**中国股市日历效应及其影响因素研究**

**为什么自行检验日历效应：**

大量的文献表明日历效应在中国股市始终存在。不同学者运用不同研究方法、选用不同研究对象往往得到不同的结论，也从侧面证明日历效应随时间不断发生变化。本文的研究日历效应所用数据较新，且探究日历效应影响因素的模型与大部分学者不同，为了确保日历效应与其影响因素判别的一致性自行检验日历效应。

**提出原假设：**

1. 存在某种周历、月历效应，属于以随机试探的形式寻找具体可能存在的该类日历效应，对该现象出现的原因在未做进一步数据分析时不作任何主观猜测；
2. 存在假日效应，休市的时间内股票暂停交易，但与股市行情高度相关的信息仍然在被随机地释放，且无法动用的购置股票所用资金的机会成本随休市时间同步增长，休市后的首个交易日股市有较高概率出现收益率的异常波动。

**所用数据与处理方法：**

本文所用的基础数据包括上证指数、深证成指与创业板指2010年XXXX至XXXX的日收盘价数据与日内五分钟高频交易数据以及其他辅助数据。本文分别对三个股指对数收益率及其已实现波动率进行R-GARCH建模分析相应的日历效应并解析其背后的影响因素。鉴于所选时间内股市行情差别较大，对收益率及已实现波动建立隐马尔科夫模型，判断其背后的股市风格（隐态），并将股市风格发生转变节点作为数据分段点，以分别研究相应的日历效应。

**模型选择：（为什么用RealGarch）**

本文将基于GARCH模型对中国股市的日历效应进行检验以及后续的影响因素分析，GARCH模型对收益率与其波动率联合建模的方法更加符合资产收益率的真实情况，采用该类方法检验日历效应更为准确。

Engle（1982）首次提出了自回归条件异方差模型（Autoregressive Conditional Heteroscedastic，ARCH），为波动率建模提供了系统框架，其认为资产收益率的扰动是序列不相关的，但不是独立的，且这类不独立性可以用扰动滞后项的简单二次函数来描述。

Bollerslev（1986）为了使ARCH模型拥有更灵活的滞后结构，类似AR过程扩展为ARMA过程将ARCH过程扩展为广义ARCH（Generalized-ARCH，GARCH）过程。

为了在模型中体现正负收益率的非对称性，Nelson（1991）与Glosten、Jagannathan和Runkle（1993）分别提出了指数GARCH模型（Exponential-ARCH，EGARCH）与门限GARCH模型（Threshold-ARCH，TGARCH）。

Hansen、Huang和Shek（2011）提出，在高频金融数据已广泛使用的背景下，现有文献介绍了许多已实现波动率的测量，例如由Andersen和Bollerslev提出的已实现波动率（1998），这类已实现波动率相比于以往收益率（扰动）平方而言，在对波动率建模以及预测上都更加实用、更具有信息性，并以此建立了已实现GARCH模型（Realized-GARCH，R-GARCH）。

结合以上各类GARCH模型的特点，本文将使用R-GARCH（Realized-GARCH）模型来检验中国股市的日历效应：首先，已实现GARCH模型由于添加了日内已实现波动率对股指收益率有更好的拟合能力（Christoffersen et al.，2010），相比于普通GARCH模型建模更加充分；其次，R-GARCH模型中的测度方程能识别波动的非对称性，与股市正、负向波动带来的收益率变动非对称性相符；再者，使用GARCH族模型进行建模在理论上使得模型贴合股票收益率的条件异方差性，且可识别波动率的聚集特点；最后，在对股指收益率拟合能力更好的模型中加入日历效应后，其检验到的日历效应更具有准确性。

**模型描述与解释：**

第一个方程为均值方程，该方程左边是股指对数收益率，方程右边是ARMA(p,q)部分加代表日历效应的其他外生变量，本文初始采用最为简单的ARMA(0,0)的形式，不对该部分作过多探究将模型复杂化，但若联合估计结果中的均值方程建模不充分则适当添加AR与MA项，是独立同分布且均值为0方差为的收益率扰动项（error项），的具体分布可根据建模需要而变化。

第二个方程为波动率方程，即对收益率扰动的对数方差进行建模，方程右边为之后q阶的对数已实现波动率以及滞后p阶的对数方差，类似GARCH模型的波动率方程，只是将收益率扰动替换为已实现波动率，该方程表明过去更大的已实现波动率与收益率的波动会导致未来更大的收益率波动，依然符合资产收益率波动率聚集现象。

第三个方程为测度方程，是对数已实现波动率表示为收益率扰动的方差与标准化收益率扰动函数的具体形式，是的杠杆函数，具体形式取，预期为负且为正，来自的负向变动导致的对数已实现波动率变动并传导至方差（将第三个方程代入第二个方程右边）的变动大于来自的正向变动。

方程之所以采用对数线性形式，是由于天然地,其中是收益率扰动项，实际也是,其期望值即日收益率的方差，自然十分接近日内已实现波动（对日收益率方差的近似），故而第三个方程扩展地写成是与的函数形式，能够更准确地测量估计。

额外加入测度方程的R-GARCH模型相比于普通GARCH模型多了已实现波动率这一桥梁，若将第三个方程代入第二个方程右边的，得到的新方程与能够反应收益率扰动正负变化带来波动非对称变化的指数GARCH模型十分接近，两者的杠杆函数略有不同。

**确定R-GARCH模型的波动方程阶数：**

相较于ARCH模型可使用均值方程残差项平方序列的偏自相关函数来确定阶数，GARCH族模型的阶数不易确定，且高阶GARCH模型会出现不稳定的情况，对较为严格的拟合或预测任务通常通过排列组合的方式来结合贝叶斯信息准则（Bayesian Information Criterion, BIC）或者赤池信息准则（Akaike Information Criterion, AIC）确定GARCH项与ARCH项的滞后阶数。本文的重点不在于精确拟合或者预测，无需建立过于复杂的R-GARCH模型，因此所有模型的波动方程初始均设定为GARCH（1,1）的形式，仅当回归结果的均值方程残差项未通过充分建模检验时才重新设定更为复杂的模型。

**选择模型中均值方程收益率扰动所服从的分布：**

标准化的收益率扰动通常被认为服从标准正态分布，由于本文建模的对象是具有尖峰厚尾特点的金融时间序列，故本文首先检验何种分布更适合所选用的数据，将从标准正态分布、具有厚尾特征的学生t-分布、GARCH模型常用的广义误差分布三种分布中选择。

分布的选择遵循以下步骤：（1）检验所选股指的对数收益率序列是否具有GARCH效应；（2）取用某一种分布，在均值方程中不加入任何代表日历效应的外生变量情况下得到回归估计结果；（3）检验R-GARCH模型是否建模充分；（4）比较三种分布下标准化残差项与所用分布的QQ图以及KS检验结果并选择最优分布。

**日历效应的检验方法：**

分别对上证指数、创业板指、深证成指各时段进行步骤相同的检验：（1）以试探的形式检验周历与月历效应，将周一至周五、一月至十二月哑变量均添加入均值方程中并回归得到估计结果，在标准化残差及其平方的Ljung-Box统计量均不显著的情况下记录均值方程中显著的周历、月历效应（称为建模充分，即通过均值方程中的ARMA项与波动方程中的GARCH项成功消除了标准化残差中的自相关与其平方的自相关），否则重新建模；（2）在代表周、月历效应的哑变量系数存在多个均为显著的情况下，仅记录系数绝对值最大的一个作为该股指该时段的周、月历效应；（3）检验假日效应，将代表本交易日与上一交易日休市天数的离散变量添加入均值方程中并回归得到估计结果，在标准化残差及其平方的Ljung-Box统计量均不显著的情况下确认休市天数变量的系数是否显著，否则重新建模，若休市天数变量的系数显著则表明存在假日效应。在建模不充分情况下，若反复重新建模至模型的波动率方程过于复杂且仍然无法通过Ljung-Box检验，则认为该股指该时段无任何日历效应。

**日历效应检验：**

收益率数据检验与分布确定：

各类检验的结果与最终分布的确定结果如下表所示：

分布选择表-来自excel

以上证综指为例，首先检验股指对数收益率与日内已实现波动率的平稳性，R-GARCH依然是线性回归模型，需避免由非平稳时间序列得到伪回归结果；其次是股指对数收益率序列的GARCH效应检验，上证综指对数收益率平方的Ljung-Box统计量在滞后阶数为8阶（ln(2147)≈8）时依然显著，说明上证综指对数收益率序列的GARCH效应在统计上是显著的，可以建立波动率模型与测度模型进行联合估计；分别用三种分布假设进行联合估计，估计结果中的标准化残差与其平方的Ljung-Box统计量均不显著，再记录下t-分布与广义误差分布优化参数中的自由度；检验联合估计结果中的标准化残差是否服从对应理论分布，运用KS检验（Kolmogorov-Smirnov是比较一个频率分布f(x)与理论分布g(x)或者两个观测值分布的检验方法，原假设为某数据服从一个理论分布）得到结果为上证综指最合适的见面破分布假设是广义误差分布。关于创业板指与深证成指与上证综指相似，结果如上表所示，均选择广义误差分布作为扰动项的分布假设。

检验过程：

各股指各时段的日历效应结果如下表所示：

日历效应记录表-来自excel

以上证综指2010年1月初至2014年10月末时段为例，在均值方程中加入代表周一、二、四、五的哑变量，周三效应可提现在均值方程常数项中；调整均值方程与波动率方程直至建模充分，最终的模型为ARMA(0,0)-GARCH(4,5)，代表周四的哑变量回归系数为-0.001834且在0.01的显著性水平下显著（P值为0.003387），则确认上证综指该时段具有负周四效应，记录如表所示的数据以及模型的具体形式；月历效应的检验与周历效应完全相同；假日效应由于仅有一个代表休市天数的离散变量无需进行第二次联合估计。检验各股指各时段的各类效应后得到上表中的结果。

结果简述：

上证综指、创业板指、深证成指各时段的日历效应并不相同，参照此结果将数据按照时段拆分是正确的，也符合其他学者所得到的结论——日历效应会随时间发生变动。其中三次出现周四效应均为负效应，即周四收益率显著为负；三次出现二月效应均为正效应，即二月收益率显著为正；二次出现十二月效应均为负，即十二月收益率显著为负；各股指各时段共检验出二次正假日效应，即该交易日之前的休市天数越长收益率越高，大部分假日效应的结果均为不显著——除了上证综指的第一时段即便令波动率方程为GARCH(8,8)的形式依然无法通过充分建模的检验。下一部分会引入辅助数据，并改进模型来检验与该类日历效应相关的因素。

**日历效应的影响因素探究：**

理论分析：

与日历效应相关的影响因素较多，现有文献表明该类异常波动来自于金融市场的风险、投资者的情绪与政府出台的金融市场相关政策，本文通过设计股市风险因素、投资者情绪因素与政策因素对上一部分得出的日历效应进行影响因素探究。

由于已实现GARCH模型中的收益率数据与已实现波动率数据为日度数据，本文选用同为日度频率的变量构造三种影响因素。

股市风险因素，投资者会通过近期股市的波动与走势以推断未来股市蕴含的风险，股市风险越大即股市异常波动频繁，说明近期投资者难以形成短期内的一致预期，依照股市常有的“波动率聚集”现象，投资者预期未来的波动较大，要求更高的收益溢价补偿，导致股票价格下跌。

本文所提到的股市风险因素指较为狭义的短期内的风险（如收益率的标准差），主要的反映方式是股指短期的波动与走势。理论上，预测短期价格是不可能的，不同投资者对近期的同一股价走势会有不同的分析与策略，相互对冲的买卖交易使得股价小幅波动；实际中，近期的股价波动异于寻常，任何理性投资者都会要求更高的回报以补偿风险，投资风格普遍变得更加保守，略悲观的一致预期会抵消原有的正收益并增大原有的负收益。

日历效应这样的异常波动是细小的市场异象，在一段时间内出现随后消失，市场从细小的无效重新变为有效，与丛集的波动十分相似。短期股市风险对日历效应的非对称影响来源于风险反馈至收益的非对称性原理——负向波动造成的损失大于正向波动带来的收益，这样的杠杆效应在异常波动中表现得更加明显。

投资者情绪因素，股票价格变化是众多投资者交易的结果。投资的主体无论是机构投资者还是个体投资者，投资的方法无论是通过主观猜测、数据分析还是量化策略，都受到投资者情绪的影响，狂热的投资情绪更容易造成股市的异常波动，一致性的狂热导致小概率事件发生。过于乐观或悲观的情绪造成显著为正或者显著为负的收益率。

政策因素，任何国家出台有关金融市场政策时均会对股市造成难以预料的波动，即便政府的出发点都是维护金融市场稳定。中国政府的相关政策对中国股市影响之深远非其他国家可比，无论是投资者或是理论学者均认为政府对股市的干预过多，从股票发行上市的审核流程到二级市场交易。中国股市是政府进行宏观调控的中间目标，政府作为市场监督者也在执行市场操纵行为，其目的是为了金融市场健康、稳定、可持续地发展，但会导致市场难以形成一致的政策预期，表现为股市涨幅或跌幅过大、波动剧烈。

导致中国出现政策市的原因是多方面的。第一，在中国股市不断发展的过程中，政府出于维护资本市场稳定、有效率的目的陆续出台新政策与新方法，叠加个体投资者容易出现“羊群效应”以获得超额收益的因素，削弱了资本市场价格决定机制的作用，对政策效果影响的测度困难直接导致均衡价格变得难以确定，加剧了股市的动荡；第二，遍布于多个行业的许多大型上市公司股权由国有资本控制，政府相当于同时扮演监管者与投资者，在该类约束下所做决策未必是有利于上市公司所有者权益的最优解；第三，政府的最终目标是多元化的，保护投资者利益是众多目标中的一个，与其余目标之间存在矛盾与取舍是必然的，这同样导致如股市过于悲观时的救市政策与股市过热时的抑制政策。

政策市会加剧股市的异常波动。第一，政策本身作为一种信息具有形成时间长、信息含量大、未发布时知悉人数多等特点，不可避免地造成了信息不对称现象，为操纵市场提供途径；第二，缺乏公平的资本市场直接导致投资者投机心理过盛，股市的资金配置功能与信息传递作用被减弱，不足以作为经济活动的晴雨表；第三，政府的政策干预理论上具有减小系统性风险的功能，这是由于不同投资者对政策的理解存在差别，由此带来不同的投资策略相互之间形成对冲，中国较低的个体投资者金融素养与机构投资者的报团取暖行为均是“羊群行为”的表现，两者对一个政策的解读具有趋同性，增大了资本市场的系统性风险。

探究方法：

添加各因素与日历效应哑变量的交互项，均值方程变为：

其中为除所研究的日历效应外的其他日历效应哑变量（如研究正二月效应，则表示除二月以外其他月份，目的是保持与原模型的一致性，假日效应没有一项），下标用以标识某一股指某一时段，下标用以标识某一日历效应，为日历效应哑变量与股市风险指标交互项，交互项系数是主要研究对象之一。

将均值方程提取公因式后可以得到：

某一日历效应即：

则：

交互项系数是各因素对日历效应的影响值，本文仅研究其正负性，具体含义需结合日历效应进行探讨。

假设与检验方法：

从上一部分对日历效应相应因素的理论分析可以得到相应的原假设：

（1）代表股市风险的指标越大，减弱正日历效应并加强负日历效应；

（2）代表投资者（乐观、狂热）情绪的指标越大，减弱负日历效应并加强正日历效应；

（3）代表政策发布数量的指标越大，加强所有日历效应。

与三个原假设对应的均值方程假设表达为：

（1），；

（2），；

（3），。

具体的检验方法与前一部分检验日历效应时相近：（1）通过加入交互项进行联合估计，适当调整模型的ARMA项与GARCH项确保建模充分；（2）记录交互项系数，结合未加入交互项时的系数与预期结果进行比较分析，得到影响日历效应的因素。

PS：

（提到了某种检验，是否需要具体说明？提到了回归估计，是否需要说明估计过程、方法？分布是否需要进一步解释？）

Ljung-Box统计量滞后阶数取的四舍五入值，为建模对象序列的总时间跨度。

解释日内已实现波动率。

Kolmogorov-Smirnov是比较一个频率分布f(x)与理论分布g(x)或者两个观测值分布的检验方法。

政策市 原因 危害 对策https://www.xzbu.com/3/view-7357328.htm