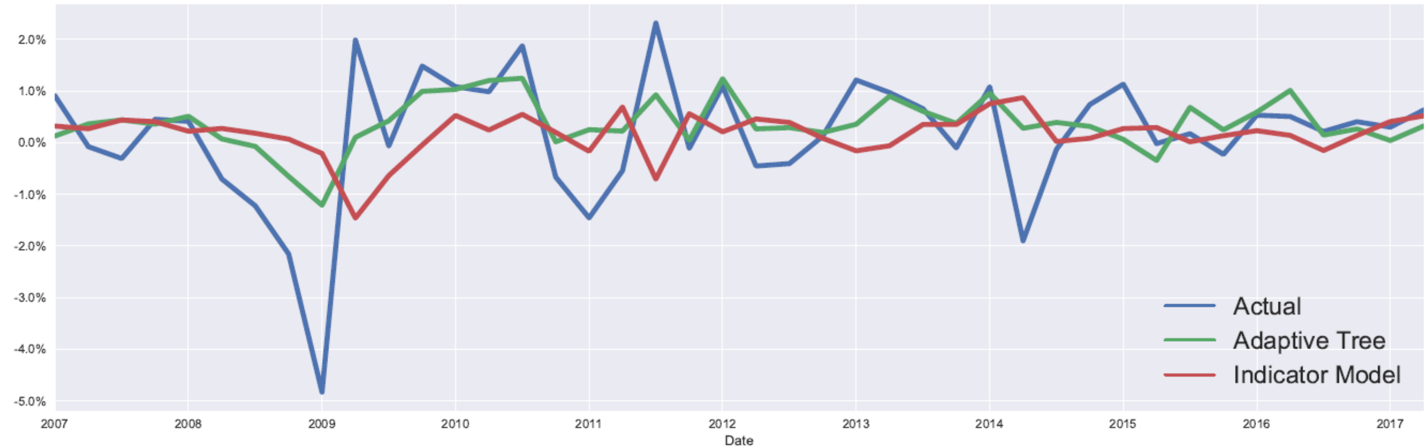
Accuracy: +9 %
Dir. Accucacy: +32 %



Comparison with OECD Indicator Model

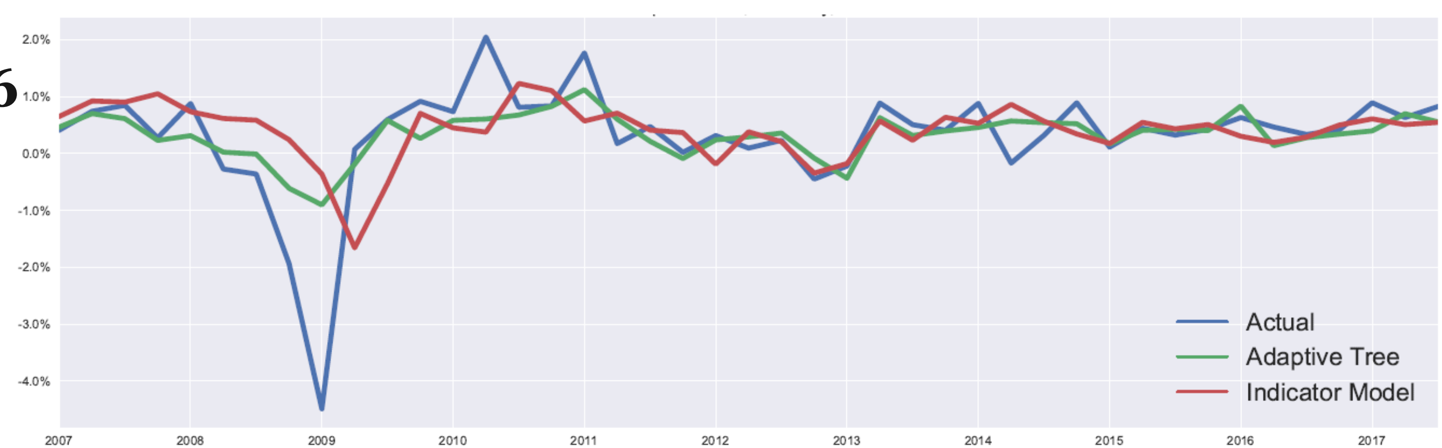
3. Japan, M+6

Accuracy: + 29 %
Dir. Accuracy: + 42 %



4. Germany, M+6

Accuracy: + 25 %
Dir. Accuracy: + 18 %



Overall improvement from Indicator Model

G6	RMSE	FDA
M+3	12%	27%
M+6	23%	32%

Adaptive Tree forecast consistently has better accuracy, and much better directional accuracy than the Indicator Model, while using the exact same data.

IV. CONCLUSION

Economics & machine learning

- Great tool to explore the complexity of the economy
- Performance: at M+6, Adaptive Trees are 23% more accurate and 32% more directionally accurate than the **Indicator Models**, using the exact same data
- Numerous possible extensions using broader set of variables

THANK YOU

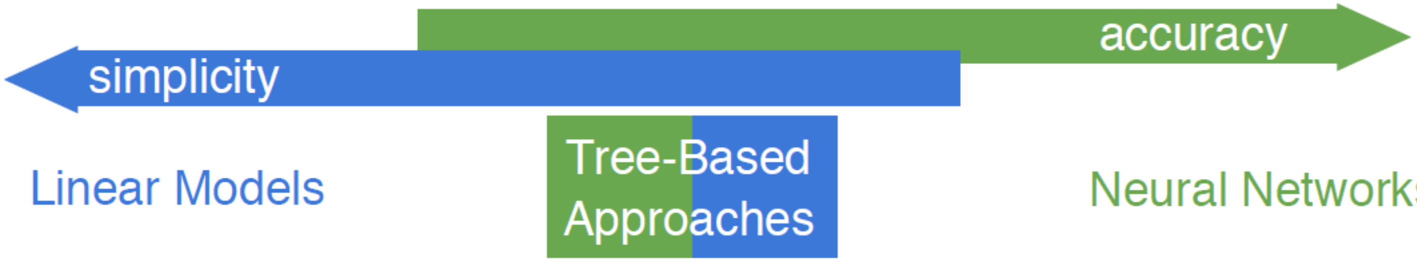
Questions ?

BETTER POLICIES FOR BETTER LIVES

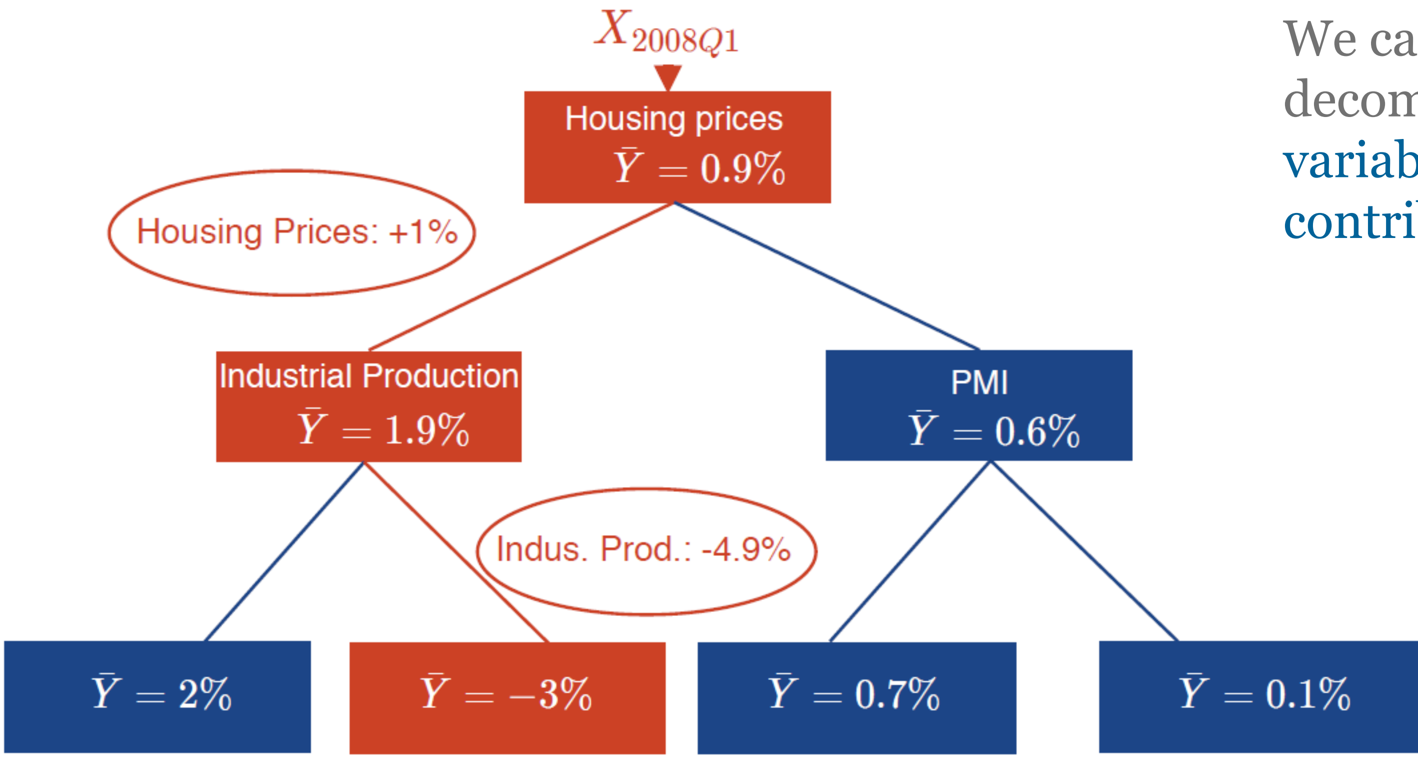
ADDITIONAL MATERIAL

Problem: interpretability

- Modelling complexity requires more complex models
- Trade off simplicity/accuracy:
  - Too much simplicity: fail to capture important variations
  - Too much complexity: fail to produce a sensible story



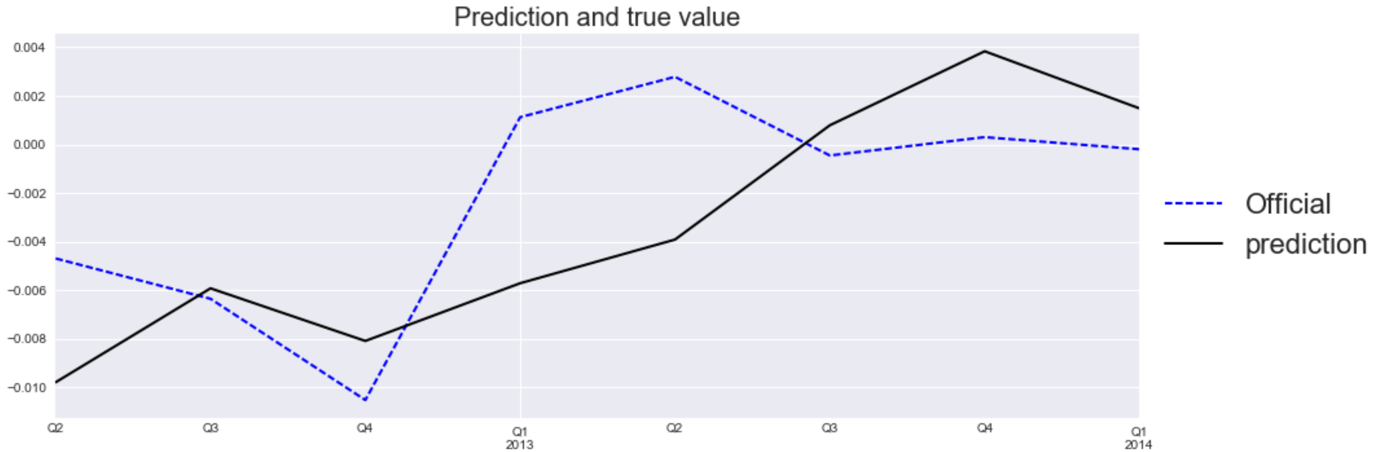
Interpretability

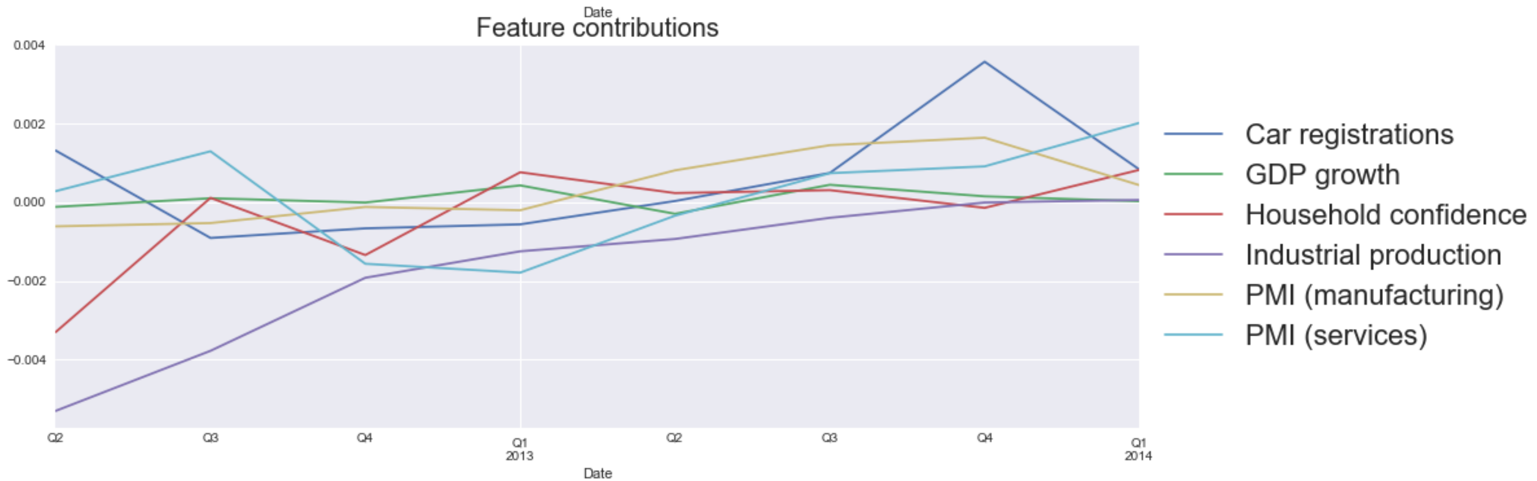


We can easily decompose in variable's contribution

$$\hat{Y} = 0.9\% + \sum Feature\ Contributions$$

Variable contributions, Italy M+3





Variable selection

- For each variable:
  - What relevant lag : M-1, M-2, M-12, M-24 ?
  - In level ? In growth rate ?

- Data-driven variable selection:
  - Based on variable importance
  - Variable importance: a variable is all the more important that it is **high in the tree**, close to the root
  - Accounts for multiple interactions (can keep a variable that is loosely correlated with the GDP but that provides relevant interactions. Ex: price of gold)

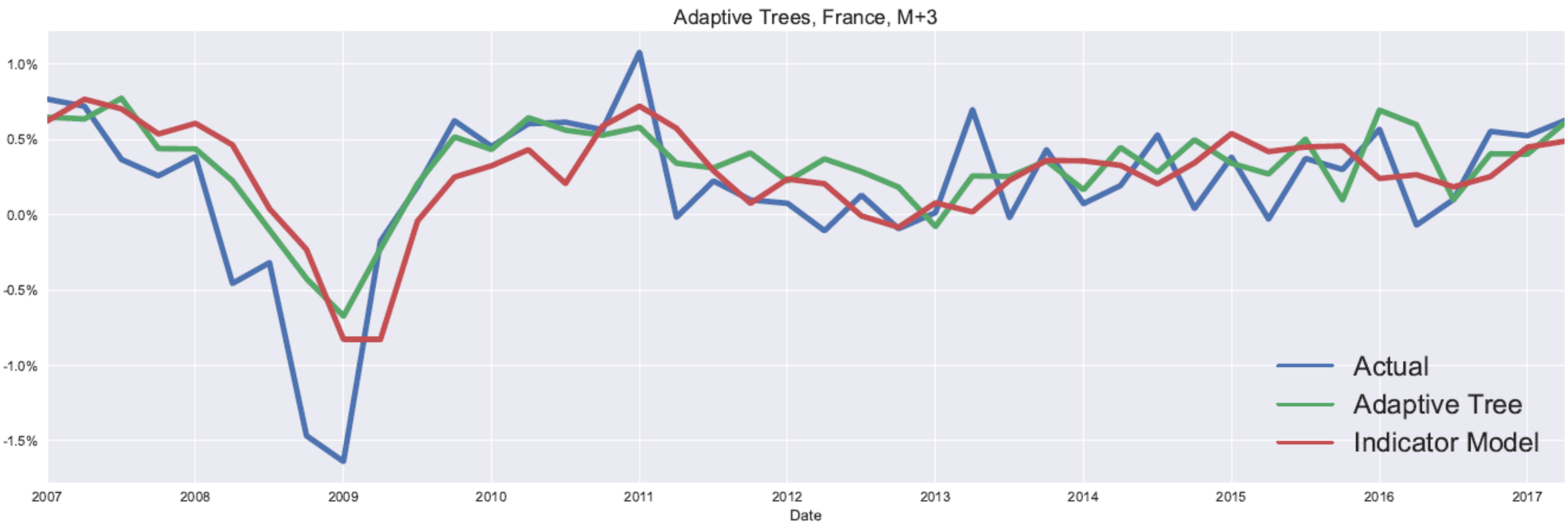
Complexity v. Bayesian econometrics

- In a regression with 10 variables, should we want to test all possible multiple interactions : 10^{10} possibilities

- With tree-based approaches, we explore all possible interactions with 120 variables

- Amount of prior knowledge:
  - Linear econometrics: we know the form of the relationship
  - Bayesian econometrics: we know the relationship can take any of the know forms
  - Machine learning: we know nothing

France, M+3



Adaptive Trees, France, M+3
1.0%
0.5%
0.0%
-0.5%
-1.0%
-1.5%
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
Date
Actual
Adaptive Tree
Indicator Model