

# 金融工程

证券研究报告

2017 年 11 月 20 日

## 海外文献推荐 第 22 期

### 估计 Beta 值

自从 Sharpe (1964), Lintner (1965) 和 Mossin (1966) 以及 Ross (1976) 的套利定价理论 (APT) 发展了关于资本资产定价模型 (CAPM) 中资产与其他风险因素的关系, 这些在金融经济学中起到了至关重要的作用。对于许多应用, 如资产定价, 投资组合选择或风险管理, 市场 beta 是唯一最重要的参数。然而, beta 因素不是直接可观察的, 因此需要估计。本文的主要贡献在于, 第一个对广泛的市场 beta 估算技术的性能提供全面和深入的实证研究, 包括几个历史的时间序列模型, 和期权隐含估计方法。此外, 我们提出了一个新的 beta 估计方法, 用于修正波动率风险溢价的期权隐含波动率。

### 价值、规模、动量、股利回报以及波动率因子在中国 A 股市场的表现

因子选股已经成为大多数机构投资者的重要投资方法, 该方法起源于 CAPM 理论, 之后又提出了著名的 Fama-French 三因子模型, 引入了价值因子和市场规模因子对股票或投资组合收益率的解释, 而且认为三因子模型对股票或投资组合收益率的解释力度要比 CAPM 模型更好。这充分说明了还存在其他因子可以更好的解释市场中股票或投资组合的收益率。但是, 不同的金融市场也许会得到不同的因子溢价, 我们分析中国市场中因子的表现, 因为中国 A 股市场有着它独特的一面。

### 作者

**吴先兴** 分析师  
SAC 执业证书编号: S1110516120001  
wuxianxing@tfzq.com  
18616029821

**阚文超** 联系人  
kanwenchao@tfzq.com  
18717948990

### 相关报告

- 1 《金融工程: 金融工程-量化择时及多因子跟踪周报》 2017-11-19
- 2 《金融工程: 金融工程-基金发行周报-2017-11-17》 2017-11-19
- 3 《金融工程: 金融工程-估值与基金重仓股配置监控半月报 2017-11-18》 2017-11-18

**风险提示:** 本报告不构成投资建议。



## 内容目录

估计 Beta 值.....	3
1. 简介 .....	3
2. 数据与方法 .....	3
3. 实证结果 .....	5
4. 结论 .....	7
价值、规模、动量、股利回报以及波动率在中国 A 股市场的分析 .....	8
1. 中国 A 股市场的独特性 .....	8
2. 数据以及因子的定义 .....	8
3. 研究方法 .....	9
4. 研究结果 .....	10

## 图表目录

图 1: 摘要统计和样本关联性 .....	5
图 2: 估计误差: 6 个月的水平下, 日数据 .....	6
图 3: 估计误差: 不同水平下的日数据 .....	6
图 4: 估计误差: 6 个月水平下, 日数据 (更深一步的时间序列模型) .....	7
图 5: MSCI 中国指数成分股和世界 MSCI 指数成分股的市值大小 .....	9
图 6: 因子的定义 .....	9
图 7: 因子的显著性检验 .....	10
图 8: 模型的解释力度 .....	10

## 估计 Beta 值

文献来源: Fabian Hollstein and Marcel Prokopczuk, Estimating Beta, Financial Analyst Journal, Vol. 68, No. 2(March/April 2012), pp.30-50

推荐原因: 自从 Sharpe (1964), Lintner (1965) 和 Mossin (1966) 以及 Ross (1976) 的套利定价理论 (APT) 发展了关于资本资产定价模型 (CAPM) 中资产与其他风险因素的关系, 这些在金融经济学中起到了至关重要的作用。对于许多应用, 如资产定价, 投资组合选择或风险管理, 市场 beta 是唯一最重要的参数。然而, beta 因素不是直接可观察的, 因此需要估计。本文的主要贡献在于, 第一个对广泛的市场 beta 估算技术的性能提供全面和深入的实证研究, 包括几个历史的时间序列模型, 和期权隐含估计方法。此外, 我们提出了一个新的 beta 估计方法, 用于修正波动率风险溢价的期权隐含波动率。

### 1. 简介

估计  $\beta$  的最基本的方法是简单地估计历史收益数据的时间序列中的协方差和方差。然而, 这种方法面临的问题是  $\beta$  系数表现出显著的时间变化 (例如, Blume (1975), Ferson 和 Harvey (1991), (1993))。为了解决这个问题, 已经开发了几种方法 (例如, 基于 GARCH 的) 来捕捉这种变化。最近有人提出, 可以纳入期权市场的信息, 即投资者可获得的所有信息都应当以当前价格进行包含, 从而克服历史估计所固有的惯性, 即使采用滚动窗口方法也是如此。

然而, 令人惊讶的是, 文献中对  $\beta$  的估计得到了相当少的关注。Faff, Hillier 和 Hillier (2000) 发现, 在样本内分析中, 时间序列模型 (尤其是那些使用卡尔曼滤波的模型) 优于历史估计值, 虽然他们没有出现任何样本外证据。

在现有文献中研究  $\beta$  估计的研究相对较少, 可能在一定程度上是由于 beta 需要相关性信息, 而不像期权的隐含波动率那样容易获得。只有最近有几位作者应用期权隐含方法来估计  $\beta$ 。

Chang, Christoffersen, Jacobs 和 Vainberg (2012) 提出了这样一种期权隐含方法, 表明它在横截面分析中往往胜过历史 beta。Baule, Korn 和 Saisson (2016) 比较了各种不同的完全隐含的  $\beta$  估计量。他们使用基于隐含方差的测试来获得最佳性能。然而, Chang 等人 (2012) 不直接比较隐含 beta 方法与其他现有方法, Baule et al (2016 年) 比较了完全隐含的估计量之间的相互关系, 以及依靠道琼斯工业平均指数 (DJIA) 30 的简单历史估计值。

### 2. 数据与方法

我们以 1996 年 1 月 1 日至 2012 年 12 月 31 日样本期标准普尔 (S&P) 500 市场指数及其构成成分为基础进行研究。此外, 我们对基于道琼斯指数的样本进行了稳健性分析。我们从 1994 年 1 月 1 日至 2012 年 12 月 31 日期间, 获取每日和每月的价格数据以及安全价格研究中心 (CRSP) 的股息和股票数据。为了能够计算历史和时间序列模型估计从研究期开始, 并使用非重叠数据进行组合分类, 这些数据在主样本期前 2 年开始。

期权数据来自 IvyDB OptionMetrics 波动面, 它直接提供标准差值水平和到期日的隐含波动率。我们使用大约 6 个月到期的期权, 因为我们想要获得 6 个月的 beta 估计值。作为一个稳健性检查, 我们也进行重复分析, 选择大约 1 个月, 3 个月和 12 个月到期。我们选择 OTM 选项, 即取看跌期权 delta 值大于 -0.05, 并使用 delta 小于 0.5 的看涨期权。因此, 我们获得了 1996 年的 438 只股票的期权数据, 在标准普尔 500 指数中, 在 2010 年增长到 493 只。平均而言, 有 472 只股票的期权数据可供使用。有关无风险利率的数据从 IvyDB 零曲线文件中收集。

在 Andersen, Bollerslev, Diebold 和 Wu (2006) 之后, 我们使用每日对数回报来计算已实现的  $\beta$  (RB)

$$\beta_{j,t}^R = \frac{\sum_{\tau=1}^N r_{j,\tau} r_{M,\tau}}{\sum_{\tau=1}^N r_{M,\tau}^2}$$

其中,  $r_{j,\tau}$  和  $r_{M,\tau}$  分别是指资产  $j$  的 (超额) 收益和市场 (超额) 收益。  $N$  是在调查期间的观察次数。

与上述方法密切相关, 我们按照通常的方式计算历史估计 (HIST), 遵循 Fama 和 MacBeth (FM) (1973) 以及其他许多方法, 将资产的 (超额) 回报率 (回报率):

$$\beta_{j,t} = \frac{\text{cov}(r_{j,\tau}, r_{M,\tau})}{\text{var}(r_{M,\tau})}$$

我们估计了动态条件  $\beta$  和 GARCH 模型（共）波动率和 AR 型模型，直接在 beta 系列上强加了一定的因子动态。我们将这两种类型称为时间序列模型。

我们考虑使用动态条件相关（DCC）GARCH 模型的动态条件测试（Engle（2014），Bali（2015））Engle（2002）和 Cappiello，Engle 和 Sheppard（2006）将经验公认的杠杆效应（通过允许正反收益创新的不对称效应）以及差异创新相关性的不对称反应。首先，根据 Glosten，Jagannathan 和 Runkle（GJR）（1993）提出的 GARCH 估计单变量波动率模型：

$$r_t = \mu + a_t$$

$$a_t \sim N(0, \sigma_t^2)$$

$$h_t^2 = \omega + (\alpha + \gamma I_{t-1}[r_{t-1} < \mu])a_{t-1}^2 + \beta h_{t-1}^2$$

其中  $r_t$  是每日（每月）资产收益， $\mu$  是平均收益，而  $a_t$  代表回报创新。

假设返回信息序列是有条件地正态分布的，均值为 0，条件协方差矩阵为  $H_t$ ，可以按下式进行分解。一旦估计了单变量模型，标准化残差就可以用  $\epsilon_{i,t} = a_{i,t}/\sqrt{h_{i,t}}$  来估计相关参数（见 Cappiello et al.（2006））：

$$H_t = D_t P_t D_t,$$

$$P_t = Q_t^{*-1} Q_t Q_t^{*-1},$$

$$Q_t = (\bar{P} - A' \bar{P} A - B' \bar{P} B - G' \bar{N} G) + A' \epsilon_{t-1} \epsilon_{t-1}' A$$

$$+ G' n_{t-1} n_{t-1}' G + B' Q_{t-1} B,$$

其中  $D_t$  是包含个别资产的标准偏差的对角矩阵； $Q_t^*$  是包含  $Q_t^*$  的各对角元素的平方根的对角矩阵，确保  $P_t$  是有效的相关矩阵。

我们还包括直接在  $\beta$  上施加因子结构并使用卡尔曼滤波器的方法（参见例如 Pagan（1980）和 Black，Fraser 和 Power（1992））。作为潜在的动力学，我们考虑随机游走（RW），漂移随机游走（RWD），AR（1）（AR）和 ARMA（1,1）（ARMA）模型。在所有四种情况下，标准 CAPM 安全市场线被作为测量方程，转换方程描述了在状态空间形式下  $\beta$  的动态演化的选择模型：

$$\beta_{j,\tau}^{RW} = \beta_{j,\tau-1} + \epsilon_{j,\tau},$$

$$\beta_{j,\tau}^{RWD} = \phi_0 + \beta_{j,\tau-1} + \epsilon_{j,\tau},$$

$$\beta_{j,\tau}^{AR(1)} = \phi_1 \beta_{j,\tau-1} + \epsilon_{j,\tau},$$

$$\beta_{j,\tau}^{ARMA(1,1)} = \phi_1 \beta_{j,\tau-1} + \epsilon_{j,\tau} + \theta_1 \epsilon_{j,\tau-1}.$$

我们估计的模型与动态条件测试类似，也使用 1 年的日收益率。

西格尔（Siegel，1995）指出，可以通过使用交换期权直接获得一个隐含的 beta 估计，这是一个交换一个公司股票换取市场指数的份额的期权。不幸的是，这些交换期权目前还不能进行交易。因此，人们必须依靠一些假设来获得一个隐含的  $\beta$ ，从而利用可以从期权价格中获得的内在的未来的信息。

我们利用前面讨论的变量，考虑选择期权隐含 beta 的几种可能性。其中包括法国，Groth 和 Kolari（1983）的混合方法，它直接结合了历史相关和期权隐含波动率，而 Buss 和 Vilkov（2012）则直接使用了“市场指数的隐含方差所有市场成分的加权组合的隐含方差相等”的结论（第一关系），并将其与隐含相关的技术条件从  $\rho_{i,j,t}^p$  转化为风险中性相关  $Q_{i,j,t}$ ，即  $\rho_{i,j,t}^p - \alpha_t(1 - \rho_{i,j,t}^p)$ ，结合这两个关系并求解  $t$ ，可以计算隐含的相关性。因此，在风险中性概率测度  $Q$  下的贝塔估计值由下式得出

$$\beta_{j,t}^Q = \frac{\sigma_{j,t}^Q \sum_{i=1}^N (\omega_{i,t} \sigma_{i,t}^Q \rho_{j,i,t}^Q)}{(\sigma_{M,t}^Q)^2}$$

其中  $\sigma_{j,t}^Q$  和  $\sigma_{M,t}^Q$  分别表示个体股票和市场指数在上述中得到的隐含波动率， $\eta$  和  $t$  分别表示市场指数中  $N$  个单项资产的权重和时间点。这种方法的一个主要缺点是，它需要关于所考虑指数的所有组成部分的信息。如果市场指数所包含的所有股票的隐含波动率不可用，估计可能会有偏差。

在风险中性概率测度下，用期权隐含方法估计  $\beta$ 。然而，在大多数情况下，我们感兴趣的是物理概率测度下的 beta。因此，我们提出了一个新的 beta 混合估计，它使用期权价格的前瞻性信息，同时对波动率风险溢价进行修正。为了得到这个估算值，我们按照 Prokopczuk 和 Wese Simen（2014）的程序对波动率风险溢价进行调整。为此，我们计算了不到 2 年的平均风险溢价：

$$ARVRP_{j,t}^2 = \frac{1}{504 - \tau} \sum_{i=t-504}^{t-\tau} \frac{\sigma_{j,i,t+\tau}^2}{RV_{j,i,t+\tau}^2}$$

其中  $ARVRP_{j,t}^2$  表示从  $t-504$  到  $t-\tau$  的平均相对方差风险溢价,  $\sigma_{j,i+\tau}^2$  是直到  $i+\tau$  的时间点  $i$  的资产  $j$  的无模型隐含方差,  $RV_{j,i+\tau}^2$  为从  $i$  到  $i+\tau$  的时间段内实现的方差,  $\tau$  表示估计的时间范围。

然后我们得到每个时间点  $t$  的风险溢价调整隐含波动率  $RMFIV_{j,t,T}$

$$RMFIV_{j,t} = \frac{\sigma_{j,t}}{ARVRP_{j,t}}$$

然后, 利用历史相关  $\rho_{i,j}$  和风险溢价调整后的隐含波动率和市场指数计算风险溢价调整后的  $\beta$  (RP ADJ)

$$\beta_{j,t}^{RP-ADJ} = \rho_{j,t} \times \frac{RMFIV_{j,t}}{RMFIV_{M,t}}$$

### 3. 实证结果

图 1 提供了 A 组中不同  $\beta$  估计技术的汇总统计数据。所有方法都利用 (如有必要) 每日收益数据并估算 6 个月的  $\beta$  值。样本期从 1996 年 1 月 (从 1996 年 2 月的估计开始) 到 2012 年 12 月。No. of Obs. 表示每月估计的数量。Mean 和 MEAN<sub>v</sub> 分别是整个样本期间估值的等值和加权平均值。STD., Dev., Median, Min. 和 Max. 表示进一步汇总总体标准差, 中位数, 最小值和最大估计值。图 B 以各个估计值为基础提出了不同的估计技术之间的样本相关系数。

图 1: 摘要统计和样本关联性

Panel A. Summary Statistics							
	No. of Obs.	Mean	MEAN <sub>v</sub>	Std. Dev.	Median	Min.	Max.
RW	98,179	1.0036	1.0026	0.5720	0.9313	-29.0756	19.8332
RW <sub>D</sub>	98,179	1.0713	1.0618	0.7650	0.9807	-36.2117	24.7961
AR	98,179	0.9662	0.9753	0.7770	0.8349	-2.0388	20.9860
ARMA	98,179	1.0547	1.0369	1.0053	0.8724	-8.6868	48.6730
DCC	98,176	0.9511	0.9304	0.8068	0.8648	-13.9973	16.7336
CCC	98,179	1.0731	1.0430	0.7009	0.9348	-0.6028	17.9049
HIST	98,243	1.0022	1.0033	0.4680	0.9360	-0.6675	4.6485
HIST <sub>6</sub>	98,630	1.0021	1.0015	0.5036	0.9344	-0.9818	7.7906
FGK	94,889	0.8473	0.8436	0.3927	0.7998	-0.8230	5.7060
RP_ADJ	90,190	1.0321	1.0206	0.5208	0.9498	-0.8554	5.5901
BV	95,043	1.0427	1.0000	0.3756	0.9870	-0.4646	6.9280
CCJV	89,530	1.2211	1.1505	0.4706	1.1579	0.0220	6.2420
SR	95,755	1.1077	1.0391	0.3693	1.0297	0.1215	6.2270
KKS1	95,755	1.1074	1.0000	0.3733	1.0260	0.1237	6.2000
KKS2	95,755	1.1038	1.0000	0.3635	1.0258	0.1341	6.3583

Panel B. Sample Correlation Coefficients														
RW	RW <sub>D</sub>	AR	ARMA	DCC	CCC	HIST	HIST <sub>6</sub>	FGK	RP_ADJ	BV	CCJV	SR	KKS1	KKS2
1.00	0.91	0.76	0.68	0.54	0.64	0.82	0.85	0.79	0.80	0.77	0.52	0.61	0.61	0.61
	1.00	0.79	0.70	0.51	0.60	0.76	0.86	0.74	0.75	0.72	0.49	0.58	0.57	0.57
		1.00	0.67	0.44	0.52	0.68	0.76	0.66	0.68	0.64	0.43	0.49	0.48	0.48
			1.00	0.41	0.48	0.59	0.68	0.57	0.60	0.57	0.40	0.48	0.47	0.47
				1.00	0.71	0.59	0.57	0.56	0.55	0.53	0.39	0.45	0.44	0.44
					1.00	0.71	0.68	0.66	0.66	0.64	0.48	0.56	0.56	0.56
						1.00	0.93	0.90	0.92	0.88	0.58	0.69	0.68	0.68
							1.00	0.87	0.88	0.84	0.57	0.67	0.66	0.66
								1.00	0.95	0.89	0.67	0.71	0.70	0.70
									1.00	0.90	0.64	0.74	0.73	0.73
										1.00	0.70	0.88	0.88	0.87
											1.00	0.77	0.77	0.77
												1.00	1.00	0.99
													1.00	1.00
														1.00

资料来源: Financial Analysts Journal, 天风证券研究所

我们将注意力转向样本外估计的准确性, 我们使用文献中最常用的损失函数, 即 RMSE 来评估不同的估计技术的性能:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (\beta_{t,T}^R - \zeta_{t,T})^2}$$

这里,  $n$  是估计窗口的长度,  $\beta_{t,T}^R$  表示在从  $t$  到  $T$  的时间段内实现的  $\beta$ 。

图 2 报告了使用每日收益数据的估计量的样本外估计误差, 用于计算每个投资组合 6 个月的实现贝塔值。我们建立了五个分位数组合, 按照排序期间的历史测试数据 (在历史测试计划的估计期间之前直接进行, 没有重叠且长度相同) 按升序分配股票。我们确定投资组合测试和回报作为加权平均。第一行报告五个投资组合的估计模型的平均均方误差。



所有方法中的最低误差用斜体表示。表格的其余部分报告估计误差的差异。上三角矩阵报告了五个投资组合中均方估计的差异。类似地，下三角矩阵报告估计误差的平均中值差异。我们计算模型[行中名称]和模型[列中名称]之间的差异。括号内的绝对数字表示差异显著的投资组合的百分比（例如，0.4 表示五个投资组合中的两个的差异具有统计显著性）。如果所有五个投资组合的差异都是显著的，则该数字将以粗体显示。重要性由修改的 Diebold Mariano 和 Wilcoxon 符号等级测试分别对上三角矩阵和下三角矩阵进行测试。标志表示重大差异的方向。

图 2：估计误差：6 个月的水平下，日数据

	HIST	RW	DCC	FGK	CCJV	KKS1	BV	RP_ADJ
Avg.	0.1381	0.1301	0.2676	0.2203	0.2783	0.1704	<i>0.1164</i>	0.1718
HIST		0.0079 (0.0)	-0.1295 (-1.0)	-0.0822 (-0.8)	-0.1402 (-0.8)	-0.0323 (-0.4)	0.0217 (0.2)	-0.0337 (-0.6)
RW	0.0002 (-0.4)		-0.1374 (-1.0)	-0.0901 (-1.0)	-0.1482 (-0.8)	-0.0403 (-0.2)	0.0138 (0.2)	-0.0416 (-0.6)
DCC	0.0639 (1.0)	0.0637 (1.0)		0.0473 (0.0)	-0.0108 (0.0)	0.0972 (0.6)	0.1512 (1.0)	0.0958 (0.8)
FGK	0.1001 (1.0)	0.0999 (1.0)	0.0362 (0.2)		-0.0581 (-0.6)	0.0499 (0.6)	0.1039 (1.0)	0.0485 (0.2)
CCJV	0.1143 (1.0)	0.1141 (1.0)	0.0504 (0.6)	0.0142 (0.4)		0.1079 (0.8)	0.1620 (1.0)	0.1066 (1.0)
KKS1	0.0113 (0.4)	0.0110 (0.6)	-0.0526 (-0.8)	-0.0888 (-0.8)	-0.1031 (-1.0)		0.0541 (0.8)	-0.0013 (-0.2)
BV	-0.0111 (-0.6)	-0.0113 (-0.6)	-0.0750 (-1.0)	-0.1112 (-1.0)	-0.1254 (-1.0)	-0.0224 (-1.0)		-0.0554 (-0.6)
RP_ADJ	0.0256 (0.8)	0.0254 (0.8)	-0.0383 (-1.0)	-0.0745 (-1.0)	-0.0888 (-1.0)	0.0143 (0.2)	0.0367 (1.0)	

资料来源：Financial Analysts Journal，天风证券研究所

我们发现，BV 总是比其他方法获得更低的平均均方误差和中值平方误差。这些差异在所有投资组合中与 DCC，FGK 和 CCJV 相比具有统计显著性。总体而言，证据表明 BV 方法获得了最好的样本外准确性，其次是 RW 和 HIST。

图 3：估计误差：不同水平下的日数据

<i>Panel A. 1 Month</i>								
	HIST	RW	DCC	FGK	CCJV	KKS1	BV	RP_ADJ
Avg.	0.1637	0.1515	0.2231	0.2193	0.4285	0.2051	<i>0.1483</i>	0.1977
HIST		0.0122 (0.0)	-0.0594 (-0.8)	-0.0556 (-0.8)	-0.2648 (-0.8)	-0.0414 (-0.2)	0.0153 (0.0)	-0.0340 (-0.8)
RW	-0.0022 (-0.2)		-0.0717 (-1.0)	-0.0678 (-1.0)	-0.2770 (-1.0)	-0.0536 (-0.4)	0.0031 (0.0)	-0.0462 (-0.8)
DCC	0.0375 (1.0)	0.0397 (1.0)		0.0038 (0.0)	-0.2053 (-0.8)	0.0181 (0.4)	0.0748 (1.0)	0.0254 (0.0)
FGK	0.0525 (1.0)	0.0547 (1.0)	0.0150 (0.0)		-0.2092 (-0.8)	0.0142 (0.4)	0.0709 (1.0)	0.0216 (0.0)
CCJV	0.1550 (1.0)	0.1572 (1.0)	0.1175 (1.0)	0.1025 (0.8)		0.2234 (0.8)	0.2801 (1.0)	0.2308 (0.8)
KKS1	0.0189 (0.4)	0.0211 (0.6)	-0.0186 (-0.4)	-0.0336 (-0.4)	-0.1361 (-1.0)		0.0567 (0.6)	0.0074 (0.0)
BV	-0.0041 (-0.4)	-0.0020 (-0.2)	-0.0417 (-1.0)	-0.0567 (-1.0)	-0.1592 (-1.0)	-0.0231 (-1.0)		-0.0493 (-0.8)
RP_ADJ	0.0340 (0.8)	0.0362 (0.8)	-0.0035 (0.0)	-0.0185 (-0.6)	-0.1210 (-1.0)	0.0151 (0.4)	0.0382 (0.8)	
<i>Panel B. 3 Months</i>								
	HIST	RW	DCC	FGK	CCJV	KKS1	BV	RP_ADJ
Avg.	0.1380	0.1281	0.2259	0.2057	0.3335	0.1771	<i>0.1191</i>	0.1791
HIST		0.0100 (0.0)	-0.0879 (-1.0)	-0.0676 (-0.8)	-0.1955 (-0.8)	-0.0391 (-0.2)	0.0189 (0.0)	-0.0411 (-0.8)
RW	-0.0031 (-0.4)		-0.0979 (-1.0)	-0.0776 (-1.0)	-0.2054 (-1.0)	-0.0491 (-0.2)	0.0089 (0.0)	-0.0511 (-0.8)
DCC	0.0494 (1.0)	0.0525 (1.0)		0.0203 (0.0)	-0.1075 (-0.8)	0.0488 (0.4)	0.1068 (1.0)	0.0468 (0.4)
FGK	0.0784 (1.0)	0.0815 (1.0)	0.0289 (0.2)		-0.1278 (-0.8)	0.0286 (0.4)	0.0866 (1.0)	0.0265 (0.0)
CCJV	0.1313 (1.0)	0.1344 (1.0)	0.0818 (0.8)	0.0529 (0.6)		0.1564 (0.8)	0.2144 (1.0)	0.1544 (0.8)
KKS1	0.0151 (0.6)	0.0182 (0.6)	-0.0343 (-0.6)	-0.0633 (-0.6)	-0.1162 (-0.8)		0.0580 (0.8)	-0.0020 (-0.2)
BV	-0.0087 (-0.6)	-0.0056 (-0.2)	-0.0581 (-1.0)	-0.0871 (-1.0)	-0.1400 (-1.0)	-0.0238 (-1.0)		-0.0600 (-0.6)
RP_ADJ	0.0305 (0.8)	0.0336 (1.0)	-0.0189 (-0.6)	-0.0479 (-0.8)	-0.1008 (-1.0)	0.0154 (0.4)	0.0392 (1.0)	

Panel C. 12 Months

	HIST	RW	DCC	FGK	CCJV	KKS1	BV	RP_ADJ
Avg.	0.1488	0.1423	0.3374	0.2307	0.2674	0.1677	0.1227	0.1895
HIST		0.0065 (0.0)	-0.1886 (-1.0)	-0.0820 (-0.8)	-0.1187 (-0.8)	-0.0190 (0.0)	0.0261 (0.2)	-0.0407 (-0.4)
RW	-0.0019 (0.0)		-0.1951 (-1.0)	-0.0884 (-1.0)	-0.1251 (-1.0)	-0.0254 (-0.2)	0.0196 (0.2)	-0.0471 (-0.6)
DCC	0.0714 (1.0)	0.0734 (1.0)		0.1067 (0.0)	0.0700 (0.2)	0.1697 (0.8)	0.2147 (1.0)	0.1479 (1.0)
FGK	0.1003 (1.0)	0.1023 (1.0)	0.0289 (-0.2)		-0.0367 (-0.2)	0.0630 (0.8)	0.1080 (1.0)	0.0413 (0.0)
CCJV	0.1188 (1.0)	0.1208 (1.0)	0.0474 (0.2)	0.0185 (0.0)		0.0997 (0.8)	0.1447 (1.0)	0.0780 (0.8)
KKS1	0.0054 (0.4)	0.0073 (0.4)	-0.0660 (-0.8)	-0.0950 (-0.8)	-0.1135 (-1.0)		0.0450 (0.8)	-0.0217 (-0.2)
BV	-0.0163 (-0.6)	-0.0144 (-0.6)	-0.0877 (-1.0)	-0.1167 (-1.0)	-0.1352 (-1.0)	-0.0217 (-1.0)		-0.0667 (-0.4)
RP_ADJ	0.0214 (0.8)	0.0234 (1.0)	-0.0500 (-1.0)	-0.0789 (-1.0)	-0.0974 (-1.0)	0.0161 (0.4)	0.0378 (0.6)	

资料来源：Financial Analysts Journal，天风证券研究所

图 3 中的 A 面板报告了我们主要方法的估计误差及其在 1 个月评估期间的意义。我们发现使用这个评估范围可以得到相同的结果，其中 BV、HIST 和 RW 是关于样本估计精度最好的方法。BV 获得最低的平均 RMSE。BV 总是获得最低的平均 RMSE，与其他方法相比，在许多情况下误差显著降低。

综上所述，将评估期改为 1,3 或 12 个月时，BV、RW 和 HIST 依然是最好的方法。

图 4：估计误差：6 个月水平下，日数据（更深一步的时间序列模型）

	HIST	HIST <sub>6</sub>	RW	RW <sub>D</sub>	AR	ARMA	DCC	CCC	BV
Avg.	0.1355	0.1283	0.1269	0.1870	0.1865	0.1899	0.2638	0.2572	0.1152
HIST		0.0072 (0.0)	0.0086 (0.0)	-0.0515 (-0.4)	-0.0510 (-0.8)	-0.0543 (-0.8)	-0.1283 (-1.0)	-0.1217 (-0.8)	0.0203 (0.2)
HIST <sub>6</sub>	-0.0041 (-0.2)		0.0014 (0.0)	-0.0587 (-1.0)	-0.0582 (-1.0)	-0.0616 (-1.0)	-0.1356 (-1.0)	-0.1289 (-0.8)	0.0131 (0.2)
RW	-0.0022 (-0.2)	0.0019 (0.0)		-0.0601 (-0.8)	-0.0596 (-1.0)	-0.0630 (-1.0)	-0.1370 (-1.0)	-0.1303 (-0.8)	0.0117 (0.2)
RW <sub>D</sub>	0.0394 (0.8)	0.0435 (1.0)	0.0416 (1.0)		0.0005 (0.0)	-0.0029 (0.0)	-0.0769 (-0.8)	-0.0702 (0.0)	0.0718 (0.8)
AR	0.0345 (1.0)	0.0385 (1.0)	0.0366 (1.0)	-0.0049 (-0.2)		-0.0034 (0.2)	-0.0774 (-0.8)	-0.0707 (0.0)	0.0713 (0.8)
ARMA	0.0369 (1.0)	0.0410 (1.0)	0.0391 (1.0)	-0.0025 (-0.2)	0.0024 (-0.2)		-0.0740 (-0.8)	-0.0673 (-0.2)	0.0746 (0.8)
DCC	0.0638 (1.0)	0.0679 (1.0)	0.0660 (1.0)	0.0244 (0.8)	0.0293 (0.8)	0.0269 (0.8)		0.0067 (0.0)	0.1486 (1.0)
CCC	0.0338 (1.0)	0.0379 (1.0)	0.0360 (1.0)	-0.0055 (0.0)	-0.0006 (0.0)	-0.0031 (0.2)	-0.0299 (-0.8)		0.1420 (0.8)
BV	-0.0137 (-0.8)	-0.0096 (-0.6)	-0.0115 (-0.6)	-0.0531 (-0.8)	-0.0482 (-1.0)	-0.0506 (-0.8)	-0.0775 (-1.0)	-0.0475 (-0.8)	

资料来源：Financial Analysts Journal，天风证券研究所

在图 4 中，我们发现 RWD、AR、ARMA，特别是 HIST6 表现相当好，而 CCC 明显跑赢大盘。此外，可以看出，与 CCC 的结果相比，在 CCC 中较不复杂的结构略有改进。HIST6 与一年以上的历史估计值相比，平均 RMSE 较小，因此短期条件估计可能更适合。RWD、AR 和 ARMA 的平均 RMSE 值仅为中等，但明显高于 BV，总体上 RMSE 平均值最低。关于显著性，我们发现 HIST、HIST6、RW 和 BV 的误差要比 AR、ARMA、DCC 和 CCC 至少有四个组合的误差要小得多。只有在极少数情况下，以前提到的方法存在显著差异，但在这些情况下，它们大多数都支持 BV。总的来说，除了 HIST 和 RW 之外，我们发现 HIST6 是 BV 方法的一个有价值的选择。

#### 4. 结论

本文考察了估算资产市场测试的各种方法的性能。具体而言，我们调查了几个依靠历史回报数据的恒定和时间变化的模型，另外还有几种方法，包括仅仅依赖于期权隐含信息的方法。总之，使用历史信息的估计只有在不做太强的结构性假设时才会表现良好，如简单的历史测试和带有随机游走参数化的卡尔曼滤波方法。相反，对波动性和相关性过程（如基于 GARCH 的 DCC）做出强烈假设的模型显示产生了非常大的误差。

总的来说，虽然 BV 方法似乎是可以选择的方法，但是必须记住这种方法（以及尝试简单的偏差校正的其他混合方法）的一个主要缺点：该方法需要关于完整市场指数的信息。因此，它不能用于没有包含在指数中的资产，或者在指数中所有资产的选项隐含信息不足

的情况。我们的结果表明，应该依赖于 RW 或基于历史回报的简单估计，因为两者总是比所有其他方法都要好。

以上感谢实习生刘佳璐的贡献。

## 价值、规模、动量、股利回报以及波动率在中国 A 股市场的分析

文献来源：CHRISTOPHER CHEUNG, GEORGE HOGUET, AND SUNNY NG, Value, Size, Momentum, Dividend Yield, and Volatility in China's A-Share Market, Journal of Portfolio Management, 2014, 41 (5): 57-70

推荐理由：因子选股已经成为大多数机构投资者的重要投资方法，该方法起源于 CAPM 理论，之后又提出了著名的 Fama-French 三因子模型，引入了价值因子和市场规模因子对股票或投资组合收益率的解释，而且认为三因子模型对股票或投资组合收益率的解释力度要比 CAPM 模型更好。这充分说明了还存在其他因子可以更好的解释市场中股票或投资组合的收益率。但是，不同的金融市场也许会得到不同的因子溢价，我们分析中国市场中因子的表现，因为中国 A 股市场有着它独特的一面。

### 1. 中国 A 股市场的独特性

多因子模型在国外的市场已经得到了很有力的验证，在 CAPM 模型提出之后，我们不断的在寻找对收益率解释更为有效的因子。接着，Fama-French 提出的三因子模型把账面价值因子和市场规模因子加入了资产定价模型中，得到了对收益率更为有利的解释。但在不同的市场环境下，因子的溢价以及因子的显著性可能也会发生改变。我们在此研究中国 A 股市场，因为随着时间的变化，A 股市场有着很大的改变，不仅规模变化了，现在 A 股市场已经成为了世界第五大金融市场。而且，在上交所和深交所成立至今，交易规则也在不断的改进。这一切的变化都使得之前对 A 股市场早期的研究成果变得不那么有说服力，我们先对当今的 A 股市场的一些特点做出解释：

- 第一、中国 A 股市场历史很短，只是从 1990 年开始，因此为我们研究者提供的数据也是相对不多的。
- 第二、有大量的散户投资者，据统计，中国 A 股市场有 90% 的投资者是散户投资者。
- 第三、很高的换手率，据统计，在 1991 年到 2004 年，A 股市场的年平均换手率为 500%。
- 第四、政府机构持有大量的资产。
- 第五、各种股票类型都有，中国市场有着 A 股，B 股还有 H 股。并且中国股票市场的资产账户是封闭的。
- 第六、更高的波动率，截止到 2013 年最近的十年，中国股市的年化波动率为 32%。
- 第七、没有资产溢价回报税，只有股利税。
- 第八、国家政策对股市的影响巨大。

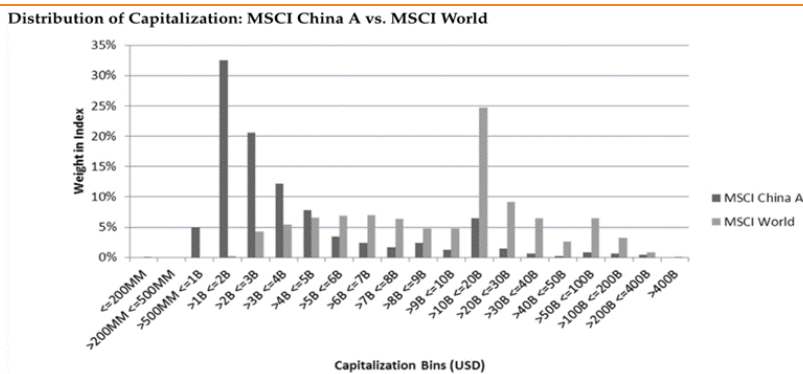
结合以上特点，我们也许可以认为，把对国外市场的因子模型的研究成果套用到中国 A 股市场也许是行不通的，比如，小市值股票一定在 A 股市场表现的好吗，动量因子一定是显著的吗，等等问题我们都要具体事情具体分析。

### 2. 数据以及因子的定义

A 股市场也存在大量的指数，比如上证综指和深证成指都分别在 1990 年和 1991 年成立，沪深 300 则在 2005 年成立。MSCI 中国指数也是在 2001 年成立，它包含了中国 A 股市场的 460 只大盘股和中盘股。我们选择 MSCI 中国指数的成分股作为我们分析投资组合的数据来源，分析时间是从 2001 年 12 月 31 日到 2013 年 12 月 31 日。采用 MSCI 中国指数的成分股，不仅仅是它包含了在上证和深证上市的大盘股和中盘股，而且最为特别的是在这些交易所的每一个 GICS 行业集团中，它捕获了 65% 的自由浮动市场资本，加上最大的 25 种证券。



图 5：MSCI 中国指数成分股和世界 MSCI 指数成分股的市值大小



资料来源：Journal of Portfolio Management，天风证券研究所

从图 5 可以看出中国 MSCI 指数和世界 MSCI 指数成分股的规模大小的对比。

之后，我们选用投资组合的月度收益率的方式作为模型的应变量，然后定义每个因子的收益率。因子的收益率是将每个因子按照大小排序分为四个不同的组，然后分别与规模因子进行组合，得到一个 4\*4 的因子收益率组合，拿高分组减去低分组的收益率作为该因子的收益率。例如，在三因子模型中，我们把规模因子分为两组，BIG 规模组和 Small 规模组，然后将价值因子分为三组，分别按照顺序取前 30%，31%-70%，71%-100%。这样就得到一个 2\*3 的组合矩阵。然后将规模为 BIG 的三个分组的平均收益率减去规模为 Small 的三个分组的平均收益率，作为市场规模因子收益率，而市值因子的收益率为前 30% 的市值的两个分组的平均收益率减去后 30% 的两组平均收益率。具体的例子见图 6：

图 6：因子的定义

### Factor Definitions

Factor (e.g., BE/ME)				
Size	Small	Small	Small	10% of
	Low	Neutral	High	Market Cap
	Big	Big	Big	90% of
	Low	Neutral	High	Market Cap
	30% of # of Securities		30% of # of Securities	

资料来源：Journal of Portfolio Management，天风证券研究所

我们这里是将每个因子和市场规模因子进行组合，得到一个 4\*4 的矩阵组合收益率，然后按照上面的方法得到因子收益率。

### 3. 研究方法

这里，我们采用每个月的投资者的收益率对市场溢价还有市场规模因子以及分别加上价值因子、动量因子、股利回报因子和波动率因子后得到的三因子模型进行线性回归，主要的模型方程如下：

$$R_p - R_f = \alpha_p + b_p(R_M - R_f) + \varepsilon_p \quad \text{-----(1)}$$

$$R_p - R_f = \alpha_p + b_p(R_M - R_f) + s_pSMB + h_pHML + \varepsilon_p \quad \text{---(2)}$$

$$R_p - R_f = \alpha_p + b_p(R_M - R_f) + s_pSMB + h_pWML + \varepsilon_p \quad \text{---(3)}$$

$$R_p - R_f = \alpha_p + b_p(R_M - R_f) + s_pSMB + \gamma_pYield + \varepsilon_p \quad \text{---(4)}$$

$$R_p - R_f = \alpha_p + b_p(R_M - R_f) + s_pSMB + v_pVolatility + \varepsilon_p \quad \text{---(5)}$$

以上五个回归方程就说不同的因子与规模因子还有市场溢价因子组成的三因子模型，我们用的是投资组合的因子回报率，而不是个股，加权的方式为市值加权的方法。其中第一个回归方程就是标准的 CAPM 模型，我们需要比较这五个方程的结果，不仅仅可以得到不同因子的回归系数的显著性，而且还能比较加入其他因子后，模型对收益率的解释力度有没有增加，也就是比较方程 2 到 5 的调整后的 R 方，与方程 1 的 R 方。如果加入因子后，R 方变大，那么说明因子的增加加大了模型对股票组合收益率的解释力度。

#### 4. 研究结果

我们对每组模型进行回归后分别得到了因子暴露的预测值以及它的显著性，而且我们还得到了五个模型后的 R 方，对比最原始的 CAPM 模型，来看看加入新的因子后模型的解释力度有没有变化。

图 7：因子的显著性检验

Summary Statistics for Factor Returns

	Market	SMB	HML	HML_s	HML_b
<b>Period: Jan 2002–Dec 2013</b>					
Mean	0.67	0.38	0.83	0.46	1.20
St Dev.	8.80	4.96	5.09	4.62	7.01
T-Mean	0.92	0.92	1.95	1.19	2.05
Significance	82%	82%	97%	88%	98%
			<b>WML</b>	<b>WML_s</b>	<b>WML_b</b>
Mean			0.20	-0.02	0.42
St Dev.			4.61	4.58	5.25
T-Mean			0.52	-0.06	0.97
Significance			70%	48%	83%
			<b>Yield</b>	<b>Yield_s</b>	<b>Yield_b</b>
Mean			0.55	0.61	0.48
St Dev.			3.36	2.97	4.99
T-Mean			1.95	2.48	1.15
Significance			97%	99%	87%
			<b>Vol</b>	<b>Vol_s</b>	<b>Vol_b</b>
Mean			0.47	0.36	0.58
St Dev.			4.24	4.05	5.60
T-Mean			1.33	1.06	1.25
Significance			91%	85%	89%

资料来源：Journal of Portfolio Management，天风证券研究所

图 8：模型的解释力度

Summary Statistics for Regressions to Explain Variability of Excess Returns on Portfolios from Sorts on Size and B/M

	Adj R <sup>2</sup>	Min. Adj R <sup>2</sup>	F-Test Significance	Intercept	p-Values		
					Market Premium	HML	SMB
CAPM	81.6%	69.4%	0.0%	0.21%	0.00		
Three-Factor	92.2%	82.0%	0.0%	0.08%	0.00	0.02	0.00

Summary Statistics for Regressions to Explain Variability of Excess Returns on Portfolios from Sorts on Size and Yield

	Adj R <sup>2</sup>	Min. Adj R <sup>2</sup>	F-Test Significance	Intercept	p-Values		
					Market Premium	Yield	SMB
CAPM	82.7%	73.2%	0.0%	0.21%	0.00		
Three-Factor	91.9%	87.1%	0.0%	0.12%	0.00	0.09	0.00

Summary Statistics for Regressions to Explain Variability of Excess Returns on Portfolios from Sorts on Size and Momentum

	Adj R <sup>2</sup>	Min. Adj R <sup>2</sup>	F-Test Significance	Intercept	p-Values		
					Market Premium	WML	SMB
CAPM	81.8%	68.5%	0.0%	0.27%	0.00		
Three-Factor	92.2%	80.3%	0.0%	0.15%	0.00	0.12	0.00

Summary Statistics for Regressions to Explain Variability of Excess Returns on Portfolios from Sorts on Size and Volatility

	Adj R <sup>2</sup>	Min. Adj R <sup>2</sup>	F-Test Significance	Intercept	p-Values		
					Market Premium	Volatility	SMB
CAPM	82.3%	69.3%	0.0%	0.22%	0.00		
Three-Factor	92.5%	87.7%	0.0%	0.12%	0.00	0.21	0.00

资料来源：Journal of Portfolio Management，天风证券研究所

从图 7 和图 8 中我们可以得出结论，价值因子和股利回报因子在对收益率的解释上是显著有效的，而且通过加入因子到 CAPM 模型，模型的 R 方都得到了提高。

以上感谢实习生丁远洋的贡献。

## 分析师声明

本报告署名分析师在此声明：我们具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，本报告所表述的所有观点均准确地反映了我们对标的证券和发行人的个人看法。我们所得报酬的任何部分不曾与，不与，也将不会与本报告中的具体投资建议或观点有直接或间接联系。

## 一般声明

除非另有规定，本报告中的所有材料版权均属天风证券股份有限公司（已获中国证监会许可的证券投资咨询业务资格）及其附属机构（以下统称“天风证券”）。未经天风证券事先书面授权，不得以任何方式修改、发送或者复制本报告及其所包含的材料、内容。所有本报告中使用的商标、服务标识及标记均为天风证券的商标、服务标识及标记。

本报告是机密的，仅供我们的客户使用，天风证券不因收件人收到本报告而视其为天风证券的客户。本报告中的信息均来源于我们认为可靠的已公开资料，但天风证券对这些信息的准确性及完整性不作任何保证。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，天风证券及/或其关联人员均不承担任何法律责任。

本报告所载的意见、评估及预测仅为本报告出具日的观点和判断。该等意见、评估及预测无需通知即可随时更改。过往的表现亦不应作为日后表现的预示和担保。在不同时期，天风证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。

天风证券的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。天风证券没有将此意见及建议向报告所有接收者进行更新的义务。天风证券的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

## 特别声明

在法律许可的情况下，天风证券可能会持有本报告中提及公司所发行的证券并进行交易，也可能为这些公司提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。因此，投资者应当考虑到天风证券及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突，投资者请勿将本报告视为投资或其他决定的唯一参考依据。

## 投资评级声明

类别	说明	评级	体系
股票投资评级	自报告日后的 6 个月内，相对同期沪深 300 指数的涨跌幅	买入	预期股价相对收益 20%以上
		增持	预期股价相对收益 10%-20%
		持有	预期股价相对收益 -10%-10%
		卖出	预期股价相对收益 -10%以下
行业投资评级	自报告日后的 6 个月内，相对同期沪深 300 指数的涨跌幅	强于大市	预期行业指数涨幅 5%以上
		中性	预期行业指数涨幅 -5%-5%
		弱于大市	预期行业指数涨幅 -5%以下

## 天风证券研究

北京	武汉	上海	深圳
北京市西城区佟麟阁路 36 号 邮编：100031 邮箱：research@tfzq.com	湖北武汉市武昌区中南路 99 号保利广场 A 座 37 楼 邮编：430071 电话：(8627)-87618889 传真：(8627)-87618863 邮箱：research@tfzq.com	上海市浦东新区兰花路 333 号 333 世纪大厦 20 楼 邮编：201204 电话：(8621)-68815388 传真：(8621)-68812910 邮箱：research@tfzq.com	深圳市福田区益田路 4068 号卓越时代广场 36 楼 邮编：518017 电话：(86755)-82566970 传真：(86755)-23913441 邮箱：research@tfzq.com