# Mean-reversion Cross-section

# 摘要

# 引言

1.加入新的影响波动率回复的特有变量

将HanBing中的财务特征依据波动率的实际情况进行精简和合并，可以考虑加入必要的适合波动率的特征

强调与HanBing只是结构相似，但研究问题不同

加入一个策略：根据公司财务特征预判未来波动率的变化情况，获取额外利润（加入其他希腊字母的锁定）

只能做mean reversion(波动率往上走，何时往下跳；而不是波动率往下走，何时往上走)

考虑使用主成分分析来寻找潜在影响原因

Tabal1 描述性统计分析

Table2 普通的均值回复

波动率不会无限上涨或下跌，总是在经历一定的运动后反向回复。随机波动率模型使用Ornstein-Uhlenbeck过程来描述波动率的运动形式（heston,1993）。其中，波动率总是围绕一个长期均值上下变化，并且下一时刻的波动率变化又取决于当前的波动率水平（Yoon,2022）。Ornstein-Uhlenbeck过程从数学上证明，波动率变化的概率分布期望与波动率水平和长期均值的差负相关，表现均值回复特征。波动率衡量资产的风险和金融市场的不确定性，然而这种不确定性又具有清晰的确定性数学统计规律。

当前已经有文献研究了波动率的均值回复特征，它们研究重点聚焦于指数期权（Cao,2020；Chiu,2019）或大宗商品指数期权（Chi,2017）。相比于个股期权，指数期权会消除诸多金融市场细节特征（Bali,2014）。本文在时间横截面上，研究了个股期权的风险中性波动率的均值回复特征，包括均值回复强度、回复概率等内容。在随机波动率的框架下，本文证明了波动率均值回复的必然性以及波动率回复概率和所需时间的关系。个股期权与个股公司的财务特征具有紧密联系（Bing Han,2022），其中做空看涨期权的delta对冲收益率与股价、盈利能力负相关，而与现金流量、分析师预测偏差等负相关。但是它将研究目标仅仅聚集于期权本身的delta对冲收益率，而delta对冲收益率与波动率风险溢价负相关（Bakish,2003），与波动率的波动率风险溢价正相关（Darien,2018）。本文重点参考了这篇论文的研究方法，在类似的个股财务分析框架中将个股期权划分为不同的分位数组合，扩展研究了波动率的均值回复特征的内容。

本文研究发现，在个股期权的时间横截面上隐含波动率均表现出显著的均值回复特征，并且均值回复强度在各个财务框架组合中表现有所差异。本文使用现金流量方差描述个股的盈利风险程度（Haugena,1996），发现经营风险较高的行业具有较强的均值回复强度，即高风险个股期权的隐含波动率往往会在短期内实现较大的均值回复程度。同时本文将个股期权按照行业类别划分为不同的组合进行测试，研究发现以苹果和特斯拉等为代表的科技行业波动率具有更强的均值回复强度，而第一产业的行业波动率均值回复程度较低甚至无明显均值回复。总之，波动率的回复特征与个股的盈利能力、资产规模、现金流动能力负相关，而与负债水平、资产负债比率等负相关。

并且，本文在个股期权的横截面上研究了均值回复概率与所需时间之间的关系。本文使用Logit模型框架估计均值回复概率的期望值，研究发现：第一，波动率的单向时间变化越久，均值回复概率越低且所需时间越长；第二，波动率的单向变化程度越高，均值回复概率越大且所需时间越短；第三，波动率的单向变化越快，均值回复概率会随着时间的延长而降低。而在依据财务类别分组的框架下，盈利能力强、资产规模大、资金流动能力强的个股期权隐含波动率往往具有较低的均值回复概率，而经营风险高、负债水平高的行业往往具有较高的均值回复概率。

考虑到期权市场同样存在羊群效应（Bernales，2016），个股期权的隐含波动率的均值回复特征可能会受到波动率持续的影响。本文以历史波动率持续作为条件前提，研究发现：当波动率持续较长时，波动率的均值回复强度会变高，而均值回复概率会下降；当波动率持续较短时，一切相反。由于期权市场和个股市场的分歧（Ruslan，2019），本文认为波动率持续会改变投资者的心里预期，甚至长期的波动率持续会改变整个金融市场的短期固定参数，进而使得波动率均值回复不再出现。同时，考虑到波动率聚集现象以及由此带来的市场环境差别（Pearson,2020），本文分别研究了高波动市场和平稳市场中的均值回复现象，发现市场越平稳，波动率均值回复强度越弱甚至不存在。

最后，本文探讨了个股期权均值回复差异形成的原因。期权市场同样存在供需关系（Cheng,2019），投资者的对冲成本会随着波动率的上升而上升，进而其对冲需求会下降。本文研究了期权购买压力与均值回复强度之间的关系，发现当波动率上升时，均值回复强度会增加。并且，本文在4/2 heston模型框架下估计了波动率回归速度和波动率长期均值（Martino,2017），用其精确估值进一步验证了该结论。

未来波动率回复程度 = 过去波动率的变化程度 + 随机项

(产业类别、高低风险不同、个股特征详细于指数)

从理论推导上数学证明可行，且实证结果符合

测算未来回归概率

波动率持续）

需要推导出一个计量模型

# 理论背景

## The Solution of Ornstein-Uhlenbeck Process

In the stochastic volatility models, the volatility follows the Ornstein-Uhlenbeck Process, which shows the existence of mean reversion about the volatility(Heston,1993). Under the physical measure ,on the trade date , the stock price and volatility evolves as[[1]](#footnote-1):

Where and represent the drift rate and volatility of the stock return, respectively. is a constant but not .is the speed of mean reversion, and is the long-term mean of volatility, is the volatility of volatility..and are the wiener process of stock and volatility under the physical measure , respectively, and their correlation is .

The option price is , thus the partial differential equation of value is:

Where is the compensation of volatility risk(Huang and Darien, 2018), thus is the drift of volatility under the risk-neutral measure .is the risk free rate, which is set to a constant. To get the risk-neutral SDE of volatility and stocks, we use the Cameron-Martin-Girsanov lemma set the wiener process of volatility under Q , in which .Thus we get the SDE of volatility under Q measure:

To get the solution of volatility SDE under Q ,we substitute and into Equation (3) so that it takes the form Ornstein-Uhlenbeck process[]:

When the are much higher than (the volatility increase higher than ), the will be much more negative, thus will become negative with a greater probability (the volatility tends to move down even more); When the are much lower than (the volatility decrease lower than ), the will be much more positive, thus will become positive with a greater probability (the volatility tends to move up even more). We call this form of movement in volatility mean reverting. The solution of Equation (4) is[]:jfdjllingking suite by WRDS,

Where and represent long-term mean and the speed of mean reversion under Q measure, respectively. We can get the risk-neutral volatility from the option price, which is called the implied volatility. The implied volatility follows the mean reversion process, and .From the Equation(5), we can get the math process of , include its possible values and the corresponding probability.

## Response of Volatility to Distance

According to the volatility SDE under Q Equation, the implied volatility fluctuates around long-term mean of volatility . For researching the response of volatility to distance between the and ,we let , thus:

Where . For the martingale properties of ITO integrals , ,thus and variance . Because the Ornstein-Uhlenbeck process belong to Markov process, .

We set the distance from implied volatility to long-term mean of volatility as and substitute it into Equation (6), then we get Equation (7):

Thus the partial derivative of with respect to is:

For , and , thus >0. This means there exists a positive relationship between and , and we call the as the intensity of mean reversion.

## Probability of Volatility Mean Reversion

Is the subsequent movement in volatility related to the current volatility? When volatility reaches a higher value, does the probability of a subsequent decline in volatility become larger? And vice versa. This section theoretically explores how the subsequent movement of volatility relates to the current level of volatility.

We discretize Equation (4), and obtain discrete SDE for volatility as:

(8)

Where under Q measure, thus according to the properties of the Wiener process that the increment of a Brownian motion is still a Brownian motion. When the volatility reaches a specified value , we study the distribution of by treating the volatility as a constant. We define the cumulative probability functions of conditional on the specified volatility as , thus the conditional probability distribution of is .

Conditional on the specified volatility , the probability of higher than is:

(9)

Where represents the conditional probability density functions for .For ,the is a normal distribution about expectation and , thus we can define:

(10)

Where and , thus:

(11)

Obviously,, this means the probability that higher than conditional on the specified will decrease when increases higher.

## Probability of Mean Reversion under Volatility cluster

We can express the analytical solution of Equation (4) in another form(Ricciardi,1988), which makes it more convenient to calculate the conditional probability of volatility movements as (12):

(12)

Equation (12) expresses the process of volatility after time increment .We set , thus we can change Equation (12) to error correction model as:

(13)

Where is conditional on a fixed , thus we can get the conditional probability of :.According to the existence of volatility cluster (GARCH effects) (Ni,2020), we need to consider the mean reversion in drastic fluctuation and slight fluctuation, respectively. Apparently, the cluster of volatility depends on the volatility of volatility , thus we study the relationship between probability of mean reversion and volatility cluster.

Suppose the volatility cluster has reached a specified level , conditional on , the probability that the absolute decline in volatility is higher than is:

(14)

Where represents conditional probability density function of , and , .Thus we get the effect of X on Y:

(15)

For , the increase of will also increase ,however will decrease , then will be less or more negative. If is more enough, will be negative. That means, if the market is drastic fluctuation, will be more, then the probability of mean reversion will decrease.

# 样本与变量

# 实证检验

## Mean Reversion of Volatility

波动率回复强度与个股本身特征密切相关，包括个股的资产规模、盈利能力和成长能力等。在对期权样本数据进行分位数分组时，如果所依据的个股财务特征不同，那么波动率的均值回复表现是否又会有所不同？本文选择账面市值比、现金流量方差、总资产、总负债等重要指标作为样本数据的分组依据，然后在各个分位数小组中使用过去波动率的变化对未来波动率的回复状态进行回归，最终结果如表一所示。同时为了研究在计算波动率时的时间跨度是否会对运行结果有影响，本文将时间跨度分别选择为10天、30天和90天。为了检验时间跨度选择的稳健性，本文在研究中选取了从5到100天的所有时间跨度进行回归拟合，并确认研究结果稳健。考虑到论文的篇幅限制，表1仅仅展示10天30天和90天。

整体来看，波动率具有明显的均值回复特征。大部分回归结果的t值都在2.5和3.5之间,这表明回归系数具有极强的显著性。而拟合优度大部分回归结果都在60%到80%，这说明各个回归方程具有较高的回归质量。在时间跨度为90天的右栏中，账面市值比所对应的回归结果拟合优度为79.7%，而现金流量方差所对应的拟合优度甚至达到99%。这都可以充分说明，波动率的均值回复在金融市场上是显著存在的。回归结果的系数均为正值，这说明刚过去的波动率发生较大的变化时，未来的波动率必然会出现相反的变化，这也就是所谓的波动率均值回复特征。大部分回归系数均为2到3，这说明每当过去的波动率变化一单位，未来的波动率平均反方向变化2到3个单位。

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1 Regression Results of Mean Reversion | | | | | | | | | |
| **Windows** | **10** | | | **30** | | | **90** | | |
|  | **coef** | **t** | **R** | **coef** | **t** | **R** | **coef** | **t** | **R** |
| **Firm** |  |  |  |  |  |  |  |  |  |
| **BM** | 2.948 | 1.948 | 0.542 | 2.734 | 2.086 | 0.752 | 2.273 | 3.353 | 0.797 |
| **CFV** | 2.941 | 3.606 | 0.783 | 2.737 | 3.572 | 0.467 | 2.497 | 3.376 | 0.99 |
| **GK** | 3.251 | 2.116 | 0.626 | 2.903 | 3.183 | 0.22 | 2.593 | 3.006 | 0.672 |
| **ROA** | 3.49 | 2.621 | 0.686 | 3.208 | 2.561 | 0.692 | 2.847 | 3.473 | 0.489 |
| **TEF** | 2.989 | 3.587 | 0.901 | 2.764 | 2.318 | 0.093 | 2.43 | 2.339 | 0.396 |
| **ZS** | 3.311 | 3.274 | 0.874 | 3 | 3.11 | 0.786 | 2.65 | 2.801 | 0.712 |
| **Assets** | 3.393 | 3.146 | 0.624 | 3.289 | 3.259 | 0.742 | 2.884 | 3.703 | 0.802 |
| **Cash** | 3.074 | 4.455 | 0.686 | 2.851 | 2.565 | 0.514 | 2.459 | 3.709 | 0.641 |
| **Debt** | 3.136 | 2.736 | 0.375 | 2.874 | 2.76 | 0.76 | 2.536 | 2.474 | 0.552 |
| **Cost** | 2.778 | 2.931 | 0.752 | 2.479 | 2.776 | 0.651 | 2.278 | 3.089 | 0.674 |

当时间跨度为10个交易日，回归系数整体最高，均接近3。ROA的回归系数为3.49，在这一列中属于最高值，这也就意味着对于盈利能力水平ROA较高的公司，他们的个股期权隐含波动率具有极强的均值回复能力。如果过去的波动率变化为一单位，那么未来的平均波动率反向回复可以达到3.49个单位。ROA较高意味着该公司具有较强的盈利能力，而这表明该公司具有较强的成长能力，显然会得到金融市场投资者的青睐。如果他的波动率较高，那么他的期权价值也会较高，而这会吸引金融市场空头来卖出他的期权。随着期权空头势力的增强，期权价格又会慢慢被压下去，进而其隐含波动率同样会下降。如果它的波动率较低，作为优质的成长性公司，他的未来期望价值会吸引金融市场的多头来投资，同时他的看涨期权也会被多头投资者以低价购入。而这些行动会持续的抬升他的期权价值，进而提升它的隐含波动率。在这一列中，经营成本较高的公司具有较差的均值回复能力，并且它的回归系数仅为2.778。如果公司的经营成本过高，那么他未来的发展能力必然会被限制。

而当时间跨度为30个交易日时，波动率均值回复的回归系数相比于10个交易日时显然有所下降。其中现金流量方差的回归系数下降为2.737，而资产的回归系数下降为3.289。由于时间跨度的拉长，进入市场上消息的及时性和有效性不断下降，因此他们对波动率的影响也会有所减弱。但是此时ROA的回归系数依然在这一列中属于最高值3.208，这说明公司的盈利能力对均值恢复的影响具有较强的稳健性。

此表展示了

## Probability of volatility

如果波动率的均值回复特征确实存在，那么测量其具体回复概率也很重要。当过去波动率上涨幅度为指定水平时，我们测量了未来的波动率回复至指定程度的发生概率。我们发现，当过去波动率上涨幅度差异较大时，未来波动率变化的分布概率具有显著差异。

如图1所示，我们分别测度了不同过去上涨幅度所对应的未来波动率变化的概率分布图。在左上图中，我们绘制了所有个股期权未来隐含波动率变化的概率分布图。其偏度为-0.087，接近正态分布的偏度0；而峰度为2.882，接近正态分布的3。整体来看，所有的波动率变化分布接近正态分布，几乎不包含有有效信息。在右上图中，当过去波动率全部为上涨时，我们绘制了未来波动率变化的分布概率图。相比于前者，该分布图的偏度微涨至-0.063，而峰度微涨至3.008，依然近似于正态分布，无明显有效信息。这意味着过去波动率上涨无法导致未来波动率的回复。由随机波动率模型可知，波动率变化依然服从几何布朗运动，其概率分布为正态分布。在左下图中，当过去波动率上涨幅度超过50%时，未来波动率变化的概率分布已经出现了极明显的变化。其偏度已经变为0.066，表现出一定的负偏现象，即未来波动率变化为负值的概率更高，这显然符合波动率的均值回复预期；而峰度明显下降为1.554，呈现出瘦尾特征，这说明出现波动率下降幅度极高的可能性很低。在右下图中，当过去波动率上涨幅度超过60%时，概率分布图已经出现非常明显的负偏。

把四个图绘制在同一个大图上

图1.波动率变化概率密度图图表, 直方图

描述已自动生成图表, 直方图

描述已自动生成

图表, 直方图

描述已自动生成图表, 直方图

描述已自动生成

偏度明显上升至0.153，这说明此时未来的波动率上涨已经有极大概率呈现负值；而峰度仅仅变化为1.627，变化幅度较小。

我们将过去波动率上涨幅度从0至1的10个分位置所对应的未来波动率变化所对应的概率分布特征进行详细计算，如表2所示。

表2. 未来波动率变化的概率分布统计

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Crireion | Mean | std | skew | Kurt |
| 0 | -0.067 | 0.281 | -0.063 | 3.008 |
| 0.1 | -0.091 | 0.298 | -0.066 | 2.608 |
| 0.2 | -0.116 | 0.315 | -0.078 | 2.074 |
| 0.3 | -0.138 | 0.329 | -0.074 | 1.785 |
| 0.4 | -0.157 | 0.339 | -0.009 | 1.582 |
| 0.5 | -0.175 | 0.348 | 0.066 | 1.554 |
| 0.6 | -0.191 | 0.354 | 0.153 | 1.627 |
| 0.7 | -0.21 | 0.361 | 0.213 | 1.717 |
| 0.8 | -0.228 | 0.372 | 0.279 | 1.676 |
| 0.9 | -0.237 | 0.383 | 0.396 | 1.679 |

图2.未来波动率变化分布特征折线图

图表, 折线图

描述已自动生成

## Days used for Mean Reversion

# 稳健性检验

# 解释

# 结论

1. Heston(1993) set the SDE of volatility as , but we can change the with the to get the SDE in this paper easily. [↑](#footnote-ref-1)