



论文编码：

RUC-BK-200600-071601-201120xxxx

中国人民大学本科学士毕业论文

美联储利率政策变化预测

——GUIDE 决策树算法应用

作者：柳在唯

院系：统计学院

专业：统计学

年级：2011 级

学号：

指导教师：

成绩：

日期：2015 年 4 月



内容提要

自美联储退出量化宽松以来，其可能的利率政策变化一直是人们议论的热点，“提前加息”一度是市场对美联储利率方针的普遍预期。然而，2015 年一季度的经济数据显示，美国的经济环境尚存有较多不确定性，各方对加息与否的预期也因此产生了较大的分歧。有关这一话题，本文从泰勒规则出发，应用决策树的 GUIDE 算法，着重回答了美联储会否加息，以及会在何时执行操作这两个问题。具体而言，本文在泰勒规则所涉及的变量之基础上，首先应用 GUIDE 分类树算法，以分析美联储会否在接下来的利率操作中升息；此后，应用 GUIDE 回归树及 GUIDE 集成算法，对美联储下一步利率操作的时点进行预测。根据决策树的预测结果，在当前的经济环境下，美联储确将升息；进一步地，预期升息时间很可能将出现在 2015 年 9 月至 2015 年 11 月之间。



Abstract

The possible change in interest rate policy of the FED has been a hot topic since the FED quitted the QE: once there was a time when the whole market was expecting that the FED would raise the target interest rate in advance. However, as the economic condition of United States in the first quarter of 2015 becomes quite uncertain these days, various opinions occur regarding to the interest rate policy of FED. For this topic, this paper focuses on answering the questions about whether and when the FED will raise the Federal Funds Target Rate. The paper is mainly built on the GUIDE algorithm and supported by the Taylor's rule. In details, the problem is decomposed into two parts. First, the GUIDE classification tree algorithm is used to analyze if raising the target rate would be the next operation of the FED. Then, the paper applies the regression tree and the integrate algorithms in GUIDE to predict when the FED will execute the next operation. The results suggest that the FED should raise the target rate at the next step; and the possible execution date might be between September and November of this year.

目录

Abstract	3
目录	4
关键词	5
一 引言	5
二 泰勒规则简述	5
(一) 泰勒规则	6
(二) 泰勒型规则	6
三 GUIDE 决策树简述	7
(一) 传统经济理论的不足	7
(二) 决策树的 GUIDE 算法	8
1. GUIDE 算法简述	8
2. 算法的伪代码表示	9
四 研究过程	9
(一) 研究框架	9
(二) 变量处理	10
(三) 引入滞后项	12
(四) 会否加息——GUIDE 分类树的应用	12
1. 分类树的训练	12
2. 分类树的预测	14
3. 分类树的解释	16
(五) 何时操作——GUIDE 回归树的应用	17
1. 回归树的训练	17
2. 回归树的预测	18
五 结语	19
作者签名	20
附录	21
参考文献	27

关键词

GUIDE 算法 决策树 利率政策 美联储 泰勒规则

一 引言

次贷危机以来，美联储 0-0.25% 的基准利率（联邦基金利率）区间已维持至今。2014 年底，随着美国经济预期逐渐复苏，美国宣布退出持续四轮的量化宽松政策，市场关于美联储开启加息的议论一时间甚嚣尘上。然而，2015 年一季度的数据显示，美国经济在当季的表现并不如预期般令人满意，这使得市场对加息的预期有所回落。

在 2015 年 3 月的 FOMC 会议期间，美联储官员多次对此做出回应，普遍仍表示出较为积极的加息观点。但与此同时，各研究机构对于加息所持的观点已日趋谨慎，摩根士丹利甚至在近日称^①，美联储下一步的利率调整“不排除降息的可能”。

美国作为世界最大的经济体国家，其利率政策的变化对各大经济体乃至全球的政治经济有着重大的影响。对于其利率政策可能的变动，各国政策制定者及投资者有必要理清局面、做出预期、提前应对。

在当前的经济形势下，美联储会否加息、何时操作？——对于这样的现实性问题，本文将从泰勒规则出发，应用决策树的 GUIDE 算法，尝试从统计学习的角度给出相应的解答。

二 泰勒规则简述

自上世纪八十年代美国政府通过预算案、着手削减财政赤字以来，政府财政政策的使用大幅受限。为此，美联储逐步放弃了以货币供应量为中介目标的货币政策，转而以利率调整作为宏观调控的主要手段。在相关的经济理论中，泰勒规则的提出是促使美联储做出

^① Benjamin Purvis. *Fed Rate Cut 'Not Impossible,' Morgan Stanley Fund Manager Says.*
<http://www.bloomberg.com/news/articles/2015-03-26/fed-rate-cut-not-impossible-morgan-stanley-fund-manager-says>

这一调整的关键。

(一) 泰勒规则

泰勒规则 (Taylor Rule) 是由经济学家 J. Taylor (1993) 根据多国经济数据总结出的货币政策规则。由于其形式简单、含义直观, 并且对历史数据有着较好的拟合结果, 因此被经济学界广泛采纳并加以研究。其基本形式可表达如下:

$$i_t = i^* + \alpha(\pi_t - \pi^*) + \beta y_t \quad (2.1.1)$$

其中, i_t 为短期实际利率, i^* 为长期均衡下的实际利率, π_t 为通货膨胀率, π^* 为目标通胀率, y_t 为产出缺口。在本文语境下, i_t 为美国联邦基金利率 (Federal Fund Rate , FFR); 长期均衡条件下, i^* 在美国联邦基金目标利率 (Federal Fund Target Rate , FFTR) 附近波动。因此, 若以 Δi_t 表示 FFTR 在简单泰勒规则下的最优变动, 则美联储利息政策的最优变化可表示为:

$$\Delta i_t = \alpha(\pi_t - \pi^*) + \beta y_t \quad (2.1.2)$$

其中, α 、 β 的符号为正。

(二) 泰勒型规则

在泰勒规则提出后的二十余年间, 经济学界从各个方面对其进行了大量实证性与探索性的研究。根据研究综述^①及相关文献^②, 经济学家们主要从以下几方面做出了修正与拓展:

首先, Clarida、Gali & Gertler (2000) 在模型中引入了政策前瞻性因素, 使 i_t 成为预期通胀 $E_t \pi_{t+h}$ 及预期产出缺口 $E_t y_{t+h}$ 的函数, 从而能够“更真实地反应美联储的实际决策过程”。

其次, 美联储对基准利率的实际操作往往会出现滞后, 这与简单泰勒规则的即时性相违。对此, Orphanides (2001) 在泰勒规则中引入了利率平滑 (interest smoothing), 使 i_t 成为前期利率 i_{t-h} 的函数; 平行地, Rudebusch (2002) 引入了持续性货币政策冲击 (persistent

^① 谭小芬, 陈昌兵. 泰勒规则研究综述[J]. 经济学动态, 2006 (4): 105-110.

^② Coibion O, Gorodnichenko Y. Why are target interest rate changes so persistent?[R]. National Bureau of Economic Research, 2011.

monetary policy shocks) 的概念，运用工具变量对非货币政策对基准利率的影响进行了刻画。

此外，潜在产出往往具有较大的测量误差，造成对产出缺口的估计不准。对此，学界一般采用产出增长率、失业率缺口等指标替代产出缺口。研究发现，经过这些处理的模型对数据的拟合要优于简单泰勒规则。

再者，Siklos、Werner & Bohl (2004) 及 Rudebusch (2006) 等学者在论文中指出，货币当局在制定政策时，除了考虑产出缺口及通胀率缺口之外，理应也会参考资产价格、货币增长率、实际汇率、市场不确定性等方面的信息，有必要将这些因素也一同纳入泰勒规则的考量范围。

在后续研究中，一些学者开始探讨货币政策的非对称性，引入平滑转换模型等非线性时间序列模型，提出了非线性泰勒规则。Petersen (2007) 的研究表明，美联储在 1985 年之后的利率政策操作是与非线性泰勒规则相符合的。

一般地，统称上述经过扩展的模型为**泰勒型规则** (Taylor-type Rule)。

三 GUIDE 决策树简述

(一) 传统经济理论的不足

泰勒型规则对一国基准利率的政策变化做出了合理的解释，其诞生与发展的过程经过了严密的宏观经济理论与计量经济学实证的论证，经验上可以用来预测美联储下一步利率操作的方向，能够回答本文提出的**会否加息**的问题。但对于美联储**具体何时操作**，传统的宏观经济学理论却无法给出一个直接的答复。

将当前时点距离下一次利率政策变动的时间 T_N 考虑为一个随机变量，可以看到 T_N 并未出现在泰勒型规则中。而在美联储何时做出利率调整的问题里， T_N 则是以内生变量的身份作为响应变量出现的。事实上，二者所关注的问题并不相同，因此必须使用不同的模型进行拟合。

目前，经济学界尚无关于加息时点的理论探讨，在这种情况下，从数据本身出发的机器学习是值得尝试的办法。在接下来的篇幅中，本文将应用决策树的 GUIDE 算法构造模型，对美联储下次利率操作的方向及时间做出预测，并加以解释。

（二）决策树的 GUIDE 算法

1. GUIDE 算法简述

决策树的思想最早是由 Morgan & Sonquist 于 1963 年提出的，但在最初并未引起重视。直到 1984 年 CART 算法 (Breiman *et al.*) 的提出，决策树才真正进入了人们的视野。在此之后，由于其决策过程直观、预测效果优异，决策树逐渐成为了机器学习领域的核心算法之一。

GUIDE (*Generalized, Unbiased, Interaction Detection and Estimation*) 是由 Wei-Yin Loh (2009) 提出的一种决策树算法。针对当前流行的 CART、C4.5 等算法中存在的不足，GUIDE 做出了相应的改进：

首先，在树的生成过程中，使用穷举搜索法的 CART、C4.5 等算法在选择分割变量时存在偏差。具体而言，定义 m 为变量的水平数，则在分割节点的过程中，穷举搜索法倾向于选择 m 较大的变量进行分割。

对此，GUIDE 在节点分割时使用了结构更为复杂的两步法，首先检验各个自变量与响应变量之前的独立性，其后再根据检验结果选择变量、进行分割。

其次，在数据存在缺失值问题时，CART 等算法使用的替代分裂点方法存在着一定的疏漏。以分类树为例，CART 等算法在节点分割时只考虑了非缺失数据的类别比例，这使得算法倾向于选择具有较多缺失值的变量进行分割^①。

GUIDE 在处理缺失值时的方法如下：对于响应变量 Y ，其观测值 y_i 为缺失值的样本点不会被用于树的生成过程；对于分类变量 X ，定义单独的“缺失分类”；对于非分类变量 X ，定义“缺失区间”。

^① 例如，当某变量 X 只具有两个观测 $(x_1, y_1) = (5, 0)$ 、 $(x_2, y_2) = (10, 1)$ 时，选择 $\{X = 7.5\}$ 作为分割可以使得当前节点的纯度最高，CART 会以其作为当前节点的分割方式。但是显然，选择只有两个观测的变量进行分割是不合适的。

由此，GUIDE 算法保证了节点分割过程的无偏性，同时增进了缺失值处理的稳健性。经验表明，GUIDE 算法较之 CART、C4.5 等一些决策树算法具有更高的预测精度。

此外，在改进上述问题的基础上，GUIDE 还具有如下的优良特性：

首先，在使用两步法进行节点分割的过程中，GUIDE 引入了对变量间交互作用的探测，这在极大程度上增加了 GUIDE 对问题的解释能力。其次，GUIDE 提供了诸多模型与选项以供选择和调整，具有较强的适用性。此外，GUIDE 还引入了 Bagging 和 Forests 两种集成方法，满足了许多问题中对预测精度的需要。

2. 算法的伪代码表示

为使表述清晰，给出 GUIDE 算法的树生成过程的伪代码如下：

1. 从根节点开始。
2. 对每个变量 X ，如果 X 是有序变量，通过划分区间的方法将其转化为无序变量（分类变量） X^c ；如果 X 是无序变量，定义 $X^0 = X$ 。
3. 对每个 X^c ，在当前节点下进行关于 Y 的 χ^2 独立性检验，使用 Wilson-Hilferty 近似求解 P 值。
4. 选择 P 值最小的 X^c 所对应的 X 作为分割变量，记为 X_{\setminus} 。
5. 记 $f_{X_{\setminus}} \cup S_g$ 为所有基于 X_{\setminus} 的分割所构成的集合，找到使得基尼不纯度最小的分割 $f_{X_{\setminus}} \cup S_g$ ，以之将当前节点划分为两个子节点。
6. 如果满足了停止准则，则停止；反之，对每个子节点应用步骤 2 至 5。
7. 以 CART 方法进行剪枝。

四 研究过程

（一）研究框架

设美联储于时刻 t 后的下次利率政策变化方向为分类变量 I_t ，时刻 t 距离下次利率政策变化的天数为顺序变量 TN_t ，则本文关注之问题的研究框架可以表述如下：

1. 从泰勒规则中的自变量出发，结合其余常用的经济变量，经过一定的变换处理，得到模型的解释变量 X_t ^①。

记当前时刻为 T ，美联储迄今为止最后一次利率变动的时刻为 S ， $S \leq T$ 。

设存在解释变量 X_t 及其余未知因素 ϵ_t 的函数 g ，使得：

$$\begin{aligned}
 & I_t = g(X_t; \epsilon_t); \quad \forall t \\
 \text{where } I_t = & \begin{cases} 1 & \text{if the 1st change after } t \text{ is raising} \\ -1 & \text{if the 1st change after } t \text{ is cutting} \end{cases}
 \end{aligned} \quad (4.1.1)$$

2. 基于已知信息 $(X_k; I_k)_{k=1}^{S-1}$ ，应用 GUIDE 分类树算法，得到对函数 g 的估计 \hat{g} 。通过分析树的结构，对其经济意义做出解释。

3. 根据当前时刻的已知信息 X_T 及分类树 \hat{g} ，对当前时刻后下一次利率调整的方向 I_T 做出预测，记为 I_T^\wedge ，即：

$$I_T^\wedge = \hat{g}(X_T) \quad (4.1.2)$$

4. 设存在解释变量 X_t 、 I_t 及 ϵ_t 的函数 h ，使得：

$$\begin{aligned}
 & TN_t = h(X_t; I_t; \epsilon_t); \quad \forall t \\
 \text{where } TN_t = & 1; 2; \dots
 \end{aligned} \quad (4.1.3)$$

5. 基于已知信息 $(X_k; I_k)_{k=1}^{S-1}$ ，应用 GUIDE 回归树算法，得到对函数 h 的估计 \hat{h} 。通过分析树的结构，对其经济意义做出解释。

6. 根据当前时刻的已知信息 X_T 、式(4.1.2)的预测值 I_T^\wedge 及回归树 \hat{h} ，对当前时刻距离下次变息的天数做出预测，记为 TN_T^\wedge ，即：

$$TN_T^\wedge = \hat{h}(X_T; I_T^\wedge) \quad (4.1.4)$$

(二) 变量处理

综合考虑泰勒型规则简述中模型部分所涉及的变量，以及宏观经济分析中常用的经济指标，本文将使用下列变量作为问题的解释变量：

^① 在时刻 t 时可以获得的经济信息。

TL_t :	t 时刻距离上一次利率调整的时间 (天)
i_t^j :	联邦基金实际利率相对 j 天前的变化率 (天)
U_t :	短期失业缺口 ^① (月)
y_t :	产出缺口 ^② (季度)
GDP_t :	实际 GDP 同比变化率 (月)
gdp_t :	实际 GDP 环比变化率 (月)
$Econ_t$:	经济政策不确定指数 (天)
$M1_t$:	货币供应量变化率 (周)
$M2_t$:	准货币供应量变化率 (周)
EX_t :	窄口径实际汇率变化 (月)
IPI_t :	工业生产指数变化率 (月)
PCE_t :	个人消费支出价格指数变化率 (月)
pce_t :	个人消费支出价格指数变化率与目标通胀率之差的示性变量 ^③ (月)
Oil_t :	原油价格指数相对 j 天前的变化率 (天)
SP_t :	标准普尔 500 指数相对 j 天前的变化率 (天)
Md_t^j :	穆迪 Baa 级债券平均利率与联邦基金实际利率之差相对 j 天前的变化率 (天)

根据 Petersen (2007) 的研究, 本文选取 1985 年后的数据进行分析, 以尽可能保证美联储利率政策在训练集上的稳定。数据来源为圣路易斯联储 FRED 数据库, 观测值序列时间间隔为天。由于多数宏观经济变量的观测间隔为月度或季度, 或有部分变量在报告期尚未得到近期观测, 在此使用末次观测值结转法进行处理, 对当月或当季各天赋以相同取值; 对于原油价格指数等序列观测中出现的删失, 使用样条插值法进行处理。特别地, 对于 2015 年一季度的 GDP 数据, 使用亚特兰大联储的 GDPNow 模型给出的预测^④以减少误差。

^① 短期失业缺口 = 实际失业率 - 短期自然失业率

^② 产出缺口 = (实际产出 - 潜在产出) / 潜在产出

^③ $pce_t = \begin{cases} 1 & PCE_t \geq \hat{\gamma}_t \\ 0 & PCE_t < \hat{\gamma}_t \end{cases}$; where $\hat{\gamma}_t$ is assumed to be 2.

^④ 一季度年增长率 0.2%, 3 月 26 日。

(三) 引入滞后项

经济时间序列一般存在着较强的序列相关性，为此，考虑在模型中引入滞后项。Chongqing (1995) 在论文中指出，引入了滞后项的决策树模型可以被视为门限自回归模型 (TAR) 的拓展^①。在此之上，利用 GUIDE 决策树对变量间交互作用的探测，可以得到结构较为复杂的非线性模型，并且免除传统时间序列方法对整体序列平稳性的假设要求。

对于时间序列变量 X_t ，分别记 i 天、周、月、季滞后项为 X_{t-d}^i 、 X_{t-w}^i 、 X_{t-m}^i 、 X_{t-q}^i 。经过讨论与在测试集上的验证，最终解释变量集合如下：

$$\begin{aligned}
 X_t = & (TL_t; i_t^1; i_t^2; i_t^3; i_t^7; i_t^{30}; U_t; U_{t-q}^1; U_{t-q}^2; U_{t-q}^3; \\
 & y_t; y_{t-q}^1; y_{t-q}^2; y_{t-q}^3; GDP_t; GDP_{t-q}^1; gdp_t; gdp_{t-q}^1; \\
 & Econ_t; Econ_{t-w}^1; Econ_{t-w}^2; Econ_{t-w}^3; \\
 & M1_t; M1_{t-w}^1; M1_{t-w}^2; M1_{t-w}^3; M2_t; M2_{t-w}^1; M2_{t-w}^2; M2_{t-w}^3; \\
 & EX_t; EX_{t-m}^1; EX_{t-m}^2; EX_{t-m}^3; \\
 & IPI_t; IPI_{t-m}^1; IPI_{t-m}^2; IPI_{t-m}^3; \\
 & PCE_t; PCE_{t-m}^1; PCE_{t-m}^2; PCE_{t-m}^3; poe_t; \\
 & Oil_t; Oil_t^1; Oil_t^2; Oil_t^3; Oil_t^7; Oil_t^{30}; \\
 & SP_t; SP_t^1; SP_t^2; SP_t^3; SP_t^7; SP_t^{30}; \\
 & Md_t; Md_t^1; Md_t^2; Md_t^3; Md_t^7; Md_t^{30})
 \end{aligned} \tag{4.3.1}$$

(四) 会否加息——GUIDE 分类树的应用

在上述变量及数据的基础上，应用 GUIDED 分类树算法，以对美联储是否会在下一步的利率操作中升息进行预测、做出解释。

1. 分类树的训练

在 1985 年至 2015 年期间，美联储共计升息 63 次，降息 65 次，在次数上基本均衡。但从幅度上看，经历了九十年代初、二十一世纪初及次贷危机的三次大规模降息，目标利率在三十年间从 8.125% 降至 0.125%^②，并已维持零利率水平五年之久。其时间序列的折线

^① 首先，决策树模型允许使用不同变量的不同滞后项作为门限，这使得模型结构更为精巧；其次，决策树算法的节点分割过程使得门限值的确定更为系统化。

^② 在 2008 年 12 月 17 日后，美联储改以区间形式提供基准利率，在此取区间均值作为基准利率点。

图如下：

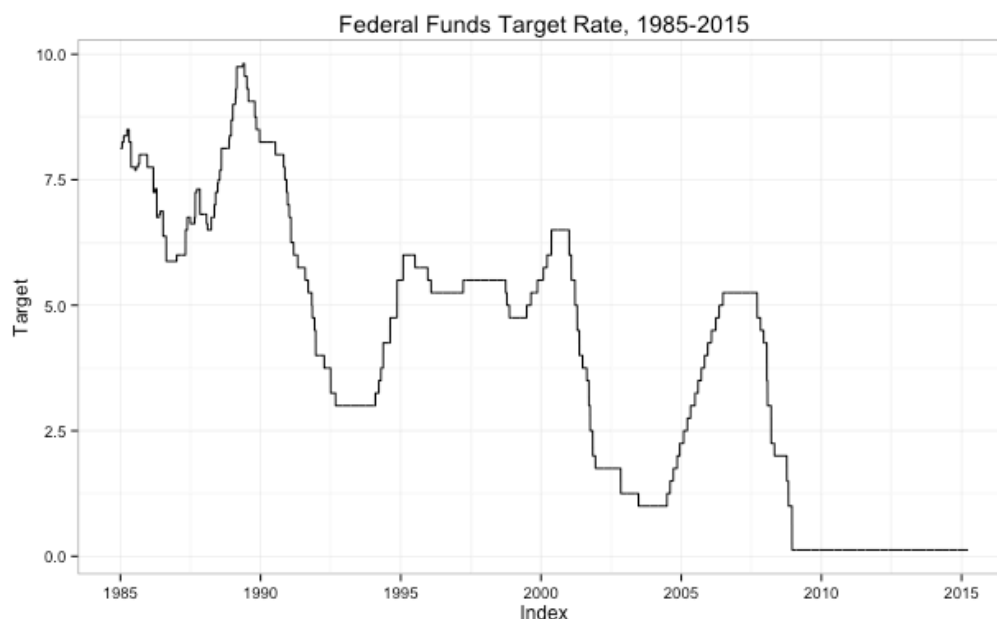


图 1 联邦基金目标利率时间序列折线图，1985-2015

为了对当前时刻（2015 年 3 月 26 日）后的下一次利率调整方向 I_T 做出预测，现利用 1985 年 1 月 1 日至 2008 年 12 月 16 日的已知样本点训练 GUIDE 分类树，其细节设定如下：

首先，由于升息、降息次数基本均衡，且二者对经济环境的重要性不分伯仲，故而假设两类别先验概率及误分成本相等；其次，在树的节点分割过程中，由于使用单变量分割得到的树结构过于简单，为了提高分类精度，转而使用线性分割，以引入变量间的交互作用；此外，为了使剪枝过程更为稳定，使用基于中位数的交叉检验作为剪枝标准；最后，为了保留模型的直观性与可解释性，暂不使用集成方法进行预测。

经此设定，得到的分类树模型 g 的树状结构如图 2 所示：

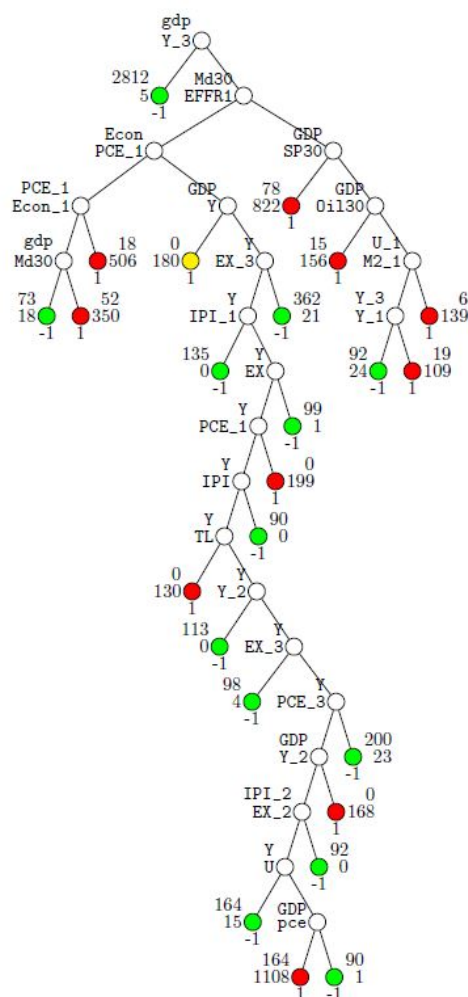


图 2 GUIDE 分类树 \hat{g}

其中，各分支节点（以空白表示）实际上皆是由两个变量构成的线性分类器^①；在各分支节点处，满足该节点线性分类器条件的数据进入左侧子节点，不满足的则进入右侧子节点。而在各叶节点处，绿色(1)表示降息，红色(1)表示升息；叶节点旁的两行数字分别表示该节点内 $I_t = 0$ 及 $I_t = 1$ 的样本个数。

根据拟合得到的分类树 \hat{g} ，GUIDE 可以自动生成用以预测的 R 代码（见附录），以便进行处理。根据代码的输出结果，分类树对 I_T 的最终预测为 $I_T^{\wedge} = 1$ ，即预测美联储下一次的利率操作为升息。预测所对应的叶节点在图 2 中以黄色表示。

2. 分类树的预测

图 3 给出了分类树 \hat{g} 做出此次预测的路径。可以发现，在得到结论前，算法共在四个

^① GUIDE 将分割变量名在节点左侧做出了标注，但省略了具体的线性分类条件。

分支节点处进行了四次线性判别：

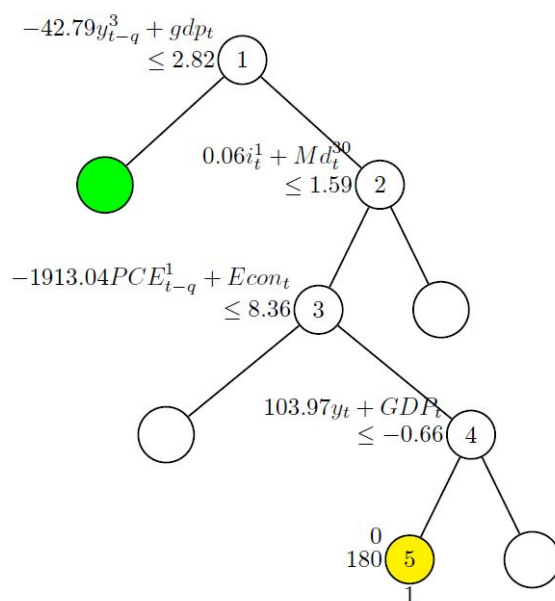


图 3 分类树 g 的细节部分

在分类树的分支节点 1 处，有：

$$-42.79 \lceil y_{t-q}^3 + gdp_t > 2.82$$

其中，三季度前的产出缺口之系数为负，当期 GDP 环比变化率之系数为正，不等式方向为大于。在此，分类树考虑了当前经济形势与早期经济形势的线性组合，直观而言，即在早期经济形势欠佳（负向产出缺口较大）时，若当期的 GDP 增长较多，则存在加息可能，但为做出决策，仍需进入下一节点进行判断。

在分支节点 2 处，有：

$$0.06 \lceil i_t^1 + Md_t^{30} \leq 1.59 \quad (4.4.1)$$

其中，当期联邦基金实际利率变化率之系数为正，穆迪 Baa 级债券平均利率与联邦基金利率之差相对一个月前的变化率之系数亦为正，不等式方向为小于。在此，分类树所考虑的因素是利率市场的变化。若当期联邦基金实际利率处于较低水平，而债券市场利率与实际利率间的利差较前期缩小，则存在加息可能，但仍需进入分支节点 3 进行判断。

在分支节点 3 处，有：

$$-1913.04 \lceil PCE_{t-q}^1 + Econ_t > 8.36 \quad (4.4.2)$$

其中，前期的个人消费支出价格指数变化率之系数为负，当期经济政策不确定性之系

数为正,不等式方向为大于。在此,分类树考虑了通胀水平及经济政策对利率政策的影响,即在前期通胀率下降,而当期经济政策不确定性较大时,存在加息可能。

最后,在分支节点 4 处,有:

$$103.97 \uparrow y_t + GDP_t \neq 0.66 \quad (4.4.3)$$

其中,当期的产出缺口之系数为正,当期 GDP 同比变化率之系数为正,不等式方向为小于。在这一节点处,考虑当期的经济形势,若满足相应条件,则做出判断,认为下一步利率变化为加息。

3. 分类树的解释

事实上,在本问题中,决策树学习的本质是对经济历史的学习,即找到历史数据中与当前经济环境相似的周期,并以当年的经验预测现在的趋势。在图 3 中,用于预测的 5 号叶节点的事实上主要是由 1993 年一季度及三季度的数据生成的。如图 4 中的标记所示,彼时美国经济开始从九十年代初的经济危机中回暖,利率维持在同一水平较长时间,此后开始回升。而分类树的结果告诉我们,这一情境是与当前的经济环境相符的。同时,加息的预测结果也符合目前市场的主流预期。

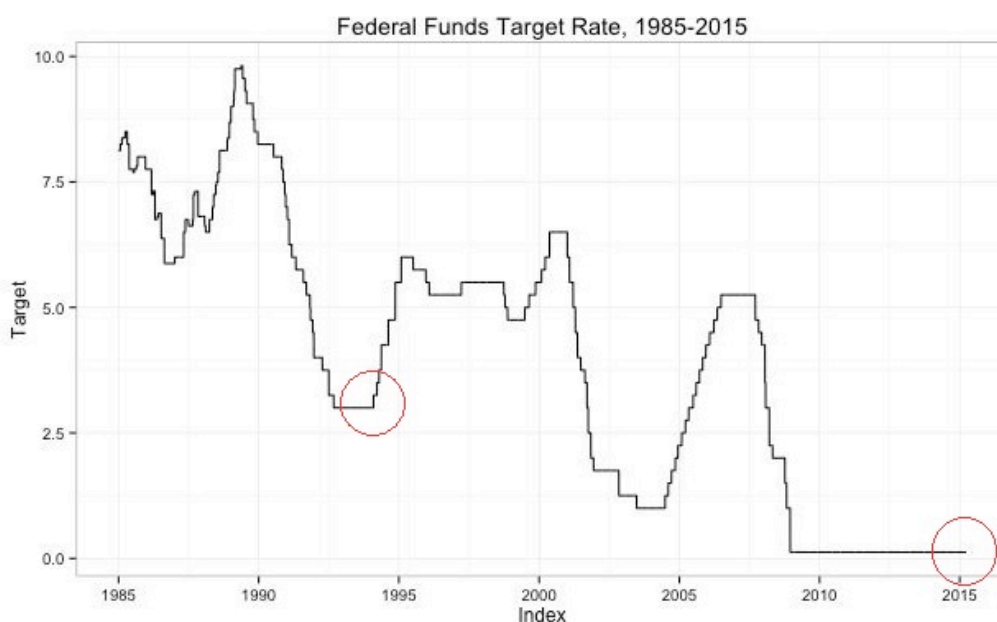


图 4 联邦基金目标利率时间序列折线图(标记), 1985-2015

观察图 2 中分类树的完整结构，可以看到其使用的变量与泰勒型规则所涉及的基本一致。值得指出的是，在图 3 中的分支节点 3 与 4 处，变量线性组合的经济意义与泰勒规则略有出入。其可能的原因之一，是美联储在进行操作时考虑了除此之外的其它因素；而从模型的角度讲，分类树在单个分支节点处的分类相当于一个条件概率，若某个节点的表现与理论方向相反，有可能是为了对此前的节点施加的过强的“条件”做出修正。事实上，如此反复修正的过程也正是树结构模型能够较处理好非线性问题的原因。

（五）何时操作——GUIDE 回归树的应用

在得到加息的分类预测结果之后，应用 GUIDED 回归树算法，以对距离下次利率政策变化的天数做出预测。

1. 回归树的训练

对于 1985 年初至 2008 年底的数据，将各时刻距离下一次利率政策变化的天数 TN 视为随机变量，可以得到其样本分布直方图如下所示：

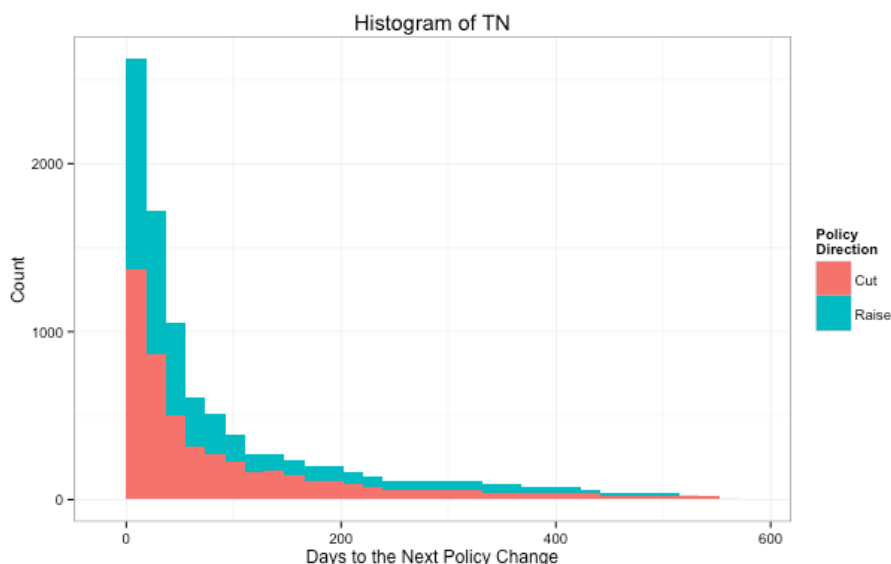


图 5 TN 分布直方图

为了对当前时刻距离下一次利率政策变化的时间 TN_T 做出预测，在此尝试训练 GUIDE 回归树，其细节设定如下：

首先，由于样本分布具有较明显的左截尾和右偏尾特征，故而使用中位数回归作为节

点处的回归模型，以增加稳健性；其次，考虑各中位数回归方程的结构，鉴于多元线性及最优多项式方程的计算复杂度过高，因此选择只含常数项的中位数回归方程，以简化模型；此外，在使用中位数交叉检验进行剪枝时，由于默认设置带来的模型过于复杂，故而对阈值进行了调整。

经此简化设定，得到的回归树模型 \hat{h} 的树状结构如下图所示：

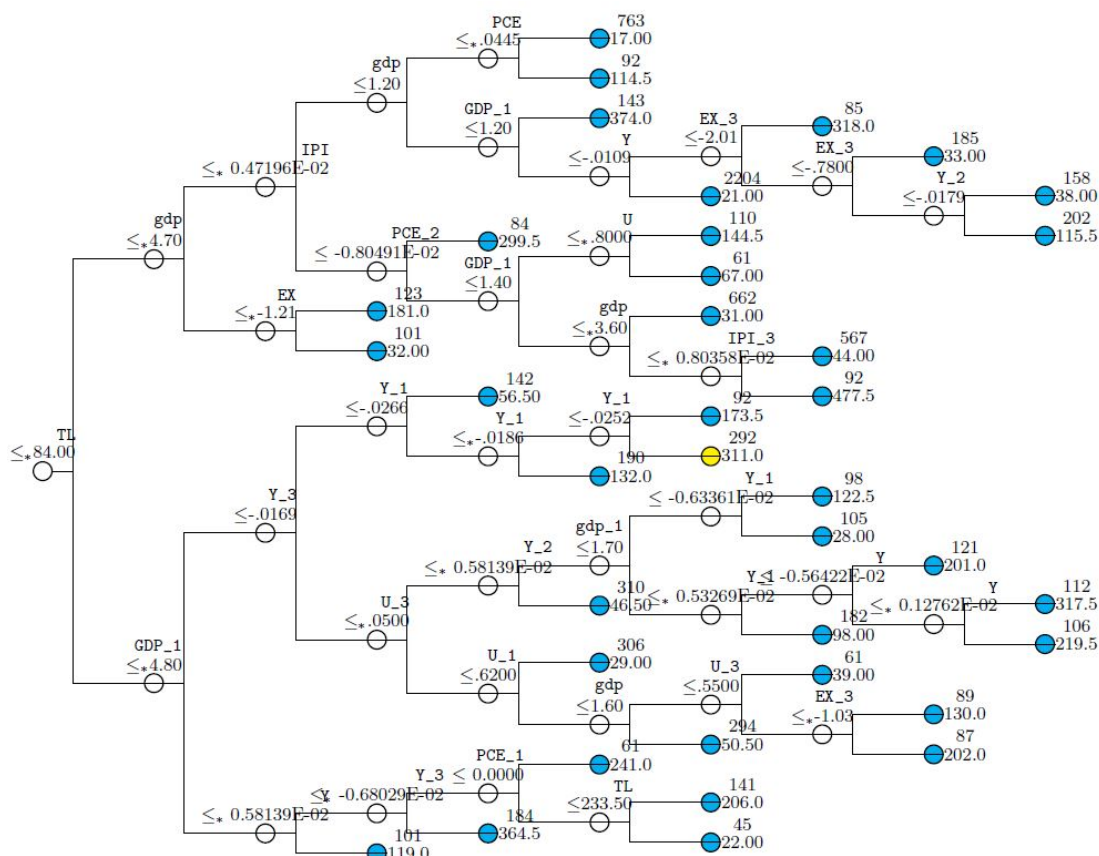


图 6 GUIDE 回归树 \hat{h}

其中，在分支节点处，满足该节点处变量条件的数据进入上侧子节点，不满足的则进入下侧子节点。预测所对应的叶节点以黄色表示，其余叶节点以蓝色表示。

2. 回归树的预测

将3月26日当期数据代入回归树 \hat{h} ，得到的分段常量中位数回归预测结果为 $T\hat{N}_T = 311$ ，即在当前的经济形势下，认为美联储在2015年内不会加息。然而，由于在训练回归树 \hat{h} 的过程中，为降低计算的复杂度，仅使用了只含常数项的回归模型，且人为调整了剪枝的阈值，因此，回归树 \hat{h} 的预测准确性在此处会受到较大影响。相应地，根据 \hat{h} 做出的经济意义

分析也难以具备健全的解释效力。

鉴于现有问题并无经济理论相依托，为了得到更为精确的预测结果，权衡决定使用 GUIDE 集成算法对 TN_T 进行预测，相应地割弃对经济意义的解释。结果显示，GUIDE Forest 的预测为 $T\hat{N}_T = 212$ ，GUIDE Bagging 的预测为 $T\hat{N}_T = 218$ ，即认为美联储将在 10 月末升息。

最后，应用 Bootstrap 方法，给出预测值的置信区间如下：

GUIDE Forest 预测值的 95%置信区间：

$$[178:6121; 244:9775] \quad (4.5.1)$$

GUIDE Bagging 预测值的 95%置信区间：

$$[191:3188; 244:4484] \quad (4.5.2)$$

为提高置信系数，在实际预测时对两个置信区间取并集，最终得到的预测结果为：预测美联储将在 9 月 20 日至 11 月 25 日间加息。

五 结语

经过统计学习与预测，本文的结论认为，在当前的经济环境下，美联储很有可能在下一步的利率操作中提高联邦基金目标利率。进一步地，在预期美联储即将升息的条件下，根据区间估计的结果，本文认为美联储执行升息操作的时点有较大可能出现在 2015 年 9 月至 11 月之间。

从经济意义出发，上述结论与泰勒规则的结果是相互印证的；从市场预期出发，目前市场的主要争议在于美联储是否将在六月提前加息，而本文预测的加息时点则为九至十一月，可见文中使用的模型在预测上是较为保守的。

就模型本身而言，GUIDE 决策树的结构可以较好地拟合历史数据中的趋势，如果认定美联储在 1985 年后的利率政策制定方针较为稳定，则 GUIDE 决策树的预测结果是值得信赖的。然而，考虑到宏观经济数据的时间跨度较长，可能涉及数据演化的问题，如果认为



美联储在近年的利率方针发生了变化，则将训练集视为静态数据的 GUIDED 决策树可能会对历史数据出现一定的过拟合，从而带来预测上的失准。由于 Petersen (2007) 的研究显示美联储在 1985 年后的利率政策操作较为一致，故而可以认为在此不存在数据演化现象；若不然，则需考虑应用演化数据的学习方法对问题进行分析。

至此，本文回答了关于美联储会否加息、何时操作的问题，其结论可作为相关宏观经济分析的参考。此外，亦可考虑应用本文的研究方法于我国的基准利率历史数据，以丰富本文的现实意义。

作者签名



附录

1. 相关链接

GUIDE 的 Fortran 实现：	http://www.stat.wisc.edu/~loh/guide.html
基准利率：	https://research.stlouisfed.org/fred2/series/DFEDTAR
EFFR：	https://research.stlouisfed.org/fred2/series/DFF#
短期自然失业率：	http://research.stlouisfed.org/fred2/series/NROUST#
实际失业率：	https://research.stlouisfed.org/fred2/series/UNRATENSA
潜在产出：	http://research.stlouisfed.org/fred2/series/GDPPOT#
实际产出：	https://research.stlouisfed.org/fred2/series/GDPC1#
Moody：	http://research.stlouisfed.org/fred2/series/BAAFF#
窄口径实际汇率：	https://research.stlouisfed.org/fred2/series/RNUSBIS
经济政策不确定性：	https://research.stlouisfed.org/fred2/series/USEPUINDXD
M1：	https://research.stlouisfed.org/fred2/series/M1
M2：	https://research.stlouisfed.org/fred2/series/M2
PCE：	https://research.stlouisfed.org/fred2/series/PCEPI#
IPI：	https://research.stlouisfed.org/fred2/series/INDPRO#
Oil：	https://research.stlouisfed.org/fred2/series/DCOILWTICO#
S&P500：	https://finance.yahoo.com/q/hp?s=%5EGSPC+Historical+Prices

2. R 代码

```
#### Transformation Functions ####
# Get a zoo (after differentiation), x is coredata and y is the index
Diff <- function(x, y, z=1) {
  diff = x[-(1:z)] - x[-((length(x)-z+1):length(x))]
  Data = zoo(diff, y[-(1:z)])
  return(Data)
}

# Get a zoo (after calculating percent change), x is coredata and y is the index
PerC <- function(x, y, z=1) {
  perc = (x[-(1:z)] - x[-((length(x)-z+1):length(x))]) / x[-((length(x)-z+1):length(x))]
  Data = zoo(perc, y[-(1:z)])
  return(Data)
}

PerC2 <- function(x, y) {
  perc = (x[,2][-1] - x[,2][-nrow(x)]) / (x[,2][-nrow(x)])
  Data = data.frame(x[,1][-1], perc)
  colnames(Data) = c("Date", y)
  return(Data)
}

# Cut the data between two date, a & b are in the format "2014-01-01"
Cut <- function(x, a, b) {
  x[which(x[,1] == a):which(x[,1] == b),]
}

# Transform Continuous Data to Categorical
CTC <- function(x, y) {
```



```
z = x[,2]
z[which(z > 0)] = 1
z[which(z < 0)] = -1
Data = data.frame(x[,1], z)
colnames(Data) = c("Date", y)
return(Data)
}
CTC2 <- function(x, y, p=0.02) {
  z = x
  z[which((z > p) | (z == p))] = 1
  z[which(z < p)] = -1
  Data = zoo(z, y)
  return(Data)
}

# Zoo NA
LOCF <- function(x) {
  a <- zoo(x[,2:dim(x)[2]], as.Date(x[,1]))
  g <- seq(start(a), end(a), "day")
  h <- na.locf(a, xout=g)
  return(h)
}
SPLINE <- function(x) {
  a <- zoo(x[,2:dim(x)[2]], as.Date(x[,1]))
  g <- seq(start(a), end(a), "day")
  h <- na.spline(a, xout=g)
  return(h)
}
#####

#### Read in data ####
FFTR <- read.csv("/Users/home/Documents/THESIS/Data/Target.csv",
na.strings="#N/A", skip=10, header=T)
EFFR <- read.csv("/Users/home/Documents/THESIS/Data/EFFR.csv", na.strings="#N/A",
skip=10, header=T)
Un <- read.csv("/Users/home/Documents/THESIS/Data/Unrate.csv", skip=10, header=T)
UnN <- read.csv("/Users/home/Documents/THESIS/Data/UnrateNature.csv", skip=10, header=T)
GDP <- read.csv("/Users/home/Documents/THESIS/Data/GDP.csv", skip=10, header=T)
GDPP <- read.csv("/Users/home/Documents/THESIS/Data/GDPPotential.csv", skip=10,
header=T)
GDPgPre <- read.csv("/Users/home/Documents/THESIS/Data/GDPgPre.csv", skip=10, header=T)
GDPgLyr <- read.csv("/Users/home/Documents/THESIS/Data/GDPgL.csv", skip=10, header=T)
Md <- read.csv("/Users/home/Documents/THESIS/Data/Moody.csv", na.strings="#N/A", skip=10,
header=T)
EX <- read.csv("/Users/home/Documents/THESIS/Data/NEExchange.csv", skip=10, header=T)
Econ <- read.csv("/Users/home/Documents/THESIS/Data/Econ.csv", na.strings="#N/A",
skip=10, header=T)
M1 <- read.csv("/Users/home/Documents/THESIS/Data/M1.csv", skip=10, header=T)
M2 <- read.csv("/Users/home/Documents/THESIS/Data/M2.csv", skip=10, header=T)
PCE <- read.csv("/Users/home/Documents/THESIS/Data/PCEIndex.csv", skip=10, header=T)
IPI <- read.csv("/Users/home/Documents/THESIS/Data/IPI.csv", skip=10, header=T)
Oil <- read.csv("/Users/home/Documents/THESIS/Data/Oil.csv", na.strings="#N/A", skip=10,
header=T)
SP <- read.csv("/Users/home/Documents/THESIS/Data/SP500.csv", na.strings="#N/A", skip=10,
header=F)
colnames(SP) <- c("observation_date", "SP500")
require(zoo)
require(quantmod)
require(ggplot2)
#####

#### Data Wrangling ####
# Target
FFTR1 <- Cut(EFFR, "1982-09-27", "2015-03-26")
FFTR1[,2] <- 0.125
FFTR1[1:9576,2] <- FFTR[,2]
FFTR_plot <- FFTR1
FFTR1 <- PerC2(FFTR1, "I")
FFTR1 <- CTC(FFTR1, "I")
index <- which(FFTR1[,2] != 0)
TN <- vector()
TL <- vector()
for (i in 1:dim(FFTR1)[1]) {
  a <- index - i + 1
  TN[i] <- min(a[a >= 0])
  TL[i] <- max(a[a < 0])
}
FFTR2 <- data.frame(FFTR1, TN, TL = abs(TL))
FFTR3 <- FFTR2[,2]
FFTR3[FFTR3==0] <- NA
FFTR3 <- na.locf(FFTR3, fromLast=T, na.rm=F)
FFTR2[,2] <- FFTR3
FFTR4 <- LOCF(FFTR2)

# Predictors
```



```
EFFR1 <- SPLINE(EFFR)
EFFR_t1 <- PerC(coredata(EFFR1), index(EFFR1), z=1)
EFFR_t2 <- PerC(coredata(EFFR1), index(EFFR1), z=2)
EFFR_t3 <- PerC(coredata(EFFR1), index(EFFR1), z=3)
EFFR_t7 <- PerC(coredata(EFFR1), index(EFFR1), z=7)
EFFR_t30 <- PerC(coredata(EFFR1), index(EFFR1), z=30)

Un1 <- zoo(Un[,2], as.Date(Un[,1]))
UnN1 <- zoo(UnN[,2], as.Date(UnN[,1]))
g <- seq(start(UnN1), end(UnN1), "month")
UnN1 <- na.locf(UnN1, xout=g)
U1 <- merge(Un1, UnN1)
U1 <- U1$Un1 - U1$UnN1
U <- U1[which(index(U1)=="1955-01-01"):which(index(U1)=="2013-01-01"),]
U_t <- Lag(U, k=1:3)

GDP1 <- zoo(GDP[,2], as.Date(GDP[,1]))
GDPp1 <- zoo(GDPp[,2], as.Date(GDPp[,1]))
y <- (GDP1-GDPp1)/GDPp1
y_t <- Lag(y, k=1:3)

GDPgP <- zoo(GDPgPre[,2], as.Date(GDPgPre[,1]))
GDPgL <- zoo(GDPgLyr[,2], as.Date(GDPgLyr[,1]))
GDPgP_t <- Lag(GDPgP, k=1)
GDPgL_t <- Lag(GDPgL, k=1)

Econ1 <- zoo(Econ[,2], as.Date(Econ[,1]))
Econ_t <- Lag(Econ1, k=1:3)
ECON <- merge(EFFR1, Econ1, Econ_t)
ECON <- na.locf(ECON, fromLast=T)[,4:7]

M1_ <- zoo(M1[,2], as.Date(M1[,1]))
M2_ <- zoo(M2[,2], as.Date(M2[,1]))
M1_ <- PerC(coredata(M1_), index(M1_))
M2_ <- PerC(coredata(M2_), index(M2_))
M1_t <- Lag(M1_, k=1:3)
M2_t <- Lag(M2_, k=1:3)

EX1 <- zoo(EX[,2], as.Date(EX[,1]))
EX1 <- Diff(coredata(EX1), index(EX1))
EX_t <- Lag(EX1, k=1:3)

IPI1 <- zoo(IPI[,2], as.Date(IPI[,1]))
IPI1 <- PerC(coredata(IPI1), index(IPI1))
IPI_t <- Lag(IPI1, k=1:3)

PCE1 <- zoo(PCE[,2], as.Date(PCE[,1]))
PCE1 <- PerC(coredata(PCE1), index(PCE1)) * 12
PCE_t <- Lag(PCE1, k=1:3)
pce <- CTC2(coredata(PCE1), index(PCE1))

Md1 <- SPLINE(Md)
Oil1 <- SPLINE(Oil)
SP1 <- SPLINE(SP)
Md_t1 <- PerC(coredata(Md1), index(Md1), z=1)
Md_t2 <- PerC(coredata(Md1), index(Md1), z=2)
Md_t3 <- PerC(coredata(Md1), index(Md1), z=3)
Md_t7 <- PerC(coredata(Md1), index(Md1), z=7)
Md_t30 <- PerC(coredata(Md1), index(Md1), z=30)
Oil_t1 <- PerC(coredata(Oil1), index(Oil1), z=1)
Oil_t2 <- PerC(coredata(Oil1), index(Oil1), z=2)
Oil_t3 <- PerC(coredata(Oil1), index(Oil1), z=3)
Oil_t7 <- PerC(coredata(Oil1), index(Oil1), z=7)
Oil_t30 <- PerC(coredata(Oil1), index(Oil1), z=30)
SP_t1 <- PerC(coredata(SP1), index(SP1), z=1)
SP_t2 <- PerC(coredata(SP1), index(SP1), z=2)
SP_t3 <- PerC(coredata(SP1), index(SP1), z=3)
SP_t7 <- PerC(coredata(SP1), index(SP1), z=7)
SP_t30 <- PerC(coredata(SP1), index(SP1), z=10)

DATA <- merge(EFFR1, EFFR_t1, EFFR_t2, EFFR_t3, EFFR_t7, EFFR_t30, U, U_t, y, y_t,
              GDPgP, GDPgP_t, GDPgL, GDPgL_t, ECON, M1_, M1_t, M2_, M2_t,
              EX1, EX_t, IPI1, IPI_t, PCE1, PCE_t, pce,
              Oil_t1, Oil_t2, Oil_t3, Oil_t7, Oil_t30,
              SP_t1, SP_t2, SP_t3, SP_t7, SP_t30,
              Md_t1, Md_t2, Md_t3, Md_t7, Md_t30)
DATA1 <- na.locf(DATA)
DATA2 <- DATA1[which(index(DATA1)=="1985-01-01"):which(index(DATA1)=="2008-12-16"),]
DATA3 <- as.data.frame(DATA2)
NAME <- c("I", "TN", "TL", "EFFR1", "EFFR2", "EFFR3", "EFFR7", "EFFR30",
          "U", "U_1", "U_2", "U_3", "Y", "Y_1", "Y_2", "Y_3", "GDP", "GDP_1", "gdp", "gdp_1",
          "Econ", "Econ_1", "Econ_2", "Econ_3", "EX", "EX_1", "EX_2", "EX_3", "M1", "M1_1", "M1_2", "M1_3",
          "M2", "M2_1", "M2_2", "M2_3", "IPI", "IPI_1", "IPI_2", "IPI_3", "PCE", "PCE_1", "PCE_2", "PCE_3", "pce",
          "Oil1", "Oil1_1", "Oil1_2", "Oil1_3", "Oil1_7", "Oil1_30",
          "SP1", "SP2", "SP3", "SP7", "SP30",
```



```
"Md1", "Md2", "Md3", "Md7", "Md30")
colnames(DATA3) <- NAME
#####

#### Plots ####
FFTR.count <- Cut(FFTR1, "1985-01-01", "2015-03-26")
table(FFTR.count[,2])
FFTR.plot <- Cut(FFTR.plot, "1985-01-01", "2015-03-26")
Target <- zoo(FFTR.plot[,2], as.Date(FFTR.plot[,1]))
autoplot(Target) + ylab("Target") + ggtitle("Federal Funds Target Rate, 1985-2015") +
theme_set(theme_bw())
ggplot(FFTR4) + geom_histogram(aes(x=TN, fill=factor(I))) + xlab("Days to the Next Policy
Change") +
  ylab("Count") + ggtitle("Histogram of TN") + theme_set(theme_bw()) +
  scale_fill_discrete(name="Policy\\nDirection", breaks=c(-1,1), labels=c("Cut",
"Raise"))
#####

#### Prediction ####
DATA.Pred <- DATA1[which(index(DATA1)=="2015-03-26"),]
DATA.Pred <- as.data.frame(DATA.Pred)
DATA.Pred$I <- NA
DATA.Pred$TN <- NA
n <- dim(DATA3)[1]
DATA3[n+1,] <- DATA.Pred
#####

#### Write Guide files ####
write(t(as.matrix(DATA3)), file="/Users/home/Documents/THESIS/GUIDE/Thesis.data",
ncolumns = dim(DATA3)[2])
IA <- 1:dim(DATA3)[2]
VA <- colnames(DATA3)
TA <- c("D", "X", rep("N", dim(DATA3)[2]-2))
write(t(matrix(c(IA,VA,TA), ncol=3)),
file="/Users/home/Documents/THESIS/GUIDE/Thesis.dsc", ncol=3)
#####

#### Prediction for I ####
colnames(DATA.Pred) <- NAME
newdata <- DATA.Pred
predicted <- function(){
  if(is.na(gdp) | is.na(Y_3) | -42.790062634429027 * Y_3 + gdp <=
2.8193113795629654 ){
    nodeid <- 2
    predict <- "-1"
  } else {
    if(is.na(Md30) | is.na(EFFR1) | 6.0037426689402706E-002 * EFFR1 + Md30 <=
1.5885512123228633E-002 ){
      if(is.na(Econ) | is.na(PCE_1) | -1913.0422729070169 * PCE_1 + Econ <=
8.3560040094112900 ){
        if(is.na(PCE_1) | is.na(Econ_1) | -3.6059350226010442E-004 * Econ_1 + PCE_1 <=
1.5247194114900511E-002 ){
          if(is.na(gdp) | is.na(Md30) | 35.711276357437839 * Md30 + gdp <=
2.8043614439205324 ){
            nodeid <- 48
            predict <- "-1"
          } else {
            nodeid <- 49
            predict <- "1"
          }
        } else {
          nodeid <- 25
          predict <- "1"
        }
      } else {
        if(is.na(GDP) | is.na(Y) | 103.96506866608914 * Y + GDP <=
-0.66164970680514767 ){
          nodeid <- 26
          predict <- "1"
        } else {
          if(is.na(Y) | is.na(EX_3) | 0.25864346326344223 * EX_3 + Y <=
0.49971554341476071 ){
            if(is.na(Y) | is.na(IPI_1) | -4.5750870024460166 * IPI_1 + Y <=
-6.1477917369106946E-002 ){
              nodeid <- 108
              predict <- "-1"
            } else {
              if(is.na(Y) | is.na(EX) | -8.5789604663866481E-003 * EX + Y <=
3.9203676727269376E-002 ){
                if(is.na(Y) | is.na(PCE_1) | -0.24505452409984912 * PCE_1 + Y <=
1.5695876446369571E-002 ){
                  if(is.na(Y) | is.na(IPI) | -1.2931855995965935 * IPI + Y <=
1.5166114260352198E-002 ){
```




```
if(is.na(Y) | is.na(TL) | -1.4600932033350688E-004 * TL + Y <=
-6.3780422551234950E-002 ){
  nodeid <- 1744
  predict <- "1"
} else {
  if(is.na(Y) | is.na(Y_2) | -1.5090025317838198 * Y_2 + Y <=
-1.4126765425308569E-003 ){
    nodeid <- 3490
    predict <- "-1"
  } else {
    if(is.na(Y) | is.na(EX_3) | 5.1027636347581494E-003 * EX_3 + Y <=
-3.4127487084069714E-002 ){
      nodeid <- 6982
      predict <- "-1"
    } else {
      if(is.na(Y) | is.na(PCE_3) | -0.51739541531184152 * PCE_3 + Y
<= -2.5988034335764259E-003 ){
        if(is.na(GDP) | is.na(Y_2) | -200.85206075717574 * Y_2 + GDP
<= 9.7433829742629730 ){
          if(is.na(IPI_2) | is.na(EX_2) | -7.5513578730608847E-003 *
EX_2 + IPI_2 <= 3.6495483879718932E-002 ){
            if(is.na(Y) | is.na(U) | 1.1155075575984283E-002 * U + Y <=
-1.1578323548405500E-002 ){
              nodeid <- 111728
              predict <- "-1"
            } else {
              if(is.na(GDP) | is.na(pce) | -0.95641902505314225 * pce
+ GDP <= 5.6999999999999993 ){
                nodeid <- 223458
                predict <- "1"
              } else {
                nodeid <- 223459
                predict <- "-1"
              }
            }
          } else {
            nodeid <- 55865
            predict <- "-1"
          }
        } else {
          nodeid <- 27933
          predict <- "1"
        }
      } else {
        nodeid <- 13967
        predict <- "-1"
      }
    }
  }
}
} else {
  nodeid <- 873
  predict <- "-1"
}
} else {
  nodeid <- 437
  predict <- "1"
}
} else {
  nodeid <- 219
  predict <- "-1"
}
}
} else {
  nodeid <- 55
  predict <- "-1"
}
}
} else {
  if(is.na(GDP) | is.na(SP30) | 17.938543297451865 * SP30 + GDP <=
4.5010152559384613 ){
    nodeid <- 14
    predict <- "1"
  } else {
    if(is.na(GDP) | is.na(Oil30) | -36.756252510816999 * Oil30 + GDP <=
3.1121163138991634 ){
      nodeid <- 30
      predict <- "1"
    } else {
      if(is.na(U_1) | is.na(M2_1) | -300.78072639314718 * M2_1 + U_1 <=
0.16020900913047098 ){
        if(is.na(Y_3) | is.na(Y_1) | -0.83982547096699589 * Y_1 + Y_3 <=
-3.0193026340389444E-003 ){
          nodeid <- 124
          predict <- "-1"
        } else {

```



```
        nodeid <- 125
        predict <- "1"
      }
    } else {
      nodeid <- 63
      predict <- "1"
    }
  }
}
}
return(c(nodeid,predict))
}
for(i in 1:dim(newdata)[1]){
  TL <- as.numeric(newdata$TL[i])
  EFFR1 <- as.numeric(newdata$EFFR1[i])
  U <- as.numeric(newdata$U[i])
  U_1 <- as.numeric(newdata$U_1[i])
  Y <- as.numeric(newdata$Y[i])
  Y_1 <- as.numeric(newdata$Y_1[i])
  Y_2 <- as.numeric(newdata$Y_2[i])
  Y_3 <- as.numeric(newdata$Y_3[i])
  GDP <- as.numeric(newdata$GDP[i])
  gdp <- as.numeric(newdata$gdp[i])
  Econ <- as.numeric(newdata$Econ[i])
  Econ_1 <- as.numeric(newdata$Econ_1[i])
  M2_1 <- as.numeric(newdata$M2_1[i])
  EX <- as.numeric(newdata$EX[i])
  EX_2 <- as.numeric(newdata$EX_2[i])
  EX_3 <- as.numeric(newdata$EX_3[i])
  IPI <- as.numeric(newdata$IPI[i])
  IPI_1 <- as.numeric(newdata$IPI_1[i])
  IPI_2 <- as.numeric(newdata$IPI_2[i])
  PCE_1 <- as.numeric(newdata$PCE_1[i])
  PCE_3 <- as.numeric(newdata$PCE_3[i])
  pce <- as.numeric(newdata$pce[i])
  Oil30 <- as.numeric(newdata$Oil30[i])
  SP30 <- as.numeric(newdata$SP30[i])
  Md30 <- as.numeric(newdata$Md30[i])
  tmp <- predicted()
  node <- tmp[1]
  rpred <- tmp[2]
  print(c("Case ",i,node,rpred))
}
#####

#### fit ####
fit <- read.table("/Users/home/Documents/THESIS/Guide/Thesis.fit", header=T)
fit <- fit[-dim(fit)[1], ]
DATA.raise <- FFTR.plot[which(fit[,2]==26),]
head(DATA.raise)
#####

#### Bootstrap ####
set.seed(30)
DATA.Train <- DATA3
n <- dim(DATA.Train)[1]
boot.forest <- vector()
for (i in 1:100) {
  index <- sample(1:n, n, replace=T)
  DATA.boot <- DATA.Train[index,]
  DATA.boot[n+1,] <- DATA.Pred
  write(t(as.matrix(DATA.boot)), file="/Users/home/Documents/THESIS/Boot/boot.data",
ncolumns=61)
  system("cd /Users/home/Documents/THESIS/Boot && ~/guide < boot.in", wait=T)
  fit <- read.table("/Users/home/Documents/THESIS/Boot/boot.fit", header=T)
  boot.forest <- c(boot.forest, fit[n+1,2])
}
boot.bag <- vector()
for (i in 1:100) {
  index <- sample(1:n, n, replace=T)
  DATA.boot <- DATA.Train[index,]
  DATA.boot[n+1,] <- DATA.Pred
  write(t(as.matrix(DATA.boot)), file="/Users/home/Documents/THESIS/bag/boot.data",
ncolumns=61)
  system("cd /Users/home/Documents/THESIS/bag && ~/guide < boot.in", wait=T)
  fit <- read.table("/Users/home/Documents/THESIS/bag/boot.fit", header=T)
  boot.bag <- c(boot.bag, fit[n+1,2])
}
217.8836 + 1.984 * sd(boot.bag)
217.8836 - 1.984 * sd(boot.bag)
211.7949 + 1.984 * sd(boot.forest)
211.7949 - 1.984 * sd(boot.forest)
#####
```

参考文献

- [1] 谭小芬, 陈昌兵. 泰勒规则研究综述[J]. 经济学动态, 2006 (4): 105-110.
- [2] 张长水, 张见闻. 演化数据的学习[J]. 计算机学报, 2013, 2: 009.
- [3] 中国人民银行营业管理部课题组, 杨国中, 姜再勇, 刘宁. 非线性泰勒规则在我国货币政策操作中的实证研究[J]. 金融研究, 2009, 12: 30-44.
- [4] Baker S R, Bloom N, Davis S J. Measuring economic policy uncertainty[J]. Chicago Booth research paper, 2013 (13-02).
- [5] Bohl M T, Siklos P L, Werner T. Do central banks react to the stock market? The case of the Bundesbank[J]. Journal of Banking & Finance, 2007, 31(3): 719-733.
- [6] Breiman L, Friedman J, Stone C J, et al. Classification and regression trees[M]. CRC press, 1984.
- [7] Castro V. Can central banks' monetary policy be described by a linear (augmented) Taylor rule or by a nonlinear rule?[J]. Journal of Financial Stability, 2011, 7(4): 228-246.
- [8] Chongqing Yan. Regression Trees and Nonlinear Time Series Modeling[D]. Madison: University of Wisconsin, 1995.
- [9] Clarida R, Gali J, Gertler M. Optimal monetary policy in closed versus open economies: An integrated approach[R]. National Bureau of Economic Research, 2001.
- [10] Coibion O, Gorodnichenko Y. Why are target interest rate changes so persistent?[R]. National Bureau of Economic Research, 2011.
- [11] Varian H R. Big data: New tricks for econometrics[J]. The Journal of Economic Perspectives, 2014: 3-27.
- [12] Loh W Y. Classification and regression trees[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2011, 1(1): 14-23.
- [13] Loh W Y. Fifty Years of Classification and Regression Trees[J]. International Statistical Review, 2014, 82(3): 329-348.



- [14] Orphanides, Athanasios, 2003. "Historical Monetary Policy Analysis and the Taylor Rule," *Journal of Monetary Economics* 50(5), 983-1022.
- [15] Petersen K. Does the Federal Reserve follow a non-linear Taylor rule[R]. 2007.
- [16] Taylor J B. Discretion versus policy rules in practice[C]//Carnegie-Rochester conference series on public policy. North-Holland, 1993, 39: 195-214.
- [17] Rudebusch G D. Term structure evidence on interest rate smoothing and monetary policy inertia[J]. *Journal of monetary economics*, 2002, 49(6): 1161-1187.
- [18] Rudebusch G D. Monetary policy inertia: fact or fiction?[J]. *FRB of San Francisco Working Paper*, 2005 (2005-19).