

Intern Record

郑骋, 香港中文大学（深圳）, 金融工程硕士
电话: 15825675534, 邮箱: zhengzc@zju.edu.cn

April 14, 2020

Contents

1 Day 1	4
1.1 Date	4
1.2 Content	4
1.3 Result	4
1.3.1 <<基本面分析在中国A股市场有用吗>>	4
1.3.2 <<分析师修正信息、基本面分析与未来股票收益>>	4
1.3.3 <<投资者如何利用财务报表盈余信息_现状、问题与启示>>	4
2 Day 2	5
2.1 Date	5
2.2 Content	5
2.3 Result	5
2.3.1 多因子研究框架	5
2.3.2 <<SEC意见信与财务造假——基于中概股危机的实证分析>>	5
3 Day 3	6
3.1 Date	6
3.2 Content	6
3.3 Result	6
3.3.1 <<中国上市公司自愿业绩预告动机研究>>	6
4 Day 4	7
4.1 Date	7
4.2 Content	7
4.3 Result	7

4.3.1	股票组问题	7
4.3.2	<<本地优势、信息披露质量和分析师预测准确性>>	7
4.3.3	<<私募股权投资与现金股利政策>>	7
5	Day 5	7
5.1	Date	7
5.2	Content	7
5.3	Result	8
5.3.1	爬虫	8
6	Day 6	9
6.1	Date	9
6.2	Content	9
6.3	Result	9
6.3.1	爬虫	9
6.3.2	<<分析师为什么选择性的发布现金流预测_基于信息需求理论的实证研究>>	9
6.3.3	<<中国上市公司现金流管理研究_兼对新兴市场和成熟市场现金流管理行为比较研究>>	9
6.3.4	<<资产减值-稳健性还是盈余管理>>	9
7	Day 7	10
7.1	Date	10
7.2	Content	10
7.3	Result	10
7.3.1	挖掘因子	10
8	Day 8	10
8.1	Date	10
8.2	Content	10
8.3	Result	10
8.3.1	挖掘因子	10
9	Day 9	11
9.1	Date	11
9.2	Content	11
9.3	Result	11
9.3.1	挖掘因子	11

10 Day 10	12
10.1 Date	12
10.2 Content	12
10.3 Result	12
10.3.1 挖掘因子	12

1 Day 1

1.1 Date

2020.03.30 Monday

1.2 Content

上午完成入职手续办理/下午阅读文献，写报告

1.3 Result

1.3.1 <<基本面分析在中国A股市场有用吗>>

创新点：六个异常因子指标的提出ABINV,ABREC,ABOREC,ABPREC,ABSGA,ABMAR；基于季度财务报表。逻辑：用季度财务报表数据更加有效，且使用“异常”指标衡量企业经营。

结论：各因子对收益率有预测作用

1.3.2 <<分析师修正信息、基本面分析与未来股票收益>>

创新点：四个因子：盈余预测修正Frev,投资评级修正Chgrec,分析师一致盈余预测水平Forecast,分析师投资评级水平Recommend,明星分析师盈余预测修正虚拟变量Fstar,明星分析师投资评级修正虚拟变量Cstar。逻辑：用分析师对股价的预测来预测股价，即相信分析师的预测是有效的。

结论：有作用

1.3.3 <<投资者如何利用财务报表盈余信息_现状、问题与启示>>

创新点：提出未预期盈余因子Sue。逻辑：股价当前水平反应了投资者的一致预期，而财报出来之后的未预期盈余则会对股价造成影响。另外，年度财务报表前后会具有反转效应，具体表现为，财报出来前股价上涨（下降），而财报出来后股价下跌（上涨），可能原因是有人提前知道财报信息。

结论：A股季度财报有正常效应，年度财报有反转效应。

2 Day 2

2.1 Date

2020.03.31 Tuesday

2.2 Content

上午开股票组分享会议/下午进行多因子研究框架培训

2.3 Result

2.3.1 多因子研究框架

制定因子—>因子回测—>因子入库—>实盘使用因子打分选股—>交易。其中一些细节说明如下：

- ①Alpha策略/多因子策略：对冲的 β 仅仅指市场收益率的系数（而不是指其他因子前面的 β 系数），因此会买多一部分股票，并卖空一部分。另外，所谓Alpha，即指具有预测效果的因子，任何数据都可以称作是因子。
- ②因子可采用单一因子，如使用基本面因子/量价因子，或者使用数据挖掘/机器学习的方法，组合因子。
- ③因子回测，首先每日使用选择出来的因子对各股票打分，其次按分数从高到低排序，并且将因子标准化，使之总和为0，且值域限制为 $[-1,1]$ （这里可采用统一减去Alpha总和/或用行业分组减去各自所在组的Alpha总和）。然后以标准化后因子值作为资金配置的比例。之后，就可以计算每日收益率，每日重复操作。最终，就能得到策略收益率的时间序列，从而计算一系列指标来评判该策略（也就是评判该Alpha因子）。
- ④实盘：实盘与因子回测不同在于，而A股难以卖空，故使用股指期货近似对冲，其他方面与回测类似。

2.3.2 <<SEC意见信与财务造假——基于中概股危机的实证分析>>

创新点：利用SEC美国证监会的信息来预测企业是否财务造假，提出因子：是否收到意见信CL，第一封意见信中问题数量Issue，审核期间意见信数量Rounds，审核时间DaySolve，文件提交的总版本数FileVersion，从事业务数量Segment，审计意见质量GoingConcern，提交文件前一年内是否解雇审计师Dismiss，是否处于频繁诉讼行业Litigation，审计师是否接受会计师协会审核Inspection，首席执行官是否持股Ceo_Own,首席执行官是否是董事会主席Ceo_Chair，首席财务官是否通过注会Cfo_CPA,流通股比例AMT%Out，公司成立至提交上市文件年数FileAge，提交上市文件是否早于第一份年报BF10K。

对于最后一个变量，可以这样思考：若能越早提交，则说明该企业的资料越正规/合规，而不需要时间去整理/审核或造假，因此提交资料的时间越快，则该企业为好企业的概率越高。

结论：这些因子对财务造假有预测作用

3 Day 3

3.1 Date

2020.04.01 Wednesday

3.2 Content

阅读文献，写报告

3.3 Result

3.3.1 <<中国上市公司自愿业绩预告动机研究>>

创新点：研究分析了上市公司自愿业绩预告的动机，包括以下因子：外部融资需求变量ExternalFinance（包含公司成长变量Growth，预期股权再融资SEO，国有控股企业State），CEO控股变量ManageShare，管理者能力信号传递变量Performance，其他控制变量（包含行业竞争指数HHI）

自愿披露指数VDI：有39个指标来衡量，该值表示企业自愿披露的程度高低。

结论：这些因素能影响公司是否自愿业绩预告

启发：可以从公司是否有自愿业绩预告，来判断该企业优劣。

4 Day 4

4.1 Date

2020.04.02 Thursday

4.2 Content

阅读文献，写报告/下午旁听了目前股票组遇到的问题

4.3 Result

4.3.1 股票组问题

①很多回测并没有进行os，导致过拟合严重

②针对突发性事件的考虑：如疫情，或其他发生的事件，因子能否捕捉这类时间的发生？还是把这类时间当作噪音？③情绪因子的收集。

4.3.2 <<本地优势、信息披露质量和分析师预测准确性>>

创新点：该文章研究了，分析师本地优势，深交所的信息披露质量信息，与分析师预测准确度的关系。

结论：本地优势up，信息披露质量up

启发：该文章中提到的深交所信息披露质量，可以作为量化因子，深交所会对每年上市公司的财报打分，分为ABCD四档，可以根据该因子给基本面因子赋权重（一定程度上反应从财报上摘取的因子的准确度）。

4.3.3 <<私募股权投资与现金股利政策>>

创新点：该文章研究了私募股权对企业发放现金股利的影响。

结论：私募股权up，外资up，国企down。启发：私募占比因子

5 Day 5

5.1 Date

2020.04.03 Friday

5.2 Content

上午写爬虫，情绪因子，百度指数/阅读文献，写报告

5.3 Result

5.3.1 爬虫

利用现有的爬虫框架，将最新A股上市股票3920+10个指数的2011年1月1日至2020年4月1日百度搜索指数爬取下来。
共3930*3379条数据。

考虑了许多细节Case:结合股票代码和名字/有名无码/无名有码/ST股票是否含星号/退市股票/搜不到等情况

6 Day 6

6.1 Date

2020.04.07 Tuesday

6.2 Content

上午将清明假期收集的百度指数进行整合，并撰写相关说明文档/阅读文献，撰写报告

6.3 Result

6.3.1 爬虫

对爬取得到的数据进行合并，并对一些遗漏股票重新爬取。

6.3.2 <<分析师为什么选择性的发布现金流预测_基于信息需求理论的实证研究>>

创新点：该文章研究了什么样的公司，分析师的报告中更有可能预测其现金流

结论：基于企业的股权性质及会计应计项目/盈余波动性/资本密集度

启发：建立以上量化因子。

6.3.3 <<中国上市公司现金流管理研究_兼对新兴市场和成熟市场现金流管理行为比较研究>>

创新点：该文章研究了企业操纵现金流的行为。

结论：在阈值点附近（0点，上一年度值，分析师预测值），企业有明显的操纵行为。

6.3.4 <<资产减值-稳健性还是盈余管理>>

创新点：该文章研究了企业资产减值的动机

结论：动机包括稳健性/经济因素/盈余管理三方面

启发：对资产减值指标进行分析。

7 Day 7

7.1 Date

2020.04.08 Wednesday

7.2 Content

开始因子量化回测旅程

7.3 Result

7.3.1 挖掘因子

熟悉回测框架，阅读GTJA191报告。（今天进度慢的原因在于linux环境配置未到位）

8 Day 8

8.1 Date

2020.04.09 Thursday

8.2 Content

从GTJA191报告中挖掘因子

8.3 Result

8.3.1 挖掘因子

挖掘前十个因子，其中部分因子效果较好，共同点：利用了Rank及成交量volume。

9 Day 9

9.1 Date

2020.04.10 Friday

9.2 Content

从GTJA191报告中挖掘因子11-20

9.3 Result

9.3.1 挖掘因子

挖掘11-20因子，部分因子胜率显著不同于50%，因此将因子反转，由于未考虑手续费，因此取得的效果较好。

10 Day 10

10.1 Date

2020.04.13 Monday

10.2 Content

从GTJA191报告中挖掘因子21-30

10.3 Result

10.3.1 挖掘因子

alpha30是目前难度最大的一个因子，因为需要用到多列回归，比较慢，这里需要攻克以下。尽量使用rolling函数，速度较快。