第20卷第1期 管理科学学报 Vol. 20 No. 1
2017年1月 JOURNAL OF MANAGEMENT SCIENCES IN CHINA Jan. 2017

基于第一HJ距离的线性因子模型设定检验①

郑振龙[1]，孙清泉[2]
(1. 厦门大学管理学院，厦门 361005; 2. 招商银行私人银行部，深圳 518040)

摘要: 模型设定检验是金融建模的重要环节，是减少模型风险的关键步骤. 本文基于Hansen和Jagannathan[1]提出的第一HJ距离模型误设测度，以台湾市场丰富的股票和指数期权数据为基础，对8种常见的线性因子模型(包括基于金融资产价格的线性因子模型)进行模型误设检验，并探究模型设定对参数检验的影响. 研究发现: 在5%的显著性水平下，所有无条件信息模型均存在模型误设问题，仅FF3、LM、VanM、SkewM的条件信息模型成为可接受的正确模型; 同时，是否考虑模型可能误设会影响SDF参数的检验，考虑模型可能误设能更有效地侦测因子的定价能力，而不考虑模型可能误设会高估模型SDF参数的t值绝对值，致使部分因子可能存在“伪”定价现象.

关键词: 第一HJ距离; 模型设定检验; 参数检验
中图分类号: F832.5 **文献标识码**: A **文章编号**: 1007-9807(2017)01-0108-19

0 引言

半个世纪以来，学术界涌现了大量新的资产定价模型. 这些模型的提出，丰富了资产定价的学术研究，为我们正确认知如何确定金融资产价格提供了众多视角. 然而，这样的结果令人喜忧参半: 可喜的是，我们总能找到对资产收益差异的合理解释; 而可忧的是，资产定价模型陷入了“富有困境”(embarrassment of riches). 总览这些资产定价模型，人们会产生如此容易解释以至于并不能较好地刻画金融资产收益差异的不安情绪，尤其当这些新的资产定价模型并无太多共同的经济含义时，这种情绪显得尤为强烈[2].

另一个值得学术界深思的问题是，如此纷繁复杂的资产定价模型中罕有被业界推广应用的模型. 消费资产定价模型显得尤为突出，它多成为解释金融异象和描述投资者消费决策行为的工具，却难以指导实务投资. 从20世纪70年代兴起的量化投资策略，在美欧成熟市场得到广泛应用，打破了理论模型与实务操作无法相辅相成的桎梏. 2009年以来，“量化型”基金产品被陆续引入国内基金市场②，线性因子模型一度成为国内券商研究报告的核心课题. 如此一来，线性因子定价模型理应大显身手，然而，业界的应用水平远远滞后于学术界的研究，仍停留在运用分散投资、财务因子组合的构建[3]及运用CAPM和FF3因子等简单模型来计算系统性风险的层面. 如何结合现有的理论研究筛选出富有经济含义的定价因子和统计意义上可接受的定价模型，为业界构建合理的量化投资策略提供参考，是学术界应重点研究的课题.

对上述问题的忧虑和关注引发了一系列的思

① 收稿日期: 2014-05-03; 修订日期: 2014-12-02.
基金项目: 国家自然科学基金资助项目(71371161); 国家自然科学基金资助项目(71471155).
作者简介: 郑振龙(1966—)，男，福建平潭人，博士，教授，博士生导师，Email: zlzheng@xmu.edu.cn

② 2009年11月11日，国内首只采用量化方法进行主动增强的指数基金，富国沪深300增强基金正式发行，沪深300基金首现“量化”概念. 随后，嘉实、中海、长盛、华商及富国基金公司又先后推出了五只量化型基金产品. 由此，2009年被业界誉为“量化基金投资元年”.

考: 任何资产定价模型都是对金融数据生成过程的一种逼近,哪些模型在统计意义上是可接受的"好"模型? 如何检测这些定价模型的模型误设,从而筛选出最接近真实数据的模型? 模型误设对参数检验有什么影响? 故而,构建一个"好"的模型误设测度,显得尤为必要. 它可用于判断特定模型的误设状况,旨在为我们评价和选择因子模型提供参考. Hansen 和 Jagannathan[1] 根据随机贴现因子定价框架构建了第一 HJ 距离,用于测度线性因子模型的误设程度. 郑振龙和孙清泉[4] 通过理论分析发现,相较传统的模型设定误差测度(如横截面回归误差加和测度(Q_C)和 J 统计量③),第一 HJ 距离极富经济含义,考虑了模型可能误设的情形,且加权矩阵具备不依赖于特定模型等良好的特征,被众多学者应用于测度特定模型的模型误设.

然而,纵观基于第一 HJ 距离的已有研究,大多简单地比较文献中常出现的几个资产定价模型(线性的和非线性的都有),对线性因子定价模型的研究仅停留在 CAPM、FF3 和 FF5 的比较,并未对线性因子资产定价模型进行系统研究,尤其针对金融资产价格隐含定价因子的线性因子模型的模型误设检验的系统性研究尚属空白. 郑振龙[5] 指出金融资产价格蕴含着极为丰富的预期信息,能及时地、高频度地反馈市场参与者对资产价格走势的判断和观点. 因而,基于金融资产价格信息构建的因子具有更好的预测力,尤其是金融衍生品隐含的信息,能为业界量化投资策略的因子构建提供思路,基于金融资产价格提取的线性因子模型理应成为最为重要的一类因子模型,值得学术界深入地探讨. 基于此,借鉴第一 HJ 距离现有的理论研究成果,利用台湾市场丰富的股票和指数期权数据资料,构建基于金融资产价格隐含的各定价因子,系统地检验了 8 种线性因子模型的模型设定问题. 同时,分别测算模型正确设定和模型误设情形下的参数估计标准误,构造两种情形下的参数检验 t 值,由此分析模型设定对参数检验带来的影响.

本文的主要贡献在于: (1) 全文从实证研究层面系统地探讨了 8 种常见的线性因子模型(包括基于金融资产价格的线性因子模型) 的模型误设问题,补充了模型设定检验的研究文献. 同时,本文考虑三种条件信息,对这 8 种模型的条件信息模型也进行了模型误设检验; (2) 通过嵌套模型的模型误设检验可以比较新增因子、条件信息等的模型定价能力,这为选取合适的定价因子提供了参考; (3) 本文比较了是否考虑模型误设对模型 SDF 参数检验的影响,发现传统的参数检验忽略了模型误设,有些因子可能存在"伪"定价现象,这为 Roll 批评提供了经验证据.

1　文献综述

Hansen 和 Jagannatha[1] 利用模型 SDF(随机贴现因子) 与对测试资产正确定价的 SDF 可行集间的最小距离,构建了第一 HJ 距离(δ). 通过对偶转换求解,他们获得了第一 HJ 距离的闭式解,并借鉴 Hansen,Heaton 和 Luttmer[6] 的结论,对多个模型进行了模型设定检验. 然而,Hansen 和 Jagannathan[1],Hansen,Heaton 和 Luttmer[6] 的模型设定检验均假设原模型存在错误设定(即总体 $\delta>0$),样本 $\hat{\delta}$ 与总体 δ 之差服从正态分布. Jagannathan和 Wang[7] 指出,在检测模型设定时,我们更关心某模型是否正确设定,故对特定模型的模型设定检验应假设原模型正确设定. 同时,当假设原模型错误设定时,δ 的总体值未知,且难以估计. 因而,Jagannathan 和 Wang 考虑假设原模型正确设定的情形,构造了样本 $\hat{\delta}$ 的统计分布,检测了 JW 模型、Chen,Roll 和 Ross[8] 的宏观因子模型、Fama 和 French[9] 的 FF3 模型的模型误设,并比较这些模型的误设程度.

也有学者从优化函数的选取视角拓展了第一 HJ 距离的定义,Almeida 和 Garcia[10] 选用 Cressie-

③　Q_C 为线性资产定价模型 Beta 设定时的横截面回归误差的加和,该指标经转变可构造出常见的 R^2、CRST 统计量、似然率检验统计量、GRS 统计量等,故本文讨论 HJ 距离与 Q_C 的差异,并不专门针对 R^2 讨论。J 统计量为 GMM 估计的模型设定误差统计量。

Read 函数集④作为优化函数定义了 CR 距离，并在原假设为模型正确设定(即 CR 距离为 0) 时，发现 CR 距离渐近服从正态分布. 同时，Almeida 和 Garcia [10] 运用股票和债券序列数据，估计幂效用函数的 CCAPM，测度不同模型的 CR 距离.

上述这些研究多考虑无条件信息模型的模型设定情形，Hodrick 和 Zhang [11] 当属条件信息模型的模型误设检验的研究典范. 他们采用月度工业增产值的 HP 滤波、一月哑变量和季度 GNP 数据作为描述经济周期变动的条件信息，加入到模型 SDF 中，以 FF25 组合和国债收益作为测试资产，分别构建了第一 HJ 距离、GMM 的 J 统计量和 Andrews [12] 提出的 SupLM 统计量，比较常数因子模型、CAPM、线性化 CCAPM、JW 模型、Campbell [13] 模型、Cochrane [14] 模型、FF3 和 FF5 共 8 个线性模型的无条件信息和有条件信息情形.

除了应用于线性因子模型的设定检验外，Buraschi 和 Jackwerth [15] 利用第一 HJ 距离探讨期权的确定性波动率定价模型的模型设定检验. 他们指出: 如果期权价格可用确定性波动率模型准确定价，则期权产品是市场冗余证券，可用标的证券和无风险利率完全复制. 这样，市场则是完备的，必然存在唯一的 SDF，可由标的证券和无风险资产的收益线性表示. 如果确定性波动率模型不能较好地刻画期权的价格，则证实随机波动率模型或跳跃模型更接近真实的期权价格生成过程. Buraschi 和 Jackwerth [15] 使用 GMM 的 J 统计量和第一 HJ 统计量，分别检验条件和无条件信息的 Black – Scholes 模型与确定性波动率模型是否模型误设. 研究发现，确定性的波动率模型并不能较好地为期权定价，证实随机波动率模型或跳跃模型可能更适合为期权产品定价.

纵观第一 HJ 距离的模型设定检验研究，大多假设原模型正确设定 [7] ，对资产定价模型进行模型设定检验. 然而，已有文献存在如下不足: 首先，现有研究多止步于常见的因子模型比较，未能对现有的模型从不同类别进行细致地研究. 由于模型构造的思路不同，基于一般均衡的消费资产定价模型和基于无套利假定的线性因子定价模型放一起进行模型设定检验有失偏颇，如何针对特

④ 其函数形式取决于参数ς，当 Cressie-Read 函数的参数ς变化时，CR 距离对应不同的测度: 若$\varsigma = 1$，即为第一 HJ 距离; 若$\varsigma = 0$，则为指数倾斜(Exponential Tilting) 距离; 若$\varsigma = -0.5$，则为林格(Hellinger) 测度; 若$\varsigma = -1$，则为经验似然(Empirical Likelihood) 测度; 若$\varsigma = -2$，则为皮尔森卡方(Pearson's Chi – Square) 距离。

定类别的定价模型系统地进行模型设定检验和比较是一项新课题. 其次,除 Kan 和 Robotti [16] 外,现有文献并未探讨模型设定对参数检验的影响. 同时,Kan 和 Robotti 的研究贡献多侧重于理论层面,并未就金融资产价格隐含因子这一类线性模型进行探讨. Roll [17] 指出,CAPM 的有效性检验实质是股指收益率为市场定价因子和模型正确设定的联合检验,只有在模型正确设定的前提下,针对定价因子的检验才是正确的. 这样,针对 CAPM 的有效性检验会受到模型是否正确设定的影响,与股指收益率为定价因子的检验并不等价. 诚然,如果模型存在模型误设,其因子是否仍具有显著的定价能力,现有的模型设定检验并未就该问题予以细致回答.

2　线性因子模型

Hansen 和 Richard [18] 指出任何资产定价模型都可以表示为随机贴现因子形式,这为计算第一 HJ 距离提供了可能. 较多研究,如 Hodrick 和 Zhang [11] 等,表明条件信息能刻画模型的时变特征,本文拟综合比较资产定价模型的无条件信息和条件信息形式(用 z_t 表示条件信息) ,探究条件信息的引入是否改善了模型设定,及条件信息的定价能力. 同时,介绍这些定价模型的因子的构造,并给出描述性统计分析.

2.1　第一 HJ 距离

Cochrane 在《资产定价》一书中指出,任何金融资产的价格都可以写成如下形式:

$$p_t = E_t [m_{t+1} x_{t+1}] \tag{1}$$

其中 p_t 为 t 时 N 维金融资产价格向量,x_{t+1} 为该资产在 $t+1$ 时相应的 N 维资产回报向量,E_t 为给定 t 时所有公开可得信息的条件期望,m_{t+1} 即为随机贴现因子(SDF) . 由于资产价格序列不平稳,故研究多采用资产收益率数据. 本文也将使用收益率形式的 SDF 定价框架,用总收益率形式表达式(1) ,并利用迭代期望法则,将 SDF 的条件信息形式简化为无条件信息形式. 假定 $\underline{\Re}$ 为 L^N(N 维实数空间) 一个子集,表示测试资产的总收益率空间,N 维资产总收益率 R 满足 $R \in \underline{\Re}$,其期望和方差均

存在,且资产收益率的二阶原点矩 $U = E[R'R]$ 可逆. 若市场无套利时,必定存在严格正的随机贴现因子(SDF) m 对 N 维资产总收益率 R 准确定价

$$1_N = E[mR],m > 0,\forall R \in \underline{\Re} \quad (2)$$

若市场是完全的,m 是唯一的,否则,存在众多的 m 满足式(2). 同时,由于任何特定模型都是对真实数据生成过程的一种近似,假定模型所用的 SDF(y)可被 K 个系统性风险因子 f 近似

$$y = \gamma_0 + \gamma_1 f = \gamma' x \quad (3)$$

其中 $x = [1,f]'$,$\gamma = [\gamma_0,\gamma_1]'$. 式(3)为任何线性因子模型 SDF 的形式,不同模型的差异在于定价因子 f 不同. Hansen 和 Jagannathan[1] 构建了模型误设的第一 HJ 测度,标记为 δ,测量特定模型 SDF(y)与可行 SDF 集合的最小平方距离,即

$$\delta = \min_{m \in M} \| y - m \| = \min_{m \in M} \sqrt{E(y - m)^2} \quad (4)$$

其中 $M = \{m: E[mR] = 1_N, \forall R \in \underline{\Re}\}$ 表示所有可行的 SDF 集合.

2.2 线性因子模型

1) CAPM 模型

资本资产定价模型(CAPM)是研究学者使用的基准模型,故也将其作为研究的基准模型. 其模型 SDF 的无条件和条件信息形式分别设定为

$$y_{t+1} = \gamma_0 + \gamma_1 r_{M,t+1}$$

$$y_{t+1} = (\gamma_0 + \gamma_{0z} z_t) + (\gamma_1 + \gamma_{1z} z_t) r_{M,t+1} \quad (5)$$

其中 $r_{M,t+1}$ 为市场组合的超额收益率; z_t 为条件信息,γ_{0z} 和 γ_{1z} 分别为相应的参数.

现实中并无确切的市场组合,仅能寻求市场组合的代理指标,本文选用台湾股票加权指数作为市场组合的代理,主要是考虑到台湾加权指数具有较好的代表性. 这主要体现在: 首先,台湾上市股票具有较好的代表性. 台湾股市大致可分为交易所上市、上柜和兴柜三个市场. 从表 1 可知,台湾股市的股票主要集中在上市和上柜交易,兴柜公司家数较少. 尽管交易所和柜台交易股票的公司数量相差不大,但就发行总股数和公司市值总值而言,上市股票是上柜股票的 10 倍左右. 这样,台湾上市股价指数能基本代表整个市场. 其次,台湾交易所的股价指数编制于 1966 年,而柜台买卖中心加权股价指数编制于 1995 年,台湾股价指数编制和使用的历史更加久远,其稳定性和认可度更高.

表 1 台湾股市的基本统计

Table 1 The basic statistics on Taiwan stock market

指标	发行总股数(百万股)			公司市值总值(百万元)		
年别	上市	上柜	兴柜	上市	上柜	兴柜
2013	687 016	66 238	30 802	24 686 575	2 327 565	653 846
2012	665 452	66 952	29 624	21 464 246	1 745 586	447 222
2011	632 636	73 304	33 501	19 288 911	1 423 200	507 732
2010	601 712	69 599	30 915	23 975 847	1 986 785	616 959
2009	602 051	76 739	25 927	21 196 306	1 924 954	371 773
2008	581 626	70 934	23 633	11 751 674	774 934	149 268
2007	573 394	71 888	23 711	21 634 710	1 878 135	363 419
2006	558 053	75 068	27 950	19 480 993	1 932 611	302 922

指标	上市(柜)公司(家)			市盈率(倍)			市净率(%)		
年别	上市	上柜	兴柜	上市	上柜	兴柜	上市	上柜	兴柜
2013	866	659	262	22.02	38.83	153.43	1.74	2.05	1.89
2012	840	639	289	21.34	31.18	−42.76	1.58	1.54	1.35
2011	824	608	283	19.29	36.97	−109.09	1.47	1.31	1.23
2010	785	563	292	16.04	25.89	21.05	1.91	1.94	1.68
2009	756	546	234	25.82	−170.20	113.44	1.84	2.05	1.49
2008	723	539	241	24.39	−10.04	−4.38	1.14	0.86	0.67
2007	705	546	255	14.60	31.25	−15.09	1.91	1.79	1.35
2006	694	531	239	18.45	16.21	34.25	1.90	1.91	0.99

数据来源: 由 Wind 数据库整理

2) FF3 模型

Fama 和 French[2] 提出规模因子和账面市值比因子是解释股票收益率差异的重要定价因子,

且三因子模型在大多成熟市场上都表现很好，常被学术界和业界用于业绩评估和资产定价的基准模型. 故考虑 FF3 因子模型，将其 SDF 的无条件和条件信息形式设定为

$$
\begin{aligned}
y_{t+1} &= \gamma_0 + \gamma_1 r_{M,t+1} + \gamma_2 r_{SMB,t+1} + \gamma_3 r_{HML,t+1} \\
y_{t+1} &= (\gamma_0 + \gamma_{0z} z_t) + (\gamma_1 + \gamma_{1z} z_t) r_{M,t+1} + \\
&\quad (\gamma_2 + \gamma_{2z} z_t) r_{SMB,t+1} + (\gamma_3 + \gamma_{3z} z_t) r_{HML,t+1}
\end{aligned}
\tag{6}
$$

其中 $r_{M,t+1}$ 为市场组合的超额收益率，$r_{SMB,t+1}$ 和 $r_{HML,t+1}$ 分别为规模因子模拟组合和账面市值比因子模拟组合收益; γ_{0z}、γ_{1z}、γ_{2z} 和 γ_{3z} 分别为相应的参数.

与市场组合的构造取样有所不同，本文综合利用上市和上柜股票构造台湾股票市场的 SMB 因子和 HML 因子，原因有二: 一方面，我们使用台湾所有上市上柜股票构造测试资产组合，为保证研究的一致性，故构造 SMB 和 HML 因子时，理应同时选择上市和上柜股票; 另一方面，如表 1 所示，上市和上柜股票虽然市值差别较大，但账面市值比指标(用市净率表示) 差别不大，故应把上柜股票纳入组合.

3) 流动性资产定价模型(LM)

从已有的学术研究来看，流动性风险的研究主要有两种思路: 一是把流动性风险因素直接加入 CAPM，构造流动性调整的 CAPM(即 LA－CAPM) ，如 Acharya 和 Pedersen[19]. 简单而言，先测算个股的流动性，再聚合成市场总体的流动性，最后提取流动性冲击的成分，代理流动性成本. 二是采用不同的流动性测度，类似规模因子(SMB) 、账面市值比因子(HML) ，构造基于流动性的模拟组合，如 Sadka[20] 等. 由于台湾股票市场采用指令驱动的交易制度，且台湾股市拥有高比例的个人投资者，投机风气浓烈，远没达到发达国家市场的完善程度，有着严苛的假设条件的 CAPM 难以适用于台湾股票市场. 同时，流动性本身是一个难以明确界定的概念，从宽带、深度、即时性和弹性四个维度可以衍生出众多的流动性指标，如买卖价差、有效价差、换手率、交易量、Amihud 非流动性指标、Amivest 流动性比率、价格反

转指标,LOT 指标等,不同的流动性指标测算了不同的流动性成本. Acharya 和 Pedersen[19] 的研究结论受制于流动性指标的选择,依赖于指标数值的大小,而基于流动性的模拟组合仅需要流动性指标的排序信息,受流动性指标选取的影响更小. 因而,我们选择构造流动性的模型组合来研究流动性风险.

Brunnermeier 和 Pedersen[21] 从理论角度阐述了 Amihud 非流动性的合理性,而多数经验研究表明 Amihud 非流动性指标,能系统和准确地衡量金融资产受流动性冲击的程度. 因而,本文采用规模和 Amihud 非流动性指标对所有样本股票进行分组,构造非流动性模拟组合收益率. 这样,得到流动性风险定价模型 SDF 的设定形式

$$\begin{aligned} y_{t+1} &= \gamma_0 + \gamma_1 r_{M,t+1} + \gamma_2 r_{SMB,t+1} + \\ &\quad \gamma_3 r_{HML,t+1} + \gamma_4 r_{MML,t+1} \\ y_{t+1} &= (\gamma_0 + \gamma_{0z} z_t) + (\gamma_1 + \gamma_{1z} z_t) r_{M,t+1} + \\ &\quad (\gamma_2 + \gamma_{2z} z_t) r_{SMB,t+1} + (\gamma_3 + \gamma_{3z} z_t) r_{HML,t+1} + \\ &\quad (\gamma_4 + \gamma_{4z} z_t) r_{MML,t+1} \end{aligned} \tag{7}$$

其中 $r_{M,t+1}$ 为市场组合的超额收益率, $r_{SMB,t+1}$、$r_{HML,t+1}$ 和 $r_{MML,t+1}$ 分别为规模因子模拟组合、账面市值比因子和非流动性因子模拟组合收益; γ_{0z}、γ_{1z}、γ_{2z}、γ_{3z} 和 γ_{4z} 分别为相应的参数.

由 ICAPM 可知,投资机会集是时变的,投资者对机会集的不利变动必然要求相应的风险补偿. 然而,ICAPM 并未明确指出这些表示投资机会集的状态变量. 除了市场流动性这种可观测的指标外,学术研究多从收益率分布引入状态变量,波动率风险和偏度风险成为研究的焦点. 由此,本文也探究波动率风险和偏度风险模型.

基于波动率的资产定价模型大致可区分为股票市场的波动率风险定价模型和衍生产品的波动率风险定价模型,由于构造的方法和利用的信息不同,本文分别考察这两类模型.

4) 基于股票市场的波动率定价模型(VM)

不同于流动性指标,市场(个股) 波动率测量了一段时间内的变动,不可直接观察. 难以寻找到合理的波动率代理变量,对所有个股进行分组,通

过构造波动率模拟组合的方法来研究股票市场波动率的横截面定价⑤. 因而,借鉴郑振龙和汤文玉[22]的做法,利用台指日收益率数据,提取市场波动率冲击信息,建立市场超额收益率、波动率信息冲击的两因子模型 SDF

$$
\begin{aligned}
y_{t+1} &= \gamma_0 + \gamma_1 r_{M,t+1} + \gamma_2 \Delta V_{t+1} \\
y_{t+1} &= (\gamma_0 + \gamma_{0z} z_t) + (\gamma_1 + \gamma_{1z} z_t) r_{M,t+1} + \\
&\quad (\gamma_2 + \gamma_{2z} z_t) \Delta V_{t+1}
\end{aligned} \tag{8}
$$

其中 $r_{M,t+1}$ 为市场组合的超额收益率，ΔV_{t+1} 表示市场波动率冲击.

5) 基于期权市场的波动率定价模型(IVM)

1987 年股灾后,方差的随机变动特征被学术界广泛认知,对方差随机建模成为资产定价模型的重要拓展方向. 本文借鉴 Neuberger[23] 的思路,通过定义加和性质和广义方差过程,设计出了公平的方差互换合约,提取期权价格隐含的方差风险价格因子,其模型设定为

$$
\begin{aligned}
y_{t+1} &= \gamma_0 + \gamma_1 r_{M,t+1} + \gamma_2 xv_{t+1} \\
y_{t+1} &= (\gamma_0 + \gamma_{0z} z_t) + (\gamma_1 + \gamma_{1z} z_t) r_{M,t+1} + \\
&\quad (\gamma_2 + \gamma_{2z} z_t) xv_{t+1}
\end{aligned} \tag{9}
$$

其中 $r_{M,t+1}$ 为市场组合的超额收益率，xv_{t+1} 表示方差风险价格.

相较而言,人们对偏度的认识滞后很多. 这有着多方面原因: 首先,市场偏度难以刻画. 同波动率指标一样,偏度在股票市场也只能用一段时间的收益率才能计算偏度⑥,市场波动具有典型的持续性特征,可被 GARCH 族模型捕捉,但市场偏度的典型特征是否可由 GARCH 族模型刻画,在学术界未有共识,故直接提取市场偏度的冲击来研究偏度的定价行为值得商榷. 其次,个股指标和市场指标的关系不明确. 个股波动率和市场波动率倾向于同增同减,而较多经验证据表明个股收益率呈现正偏,市场收益率负偏. 市场波动在市场行情下挫时大于在市场行情上行时,而市场偏度并无明显特征. 最后,偏度为波动的非对称性,致使对偏度的经济含义的理解有所差异. 传统金融理论对金融资产收益的波动性有明确的态度区分,投资者厌恶波动(风险) ,但对偏度的理解较为复杂,有个股偏度和协偏度两个视角. 个股高正偏度常与“大概率获小亏损,小概率博大收益”的彩票特性相联系,投资者倾向于偏好高偏度. 然而,如何界定协偏度⑦,如何理解协偏度的定价机理较为困难. 本文综合协偏度和偏度的已有研究,构造如下三个模型.

6) 市场协偏度模型(HSM)

Harvey 和 Siddique[24] 认为偏度水平会影响个股的收益率,他们通过定义协偏度,引入了资产收益的高阶矩信息,构建了市场协偏度模型

$$
\begin{aligned}
y_{t+1} &= \gamma_0 + \gamma_1 r_{M,t+1} + \gamma_2 r_{M,t+1}^2 \\
y_{t+1} &= (\gamma_0 + \gamma_{0z} z_t) + (\gamma_1 + \gamma_{1z} z_t) r_{M,t+1} + \\
&\quad (\gamma_2 + \gamma_{2z} z_t) r_{M,t+1}^2
\end{aligned} \tag{10}
$$

其中 $r_{M,t+1}$ 为市场组合的超额收益率，$r_{M,t+1}^2$ 表示市场协偏因子,由市场组合超额收益的平方引入.

值得注意的是,Harvey 和 Siddique 将个股 i 的协偏度定义为 $Cov(r_i, r_M^2)$,刻画个股收益率 r_i 与市场波动 r_M^2 的关系. 其经济直觉并不如 $Cov(r_i^2, r_M)$ 来得直接,因为经验证据表明个股波动率随同市场行情变差而增大,这种设定更多的是为了配合 SDF 的设定形式.

7) 期权协偏度模型(VanM)

Buraschi 和 Jackwerth[15] 等研究表明期权并非冗余证券; Vanden[25] 指出,Harvey 和 Siddique[24] 的模型设定忽略了衍生证券对权益收益解释的可能性,可以用期权回报的天然非对称性刻画市场波动的非对称性. Vanden 利用分段指数效用函数导出金融资产的定价表达式,发现 SDF 可表示为看涨期权和市场收益的线性函数,其 SDF 形式为

---

⑤ 当然,如果可以用某个指标直接衡量或提取出某一时点的个股波动率,仍然可以采用因子分组、构造模拟组合的形式研究波动率的定价.

⑥ 衍生产品市场有所不同,可以利用衍生品的不同执行价格、不同到期日,构造方差(或) 偏度互换合约来提取各时点的隐含方差和隐含偏度.

⑦ 在统计上,两个随机变量 X 和 Y 的协偏度可定义为 $Cov(X, Y^2)$ 或 $Cov(X^2, Y)$,然而,在金融资产定价中,这两个定义并不等价,$Cov(r_i, r_M^2)$ 表示个股收益率 r_i 与市场波动的关系,而 $Cov(r_i^2, r_M)$ 表示个股收益波动率与市场行情的关系.

$$y_{t+1} = \gamma_0 + \gamma_1 r_{M,t+1} + \gamma_2 r_{O,t+1} + \gamma_3 r_{M,t+1}^2 + \gamma_4 r_{O,t+1}^2 + \gamma_5 r_{M,t+1} r_{O,t+1}$$

$$y_{t+1} = (\gamma_0 + \gamma_{0z} z_t) + (\gamma_1 + \gamma_{1z} z_t) r_{M,t+1} + (\gamma_2 + \gamma_{2z} z_t) r_{O,t+1} + (\gamma_3 + \gamma_{3z} z_t) r_{M,t+1}^2 + (\gamma_4 + \gamma_{4z} z_t) r_{O,t+1}^2 + (\gamma_5 + \gamma_{5z} z_t) r_{M,t+1} r_{O,t+1} \tag{11}$$

其中 $r_{M,t+1}$ 为市场组合的超额收益率，$r_{O,t+1}$ 表示期权的收益率；$r_{M,t+1}^2$、$r_{O,t+1}^2$ 和 $r_{M,t+1} r_{O,t+1}$ 刻画了 SDF 随市场收益和市场波动的非线性变动成分；γ_{0z}、γ_{1z}、γ_{2z}、γ_{3z}、γ_{4z} 和 γ_{5z} 分别为相应的参数. 对比式(9)和式(10)可以发现，Vanden 模型为 Harvey 和 Siddique 的市场协偏度模型的一般形式.

借鉴 Vanden[25] 的方法，利用台指看涨期权数据，构造四个不同在值程度的期权月度收益时间序列. 首先，为构造第 n 月的期权收益，挑选所有在 $n+1$ 月到期的期权，组成 n 月近到期期权集合. 考虑到交易的活跃性，排除在值程度不在 ±4% 范围内的台指看涨期权合约. 其次，界定在值程度. 标的资产价格为执行价格 ±2% 范围内的，可视为近平价情形，而虚值(OTM)和实值(ITM)看涨期权由标的资产价格为执行价格的 −4% 到 −2% 范围和 2% 到 4% 范围界定. 最后，计算各期权组合的月度收益序列. 近平价(Near-the-money, NTM)看涨期权组合收益，由包含近平价情形的所有看涨期权的价值加权收益计算得到. 相应地，虚值(OTM)期权收益和实值(ITM)期权收益由所有虚值和实值期权的收益价值加权计算，而多空组合(LS)收益由多头 OTM 和空头 ITM 构造，即为熊市差价组合收益.

8) 期权隐含偏度因子模型(SkewM)

随着金融市场每日的信息更新，市场偏度随着指数收益分布的变化而变化，市场偏度本身的变动引发了市场投资机会集的改变，这种时变的市场偏度是否带来了风险？本文借鉴 Neuberger[23] 的思路，在互换合约的统一框架下构建偏度互换合约，利用合约的固定端提取中性测度偏度，合约的浮动端提取已实现偏度，将两者之差与隐含偏度的比值界定为偏度风险价格. 由此，将期权隐含偏度因子模型设定为

$$y_{t+1} = \gamma_0 + \gamma_1 r_{M,t+1} + \gamma_2 xs_{t+1}$$

$$y_{t+1} = (\gamma_0 + \gamma_{0z} z_t) + (\gamma_1 + \gamma_{1z} z_t) r_{M,t+1} + (\gamma_2 + \gamma_{2z} z_t) xs_{t+1} \tag{12}$$

其中 xs_{t+1} 表示方差风险价格.

2.3 描述性统计分析

1) 样本数据

本文选用台湾市场的金融数据进行定价模型设定检验研究，主要考虑到本文需要从股指期权价格中提取市场定价因子，台湾金融市场于 2001 年 12 月 24 日推出指数期权，为提取定价因子提供了可能. 为保持全文研究的一致性，本文选择的股票样本为 2002 年 1 月 1 日至 2013 年 4 月 30 日所有上市、上柜个股股票和台湾加权股价指数. 所有数据均来自台湾经济新报(TEJ)数据库，主要包括个股的开盘价、最高价、最低价、收盘价、成交值、成交量、周转率、市值、市净率、TSE 产业别、TEJ 产业别信息. 为构造组合的需要，本文只考虑 2013 年之前上市上柜的股票，共 1 473 只个股. 其中，839 只上市股票，634 只上柜股票. 同时，样本剔除所有股票上市首日的观测值. 另外，本文的宏观变量选用消费物价指数(CPI)、信用价差和期限价差，均来源于 TEJ 数据库.

同时，本文选取台湾股价指数期权提取定价因子，时间跨度为 2002 年 1 月至 2013 年 4 月，包括了每只期权的期权代码、简称、买/卖权、履约方式、履约价、现货码、现货名、上市日、到期月、最后交易日、最后结算价、市场别、交易日、期权当日开盘价⑧、期权当日最高价、期权当日最低价、期权收盘价、期权结算价、当日成交量、未冲销契约数、目标证券价格、波动率、隐含波动率、剩余期间、剩余交易日、杠杆倍数、一年定存利率、期权理论价值、期权高估价、Delta、Gamma、Vega、Theta、Rho 等变量，所有数据均源于 TEJ 数据库. 同时，按下

⑧ 值得注意的是，台湾期货市场虽然引入了做市商制度，但与西方纯做市商制度下报价驱动市场不同，台指期权市场做市商制度并非由做市商报出证券的买卖价格，而是由交易者的限价指令委托，通过高效快速的电子交易系统撮合交易，因此，台湾期权市场仍属于指令驱动市场.

列原则筛选数据: 1) 台湾期权交易所从 2012 年 11 月起，发行交易期限较短的期权，一般存续期不足 7 日，在期权代码中用 W 予以标示. 这类期权交易并不活跃，且不利于提取月度定价信息，故删除这类期权; 2) 去掉 delta 大于 1 或小于 −1 的观测值; 3) 删除无交易(成交量为 0) 的观测值; 4) 删除不符合无套利条件的期权观测值. 根据无套利条件，看涨期权价格在 $[\max(S_t - Ke^{-r(T-t)}, 0), S_t]$，而看跌期权价格在 $[\max(Ke^{-r(T-t)} - S_t, 0), Ke^{-r(T-t)}]$，剔除不在这些区间的期权观测值. 这样，得到 305 565 个观测值.

2) Fama-French 组合

25 个 Fama-French 组合常用于资产定价模型的检验基准，故也选择利用台湾股票，基于规模和账面市值比指标，构造 25 个 Fama-French 组合⑨. 具体如下分组计算方法为: 首先，借鉴洪永淼和林海[26]，吴世农和许年行[27]等的研究，规模选用 T 年 6 月末的市值代表，账面市值比用 $T-1$ 年会计年度股权的账面价值除以 $T-1$ 年 12 月末的市值; 其次，对所有股票，分别取市值及账面市值比的五等分点，两个五等分点相交，将所有的股票分为 25 组; 最后，以划分好的组为依据，计算每组 T 年 7 月到 $T+1$ 年 6 月每月用市值加权的平均收益率.

表 2　FF25 组合收益的描述性统计

Table 2 The descriptive statistics on FF25 portfolio return

	按照 BM 分组					
组别	H	2	3	5	L	H − L
Size	均值					
S	0.015 6	0.014 8	0.015 1	0.015 9	0.015 3	0.000 3
2	0.012 8	0.015 9	0.014 9	0.013 3	0.010 3	0.002 5
3	0.013 6	0.013 8	0.014 3	0.011 8	0.011 1	0.002 5
4	0.010 3	0.011 8	0.013 1	0.010 0	0.007 2	0.003 1
B	0.008 8	0.013 0	0.012 0	0.009 8	0.007 3	0.001 5
S-B	0.006 8	0.001 9	0.003 1	0.006 2	0.008 0	
Size	标准差					
S	0.108 8	0.109 0	0.108 8	0.111 6	0.113 0	0.032 2
2	0.109 1	0.112 8	0.106 8	0.105 2	0.106 7	0.035 2
3	0.106 4	0.106 5	0.104 9	0.103 8	0.102 9	0.038 8
4	0.103 5	0.103 6	0.107 9	0.100 4	0.095 7	0.040 7
B	0.095 1	0.095 8	0.103 3	0.097 7	0.087 2	0.049 1
S-B	0.071 3	0.059 8	0.061 2	0.063 3	0.068 2	
Size	偏度					
S	0.406 0	0.311 5	0.381 7	0.291 7	0.402 2	−0.585 1
2	0.382 0	0.380 7	0.340 8	0.294 4	0.160 0	−0.803 2
3	0.251 6	0.281 2	0.180 5	0.105 6	0.226 5	0.013 4
4	0.246 9	0.201 6	0.218 0	0.100 2	−0.069 0	0.498 8
B	−0.249 3	0.263 4	0.649 4	−0.007 1	−0.102 5	0.151 2
S-B	−0.214 5	0.053 8	−0.152 8	−0.303 6	0.244 8	
Size	峰度					
S	0.722 7	0.604 2	0.430 8	0.656 0	0.506 8	1.937 0
2	1.123 3	0.839 1	0.786 7	0.442 1	0.315 4	6.696 4
3	0.897 1	0.542 1	0.686 2	0.539 7	0.722 5	1.111 1
4	1.096 7	0.478 6	0.758 8	0.423 1	0.384 9	0.516 9
B	1.067 3	1.760 3	3.036 3	1.413 2	0.497 4	0.231 5
S-B	0.242 7	0.090 2	0.066 7	2.788 0	1.382 7	

从表 2 可以看出，FF25 组合收益呈现明显的因子结构特征: 低规模组合平均收益率高于高规

⑨ Lewellen，等[2]指出，FF25 组合具有明显的因子分组结构，可能影响资产定价模型的检验，因而，本文也利用行业组合作为测试资产组合进行所有的分析，结论基本一致。限于篇幅未能列出，若有需要可联系作者。

模平均收益率，高账面价值比组合平均收益率高于低账面价值比组合，这从 H - L 组合收益和 S-B 组合收益的均值可以看出. 同时，几乎所有组合都呈现正偏性，但从峰度数值来看，样本内 25 个股票组合的尖峰特征并不突出，峰度数值均在 1 左右，极个别组合具有明显的尖峰特征.

3) 定价因子

利用台湾股票市场和期权市场的资料，提出本文的资产定价模型的定价因子，其基本描述性统计量如表 3 所示. 市场指数收益率(rmrf) 为正，而多空组合超额收益率(rorf) 为负，表明整个样本期内，股票市场是牛市行情，而通过构造熊市价差组合(LS) 的期权策略，获得负的收益，符合经济直觉.

表 3 定价因子描述性统计分析

Table 3 The descriptive statistic analysis on pricing factors

统计量	rmrf	smb	hml	mml	Dvol	xv	srmrf	rorf	srorf	rmro	xs
Panel A 描述性统计分析											
均值	0.007 2	0.004 4	0.002 1	0.004 1	-0.006 2	-0.234 5	0.008 1	-0.005	0.017 9	-0.000 5	0.438 1
标准差	0.090 2	0.040 8	0.027 6	0.052 8	0.226 9	0.424 8	0.013 5	0.134 1	0.088 7	0.009 3	0.200 9
偏度	0.021 4	-0.186 2	-0.005 5	-1.745	0.707 3	1.395 7	3.208 1	4.491 5	6.845 4	-1.823 9	-2.083 2
峰度	0.804 1	0.270 8	0.678	5.543	1.173	4.683 1	12.625 1	24.045	48.915 6	24.534 8	9.745 4
Panel B 相关性结构											
smb	0.148 0										
hml	0.171 7	-0.285 8									
mml	-0.554 5	-0.507 6	0.049								
dvol	-0.098 9	0.163 2	-0.089 8	-0.167							
xv	0.097 1	-0.000 5	-0.009 3	-0.102 4	-0.050 9						
srmrf	0.109 4	0.141 8	0.054 4	-0.340 2	0.047 1	0.061 7					
rorf	-0.035 2	-0.032 3	-0.055 3	-0.000 4	0.068 4	-0.045 4	-0.064 1				
srorf	-0.042 7	-0.018 2	-0.036 6	0.015 3	0.080 4	-0.003 5	-0.050 2	0.882 9			
rmro	-0.182 7	0.141 6	-0.062 6	0.094 8	0.080 1	-0.073 6	-0.052 2	-0.171 6	-0.156 1		
xs	0.092 0	0.069 4	-0.036	-0.021 8	0.038 2	-0.468 5	-0.064 3	-0.017 6	-0.010 5	-0.027 8	1

注: rmrf 为台指超额收益率，smb 为规模因子模拟组合收益率，hml 为账面市值比因子模拟组合收益率，mml 为非流动性因子模拟组合收益率，dvol 为市场波动率冲击，xv 为期权价格中提取的市场波动率价格，srmrf 为股票市场协偏指标，rorf 为期权多空组合(LS) 超额收益率，srorf 为期权市场协偏指标，rmro 为股票市场超额收益率和期权市场超额收益率的交乘项，xs 为期权价格中提取的市场偏度价格.

构造的三个模拟组合收益率收益率为正，表明利用这些因子，构造对冲策略，的确能获得超额收益率. 除了非流动性因子组合收益率呈现明显的有偏尖峰外，rmrf、smb 和 hml 较为接近正态分布. 方差风险价格 XV 为 -23.45% ，偏度风险价格 XS 高达 43.81% ，表明台湾市场存在显著为负的方差风险价格和正的偏度风险价格. 这与投资者爱好方差的随机波动和厌恶偏度的随机波动的经济直觉相吻合，且 XV 和 XS 的偏度和峰度表明隐含风险价格均呈现尖峰厚尾特征. 从相关性结构来看，除了期权超额收益率和期权超额收益率平方相关性较高外，其余定价因子间的相关系数均较低，在某种程度上表明定价因子间携带不同的定价信息. 同时，在对多因子线性模型检验时，不受共线性问题的干扰.

3) 条件信息

Harvey [28] 指出，计量分析中使用不同的前定条件信息变量会影响风险收益关系的估计，而随意使用某种条件信息可能导致研究结论的信服力不强，故本文综合考虑多种条件信息，一是台湾市场的消费者物价指数，二是台湾债券市场的信用溢酬，三是台湾债券市场的期限溢酬. 其描述性分析如表 4 所示，可以发现，台指市场的通货膨胀率并不太高. 相较而言，台湾市场的信用溢酬明显高于期限溢酬，且信用溢酬的波动都较大. 信用溢酬的波动接率近期限溢酬波动率的 4 倍，且呈现明显的正偏特征. 从相关性来看，CPI 数据与信用溢酬和期限溢酬相关性并不明显，均高于 -0.3，呈负相关，而期限溢酬和信用溢酬的相关性高达

0.650 8，这可能与信用溢酬的界定有关，难以排除期限溢酬的影响.

表 4　条件信息的描述性统计分析

Table 4 The descriptive statistic analysis on conditional information

变量	均值	标准差	偏度	峰度	cpi	crepr	terpr
cpi	1.196 0	1.573 7	0.283 7	0.177 6	1.000 0		
crepr	2.198 0	2.327 8	1.021 1	−0.438 6	−0.149 5	1.000 0	
terpr	1.332 2	0.608 0	0.012 6	−1.174 3	−0.291 5	0.650 8	1.000 0

注：cpi 为消费者物价指数，crepr 为台湾债市的信用溢酬，terpr 为台湾债市的期限溢酬.

值得注意的是，本文并未使用部分文献，如 Hodrick 和 Zhang(2001)，所采用的 GDP 数据，主要考虑到 GDP 数据的频率为季度，与本文的数据研究不匹配. 如果借鉴部分文献的方法，生硬地采用 HP 滤波估计出 GDP 的月度数据，由于季度数据较少，其准确性难以保证.

3　模型设定检验

3.1　模型设定检验的理论分析

从文献综述可知，针对某一特定模型的模型误设检验，有原模型正确设定和原模型错误设定两类情形. Jagannathan 和 Wang[7] 指出，当假定原模型错误设定时，第一 HJ 距离平方(δ^2)的总体值未知，且难以估计. 因而，考虑原假设为模型正确设定的情形，对样本 δ^2 进行检验，判断模型是否存在误设. 对于线性 SDF 而言，借鉴 Jagannathan 和 Wang[7] 的结论，在原模型正确设定($\delta = 0$)的假设下有

$$T\hat{\delta}^2 : \sum_{i=1}^{N-K-1} \xi_i x_i \tag{13}$$

在式(13)中，x_i 为服从自由度为 1 的卡方分布，即 $x_i : \chi_1^2$，而权重 ξ_i 为下列矩阵 A 的正的特征根

$$A = S^{-\frac{1}{2}} V_{22}^{-1} S^{\frac{1}{2}} - S^{-\frac{1}{2}} V_{22}^{-1} D (D' V_{22}^{-1} D)^{-1} D' V_{22}^{-1} S^{\frac{1}{2}}$$

其中 $V_{22} = U = E(R'R)$，为收益率的原点二阶矩. S 为定价误 e 的方差协方差矩阵，$\sqrt{T}\hat{e}(\gamma) \to N(0_N, S)$. 考虑到序列的自相关和异方差，$S$ 通常用 Newey-West 方差协方差矩阵代替，而 $S^{\frac{1}{2}}$ 则为 SCholesky 分解的上三角矩阵. 若加权矩阵为 Hansen 的最优矩阵，则 GMM 估计的 J 统计量渐近服从自由度为 $N-K-1$ 的 χ^2 分布. 然而，第一 HJ 距离的加权矩阵 U^{-1} 不是最优的，则 $T\hat{\delta}^2$ 不再渐近服从 χ_{N-K-1}^2，而是由一组服从 χ_1^2 分布的随机变量的线性加权和所构成的分布确定，其权重为矩阵 A 的正特征根.

当然，由 χ_1^2 分布的性质可知，若对所有的 i，都有 $\xi_i = 1$，则 $\sum_{i=1}^{N-K-1} x_i$ 仍然服从卡方分布，即 $\sum_{i=1}^{N-K-1} x_i : \chi_{N-K-1}^2$；若 $\xi_i \neq 1$，且 $\xi_i > 0$，则 $\xi_i x_i$ 服从 Gamma 分布，即 $\xi_i x_i : \Gamma(k = 1/2, \theta = 2\xi_i)$，这样，$\sum_{i=1}^{N-K-1} \xi_i x_i$ 则为 Gamma 分布的加和. 但 Gamma 分布并不具有加和性质，使得 $\sum_{i=1}^{N-K-1} \xi_i x_i$ 的具体分布不可知，可以通过截断界来获得小于某一数值 c 的概率，即 $p(\sum_{i=1}^{N-K-1} \xi_i x_i \leq c)$.

3.2　经验研究

对于特定模型，如果其第一 HJ 距离的值为“0”，表明该模型正确设定，否则该模型存在模型误设. 这样，可以通过式(13)构造测试资产组合，对特定模型的第一 HJ 距离进行样本检验，以证实该模型是否误设. 下文就选择的 8 个定价模型的无条件信息和条件信息版本共 16 个版本，利用数据样本分别对其进行模型设定检验.

鉴于 FF25 组合常被用作资产定价模型研究的测试资产组合，本文也以它为基本检验工具，对 CAPM、FF3 因子模型(FF3)、流动性因子模型(LM)、股票市场波动率模型(VM)、期权价格提取波动率模型(IVM)、Harvey 和 Siddique[24] 的股票协偏模型(HS)、Vanden[25] 的期权协偏(VanM)和期权价格提取的偏度模型(SkewM)进行特定模型的误设检验. 为了比较 GMM 和第一 HJ 距离优化的差异，也报告 GMM 的模型设定检验结果. 同时，Ghysels[29] 指出，使用条件信息虽能改善资产定价模型，但也可能造成参数的不稳定性，以至于弱化了定价模型的样本外表现. 因而，借鉴 Ghysels 的思路，使用 Andrews [12] 提出的 supLM 检验，甄别参数的不稳定性，分析结果如表 4 所示.

表 4 FF25 组合的模型设定检验

Table 4 The model specification test on FF25 portfolio

模型	CAPM	FF3	LM	VM	IVM	HSM	VanM	SkewM
Panel A 无条件信息情形								
$\hat{\delta}$	0.371	0.328	0.321	0.370	0.365	0.367	0.358	0.368
$p(\delta=0)$	0.001	0.022	0.025	0.004	0.007	0.007	0.014	0.012
$se(\hat{\delta})$	0.083	0.086	0.086	0.084	0.083	0.083	0.091	0.083
2.5% CI(δ)	0.248	0.230	0.230	0.242	0.248	0.225	0.201	0.240
97.5% CI(δ)	0.579	0.578	0.577	0.576	0.579	0.558	0.574	0.573
J_{GMM}	105.130	47.090	38.440	74.320	57.690	67.020	32.730	67.750
$p(J_{GMM}=0)$	0.000	0.001	0.011	0.000	0.000	0.000	0.036	0.000
SupLM	9.870	20.390	20.050	13.920	8.960	10.950	13.010	7.660
E[m]	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.984
Panel B CPI 为条件信息								
$\hat{\delta}$	0.334	0.255	0.248	0.327	0.339	0.311	0.271	0.331
$p(\delta=0)$	0.000	0.237	0.249	0.009	0.012	0.013	0.187	0.095
$se(\hat{\delta})$	0.089	0.087	0.088	0.090	0.083	0.096	0.102	0.091
2.5% CI(δ)	0.177	0.167	0.166	0.170	0.135	0.151	0.089	0.172
97.5% CI(δ)	0.536	0.519	0.520	0.533	0.463	0.550	0.502	0.539
J_{GMM}	71.890	27.730	24.940	51.400	49.470	51.340	23.460	42.560
$p(J_{GMM}=0)$	0.000	0.185	0.250	0.001	0.001	0.001	0.267	0.008
SupLM	35.940	46.150	45.640	35.900	32.740	37.550	25.350	44.380
E[m]	0.984	0.985	0.985	0.984	0.985	0.984	0.985	0.985
Panel C 信用价差为条件信息								
$\hat{\delta}$	0.337	0.261	0.249	0.335	0.336	0.309	0.274	0.324
$p(\delta=0)$	0.000	0.201	0.238	0.008	0.012	0.011	0.167	0.087
$se(\hat{\delta})$	0.089	0.089	0.089	0.079	0.078	0.081	0.075	0.077
2.5% CI(δ)	0.216	0.157	0.157	0.215	0.216	0.189	0.142	0.206
97.5% CI(δ)	0.529	0.513	0.515	0.53	0.528	0.512	0.439	0.515
J_{GMM}	77.23	28.31	27.56	47.32	47.14	54.07	25.75	46.76
$p(J_{GMM}=0)$	0.000	0.166	0.153	0.002	0.002	0.000	0.174	0.002
SupLM	116.86*	107.06*	100.65*	117.63*	117.85*	72.74*	60.52*	102.65*
E[m]	0.984	0.984	0.984	0.984	0.984	0.984	0.985	0.984
Panel D 期限价差为条件信息								
$\hat{\delta}$	0.333	0.279	0.253	0.328	0.318	0.308	0.282	0.326
$p(\delta=0)$	0.000	0.171	0.249	0.009	0.013	0.011	0.145	0.087
$se(\hat{\delta})$	0.081	0.087	0.087	0.083	0.083	0.086	0.09	0.081
2.5% CI(δ)	0.212	0.177	0.176	0.204	0.196	0.183	0.132	0.205
97.5% CI(δ)	0.537	0.533	0.53	0.537	0.53	0.529	0.489	0.532
J_{GMM}	81.17	35.98	34.6	48.84	46.4	52.57	32.08	56.76
$p(J_{GMM}=0)$	0.000	0.031	0.031	0.001	0.003	0.000	0.043	0.000
SupLM	60.74*	64.54*	64.06*	52.1	53.54	38.77	38.12	44.97
E[m]	0.985	0.985	0.985	0.985	0.985	0.985	0.985	0.985

注: CAPM、FF3 为基准模型，LM 为加入非流动性因子模型，VM 为波动率冲击模型，IVM 为期权价格提取的波动率模型，HSM 为 Harvey 和 Siddeque [24] 模型，VanM 为 Vanden [25] 模型，SkewM 为期权价格提取的偏度模型. 条件信息分别为台湾市场的 CPI、信用溢酬和期限溢酬变量. 报告特定模型的第一 HJ 距离样本值($\hat{\delta}$)，$p(\delta=0)$ 为原假设 $H_0: \delta=0$ 的 p 值，$se(\hat{\delta})$ 为样本第一 HJ 距离的标准误，CI(δ) 为 δ 95% 置信区间. J_{GMM} 为 GMM 方法下的估计值，$p(J_{GMM}=0)$ 为相应检验的 p 值. SupLM 为 Andrews [12] 的 SupLM 统计检验值，右上角带星号表明该模型在 5% 显著性水平，未能通过 supLM 检验. E[m] 为 $1/(1+r_f)$ 的估计值，其中，r_f 为无风险利率.

表 4 的 Panel A 为无条件信息模型的估计结果. 表中所有无条件信息模型的第一 HJ 距离数

值均大于 3.20. 以第一 HJ 距离的数值作判断排序,上述 8 个模型的定价能力强弱顺序为: LM、FF3、VanM、IVM、HSM、SkewM、VM、CAPM. 模型设定检验 p 值显示,所有无条件信息模型的第一 HJ 距离在 5% 显著性水平下均显著异于 0,即所有无条件信息的定价模型均存在模型误设问题. 该结论与 GMM 估计的 J_{GMM} 统计量检验所得到的结论相一致. 同时,各模型的第一 HJ 距离的标准误相差不大,均在 0.08 附近,表明不同模型 $\hat{\delta}$ 的置信区间存在较大重叠.

Panel B、Panel C 和 Panel D 分别表示 CPI、信用价差和期限价差作为条件信息的条件资产定价模型的模型设定检验结果. 对比数值显示,条件信息模型的第一 HJ 距离(δ)均大幅下降,且流动性因子模型(LM)的第一 HJ 距离(δ)最小,FF3 模型的第一 HJ 距离(δ)次之. 可以粗略地认为 SMB 和 HML 因子有利于解释基于规模和账面市值比因子分组的 FF25 组合. 由第一 HJ 距离的数值判断,前四个模型的优劣顺序一致,后四个排序略有差异. 在 5% 显著性水平下,条件 FF3 模型、条件流动性因子模型(LM)、条件期权协偏度模型(VanM)、期权隐含偏度模型(SkewM)成为不能被拒绝的正确模型. 而且,第一 HJ 距离(δ)的标准误 $se(\hat{\delta})$ 并未随条件信息的引入发生较大的变动,不同模型 $\hat{\delta}$ 的置信区间仍存在较大重叠,不同模型的 $E[m]$ 估计值相差不大.

然而,区别于 Panel A,在 Panel B、Panel C 和 Panel D 中需要注意两点: (1) GMM 的 J_{GMM} 统计量的甄别效果显示,在条件信息下,绝大多数模型仍然是误设模型. 显然,GMM 的 J_{GMM} 统计量较第一 HJ 距离(δ)更为严苛,这与 GMM 估计同时对定价误和加权矩阵优化有关⑩. (2) 以信用价差和期限溢酬作为条件信息时,Andrews [12] 提出的 Sup-LM 检验甄别出条件参数的高度不稳定性,表明条件信息可能在样本内被过度拟合,其定价模型的样本外表现却有所弱化,故应谨慎地看待这样的结论.

4　模型设定假设对参数检验的影响

4.1　对模型误设稳健的参数检验统计量

传统的资产定价研究假定模型正确设定,通过因子的参数检验来判断因子是否具有显著的系统性定价能力. 然而,若定价模型本身可能误设时,传统计量方法下因子定价能力的结论理应受到质疑[17]. 那么,模型设定假设是否影响参数检验,进而影响因子的定价能力? 如何更有效地甄别因子的定价能力? 为此,借鉴 Kan 和 Robotti [16] 的研究,给出并比较是否考虑模型可能误设的参数检验,试图回答上述问题.

假定 $Y_t = [f_t', R_t']'$ 为独立同分布的多元球形分布,具有有限的四阶矩,且多元峰度参数为 κ. 为了表述方便,令 $H = (D'V_{22}^{-1}D)^{-1}$,$u_t = e'V_{22}^{-1}R_t$,且模型 SDF($y$)的均值和方差为 $\mu_y = \gamma_0 + \gamma_1'\mu_1$,$\sigma_y^2 = \gamma_1'V_{11}\gamma_1$. 这样,如果考虑模型可能存在误设时,针对模型 SDF 的参数检验可由下式(14)构造

$$\sqrt{T}(\hat{\gamma} - \gamma) : N(0_{K+1}, V(\hat{\gamma})) \tag{14}$$

式(14)中,$V(\hat{\gamma})$ 为 $\sqrt{T}(\hat{\gamma} - \gamma)$ 的渐近方差,由 Newey-West 估计量($\sum_{-\infty}^{\infty} E[h_t h_{t+j}']$)替代. 具体而言,残差($h_t$)和方差($V(\hat{\gamma})$)由下式给出

$$h_t^{ms} = -HD'V_{22}^{-1}R_t y_t + H[D'V_{22}^{-1}(R_t - \mu_2) - x_t]u_t + \gamma$$

$$V^{ms}(\hat{\gamma}) = [\mu_y^2 + (1+\kappa)\sigma_y^2]H + \begin{pmatrix} \sigma_y^2 - \mu_y^2 + \gamma_0^2 + 2\kappa(\mu_1'\gamma_1)^2 & (\gamma_0 - 2\kappa\mu_1'\gamma_1)\gamma_1' \\ (\gamma_0 - 2\kappa\mu_1'\gamma_1)\gamma_1 & (1+2\kappa)\gamma_1\gamma_1' \end{pmatrix} + \delta^2 H\left([1 + (1+\kappa)\mu_2'V_{22}^{-1}\mu_2]\begin{bmatrix}1\\ \mu_1\end{bmatrix}[1, \mu_1'] + \begin{pmatrix} 0 & 0_K' \\ 0_K & (1+\kappa)(V_{11} - V_{12}V_{22}^{-1}V_{21}) \end{pmatrix}\right)H \tag{15}$$

式(15)为考虑原模型可能存在误设时(mis-

⑩　值得注意的是,在此并非比较第一 HJ 距离和 GMM 两种优化方法的优劣,仅指出两种优化方式下的结论有差异. 第一 HJ 距离和 GMM 的优化本质差别在于加权矩阵. 第一 HJ 距离的加权矩阵并不是最优的,从统计效力而言,GMM 的加权矩阵更好,也更严格. 然而,郑振龙和孙清泉[4] 指出,采用第一 HJ 距离优化,除了有经济含义外,可以比较不同模型,也可以探讨模型设定对参数检验的影响.

specified)，即 $e \neq 0_N$，$u_t \neq 0$，$\delta \neq 0$，残差(h_t)和方差($V(\hat{\gamma})$)的计算式. 而传统的参数检验方法不考虑原模型可能误设，即默认模型正确设定(correctly specified)，则 $e = 0_N$，$u_t = 0$，$\delta = 0$，式(15)中的 h_t 和 $V(\hat{\gamma})$ 可简化为

$$h_t^{cs} = -HD'V_{22}^{-1}R_t y_t + \gamma$$

$$V^{cs}(\hat{\gamma}) = [\mu_y^2 + (1+\kappa)\sigma_y^2]H + \begin{pmatrix} \sigma_y^2 - \mu_y^2 + \gamma_0^2 + 2\kappa(\mu_1'\gamma_1)^2 & (\gamma_0 - 2\kappa\mu_1'\gamma_1)\gamma_1' \\ (\gamma_0 - 2\kappa\mu_1'\gamma_1)\gamma_1 & (1+2\kappa)\gamma_1\gamma_1' \end{pmatrix} \tag{16}$$

对比式(15)和式(16)的 h_t^{ms} 和 h_t^{cs} 可以看出，如果原模型误设时，h_t^{ms} 有额外的一项 $H[D'V_{22}^{-1}(R_t - \mu_2) - x_t]u_t$，其取值是否为0，受 u_t 和 $D'V_{22}^{-1}(R_t - \mu_2) - x_t$ 两项影响. 在以下两种情形下，这一额外项对 h_t^{ms} 的影响很小：一是模型误设测度 δ^2 相对较小. 由于 $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t : N(0,\delta^2)$，若 δ^2 很小，则 u_t 有非常小的方差，这样，额外项倾向于较小；二是测试资产收益能较好地捕捉定价因子. $D'V_{22}^{-1}(R_t - \mu_2) - x_t$ 可看成定价因子对测试资产收益线性回归的误差，若该回归误差较小，额外项的影响也就较小.

这些结论也可以从 $V^{ms}(\hat{\gamma})$ 相较 $V^{cs}(\hat{\gamma})$ 的额外项得到佐证，$V(\hat{\gamma})$ 除了随 δ^2 的增加而增加外，还取决于定价因子对测试资产收益线性回归残差的协方差矩阵，$V_{11} - V_{12}V_{22}^{-1}V_{21}$. 当定价因子是测试资产组合收益率时，这一项倾向于较小，则模型误设的调整项相对可忽略. 然而，若定价因子的信息并不在资产回报空间时，$V_{11} - V_{12}V_{22}^{-1}V_{21}$ 倾向于非常大，考虑模型误设对 $\hat{\gamma}$ 的渐近方差会有大的影响. 因而，在忽略模型可能误设的情形下进行参数检验，并不能有效地侦测因子的定价能力，可能错误地认为这个因子被定价.

4.2 模型设定假设对参数检验的影响

由第4节分析可知，无条件信息模型均存在模型误设. 而在条件信息模型下，只有部分定价模型，如条件 FF3、条件 LM、条件 VanM 和条件 SkewM，能通过第一 HJ 距离的模型设定检验. 一言以蔽之，大多数模型均存在模型误设. 那么，如何在模型可能误设的情形下，检测因子的定价能力将是破解 Roll 批评的重要突破口. 为此，将考察模型设定对模型 SDF 参数检验的影响，关注模型可能误设情形下的因子定价能力. 具体而言，我们使用 FF25 组合作为测试资产，在第一 HJ 距离的优化下，估计无条件信息模型 SDF 的参数数值，并利用式(15)和式(16)分别报告考虑原模型可能误设时的参数 t 值(t_{ms})和不考虑原模型可能误设时的参数 t 值(t_{cs})，进而探讨不同模型设定对 SDF 参数检验的影响. Kan 和 Robotti[16] 发现传统的资产定价研究忽略了模型可能误设的情形，致使宏观因子可能存在伪定价现象. 而本文使用了宏观因子作为条件信息加入到资产定价模型中，须谨慎地判断这些条件信息和条件信息调整因子的定价能力. 若考虑了模型可能误设的参数检验，这些条件信息和条件信息调整因子是否仍然是显著的定价因子？同时，本文构造 Wald 统计量，检验引入条件信息是否有助于改进资产定价模型.

1) 无条件信息情形

表5给出了以 FF25 组合为测试资产时，在第一 HJ 距离优化下，无条件信息模型的 SDF 参数估计值和相应的检验 t 值. 可以发现，对于所有模型 SDF 参数的检验 t 值而言，考虑模型可能误设时的 t_{ms} 值较不考虑原模型可能误设的 t_{cs} 值的绝对值有显著下降. 这表明模型设定不同，对 SDF 参数检验的结论存在显著差异. 如果原模型设定错误而假定模型设定正确，会高估 SDF 参数的检验 t 值，得出因子被显著性定价的错误结论. 在众多无条件信息模型中，无论模型是否误设，r_{rmrf}、r_{smb}、r_{hml}、r_{xv}、r_{rorf}、r_{xs} 都是显著的系统性定价因子. 它们的 t 值绝对值均较大，能显著地拒绝原假设. 然而，r_{dovl} 和 r_{mml} 在不考虑原模型可能误设时，t_{cs} 值的绝对值很大，是显著的定价因子. 若考虑模型

误设时，检验 t_{ms} 值的绝对值很小，其定价能力变得不再显著，这意味着：在考虑模型可能误设时，股票市场波动因子和流动性因子并不是显著的定价因子，而忽略模型误设可能误认为该因子有显著的定价能力. 这些证据充分地表明，对于真正的系统性因子，是否考虑模型误设并不会影响其被定价的结论. 但对一些因子而言，忽略模型误设可能得到错误的结论. 可见，考虑模型可能误设的参数检验能有效地侦测出系统性定价因子.

从单个模型来看，由 IVM 和 SkewM 的 SDF 参数值可知，r_{xv} 和 r_{xs} 的符号相反，这与 Kozhan、Neuberger 和 Schneider [31] 研究结论相一致. 同时，本文针对所有测试资产组合检验 VanM 的 SDF 参数. 从市场整体而言，r_{rmrf}、r_{rorf} 有显著的定价能力且符号相反，而 r_{srmrf}、r_{srorf}、r_{rmro} 的定价能力并不明显，因为它们仅对于部分组合有不同的定价能力. 这与 Vanden [26] 在美国市场发现的结论也相一致.

表 5　是否考虑原模型可能误设的参数检验：无条件信息情形

Table 5 The parameter test considering whether the original model misspecification: No conditional information

	CAPM		FF3			
	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{smb}$	$\hat{\gamma}_{hml}$
估计值	0.38	-4.48	0.39	-4.33	-1.57	-1.49
t_{cs}	42.11	-8.29	30.07	-7.86	-5.09	-7.98
t_{ms}	40.06	-6.30	29.93	-5.87	-2.08	-4.97
	LM					
	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{smb}$	$\hat{\gamma}_{hml}$	$\hat{\gamma}_{mml}$	
估计值	0.39	-3.33	-1.48	-1.44	-3.41	
t_{cs}	21.62	-7.53	-5.76	-5.98	-3.21	
t_{ms}	19.61	-5.48	-3.68	-4.97	-1.78	
	VM			IVM		
	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{dvol}$	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{xv}$
估计值	0.38	-4.53	-0.19	0.48	-4.48	-0.58
t_{cs}	39.97	-7.43	-2.66	30.26	-5.24	-4.01
t_{ms}	41.03	-5.43	-0.53	27.72	-4.34	-3.27
	HS			SkewM		
	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{srmrf}$	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{xs}$
估计值	0.43	-3.38	-5.83	0.09	-3.56	0.31
t_{cs}	14.57	-6.28	-1.86	0.26	-5.44	5.81
t_{ms}	13.17	-4.27	-1.66	0.18	-4.36	4.57
	VanM					
	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{srmrf}$	$\hat{\gamma}_{rorf}$	$\hat{\gamma}_{srorf}$	$\hat{\gamma}_{rmro}$
估计值	0.37	-2.36	-5.49	-0.89	2.89	-2.45
t_{cs}	5.20	-4.85	-1.84	-4.49	0.87	-0.24
t_{ms}	4.08	-2.77	-1.67	-3.35	0.69	-0.19

注：CAPM、FF3 为基准模型，LM 为加入非流动性因子模型，VM 为波动率冲击模型，IVM 为期权价格提取的波动率模型，HSM 为 Harvey 和 Siddeque [24] 模型，VanM 为 Vanden [25] 模型，SkewM 为期权价格提取的偏度模型. t_{CS} 表示不考虑原模型可能误设的 t 统计量，t_{ms} 表示考虑模型可能误设下的 t 统计量.

2) 条件信息

Kan 和 Robotti [16] 研究发现宏观因子或不可交易因子，在模型正确设定下是显著的定价因子. 然而，考虑模型可能误设时，这些因子的定价能力

消失，表明在传统的资产定价研究中，这些因子存在“伪”定价现象，为 Roll[17] 批评找到了证据支持. 显然，考虑模型误设的参数检验能更有效地甄别出这些因子的定价能力. 本文使用的 CPI、信用价差和期限价差作为条件信息，以及条件信息和定价因子组成的新因子是否也具有“伪”定价现象，表 6 对该问题予以回答. 它报告了引入条件信息时，模型设定对参数检验的影响. 由于不同条件信息影响 SDF 参数检验的结论大致相同，仅以 CPI 为条件信息的情形为例分析，如表 6 所示.

表 6　是否考虑原模型可能误设的参数检验：CPI 为条件信息

Table 6 The parameter test considering whether the original model misspecification: CPI as conditional information

	CAPM				FF3							
	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_z$	$\hat{\gamma}_{rmrf^*z}$	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{smb}$	$\hat{\gamma}_{hml}$	$\hat{\gamma}_z$	$\hat{\gamma}_{rmrf^*z}$	$\hat{\gamma}_{smb^*z}$	$\hat{\gamma}_{hml^*z}$
估计值	0.50	-1.69	-0.07	0.64	0.51	-1.93	-1.09	-1.17	-0.07	0.75	-0.79	-1.11
t_{cs}	6.49	-5.03	-2.45	1.48	6.14	-4.86	-3.45	-3.44	-2.33	1.48	-2.66	-3.66
t_{ms}	6.08	-4.84	-1.35	1.26	5.68	-3.63	-2.39	-2.37	-1.23	1.24	-2.55	-2.53
	LM											
	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{smb}$	$\hat{\gamma}_{hml}$	$\hat{\gamma}_{mml}$	$\hat{\gamma}_z$	$\hat{\gamma}_{rmrf^*z}$	$\hat{\gamma}_{smb^*z}$	$\hat{\gamma}_{hml^*z}$	$\hat{\gamma}_{mml^*z}$		
估计值	0.52	-2.24	-0.69	-1.28	-0.83	-0.08	0.85	-0.63	-1.08	0.32		
t_{cs}	5.42	-4.48	-4.25	-4.47	-0.27	-3.16	1.20	-3.39	-2.62	0.15		
t_{ms}	5.03	-3.34	-3.21	-3.41	-0.25	-1.06	1.07	-2.31	-2.49	0.14		
	VM						IVM					
	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{dvol}$	$\hat{\gamma}_z$	$\hat{\gamma}_{rmrf^*z}$	$\hat{\gamma}_{dvol^*z}$	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{xv}$	$\hat{\gamma}_z$	$\hat{\gamma}_{rmrf^*z}$	$\hat{\gamma}_{xv^*z}$
估计值	0.50	-1.72	-0.35	-0.07	0.69	0.19	0.41	-2.05	-0.25	-0.02	0.73	-0.14
t_{cs}	6.13	-3.92	-0.73	-3.35	1.43	0.82	2.71	-2.35	-5.81	-3.26	1.54	-4.14
t_{ms}	5.82	-3.77	-0.58	-1.29	1.31	0.66	2.39	-2.24	-5.73	-0.24	1.37	-3.08
	HS						SkewM					
	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{srmrf}$	$\hat{\gamma}_z$	$\hat{\gamma}_{rmrf^*z}$	$\hat{\gamma}_{srmrf^*z}$	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{xs}$	$\hat{\gamma}_z$	$\hat{\gamma}_{rmrf^*z}$	$\hat{\gamma}_{xs^*z}$
估计值	0.59	-1.10	-9.95	-0.13	0.62	5.06	0.29	-1.69	0.23	-0.05	0.59	-0.03
t_{cs}	6.49	-4.16	-3.45	-2.04	1.23	1.34	0.53	-4.99	4.37	-4.24	1.30	-6.11
t_{ms}	5.84	-3.06	-2.26	-1.77	1.20	1.14	0.33	-3.81	2.23	-0.15	1.16	-4.07
	VanM											
	$\hat{\gamma}_0$	$\hat{\gamma}_{rmrf}$	$\hat{\gamma}_{srmrf}$	$\hat{\gamma}_{rorf}$	$\hat{\gamma}_{srorf}$	$\hat{\gamma}_{rmro}$	$\hat{\gamma}_z$	$\hat{\gamma}_{rmrf^*z}$	$\hat{\gamma}_{srmrf^*z}$	$\hat{\gamma}_{rorf^*z}$	$\hat{\gamma}_{srorf^*z}$	$\hat{\gamma}_{rmro^*z}$
估计值	0.49	-0.95	-11.51	-2.23	6.54	-5.21	-0.09	-0.02	4.63	1.09	-1.13	-11.44
t_{cs}	3.07	-3.62	-1.34	-7.71	1.28	-0.33	-2.52	-0.02	0.97	6.43	-0.22	-0.93
t_{ms}	2.75	-2.53	-1.23	-6.64	1.05	-0.29	-0.43	-0.02	0.97	5.33	-0.18	-0.74

注：CAPM、FF3 为基准模型，LM 为加入非流动性因子模型，VM 为波动率冲击模型，IVM 为期权价格提取的波动率模型，HSM 为 Harvey 和 Siddeque[24] 模型，VanM 为 Vanden[25] 模型，SkewM 为期权价格提取的偏度模型. t_{CS} 表示不考虑原模型可能误设的 t 统计量，t_{ms} 表示模型可能误设下的 t 统计量.

同无条件信息情形一致，r_{rmrf}、r_{smb}、r_{hml}、r_{xv}、r_{rorf}、r_{xs} 仍然是显著的定价因子，它们的 t 值的绝对值均较大，而 r_{srmrf}、r_{srorf}、r_{rmro} 不具有显著的定价能力. r_{mml} 和 r_{dovl} 的定价能力不显著，这可能意味着 FF3 模型和 LM 模型、CAPM 和 VM 模型的解释能力差异不大，这从表 5 中 Panel A 的第一 HJ 距离的数值差异较小可见一斑.

从表 6 可见，所有模型的条件信息(z) 的参数检验，在模型正确设定时，t_{cs} 的绝对值均较大，表明存在显著的定价能力. 而如果模型误设，t_{ms} 绝对值变小，不存在显著的定价能力. 这种变化表明，在传统的计量检验下(即不考虑模型是否误设) ，这些宏观因子可能存在“伪”定价能力，且考虑模型可能误设的参数检验能更有效地侦测出因

子的定价能力，这为宏观因子定价能力的稳健性研究的必要性提供了证据．同时，部分定价因子，如 r_{smb}、r_{hml}、r_{xv}、r_{rorf}、r_{xs} 和条件信息的交乘项组建的新因子，仍具有显著的定价能力．但有所区别的是，r_{smb}、r_{hml}、r_{xv} 的交乘项与本身的参数符号一致，而 r_{rorf}、r_{xs} 同条件信息的交乘项与 r_{rorf}、r_{xs} 的符号相反．

4.3　Wald 检验

表 6 估计了条件信息为 CPI 时的各 SDF 参数值，计算了是否考虑模型可能误设时的 t 值，提供了单个条件信息及条件信息和定价因子组建的新因子是否具有显著的定价能力的结论．那么，这些因子的联合定价行为又如何？这一问题的回答事关条件信息引入的必要性．表 7 考察是否考虑模型误设对条件信息与条件信息和定价因子的交乘项因子参数的联合检验（Wald 检验）的影响，以探究条件信息调整因子的定价能力．Wald 检验的原假设为条件信息调整因子的参数联合为 0，即 $H_0: \hat{\gamma}_z = \hat{\gamma}_{factor^* z} = 0$，其中 $factor$ 为模型的定价因子．当然，Wald 检验等价于检验两模型的样本第一 HJ 距离相同，即 $H_0: \hat{\delta}_1^2 = \hat{\delta}_2^2$，其中，模型 1 为无条件信息模型，模型 2 为条件信息模型．模型 2 嵌套模型 1，但两者构造的原理不同．

表 7　条件信息模型的 SDF 参数的 Wald 检验

Table 7 Wald test on SDF parameter of conditional information model

Panel A CPI 为条件信息								
Model	CAPM	FF3	LM	VM	IVM	HS	VanM	SkewM
$Wald_{CS}$	3.70	8.27	8.24	4.48	6.59	6.05	7.82	4.33
p_{CS}	0.035	0.007	0.016	0.014	0.086	0.109	0.021	0.028
$Wald_{MS}$	1.24	5.26	4.24	1.51	5.65	5.01	4.13	3.39
p_{MS}	0.371	0.015	0.064	0.420	0.130	0.171	0.059	0.036
Panel B 信用价差为条件信息								
$Wald_{CS}$	3.35	5.61	5.56	3.58	3.43	2.81	6.19	3.67
p_{CS}	0.047	0.023	0.052	0.011	0.131	0.122	0. 043	0.019
$Wald_{MS}$	2.24	3.99	4.04	1.89	2.43	2.19	4.87	2.42
p_{MS}	0.326	0. 047	0.044	0.295	0.489	0.234	0.056	0.029
Panel C 期限价差为条件信息								
$Wald_{CS}$	4.12	4.62	4.93	3.87	4.4	3.3	4.86	4.31
p_{CS}	0.027	0.028	0.024	0.075	0.222	0.347	0.012	0.029
$Wald_{MS}$	1.03	3.44	3.6	2.84	3.35	2.80	3.66	3.06
p_{MS}	0.221	0.087	0. 069	0.217	0.340	0.423	0.022	0.038

注：CAPM、FF3 为基准模型，LM 为加入非流动性因子模型，VM 为波动率冲击模型，IVM 为期权价格提取的波动率模型，HSM 为 Harvey 和 Siddeque[24] 模型，VanM 为 Vanden[25] 模型，SkewM 为期权价格提取的偏度模型．上表报告的是条件信息调整因子的 SDF 参数同为 0 的 Wald 检验，$Wald_{CS}$ 和 p_{CS} 表示不考虑原模型可能误设的 $Wald$ 统计量和相应的 p 值，$Wald_{MS}$ 和 p_{MS} 表示考虑模型可能误设的 $Wald$ 统计量和相应的 p 值．

由表 7 的 Panel A、Panel B 和 Panel C 可以发现，是否考虑模型误设对 Wald 检验的数值和 p 值有重大影响．当使用传统的 Wald 检验，即不考虑模型可能误设时，CAPM、FF3、LM、VM、IVM、VanM、SkewM 均拒绝了 $\hat{\gamma}_z = \hat{\gamma}_{factor^* z} = 0$ 的原假设，也即是说，条件信息的加入提升了模型的解释能力．然而，在考虑模型可能误设时，CAPM、VM、IVM 三个模型的 Wald 检验 p 值增加，不再拒绝其条件信息模型和无条件信息模型表现相当的原假设．可见，尽管条件信息模型相较无条件信息模型能获得更小的样本第一 HJ 距离，但在考虑模型误设后，并无强有力的证据支持这些条件信息模型较无条件信息模型表现更好．

5 结束语

利用台湾市场的数据资料，构造 FF25 组合，在第一 HJ 距离的优化下，探讨特定模型的模型误设检验，并是否考虑模型可能误设对模型 SDF 参数检验的影响，得出以下主要结论：

首先，仅部分条件信息模型能通过模型设定检验. 在 5% 显著性水平下，所有无条件信息模型均存在模型误设问题. 考虑条件信息时，所有模型的样本第一 HJ 距离数值均大幅下降，仅条件 FF3、条件 LM、条件 VanM 和条件 SkewM 成为不能拒绝的正确模型. 同时，根据样本第一 HJ 距离的数值，不同条件信息给出了不同的排序，SupLM 检验表明条件信息模型的 SDF 参数估计高度不稳定.

其次，多数因子具有显著的定价能力. 第一 HJ 距离优化下的参数检验表明，r_{rmrf}、r_{smb}、r_{hml}、r_{xv}、r_{rorf}、r_{xs} 是显著的定价因子，它们的 t 值绝对值均较大. 而 r_{mml} 和 r_{dovl} 定价能力不显著，这意味着 FF3 模型和 LM 模型、CAPM 和 VM 模型的解释能力差异不大. 部分定价因子，如 r_{smb}、r_{hml}、r_{xv}、r_{rorf}、r_{xs} 和条件信息的交乘项构成新的定价因子仍有显著的定价能力. 区别在于，r_{smb}、r_{hml}、r_{xv} 的交乘项与本身的参数符号一致，而 r_{rorf}、r_{xs} 同条件信息的交乘项与 r_{rorf}、r_{xs} 的符号相反.

第三，模型设定假设影响模型 SDF 参数的统计检验，考虑模型误设的参数检验能更有效地侦测因子的定价能力. 对比模型正确设定和误设两种情形给出的模型 SDF 参数的 t 值，可以发现，模型正确设定假设会高估模型 SDF 参数的 t 值绝对值. 在模型正确设定和模型误设两种情形下，对比 r_{mml}、r_{dovl}、条件信息 z，及条件信息和部分定价因子组成的新的因子的定价能力，发现这些因子可能存在“伪”定价现象. 同样，构造 Wald 统计量检验条件信息及条件信息和定价因子组成的新因子的参数是否联合显著异于 0 后发现：如果不考虑模型可能误设时，多数模型均拒绝了条件信息调整因子的参数联合无异于 0 的假设. 但在考虑模型可能误设时，CAPM、VM、IVM 的 Wald 统计量数值增加，不再拒绝其条件信息模型和无条件信息模型表现无差的原假设.

参考文献：

[1] Hansen L P, Jagannathan R. Assessing specification errors in stochastic discount factor models [J]. The Journal of Finance, 1997, 52(2): 557 – 590.

[2] Lewellen J, Nagel S, Shanken J. A skeptical appraisal of asset pricing tests [J]. Journal of Financial Economics, 2010, 96 (2): 175 – 194.

[3] Tortoriello R. Quantitative Strategies for Achieving Alpha [M]. New York: McGraw Hill, 2009.

[4] 郑振龙，孙清泉. 第一 HJ 距离的理论分析：基于模型设定检验视角 [J]. 统计研究，2014，31(6)：98 – 106.
Zheng Zhenlong, Sun Qingquan. The theoretic analysis of the first distance: Based on the model specification angle [J]. The Statistic Research, 2014, 31(6): 98 – 106. (in Chinese)

[5] 郑振龙. 资产价格隐含信息分析框架：目标，方法与应用 [J]. 经济学动态，2012，(3)：33 – 40.
Zheng Zhenlong. Analyzing framework of asset price implied information: Goal, method and application [J]. The Economic Dynamics, 2012, (3): 33 – 40. (in Chinese)

[6] Hansen L P, Heaton J Luttmer E G J. Econometric evaluation of asset pricing models [J]. Review of Financial Studies, 1995, 8(2): 237 – 274.

[7] Jagannathan R, Wang Z. The conditional CAPM and the cross-section of expected returns [J]. The Journal of Finance, 1996, 51(1): 3 – 53.

[8] Chen N-F, Roll R, Ross S A. Economic forces and the stock market [J]. Journal of Business, 1986, 59(3): 383.

[9] Fama E F，French K R. Common risk factors in the returns on stocks and bonds [J]. Journal of Financial Economics，1993，33(1)：3－56.
[10] Almeida C，Garcia R. Assessing misspecified asset pricing models with empirical likelihood estimators [J]. Journal of Econometrics，2012，170(2)：519－537.
[11] Hodrick R J，Zhang X. Evaluating the specification errors of asset pricing models [J]. Journal of Financial Economics，2001，62(2)：327－376.
[12] Andrews D W. Tests for parameter instability and structural change with unknown change point：A corrigendum [J]. Econometrica，1993，71(1)：395－407.
[13] Campbell J Y. Understanding risk and return [J]. Journal of Political Economy，1996，104(2)：298－345.
[14] Cochrane J H. A cross-sectional test of a production-based asset pricing model [J]. National Bureau of Economic Research，1996.
[15] Buraschi A，Jackwerth J. The price of a smile：Hedging and spanning in option markets [J]. Review of Financial Studies，2001，14(2)：495－527.
[16] Kan R，Robotti C. Model comparison using the Hansen-Jagannathan distance [J]. Review of Financial Studies，2009，22(9)：3449－3490.
[17] Roll R. A critique of the asset pricing theory's tests Part I：On past and potential testability of the theory [J]. Journal of Financial Economics，1977，4(2)：129－176.
[18] Hansen L P，Richard S F. The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models [J]. Econometrica：Journal of the Econometric Society，1987，587－613.
[19] Acharya V，VPedersen L H，Asset pricing with liquidity risk [J]. Journal of Financial Economics，2005，77(3)：375－410.
[20] Sadka R. Liquidity risk and the cross-section of hedge-fund returns [J]. Journal of Financial Economics，2010，98(1)：54－71.
[21] Brunnermeier M K，Pedersen L H. Market liquidity and funding liquidity [J]. Review of Financial Studies，2009，22(6)：2201－2238.
[22] **郑振龙，汤文玉. 波动率风险及风险价格——来自中国 A 股市场的证据 [J]. 金融研究**，2011，(4)：143－157.
Zheng Zhenlong，Tang Wenyu. Volatility risk and risk price：Evidence from Chinese A stock market [J]. Journal of Financial Research，2011，(4)：143－157. (in Chinese)
[23] Neuberger A. Realized skewness [J]. Review of Financial Studies，2012，25(11)：3423－3455.
[24] Harvey C R，Siddique A. Conditional skewness in asset pricing tests [J]. The Journal of Finance，2000，55(3)：1263－1295.
[25] Vanden J M. Option coskewness and capital asset pricing [J]. Review of Financial Studies，2006，19(4)：1279－1320.
[26] **洪永淼，林　海. 中国市场利率动态研究——基于短期国债回购利率的实证分析 [J]. 经济学(季刊)**，2006，(1)：511－532.
Hong Yongmiao，Lin Hai. Dynamic Research on Chinese market rate：The empirical study on short bond repo rate [J]. Quarterly of Economics，2006，(1)：511－532. (in Chinese)
[27] **吴世农，许年行. 资产的理性定价模型和非理性定价模型的比较研究——基于中国股市的实证分析 [J]. 经济研究**，2004，(6)：105－116.
Wu Shinong，Xu Nianxing. The comparative research on rational pricing model and irrational pricing modes of assetshe empirical study on Chinese stock market [J]. Economic Research Journal，2004，(6)：105－116. (in Chinese)
[28] Harvey C R. The specification of conditional expectations [J]. Journal of Empirical Finance，2001，8(5)：573－637.
[29] Ghysels E. On Stable Factor structures in the pricing of risk：Do time-varying betas help or hurt? [J]. The Journal of Fi-

nance, 1998, 53(2): 549 - 573.

[30] Kozhan R, Neuberger A, Schneider P. The skew risk premium in the equity index market [J]. Review of Financial Studies, 2013, 26(9): 2174 - 2203.

Linear factor models' model specification test based on the first HJ distance

ZHENG Zhen-long[1], *SUN Qing-quan*[2]

1. School of Management, Xiamen University, Xiamen 361005, China;

2. Department of Private Banking, China Merchants Bank, Shenzhen 518040, China

Abstract: Model Specification Test is a key step in financial modeling to reduce the model risk. Based on the first HJ distance proposed by Hansen and Jagannthan(1997), Taiwan market data are used to test the model specifications of eight linear factor models (including models based on financial asset prices), and the impacts of model specification assumptions on parameter tests are discussed. The paper finds that, under the 5% significance level, there exists model misspecification problems for all unconditional models and only the conditional versions of FF3, LM, VanM and SkewM are acceptable right models. Meanwhile, taking potential model misspecification into account may detect the factors' pricing ability more efficiently. Assuming a model is rightly specified overestimates the t absolute values of SDF parameters, resulting in a "pseudo-pricing" for some parts of the factors.

Keywords: the first HJ distance; model specification test; parameter test