

事件模拟与非正常收益模型的检验力^[1]

——基于中国 A 股市场的经验检验

陈信元 江 峰

(上海财经大学 会计与财务研究院 上海 200433)

摘要: 本文分别以 1990 年 12 月至 2003 年 12 月沪市、深市和两市所有 A 股公司为抽样总体，检验了均值调整模型、市场调整模型和市场模型为基础的多种检验方法的检验力。研究发现，无论事件研究中各公司事件是否相近或重叠，都应采用市场模型为基础的非参数秩检验法。而累积非正常收益的检验也应采用市场模型为计算基础。若样本公司事件日相近或重叠，统计量的设置要考虑累积非正常收益截面数据的相关性。均值调整模型在本文所定义的各种检验方法中，均无明显优势。经敏感性测试，本文结论不变。

关键词: 事件研究、非正常收益、检验力

中国图书分类号: F2 文献识别码: A

1 引言

事件研究法是指运用股票收益率数据来测定某一特定经济事件对公司价值的影响。该方法是一个重要的分析工具，已被广泛应用于社会科学研究中。早在 1933 年，Dolly(1933)就发表了有关事件研究法的论文。但该方法的兴起和完善则是以 Ball 和 Brown(1968)对会计盈余有用性经验检验以及 Fama 等(1969)利用 CRSP 数据库针对股票分割的市场反应所进行的研究为标志，他们所提出的研究方法更是一直沿用至今。

事件研究法先利用估计期，估计出事件日的期望收益，由事件期的实际收益扣除期望收益得到非正常收益，再检验样本平均非正常收益是否显著区别于原假设。因此，估计出事件日的期望收益显得尤其重要。事件日的期望收益可以由均值调整模型、市场调整模型和市场模型来估计。Brown 和 Warner(1985)以美国 NYSE/ASE 的日收益率数据模拟发现，在某些情形下，采用均值调整模型、市场调整模型这些简单的方法也能取得很好的检验效果。Brown 和 Warner(1980)指出，在事件研究中用于计算非正常收益的不同模型所体现的特性，虽然可以通过数学推导来比较，但由此得出的差异并不一定能在实证检验结果中反映出来。因此，如何选择合适模型需要在不同时期、不同市场进行比较，才能通过针对不同情况对模型加以选择并进行改进，从而确保事件研究的严谨性。

作为新兴的证券市场，我国证券市场结构与成熟证券市场存在若干不同之处。例如，我国上市公司在上海、深圳两个分割的证券交易所(以下简称“沪市”、“深市”)上市交易，而交易规则中有针对个股的涨跌幅及卖空的限制等。特别地，Morck 等(2000)研究发现，在新兴市场经济国家(中国、马来西亚和波兰等)，80%以上的股票表现出跟随大盘的同涨同跌，远高于发达市场经济国家^[2]。那么，在具有这些特征的我国证券市场，事件研究的结果会有何种特性？与国外的研究相比，有哪些系统的差异？另外，事件研究中样本量大小、公告聚集等问题是否会影响模型与统计量的选取，目前还没有明确一致的结论。本文对这些问题进行了系统的研究。

本文第二部分介绍非正常收益模型有效性研究的理论背景以及相关文献，第三部分研究设计给出了非正常收益的度量与相关统计量的设定，第四部分是检验结果，第五部分进行敏感性检验，最后总结全文。

2 理论背景与文献回顾

Brown 和 Warner(1980)最先利用模拟方式进行事件研究法的分析。此法的优点在于研究人员可以在重复抽取实验中，观察各种非正常收益模型及检验方法的检验力^[3]。其原理是先随机抽取一组股票及一段期间，同时在事件日引入不同水平的非正常收益(即制造一个事件)，再考察各种方法的检验力。由于采用了不同的非正常收益模型和统计量，各检验方法的检验力也会有所不同。非正常收益的统计检验可分为参数检验和非参数检验，在检验过程中，统计量的设定是关键的步骤。传统上，研究人员假设事件日非正常收益的横截面数据之间独立、同分布且方差有限，根据中心极限定理，如果样本量足够大(大于 30)，非正常收益的均值便近似服从正态分布，从而可以利用统计量检验事件日非正常收益是否显著区别于零。Brown 和 Warner(1980)发展出一种新的检验统计量——原始相关调整(Crude Dependence Adjustment)。该统计量考虑了事件日非正常收益横截面数据之间的相关性。此统计量在后来的事件研究中得到了广泛的应用(例如 Masulis(1980), Dann(1981), Holthausen(1981)等学者的研究)。但究竟应该采用哪种统

计量，是否需要调整非正常收益截面数据之间的相关性，国内尚无相关文献，这也是本文的目的之一。

Brown 和 Warner(1980)利用 CRSP 数据库提供的 NYSE/ASE 股票月收益率数据，采用随机抽取的方法检验了均值调整模型、市场调整模型和市场模型的有效性。研究发现，市场模型具有显著的优势，但另两种简单的模型，尽管没有调整市场风险，但也能取得较好的效果。Brown 和 Warner(1985)采用同样的方法，检验以上三种模型在日收益率数据中的有效性，结论与月度数据类似。Campbell 和 Wasley(1993)利用 NASDAQ 日收益率数据进行模拟分析后，发现以市场模型为基础，采用秩检验方法的检验力最佳。

陈汉文和陈向民(2002)以我国 1990-2000 年沪市、深市交易所上市交易的 A 股公司为抽样总体，采用模拟方法对上述三种非正常收益模型进行了经验比较，结果显示市场模型的局限和均值调整模型的某些优势。本文在第五部分敏感性检验中以同一抽样总体重复这一过程，却未发现这一现象。

3 研究设计

3.1 数据与样本构建

3.1.1 数据

本文所设定的抽样总体为 1990 年 12 月-2003 年 12 月在沪市、深市上市交易的 A 股公司。其中沪市 774 家，深市 504 家，共计 1278 家。本文所用股票收益率与指数日收益率数据来自香港理工大学中国会计与金融研究中心和深圳市国泰安信息技术有限公司开发的 CSMAR 数据库。

3.1.2 样本构建

我们沿用 Brown 和 Warner(1985)随机抽样方法，在所设定的总体中进行 250 次有放回的随机抽样，每次抽取 50 只股票。每取一只股票，同时随机赋予一个假想的事件日，据此构成估计期和事件期。

我们定义第 0 日为所抽取股票的事件日。对于抽取的股票，估计期与事件期之和最大为 250 日(开始于-244 日，结束于+5 日)，估计期为-244 日至-6 日，事件期为-5 日至+5 日。进入样本的股票在 250 个交易日中，至少应存在 30 个日收益率，且最后 20 日没有缺失数据(例如股票暂停交易)。另外，上市首日的收益率与上市后日收益率缺乏可比性，因此，估计期均未包括该日。

本文对每个模型的检验都从三个总体分别进行了有放回的随机抽样。这三个总体分别是沪市 A 股、深市 A 股和沪市深市所有 A 股总和(以下简称“两市”)。

3.2 非正常收益的度量

令 $R_{i,t}$ 表示股票 i 在 t 日的日收益率, $A_{i,t}$ 为股票 i 在 t 日的非正常收益率, 样本中的每一只股票在事件期内每一交易日的非正常收益由以下三种模型决定:

(1) 均值调整模型

$$A_{i,t} = R_{i,t} - \bar{R}_i$$

$$\bar{R}_i = \frac{1}{239} \sum_{t=-244}^{-6} R_{i,t}$$

其中 \bar{R}_i 为股票 i 在估计期(-244,-6)内的日平均收益率。^[4]

(2) 市场模型

$$A_{i,t} = R_{i,t} - \hat{\alpha}_i - \hat{\beta}_i R_{m,t}$$

其中 $\hat{\alpha}_i$ 和 $\hat{\beta}_i$ 是估计期内的 OLS 估计值。

(3) 市场调整模型

$$A_{i,t} = R_{i,t} - R_{m,t}$$

其中 $R_{m,t}$ 为 t 日的指数收益率。如果随机抽取的对象是沪市 A 股, 取上证指数作为指数收益率的度量; 如为深市 A 股, 取深圳成分指数; 如为两市 A 股, 则取考虑现金再投资的综合日市场收益率(等权平均法)。

3.3 待检验的统计量

我们检验了每个样本运用以上三种方式计算出的非正常收益的统计显著性。待检验的原假设为事件日的平均非正常收益等于零。第一个检验统计量根据样本事件日平均非正常收益除以样本均值的标准差计算得出，而样本均值的标准差则是根据估计期非正常收益的数据估计而来。该统计量由 Brown 和 Warner(1980)首先提出，考虑了事件日非正常收益横截面数据之间的相关性，表达

$$\text{式为: } t_1 = \frac{\bar{A}_t}{\hat{S}(\bar{A}_t)}$$

$$\text{其中, } \bar{A}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} A_{i,t} \quad \hat{S}(\bar{A}_t) = \sqrt{\frac{\sum_{t=-244}^{-6} (\bar{A}_t - \bar{\bar{A}})^2}{238}} \quad \bar{\bar{A}} = \frac{1}{239} \sum_{t=-244}^{-6} \bar{A}_t$$

如果 \bar{A}_t 满足独立、同分布和正态分布的条件，那么在原假设成立的情况下，待检验的统计量服从 t 分布。由于该例的自由度大于 200， t_1 近似服从标准正态分布。

若不考虑事件日非正常收益之间的相关性，并且假设事件日非正常收益满足独立、同分布有限方差的条件，根据中心极限定理，

$$\text{在样本容量大于 30 时, 非正常收益的均值便近似服从正态分布, 从而有第二个检验统计量: } t_2 = \frac{\bar{A}_t}{\hat{S}'(\bar{A}_t)}$$

$$\text{其中, } \bar{A}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} A_{i,t} \quad \hat{S}'(\bar{A}_t) = \sqrt{\frac{\frac{1}{N_t-1} \sum_{i=1}^{N_t} (A_{i,t} - \bar{A}_t)^2}{N_t}}$$

以上两个检验方法是参数检验，本文还检验了非参数方法——秩检验的检验力，该方法由 Corrado(1989)提出，它根据非正常收益在时间序列数据中相对顺序关系来考察非正常收益是否显著区别于零。在此方法下，首先要将每只股票非正常收益时间序列数

据转为秩, $k_{i,t} = \text{rank}(A_{i,t}, t = -244, \dots, 5)$, 统计量表达式为: $t_3 = \frac{\frac{1}{N_t} \sum_{i=1}^{N_t} (k_{i,0} - E(k_i))}{s(k)}$

其中 $E(k_i)$ 表示第 i 只股票秩期望值, 等于 $0.5T_i + 0.5$, T_i 为第 i 只股票估计期和事件期非缺失日收益率的个数。

$$s(k) = \sqrt{\frac{1}{250} \sum_{t=-244}^{+5} \left(\frac{1}{N_t} \sum_{i=1}^{N_t} (k_{i,t} - E(k_i)) \right)^2}$$

涉及累积非正常收益(CAR)的统计检验时, 本文采用了两个统计量。第一个统计量也是由 Brown 和 Warner (1985) 提出的, 与事

件日非正常收益的检验类似, 该统计量考虑了累积非正常收益横截面数据之间的相关性。统计量表达式为: $t_4 = \frac{\sum_{t=-5}^{+5} \bar{A}_t}{\left(\sum_{t=-5}^{+5} \hat{S}^2(\bar{A}_t) \right)^{\frac{1}{2}}}$,

其中 \bar{A}_t 、 $\hat{S}(\bar{A}_t)$ 与 t_1 的计算方法一致。

若不考虑累积非正常收益之间的相关性, 并且假设累积非正常收益满足独立、同分布和有限方差的条件, 根据中心极限定理,

在样本容量大于 30 时, 累积非正常收益的均值便近似服从正态分布, 从而有: $t_5 = \frac{\sum_{t=-5}^{+5} \bar{A}_t}{\hat{S}(\sum_{t=-5}^{+5} \bar{A}_t)}$

$$\text{其中, } \hat{S}\left(\sum_{t=-5}^{+5} \bar{A}_t\right) = \sqrt{\frac{1}{N_t - 1} \sum_{i=1}^{N_t} \left(\sum_{t=-5}^{+5} A_{i,t} - \sum_{t=-5}^{+5} \bar{A}_t \right)^2}$$

3.4 第一类错误(Type I Error)与第二类错误(Type II Error)

由于每一个样本中的每一只股票的事件日均是随机抽取而来,若事件日未引入非正常收益,则样本事件日平均非正常收益应等于零,即原假设为真。如果原假设被拒绝,则为第一类错误——拒真错误。

当样本中每只股票在事件日引入正的非正常收益,事件日平均非正常收益应大于零,如果原假设没有被拒绝,则为第二类错误——取假错误。检验力则是指检验方法正确地拒绝原假设的概率,在事件日人为地引入非正常收益即是为了判断检验方法的检验力,当犯第一类错误概率较低时,若检验力愈强,则该检验方法愈好。

4 检验结果

4.1 非正常收益的统计特征

我们首先考察了未引入非正常收益情况下各种模型计算结果的统计特征。表 1 Panel A, Panel B, Panel C 分别给出了以两市、沪市和深市为抽样总体所有样本观测(共计 12500 个)汇总后的估计期收益和非正常收益的分布状况。Panel D 显示了样本量为 5000 的标准正态分布随机数的数字特征。

表 1 样本汇总后的估计期收益和非正常收益分布状况

Panel A 两市估计期收益和非正常收益分布状况(观测数: 12500)

变量	均值	标准差	偏度	峰度	学生全距
日收益率	0.0001	0.0243	0.37	5.84	7.28

均值调整非正常收益	0.0000	0.0242	0.39	5.95	7.32
市场模型非正常收益	0.0000	0.0176	0.68	7.12	7.89
市场调整非正常收益	-0.0001	0.0178	0.67	7.02	7.86

Panel B 沪市估计期收益和非正常收益分布状况(观测数: 12500)

变量	均值	标准差	偏度	峰度	学生全距
日收益率	0.0000	0.0236	0.42	6.00	7.39
均值调整非正常收益	0.0000	0.0234	0.42	6.04	7.41
市场模型非正常收益	0.0000	0.0169	0.59	7.36	8.05
市场调整非正常收益	-0.0001	0.0172	0.57	7.24	8.03

Panel C 深市估计期收益和非正常收益分布状况(观测数: 12500)

变量	均值	标准差	偏度	峰度	学生全距
日收益率	0.0002	0.0251	0.35	5.60	7.12
均值调整非正常收益	0.0000	0.0250	0.34	5.70	7.19
市场模型非正常收益	0.0000	0.0182	0.69	7.01	7.82
市场调整非正常收益	0.0001	0.0184	0.66	6.76	7.75

Panel D 标准正态分布随机数的数字特征(样本容量: 5000)

均值	偏度	峰度	学生全距
0.000102	0.03	3.06	7.16

观察表 1 Panel A, B, C, 在未引入非正常收益时, 日收益率数据和用各种模型计算得出的非正常收益都偏离了标准正态分布^[5]。用市场模型计算得出的非正常收益偏度、峰度和学生全距在三种模型中最大, 因此偏离正态分布也最严重; 而均值调整模型产生的非正常收益偏离程度最弱, 这一现象同时存在于沪市和深市。

我们注意到, 无论是沪市、深市, 还是两市汇总后的估计期日收益率标准差均小于 Brown 和 Warner(1985)的表 1 对应结果。这一结果并不反常, 日收益率的标准差可作为股票总风险的度量, 而宋逢明、江婕(2003)的研究发现中国股票市场的整体波动性在 1998、1999 年开始与 S&P500 的整体波动性持平, 2000 年及之后低于 S&P500 的整体波动性, 此间中国股票市场的总体风险并不比海外成熟市场高。

需特别指出的是，三种方法计算的非正常收益标准差，均小于 Brown 和 Warner(1985)的计算结果。其中市场模型和市场调整模型的计算结果与均值调整模型相比，更显偏低。此发现非常重要，因为统计值与非正常收益的标准差负相关，若某一模型计算出的非正常收益标准差较低，在事件日引入非正常收益时，会导致统计值偏高，原假设就比较容易被拒绝，此方法检验力也就越强(前提是犯第一类错误的概率在允许范围内)。

我们认为，产生上述现象的原因在于我国的股票同涨同跌的特性。Morck 等(2000)研究发现，在新兴市场国家(中国、马来西亚和波兰等)，80%以上的股票表现出跟随大盘的同涨同跌，远高于发达市场国家。宋逢明、江婕(2003)的研究也发现系统风险是中国股票市场的主要风险。由于市场因素对股票收益率具有较强的影响力，那么扣除市场因素得到的非正常收益离散程度应小于未扣除市场因素的模型——即均值调整模型，并且小于发达市场国家。

表 2 两市 250 个样本均值的统计特征

样本量	变量	均值	标准差	偏度	峰度	学生全距
5	日收益率	0.0002	0.0080	0.24	3.58	7.75
	均值调整非 正常收益	0.0001	0.0107	0.20	3.64	8.83
	市场模型非 正常收益	0.0002	0.0079	0.37	3.84	8.25
	市场调整非 正常收益	0.0000	0.0082	0.27	3.74	8.32
20	日收益率	0.0002	0.0041	0.32	3.05	5.45
	均值调整非 正常收益	0.0001	0.0054	0.18	3.28	6.55

	市场模型非 正常收益	0.0002	0.0040	0.00	3.01	5.67
	市场调整非 正常收益	0.0000	0.0041	0.02	2.96	5.90
50	日收益率	0.0002	0.0025	-0.27	2.86	4.78
	均值调整非 正常收益	0.0001	0.0035	0.18	2.65	5.00
	市场模型非 正常收益	0.0002	0.0025	0.22	2.88	5.53
	市场调整非 正常收益	0.0000	0.0026	0.26	2.88	5.55

表2显示了未引入非正常收益情形下每个样本事件日平均收益与平均非正常收益的统计特征。首先，样本均值的偏度与峰度明显小于个股，说明样本均值偏离正态的程度较个股已经有了明显的改善，这与中心极限定理的预期相符。第二，事件日平均非正常收益接近于零，并且三种模型之间无明显差异，这与陈汉文和陈向民(2002)的发现并不一致，因此，不能得出哪个模型犯第一类错误可能性较大的结论^[6]。第三，均值调整模型所计算的事件日样本均值标准差仍然大于其他两个模型，如前所述，均值模型的这一特征会导致检验力低于其他两种模型。

4.2 待验统计量的特征

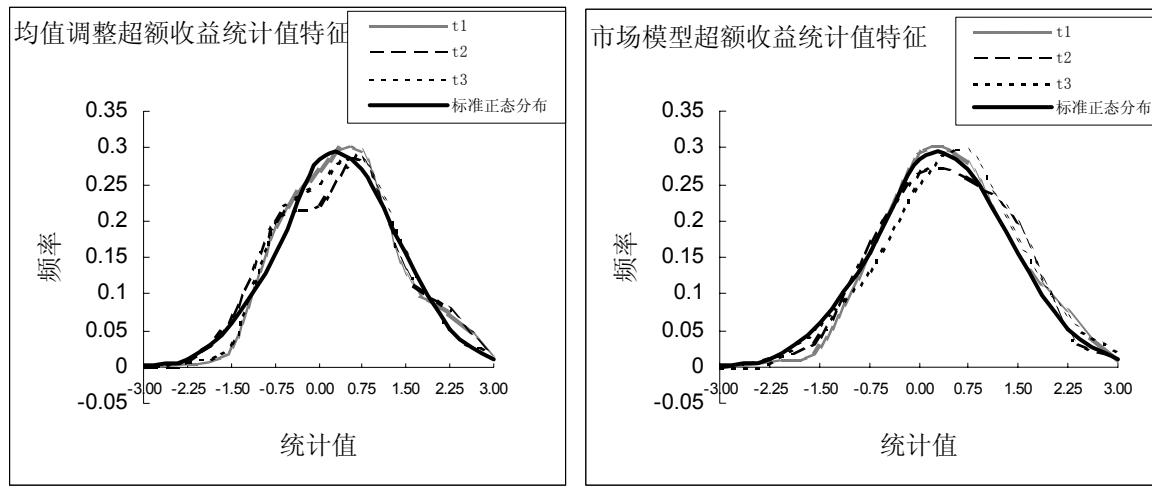


图 1 以均值调整模型和市场模型为基础的待验统计量分布状况

图 1 比较了以均值调整模型与市场模型为基础的待验统计量分布状况。直观上，均值调整模型计算出的统计值偏离正态分布更为严重。表 3 描述了三种模型的待验统计量统计特征。其中均值调整模型计算出的每一个待验统计量的偏度均大于其他模型，而峰度都小于其他模型对应统计量。本文的这一现象与陈汉文和陈向民(2002)的研究结果存在差异。^[7]

表 3 待验统计量统计特征

模型	统计量	均值	标准差	t 值	偏度	峰度	学生全距
均值调整	t1	0.0337	0.9714	0.55	0.23	2.75	5.26
	t2	0.0022	1.0352	0.03	0.13	2.52	5.06
	t3	0.0307	0.9871	0.49	0.15	2.73	5.04

市场模型	t1	0.0692	0.9285	1.18	0.19	2.87	5.55
	t2	0.0294	0.9700	0.48	-0.03	2.83	5.22
	t3	0.1173	0.9868	1.88	0.03	2.97	5.09
市场调整	t1	0.0107	0.9510	0.18	0.21	2.87	5.66
	t2	-0.0462	0.9910	-0.74	-0.02	2.71	5.21
	t3	0.1072	1.0068	1.68	-0.05	2.79	5.25

注：1. 抽样总体为 1990 年 12 月-2003 年 12 月在两市交易的所有 A 股

2. t1、t2 和 t3 分别代表本文第三部分定义的三种检验事件日非正常收益的统计量(下同)。

4.3 非正常收益检验方法的检验力

表 4 列出了事件日分别引入 0.5%, 1%, 2% 非正常收益时各种检验方法的检验力。其中，所设计检验均为显著性水平为 5% 的单尾检验(下同)。检验力是根据检验方法拒绝原假设的频率计算而来。以表 4 Panel A 为例，在事件日引入 0.5% 非正常收益时，利用均值调整模型和 t1 统计量检验发现 250 个样本中，111 个样本拒绝了原假设，我们便将拒绝原假设的样本数占样本总数的百分比作为拒绝频率(此处为 44.4%)。

表 4 引入非正常收益后各检验方法的检验力 (%)

Panel A 以 1990 年 12 月-2003 年 12 月在两市交易的所有 A 股为抽样总体

模型	统计量	事件日引入的非正常收益			
		0	0.005	0.01	0.02

均值调整	t1	5.6	44.4	88.4	100
	t2	6.4	43.6	92.0	100
	t3	5.2	59.6	99.2	100
市场模型	t1	4.8	62.4	99.6	100
	t2	4.0	66.0	99.6	100
	t3	7.6	87.6	100	100
市场调整	t1	5.2	56.8	99.2	99.6
	t2	4.4	64.4	99.2	100
	t3	7.2	84.8	100	100

Panel B 以 1990 年 12 月-2003 年 12 月在沪市交易的 A 股为抽样总体

模型	统计量	事件日引入的非正常收益			
		0	0.005	0.01	0.02
均值调整	t1	5.2	46.0	90.4	100
	t2	5.2	48.0	92.0	100
	t3	4.0	58.8	98.0	100
市场模型	t1	6.0	61.2	98.0	100

	t2	4.4	66.0	98.4	100
	t3	4.8	86.8	100	100
市场调整	t1	6.8	57.2	98.4	100
	t2	6.0	61.2	97.6	100
	t3	6.4	79.2	100	100

Panel C 以 1991 年 4 月-2003 年 12 月在深市交易的 A 股为抽样总体

模型	统计量	事件日引入的非正常收益			
		0	0.005	0.01	0.02
均值调整	t1	4.8	37.6	84.8	100
	t2	3.6	39.6	88.0	100
	t3	4.8	51.6	95.2	100
市场模型	t1	3.6	52.8	97.6	100
	t2	2.8	58.4	96.4	100
	t3	4.8	82.0	100	100
市场调整	t1	4.8	58.0	97.6	100
	t2	2.8	66.0	98.8	100

t3	5.2	79.2	100	100
----	-----	------	-----	-----

观察表 4 Panel A 不难发现，在未引入事件时，各模型拒绝原假设的频率在 4%至 7.6%之间，三种模型及各检验方法之间并无明显差异。当事件日引入了非正常收益时，市场模型、市场调整模型均能以较高的检验力拒绝原假设。同时，检验力还高于 Brown 和 Warner(1985)的研究结果。这与本节第 1 部分的预期相一致。另外，无论采用何种统计方法，市场模型在绝大部分情况下表现均为最佳。市场调整模型稍逊一筹，但仍明显优于均值调整模型。^[8]对于每个模型，我们都采用了三种统计检验方法，发现不考虑事件日非正常收益之间的相关性的检验方法(t2)检验力高于相对复杂的方法——原始相关调整法(t1)。而秩检验(t3)明显优于其他两种检验方法。上述所有现象均分别存在于沪市、深市(见表 4 Panel B,C)。

4.4 累积非正常收益检验方法的检验力

对累积非正常收益进行检验时，市场模型和市场调整模型犯第一类错误的频率稍高于均值调整模型，但并未达到无法容忍的地步。引入事件后，市场模型的表现依然稳定，检验力在绝大多数情况下均大于其他两类模型。同时，不考虑累积非正常收益之间的相关性的检验方法(t5)检验力高于相对复杂的方法(t4)。这一现象也分别存在于沪市、深市(见表 5 Panel B,C)。

表 5 累积非正常收益检验方法的检验力 (%)

Panel A 以 1990 年 12 月-2003 年 12 月在两市交易的所有 A 股为抽样总体

模型	统计量	事件日引入的非正常收益			
		0	0.005	0.01	0.02
均值调整	t4	4.8	12.0	18.8	48.0
	t5	3.2	12.0	22.0	52.4
市场模型	t4	7.2	20.0	36.8	81.2

	t5	6.8	20.0	41.6	85.2
市场调整	t4	7.6	16.8	33.2	79.6
	t5	7.6	20.8	38.8	85.2

Panel B 以 1990 年 12 月-2003 年 12 月在沪市交易的 A 股为抽样总体

模型	统计量	事件日引入的非正常收益			
		0	0.005	0.01	0.02
均值调整	t4	2.4	9.6	21.2	55.6
	t5	3.2	10.4	20.8	57.2
市场模型	t4	3.6	14.0	26.4	66.4
	t5	3.6	10.8	30.4	66.0
市场调整	t4	2.4	8.0	19.6	60.0
	t5	2.8	7.6	23.2	61.6

Panel C 以 1991 年 4 月-2003 年 12 月在深市交易的 A 股为抽样总体

模型	统计量	事件日引入的非正常收益			
		0	0.005	0.01	0.02
均值调整	t4	3.6	6.4	13.2	42.8

	t5	3.2	5.6	13.2	44.8
市场模型	t4	5.6	14.0	28.0	70.4
	t5	5.2	13.2	30.4	74.0
市场调整	t4	4.8	12.8	30.8	73.2
	t5	5.6	14.4	35.6	78.0

注: t4 和 t5 分别代表本文第三部分定义的两种检验累积非正常收益的统计量(下同)。

5 敏感性检验

为使本文的结论更具可靠性, 我们改变了一些抽取条件, 作如下敏感性测试。

(1) 小样本情形

首先, 我们将 250 个样本容量由 50 缩小为 20, 表 6 显示了各检验方法的检验力。

表 6 样本量为 20 时各检验方法的检验力 (%)

模型	统计量	事件日引入的非正常收益			
		0	0.005	0.01	0.02
均值调整	t1	6.8	22.8	59.2	98.8
	t2	5.2	24.4	62.8	97.2
	t3	5.6	31.6	75.2	99.6

	t4	4.4	8.8	15.2	31.2
	t5	4.4	9.2	14.0	33.6
市场模型	t1	4.0	34.8	80.8	100
	t2	1.6	39.6	84.4	100
	t3	4.4	55.2	96.0	100
	t4	7.6	13.2	21.6	44.8
	t5	6.8	11.6	21.6	50.0
市场调整	t1	3.6	32.4	81.2	99.6
	t2	2.8	34.8	78.8	99.6
	t3	4.4	52.0	94.8	100
	t4	7.6	11.6	20.4	44.0
	t5	6.8	15.6	24.8	51.2

注：以 1990 年 12 月-2003 年 12 月在两市交易的所有 A 股为抽样总体

观察表 6 可以发现，小样本的检验力相对于大样本有较为明显的下降，但市场模型及秩检验在各种检验事件日非正常收益的方法中仍然表现最好。而累积非正常收益的检验结果，也与本文第四部分结论一致。因此，样本容量的变化并未改变本文的结论。

(2) 公告聚集

至此，本文所有结论均建立在事件日随机抽取的基础之上，然而时间的发生可能存在聚集的情况，例如某些上市公司可能在同

一交易日公布盈余信息。由于事件日相隔很近或发生重叠，事件日非正常收益之间就可能存在相关关系。中心极限定理的前提条件之一——独立性便不复存在。那么，不同检验方法的检验力会受到不同程度的影响，表 7 显示了样本事件日发生重叠时各检验方法的检验力。

表 7 样本事件日重叠时各检验方法的检验力(%)

模型	统计量	事件日引入的非正常收益			
		0	0.005	0.01	0.02
均值调整	t1	3.6	5.2	8.4	20.8
	t2	25.2	48.8	67.2	83.2
	t3	4.0	6.8	11.2	30.0
	t4	4.0	4.0	4.4	5.2
	t5	34.0	43.6	48.4	52.8
市场模型	t1	4.4	48.8	78.0	99.6
	t2	8.0	63.6	92.4	99.6
	t3	4.4	53.6	88.4	99.2
	t4	6.0	12.4	26.4	56.4
	t5	15.2	28.4	47.2	75.6
市场调整	t1	3.6	40.8	72.4	88.4

t2	4.4	62.0	90.8	99.6
t3	4.4	50.8	88.4	98.4
t4	2.0	9.6	22.4	48.0
t5	6.8	16.0	38.0	70.4

注：以 1990 年 12 月-2003 年 12 月在两市交易的所有 A 股为抽样总体

从表 7 中可见，对均值调整模型而言，t2、t5 方法犯第一类错误频率分别为 25.2% 和 34%，远超过 5% 的显著性水平。尽管 t1、t3 和 t4 犯第一类错误频率较低，但检验力也大幅降低，因此，均值调整模型已不适用于公告聚集的事件研究。对市场模型而言，t2、t5 犯第一类错误频率有所上升，但其检验力依然保持在较高水平，虽然 t1、t3 和 t4 犯第一类错误频率较低，但检验力却有一定程度的下降。同样，市场调整模型的检验力也有所下降。

尽管公告聚集导致各模型检验力有所下降，但市场模型在其中表现依然最佳。因此，一个比较可靠的结论是，事件研究中各公司事件相隔很近或者重叠时，应当采用市场模型为基础的秩检验法检验事件日市场是否对信息作出反应。而累积非正常收益的检验也应采用市场模型为基础，且统计量的设置要考虑累积非正常收益截面数据之间的相关性(t4)以降低犯第一类错误的概率。

(3) 抽取期间

本文以 1990 年 12 月-2003 年 12 月在沪市、深市上市交易的 A 股公司为抽样总体，采用类似的方法，结论却与陈汉文和陈向民(2002)不尽相同。是否是抽取期间的差异导致结论的差异呢？我们以同一抽取期间，计算了不同检验方法的检验力。表 8 显示了计算结果。

表 8 引入非正常收益后各检验方法的检验力(%)

模型	统计量	事件日引入的非正常收益			
		0	0.005	0.01	0.02

均值调整	t1	2.8	34.0	80.4	100
	t2	3.6	35.6	86.0	100
	t3	4.8	47.2	92.8	100
	t4	3.6	8.0	14.4	40.4
	t5	3.2	7.6	18.0	45.2
市场模型	t1	6.0	49.6	92.8	100
	t2	5.2	52.8	94.4	100
	t3	8.0	68.0	98.8	100
	t4	4.8	14.0	25.2	62.0
	t5	4.8	15.2	29.2	67.6
市场调整	t1	4.8	45.2	90.8	98.4
	t2	4.8	50.8	93.6	100
	t3	7.6	66.0	99.2	100
	t4	3.6	9.2	20.0	57.2
	t5	4.4	10.0	23.2	66.4

注：以 1990 年 12 月-2001 年 12 月在两市交易的所有 A 股为抽样总体

表 8 中各检验方法的检验力较表 4 Panel A 和表 5 Panel A 对应方法有所下降，但本文的基本结论不变，即市场模型为基础的

各检验方法检验力高于其他两种模型，均值调整模型的检验力最弱。在检验事件日非正常收益的各种方法中，秩检验的检验力最强。而累积非正常收益的检验也应以市场模型为基础。

(4) 其他敏感性检验

我们测试了 1% 的显著性水平下的各检验方法的检验力，还测试了采用以市值加权市场收益率的非正常收益模型，但测试结果均未改变本文的基本结论。

6 总结

尽管我国股票市场仍处于发展的初级阶段，但选择恰当的非正常收益模型和检验统计量可明确地判断市场是否对信息作出了反应。抽样时段的变化、公告聚集效应以及样本容量的改变均会导致各检验方法检验力的变化。尽管如此，我们仍然可以得出一个基本结论：无论事件研究中各公司事件是否相近或重叠，都应采用市场模型为基础的非参数秩检验法以检验事件日市场是否对信息作出反应。而累积非正常收益的检验也应以市场模型为计算基础，若样本公司事件日相近或重叠，统计量的设置要考虑累积非正常收益截面数据之间的相关性。市场调整模型检验力稍弱于市场模型，可以作为市场模型的替代。但均值调整模型在本文所定义的各种检验方法中，均无明显优势，不可用于计算非正常收益。

参考文献

- [1] Brown, S.J. and J.B. Warner, 1980, Measuring security price performance, *Journal of Financial Economics* ,8, 205-258.
- [2] Brown, S.J. and J. B. Warner, 1985, Using daily stock return: The case of event studies, *Journal of Financial Economics* 14, 3-31.
- [3] Campbell, C.J. and C. E. Wasley, 1993, Measuring security price performance using daily NASDAQ returns, *Journal of Financial Economics* 33, 1993, 73-92.
- [4] Dann,L.,1981,Common stock repurchases: An analysis of returns to bondholders and stockholders, *Journal of Financial Economics* 9, 113-138.

- [5] Fama E. F., L. Fisher, M. C. Jensen and R. Roll, 1969, The adjustment of stock prices to new information, International Economic Review 10, 1-21.
- [6] Holthausen, R., 1981, Evidence on the effect of bond covenants and management compensation contracts on the choice of accounting techniques: The case of the depreciation switchback, Journal of Accounting and Economics 3, 73-109.
- [7] Masulis, R., 1980, The effects of capital structure change on security prices: A study of exchange offers, Journal of Financial Economics 8, 139-177.
- [8] Morck Randall., Yeung Bernard, and Wayne Yu, 2000, The information content of stock markets: why do emerging markets have synchronous stock price movements?. Journal of Financial Economics 58, 215-260.
- [9] 陈汉文、陈向民, 2002:《证券价格的事件性反应——方法、背景和基于中国证券市场的应用》,《经济研究》第1期。
- [10] 宋逢明、江婕, 2003:《中国股票市场波动性特性的实证研究》,《金融研究》第四期。

Event simulation and the test power of abnormal return model

——Empirical research on China A-share stock market

Jiang Feng Chen Xin-yuan

(Shanghai University of Finance and Economics, Shanghai 200433, China)

Abstract: We evaluate the performance of alternative test statistics in event studies which include Chinese daily security returns. Whether there is clustering in event dates, we find that the nonparametric rank statistic performs the best overall in our samples; we recommend its use with market model abnormal returns to test event day abnormal returns. For tests over the long interval, the test statistic should consider the cross-sectional dependence in the security abnormal returns. Sensitivity tests do not change our conclusions.

Keywords: Event study, Abnormal returns, Power

作者简介: 江峰 (1979-), 男 (汉族), 江苏淮阴人, 上海财经大学会计学院博士研究生。

陈信元(1964-), 男(汉族), 浙江宁波人, 上海财经大学会计与财务研究院教授。

[¹] 本文是教育部人文社会科学研究基地重大项目的后续成果。本文的写作得益于香港科技大学陈建文教授的启发和指导, 会计与财务研究院张人骥教授给出了有益的建议。但错误与疏漏由作者负责。

[²] 同涨同跌的现象可能会导致非正常收益的离散程度小于发达市场国家。因此, 在事件日非正常收益相等的情形下, 利用我国 A 股数据, 能够相对容易检验出非正常收益是否显著区别于零。本文的检验结果最终证明了此假说。

[³] 严格地讲, 检验力是检验方法正确拒绝原假设的概率。本文所述的非正常收益模型的检验力是针对同一种检验方法而言的。当我们指出某种非正常收益模型的检验力最强, 是指某种检验方法利用该非正常收益模型, 正确拒绝原假设的概率高于利用其他非正常收益模型。

[⁴] 等式(2)及随后的等式仅当股票在估计期无缺失数据才成立。如估计期缺失了数据, 我们在计算模型参数时剔除了该日。

[⁵] 标准正态分布偏度为 0, 峰度为 3。

[⁶] 作者在文中指出: “事件日的期望非正常收益接近与 0, 但是由市场模型计算出的结果较大一些, 那么在实际采用市场模型的事件研究中, 犯一次类错误的可能性相应会大一些”。见《经济研究》2002 年第 1 期, 第 44 页。

[⁷] 作者在文中指出: “相比较而言, 由均值调整模型计算出的待验统计量更接近于标准正态分布”。见《经济研究》2002 年第 1 期, 第 45 页。

[⁸] 陈汉文和陈向民(2002)的研究结果显示以市场模型为基础的检验力最强, 均值调整模型次之, 最后是市场调整模型。同时还需指出, 该文表 1 显示, 用均值调整模型计算的非正常收益标准差大于其他两种模型。因此, 均值调整模型检验力就很可能弱于其他两类模型, 但其研究结果却未能与之保持内一致。见《经济研究》2002 年第 1 期, 第 46 页表 1, 第 46 页表 4 与表 5。