

**实 验 报 告**

**实验课程： 量 化 投 资**

姓 名： 曹宇

学 号： 3190105924

专业名称： 英语

所在学院： 外国语学院

**A股市场融券卖出量占比（rqmclzb）因子的预测效力研究**

**摘要：**本报告研究了融券卖出量占比（**rqmclzb**）因子的预测效力。本文以A股中487只股票从2017年初至2018年底的面板数据作为样本，对**rqmclzb**因子进行了分层回测和信息系数（IC）分析。分层回测结果显示，该因子具有反向特征，并且在空头筛选上展示了较强的预测能力；然而，在多头筛选方面，**rqmclzb**因子的表现并不突出。此外，IC分析进一步揭示了因子的不稳定性和有限的预测能力。报告在最后提出了对因子未来研究方向的思考。

**关键词**：融券卖出量占比，分层回测，IC分析

1. **课题具体说明**

在春节前夕，A股市场遭遇连续下挫，期间上证综指一度跌至2700点以下。面对市场的颓势，中国证监会迅速出台了一系列稳定股市的措施。这些措施主要包括：

（1）中央汇金公司对A股市场进行增持，以此直接支持市场资金面，增强市场流动性。

（2）对于融资账户的监管，出台“窗口指导”政策，即当融资账户抵押率降至130%以下时，不强制进行平仓，从而减少市场上因强制平仓带来的抛售压力。

（3）暂停新增证券转融资业务规模，防止股市的过度空头行为，减轻市场的下行压力。

这些举措自2月7日起开始实施，并迅速对市场产生了积极影响。在春节前的最后几个交易日，全市场股指实现了显著的反弹。在这一过程中，全市场整体融券规模的缩减一定程度上导致了股票整体价格的上扬。倘若从整个市场的角度来看存在这一规律，那么从个股来看是否存在同样的规律呢？基于这一好奇，作者试图在本文构建一个融券因子，并通过实证分析来考察该因子对股票未来价格走势的预测效力。

1. **数据描述**

本研究使用了从2017年初至2018年底共487天（其中因为Tushares网站的原因有9天的数据缺失）的487只股票的数据，包括这些股票的行业、流通市值、日度收益率、股本量以及日度融券卖出量（**rqmcl**）。其中，前四项数据来自何康老师所给的数据文件，日度融券卖出量来源于tushare.pro网站，数据具有权威性。

各项数据的作用如下：股本量以及日度融券卖出量（**rqmcl**）用于生成因子——融券卖出量占比 (**rqmclzb**)，股票的行业和流通市值用于对原始因子值进行中性化处理，股票的日度收益率用于IC值分析与分层回测检验。

以下部分说明选取2017年初至2018年底这一段时间的原因以及选取487只股票的原因。 首先，以两年时间段作为样本的决定主要是考虑到A股市场在过去十年间融资融券业务的逐步开放与普及。如果时间跨度设置过宽，例如自2010年起，会遇到许多股票在早期阶段尚未开展融资融券业务的情况。举例来说，根据Tushare平台的数据，2010年3月31日只有40支股票记录了融资融券交易数据，而到2017年初，该数字已激增至近1000支。如果分析区间扩展至2010年至2017年，会发现在时间序列的早期阶段，大量股票的融券卖出量比例（**rqmclzb**）因子值为零，这将对进行分层回测和信息系数（IC）分析产生不利影响。基于这些考量，本研究最终仅使用两年的时间跨度作为样本区间。

其次，选取2017年初至2018年底这一时间段的原因是这一段时间的行情较为稳定，所得结论更具有普遍性。

最后，之所以样本中有487只股票是因为，在选定的研究期间，即2017年初至2018年末，共有487只股票在此期间的非零融券卖出天数超过200天。尽管在同一时段内，有多达883只股票记录到融券卖出数据，但计算后发现，平均每天有大约40%的股票融券卖出量为零。这一现象若不加以处理，在后续进行分层回测时，会发现相当一部分股票的**rqmclzb**因子值实际上为零，这会对回测结果的准确性和解释力产生不利影响。基于以上思考，本研究对股票样本进行了筛选。在尝试设置不同的天数阈值（如100、200、250、300、400天）后，发现当阈值定为200天时，样本股票数量降至487只，同时平均每天约有88只股票的融券卖出量为零，占比下降至18%。这一比例与分层回测中最低因子值组（第1组）的股票数量相近，使得我们可以将第1组近似视为每天都没有融券卖出量的股票群体。通过这样的样本筛选标准，不仅能够更细致地评估因子在不同分层的表现，还能在回测分析中清晰地展示出当日无融券卖出量的股票与有融券卖出量股票之间的表现对比，从而增强了研究的实用性。

1. **因子构造逻辑与公式**

本研究所采用的因子是“融券卖出量占比”（**rqmclzb**）。**rqmclzb**是通过以下公式构建的：

***rqmclzb i,t = rqmcl i,t / shares i,t***

其中，***rqmcl i,t***是第t期第i个股票当日的融券卖出量，***shares i,t*** 是第t期第i个股票当日的流通股本数量。***rqmclzb i,t***反映了每只股票在当日被融券卖出的股本比例。由于该指标的大小不依赖于股本规模或者市值规模（因为是比例），因此它可以同来代表在第t期每一支股票的做空力量。

本文在此提出假说：**rqmclzb**是一个反向因子。这一假说的提出基于以下经济理论与市场观察：当一只股票前一日被融券卖出的股本比例较大时，往往会经历较大的跌幅。而若股票价格已经出现明显下跌，往往在后一个交易日仍然会呈现下跌的态势。因此，前一日**rqmclzb**越大，则后一日收益率越低。

1. **因子有效性分析**

本报告采用分层回测和IC分析对因子的有效性进行了检验。

**4.1分层回测**

**（1）反向因子**

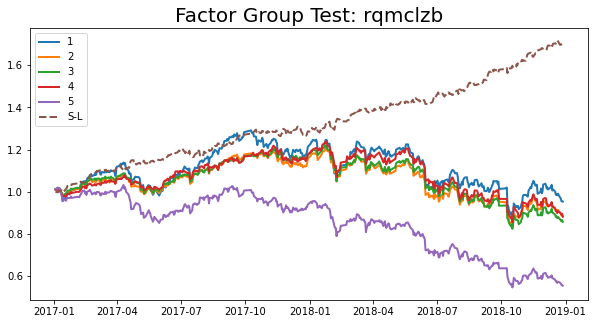
可见，图中反映了**rqmclzb**是一个反向因子。前一个交易日融券卖出的股本比例越大，则后一个交易日的收益率（也可以是累计收益率）越低。

**（2）多头筛选能力不足**

从图中可以看出，第一组、第二组、第三组、第四组（前一个交易日融券卖出量占比从小到大排列，位列0%-80%的股票）之间的后一个交易日的收益率差异并不明显，因此用该因子进行多头筛选效果不佳；与之相对，前四组和第五组后一个交易日的收益率差异（也可以是累计收益率）明显，因此**rqmclzb**空头筛选能力较好。

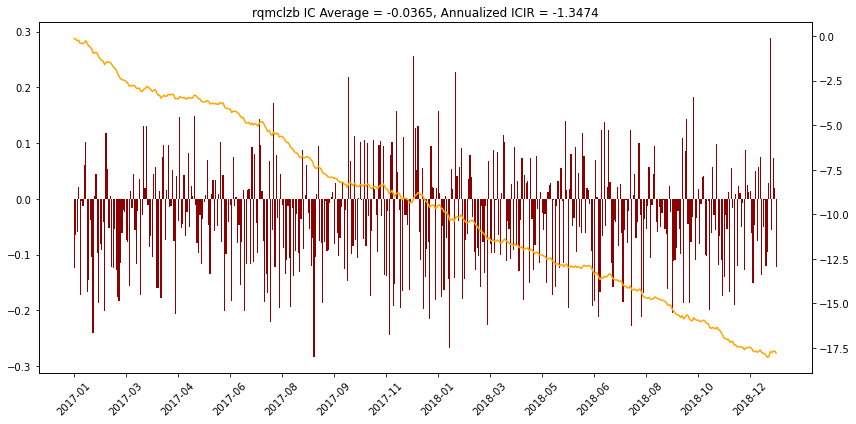
1. **整体结论**

总的来说，虽然第一组（在当日没有融券卖出量的股票）与第二三四组的收益率差异并不明显，但是总体上第一组的累计收益率曲线始终在其他组别之上。因此可以得出以下结论：在前一日倘若某一支股票有巨大的融券卖出量（相对于其股本来说），则该股票后一日的表现将会明显差于其他没有巨大的融券卖出量的股票；在前一日倘若某一支股票有较小的融券卖出量（相对于其股本来说），则该股票后一日的表现将会小幅差于那些当日没有融券卖出量的股票。



**4.2 IC分析**

如图所示，**rqmclzb**的IC均值是-0.0365，当考虑因子不稳定性的时候（即用ICIR来衡量），**rqmclzb**年化ICIR是-1.3474。与期中作业的四个因子做对比，**rqmclzb**的 IC绝对值和ICIR绝对值均是所有五个因子中最小的，同时**rqmclzb**也呈现出更大的不稳定性（很多交易日IC值为正）。因此，倘若以IC与ICIR来评价因子好坏，可以认为**rqmclzb**的预测能力弱于**return\_1m、std\_1m、turn\_1m、std\_FF3factor\_1m**。



1. **因子思考**

**rqmclzb**因子展示出了较强的空头筛选能力。作者认为，该因子的检验还可以采用另外一种思路，即在每一个交易日仅使用在当日有融券卖出量的股票样本（即不同交易日采用不同的股票池）来进行分层回测与IC分析。这样做不仅能够拉长时间跨度来检验**rqmclz**b因子，还能够将更多的股票样本纳入检验范围，以弥补本研究所具有的两大缺陷。但是这样做的一个潜在缺点是无法直观地比较没有融券卖出量和有融券卖出量股票间收益率的差异。

1. **结论**

本研究报告对A股市场的487只股票在2017年至2018年的数据进行了深入分析，探讨了融券卖出量占比（**rqmclzb**）因子的预测效力。通过采用分层回测和信息系数（IC）分析的方法，研究表明**rqmclzb**因子具有较强的空头筛选能力以及较弱的多头筛选能力。同时，**rqmclzb**因子也存在较大的不稳定性。