MDH假说下股指期货交易量和波动率动态关系研究

张苏林

（重庆理工大学经济与贸易学院 重庆 400050）

【摘要】本文利用EGARCH模型和VAR模型研究了我国沪深300股指期货开市一年来交易量和波动率的动态和因果关系。实证结论表明股指期货的非预期交易量是信息量的有效代理变量，可以很好地解释波动率，同时二者没有显著的Granger因果关系，因此我国股指期货市场支持分布混合假说（MDH），证实了该市场是有一定效率的。

【关键词】交易量 波动率 分布混合假说

一、研究背景

我国沪深300股指期货诞生已经两年了，作为我国金融市场的首只金融衍生工具，股指期货越来越被市场接受，截至2011年4月12日，开户数已经超过7万户，累计成交量4 730万手，累计成交额42.9万亿。同时股指期货基本功能发挥良好，一些机构开始利用股指期货进行套期保值，而且已经取得了良好的套期保值效果；一些研究也证实了股指期货具有价格发现和价格引导功能，期货市场为投资者提供了不少套利的好机会。

金融市场最直接的行为表现是资产价格与交易量，大多技术分析者或者说个人投资者经常将交易量所隐含的信息作为预测资产价格走势的重要依据，因此量价关系的最初研究集中于交易量与收益率之间的关系。上世纪90年代以来，一些研究者认为收益率的变动或许是市场参与者、研究者以及政策制定者更为关注的变量，近年来量价关系方面的文献集中于交易量与资产波动率的动态关系，成为了金融领域研究的热点问题。

交易量与资产波动率关系的研究被应用于众多国家地区的股票市场、期货市场、货币市场，具有广泛的研究价值。本文以我国沪深300股指期货为研究对象，考察交易量与资产波动率的动态和因果关系。

本文有三个方面的应用价值：①揭示沪深300股指期货的内部动态关系，将有助于该产品的价格确定；②填补沪深300股指期货交易量对波动持续性冲击的文献空白；③帮助市场的投资者深层次地理解市场价格形成过程。

二、文献回顾

Karpoff（1987）对量价关系的文献进行了梳理和总结，认为量价关系之所以重要有四个原因：①量价关系有助于我们理解信息对市场传导作用，进而分别理解市场微结构的不同假定；②量价关系在利用双变量的事件研究发中具有重要作用，因为检验的构造和推论的有效性取决于量价关系的联合分布；③量价关系在研究投机价格的经验分布中具有决定性作用；④量价关系在期货研究中具有重要意义，在决定投资者需求方面，依据量价关系的私有信息要比公开信息重要。

Schwert（1989）利用分布滞后模型和VAR模型对标准普尔500指数实证分析，证实了波动率与交易量以及滞后期交易量有确定的关系。Lamoureux和Lastrapes（1990）基于分布混合假说（MDH），利用波动率测量最常用的工具GARCH模型研究了交易量和波动率的动态关系，开辟了该领域研究新的方向，结果证实了作为信息代理变量的交易量被作为解释变量加入GARCH模型，降低了波动的持续性，同时有些波动率指标的系数不再显著，支持了分布混合假说。

Bessembinder和Seguin（1993）把研究对象从股票市场转向了期货市场，同时对交易量进行了分解，得到了私有信息和公有信息的代表，结论是波动率与同期交易量有显著的正相关，未预期交易量是比预期交易量更好的解释波动率的指标。Sharma等（1996）以纽约证交所指数代表市场指数为研究对象，得到的结论是交易量只能解释部分的GARCH效应，因此认为在影响资产波动率除了成交量以外，还有其他因素。而Arago和Nieto（2005）考虑到波动率的非对称性，并将成交量分成可预期和未预期两部分，然后与单纯的成交量做比较，结论显示单纯的成交量并不显著，而未预期的交易量影响显著，但也只能减低部份波动的持续性，波动的持续性仍然不会消失。同时波动性对称性并不会随着交易量的变化而有所影响。

Marilyn和Robert（1999）将有关量价关系的理论模型分为信息理论模型、交易理论模型、信息非对称模型和理念分散模型四种。认为信息是决定资产价格波动和交易量的共同动力，信息理论模型与市场微结构基本结论暗合，新的信息不断到达市场并不断融合到市场价格中去是金融市场价格变动的主要原因，因此被认为是主流理论模型。而该模型中的分布混合假说（MDH）得到了大多经验实证的支持，成为了本文的理论依据。

Clark（1973）提出的金融资产日价格波动的混合分布假说模型，认为资产的回报和交易量是由一个潜在的、不可观测

的信息流过程共同趋动的。新信息流进入市场，冲击市场产生新的价格波动与交易量。投机资产的日收益波动表现为相对于正态分布呈现尖峰厚尾的分布，这种尖峰态是因为每日价格波动序列可看作来自不同方差的分布集，即“混合分布”。Tauchen(1983)在中心极限定理的推导下得出每日价格的变动和交易量是信息流的条件分布，有：

$\triangle P_t|n_t \sim N(0, an_t)$

$V_t|n_t \sim N(cn_t, dn_t)$

$Cov(\triangle P_t, V_t|n_t)>0$

其中：P_t和V_t是资产的价格和交易量，n_t是信息流的数量序列，参数a、c、d为非负，得到重要结论是日价格波动与交易量是正相关的。

同时支持混合分布假说的研究者认为所有投资者同时接受新信息，并做出反应，瞬间即可在新价格水平上达到平衡，因此价格的变动与交易量之间不存在任何方向上的互为引导关系，即格兰杰因果关系。交易量与价格波动之间的正相关关系是因为市场信息把二者联系起来，信息不可观测，因此交易量或交易次数被作为信息流的替代指标。如Lamoureux和Lastrapes(1990)利用每日交易量作为每天流入市场信息的代理变量，被证明是有效的。

Andersen(1996)将混合分布模型与市场微观结构理论结合起来，建立了日价格回报与交易量关系的修正混合分布模型。利用自回归AR(p)模型把交易量分解为流动性需求交易和信息趋动性交易两部分：

$$v_t=\varnothing_t+\sum_{j}^{p}\varnothing_j v_{t-j}+\varepsilon_t \qquad (1)$$

其中：$\varnothing_j$为参数，j=0，1，…，p，利用历史成交量的信息来估计当期成交量，经由回归式得到的$v_t-\varepsilon_t$即为可预期成交量，另外残差部分ε_t即为未预期成交量，记为v_u。

三、数据及相关处理

本文研究的对象是沪深300股指期货，而该股指期货合约有当月合约、下月合约和随后两个季度四种合约品种，根据沪深300股指期货实际的交易情况以及其他研究者的惯例，选择交易最活跃的新近到期合约为研究对象，并形成滚动历史数据。

本文研究的时间区间为2010年4月16日到2011年4月15日，正好一年的历史交易数据，选择的时间间隔为日，共有241个价格数据，对应着240个收益率数据和交易量数据。

从表1的描述统计可以看出，我国沪深300股指期货的收益率的均值很小，接近于0，分布呈现出有偏尖峰的特征，且Jarque-Bera检验显著拒绝了收益率的无条件正态分布假设，符合一般金融资产的分布特征。同时从收益率序列的自相关系数和Q统计量来看，样本之间的相关性很小，拒绝了存在自相关性的零假设。

而沪深300股指期货的交易量序列滞后一阶、二阶、五阶直到十二阶的自相关系数均为正数，且在1%的置信水平上统计显著，因此股指期货的交易量序列具有很强的自相关性，可以通过建模进行一定的预测。股指期货的交易量和收益率序列的ADF检验都拒绝存在单位根的假设，是平稳的时间序列，可以直接用于建模。

交易量序列大多具有一定的时间趋势，因此必须首先对交易量序列进行去势处理。根据Gallant等（1992）和Chen等(2001)的做法，我们利用下面回归方程对沪深300股指期货的交易量序列进行取势处理：

$$V_t=\eta_0+\eta_1 t+\eta_2 t^2+\varepsilon_t \qquad (2)$$

其中：V_t为原始交易量，t为时间趋势项，t^2为时间趋势项的平方，所得的残差项ε_t为不含时间趋势的交易量，定义为v_t，再利用式(1)对取势的交易量进行分解，得到预期交易量v_e和非预期的交易量v_u。

四、研究方法和模型

1. 引入交易量的EGARCH（1,1）模型。由Engle提出ARCH模型，后来由Bollerslev等人加以扩充为GARCH族模型已成为金融计量研究中最受欢迎的波动率模型之一。随后一些学者提出了很多其他类型的非线性GARCH模型，如GJR模型、EGARCH模型、APGARCH模型等。大多实证研究表明EGARCH(1,1)模型是最有效的估计资产波动率的模型，考虑到便于比较，本文也选择该模型。

EGARCH(1,1)模型的具体设定为：

$$R_t=\varepsilon_t$$

$$\ln(\delta^2_t)=\alpha+\gamma\frac{\varepsilon_{t-1}}{\delta^2_{t-1}}+\beta\left|\frac{\varepsilon_{t-1}}{\delta^2_{t-1}}\right|+\lambda\ln(\delta^2_{t-1}) \qquad (3)$$

其中：R_t表示t时刻股票价格P_t取自然对数后的收益率，计算公式为：

$$P_t=InP_t^{close}-InP_{t-1}^{close}$$

模型满足平稳性的条件是λ<1，资产价格的波动持续性(persistence)可以λ来反映，λ越接近1，就说明波动的持续性越强。ε_t可以作为好消息和坏消息的指示器，$\varepsilon_t<0$表示坏消息，而$\varepsilon_t>0$表示好消息，因此参数γ可以检验条件方差对好坏信息的反应，即金融领域的“杠杆效应”。在模型的参数估计过程中我们假定残产差项服从学生t分布。

本文研究我国沪深300股指期货交易量和波动率的动态关系，而根据Lamourex和Lastrapes(1990)的研究，以往存在的GARCH模型反映的波动持续性是由于信息的冲击，但信息无法观测，而交易量可以作为信息的代理变量，因此将交易量

表1　该区间收益率序列和交易量序列的基本统计描述

	Mean	S.D.	Skewness	Kurtosis	J-B	ρ1	ρ2	ρ5	ρ12	Q(12)	ADF
收益率序列	-9.25E-05	0.017	-0.434	4.887	43.1*** 63	-0.049	0.021	0.015	0.038	13.425	-16.70***
交易量序列	1.98E+05	9.52E+04	-0.227	2.584	3.807	0.710***	0.451***	0.150***	0.082***	216.90***	-6.133***

注：ρ_k是滞后k阶的序列相关系数，表明序列低阶自相关性，Q(n)是滞后n阶的Ljung-BoxQ 统计量，用于检验序列高阶自相关性，ADF是修正的单位根检验值，*表示在10% 水平上显著，**表示在5% 水平上显著，***表示在1% 水平上显著

加入到波动方程中，研究GARCH效应和交易量两者是否有替代关系。Bessembinder和Seguin(1993)更进一步将交易量分解为预期交易量和未预期交易量，研究两者对于条件波动率的影响。

因此本文将式(1)、(2)得到的取势交易量和非预期交易量分别加入到EGARCH(1,1)模型中,得到式(4)、式(5)：

$$R_t=\varepsilon_t$$

$$\ln(\delta^2_t)=\alpha+\gamma\frac{\varepsilon_{t-1}}{\delta^2_{t-1}}+\beta\left|\frac{\varepsilon_{t-1}}{\delta^2_{t-1}}\right|+\lambda\ln(\delta^2_{t-1})+\delta_1 v_t \quad (4)$$

$$R_t=\varepsilon_t$$

$$\ln(\delta^2_t)=\alpha+\gamma\frac{\varepsilon_{t-1}}{\delta^2_{t-1}}+\beta\left|\frac{\varepsilon_{t-1}}{\delta^2_{t-1}}\right|+\lambda\ln(\delta^2_{t-1})+\delta_2 v_u \quad (5)$$

根据MDH假说,如果该假说成立,那么交易量可以作为信息的代理变量，具体表现为交易量的待估参数δ显著为正，表明资产的波动率与交易量是正相关关系，波动率较好地为交易量所解释，而表示波动持续性的参数λ将变小并且变得不显著。

2. 资产波动率与交易量的Granger因果关系。MDH假说的另一个结论，资产价格的变动与交易量不存在相互的引导关系,二者的滞后期对同期对方变量不具有解释作用。我们通过双变量VAR模型及Granger因果关系方法检验资产波动率和交易量的动态相互关系。本文对交易量进行了分解,因此这里有两个VAR模型,记做式(6)、式(7),具体如下：

$$\delta_t=\alpha_0+\sum_{i=1}^{m}\alpha_i\delta_{t-1}+\sum_{i=1}^{n}\beta_i v_{tt-1}+\varepsilon_t$$

$$v_{tt}=\alpha_0+\sum_{i=1}^{m}\gamma_i v_{tt-1}+\sum_{i=1}^{n}\tau_i\delta_{t-1}+\eta_t \quad (6)$$

$$\delta_t=\alpha_0+\sum_{i=1}^{m}\alpha_i\delta_{tt-1}+\sum_{i=1}^{n}\beta_i v_{ut-1}+\varepsilon_t$$

$$v_{ut}=\alpha_0+\sum_{i=1}^{m}\gamma_i v_{ut-1}+\sum_{i=1}^{n}\tau_i\delta_{tt-1}+\eta_t \quad (7)$$

其中:δ_t、v_{tt}、v_{ut}分别代表资产t时刻的波动率、取势交易量和非预期交易量。δ_t来自于上步的EGARCH(1,1)的估计,但注意是不含交易量的估计。如果待估参数β和τ显著异于0,那么我们就可以认为资产波动率与交易量有相互的反馈作用，一般我们用F统计量来检验变量参数是否联合等于零假设。

五、实证结果分析

1. EGARCH(1,1)模型和加入交易量变量的EGARCH(1,1)模型参数估计。

表2是我国沪深300股指期货在开市一年来利用EGARCH（1,1）和引入取势交易量v_t、非预期较量v_u的EGARCH(1,1)模型的参数估计结果,以比较波动的持续性变化情况。同时还列出了各个模型滞后12期的Ljung-BoxQ统计量、Q^2统计量以及ARCH统计量和AIC值,以考察模型估计结果的稳健。

表2EGARCH(1,1)和引入取势交易量v_t、非预期较量v_u的EGARCH(1,1)模型的参数估计结果。

表2

	EGARCH(1,1)	EGARCH(1,1)-v_t	EGARCH(1,1)-v_u
α	-0.113 475 *	-0.135 723 ***	-5.591 173 ***
γ	0.078 671 ***	0.081 490 ***	-0.077 102
β	-0.104 106 ***	-0.136 607 ***	-0.007 054
λ	0.978 210 ***	0.972 859 ***	0.311 360
δ_1		-3.11E-07 ***	
δ_2			4.21E-06 **
Q(12)	13.023(0.367)	14.456(0.273)	14.081(0.296)
Q^2(12)	5.221 5(0.950)	4.519 7(0.972)	5.601 7(0.935)
ARCH(3)	0.030 631(0.992 8)	0.331 23(0.565 5)	0.122 162(0.727 0)
AIC值	-5.396 559	-5.406 504	-5.374 545

注:括号中是对应统计量的p值。*表示在10%水平上显著，**表示在5%水平上显著，***表示在1%水平上显著。Q(12)表示标准化残差(残差除于条件标准差)滞后12阶的Ljung-Box Q统计量,Q^2(12)表示标准化平方残差(平方残差除于条件方差）滞后12阶的Ljung-BoxQ统计量。ARCH(3)是滞后3阶的ARCH统计量

EGARCH(1,1)模型的各个参数在统计上都是显著,λ<1保证了模型的平稳性。在数值上接近1说明了我国股指期货具有很强的波动持续性,反映“杠杆效应”的参数大于0且显著,意味着股指期货具有明显的非对称性，即说明利空消息对收益波动的冲击大于利好消息对波动率的冲击，我国股指期货的这种特征与一般国际上的经验和金融理论相符合。标准化残差滞后12阶的Ljung-BoxQ统计量均对应较大的p值，表明不可以拒绝残差之间无自相关的虚拟假设,Q^2(12)统计量也不显著,这意味着经由EGARCH(1,1)模型滤波后的残差序列也不再存有时变方差现象,滞后3阶的残差ARCH统计量也对应较大的p值表明也不可以拒绝残差无异方差的虚拟假设。因此总体来看,EGARCH(1,1)模型很好地拟合了我国股指期货的波动率,结论具有稳健性。

在EGARCH(1,1)模型中加入取势交易量后,各项参数也在统计上显著,模型的稳健性也得到了证实。反映波动持续性的参数相比在EGARCH(1,1)模型只是稍微的下降,下降幅度不超过6‰，即取势交易量的引入没有减少波动的持续性，因此作为信息代理变量的取势交易量不能完全地解释波动率。同时参数的符号为负,即交易量和波动率呈负相关,这些都与MDH假说。得到的结论,从现有的文献看与Berna对土耳其ISE-30股指期货的研究结论一致,Berna的认为连续信息到达假说(SIAH)能够解释上述结论。

在EGARCH(1,1)模型中加入非预期交易量后,反映波动率持续性的参数和“杠杆效应”的参数均在数值上降低,同时在统计上不显著，而非预期交易量的参数在统计上显著且大于0。这说明了股指期货波动的持续性主要由非预期交易量来解释,这与修正MDH假说的基本结论是一致的,即非预期交易量作为日信息量的替代指标比交易量本身更合适，非预期交易量所揭示的新信息冲击是产生价格波动的根源。该结论得到了大多研究者的支持,如Arag和Arag(2005)对世界主

要股票市场指数的量价研究以及李双成(2006)对我国沪深股市股指的量价研究,说明了我国股指期货市场是有效率的。

2. 双变量VAR模型和Granger因果关系检验。利用上面介绍的双变量VAR模型方程(6)、(7)来检验波动率与交易量的动态因果关系,其中波动率来自于单纯EGARCH(1,1)模型计算后生成的Garch方差序列。VAR模型滞后期的确定本文使用了LR检验统计量、信息准则(AIC、SC、HQ)和最终预测误差FPE,最终确定方程(6)、(7)的滞后阶数为1阶。表3给出了取势交易量和非预期交易量分别不是波动率互为Granger因果原因零假设条件下Granger因果检验的结果:

表3　　Granger 因果检验的结果

Null Hypothesis Obs	Obs	F-Statistic	Probability
V_t does not Granger Cause Volatility	239	0.0959 2	0.757 1
Volatility does not Granger Cause V_t		2.9147 5	0.089 1
V_u does not Granger Cause Volatility	239	0.0878 3	0.767 2
Volatility does not Granger Cause V_u		1.5921 6	0.208 3

从F统计量伴随概率来看,原假设为"V_t不是Volatility的格兰杰原因"的伴随概率为0.757 1,大于0.1,这时接受原假设,表明V_t不是Volatility的格兰杰原因,取势交易量的变化并不影响波动率的变化,而原假设为"Volatility不是V_t的格兰杰原因"的伴随概率为0.089 1,小于0.1,这时接受备则假设,表明在10%的置信水平上Volatility是V_t的格兰杰原因,波动率前期变化影响着取势交易量同期变化。这点与一般股票市场上所谓"价随量涨"规则相反,但我们应注意到伴随的概率是在很高的置信水平上,也可持否定的态度。基本结论和与Hiemstra和Jones(1994)的分析结果一致,但与MDH假说矛盾。

如果利用非预期交易量观测的话,交易量与波动率都不互为Granger因果关系,且只存在非预期交易量与波动率同时期的因果关系,该结论与MDH假说一致。因此非预期交易量是比交易量更好的信息代理变量。从笔者掌握的文献来看,还没有发现相似的论述。

六、结论及建议

本文运用EGARCH(1,1)模型,以MDH假说为理论基础,对我国沪深300股指期货的交易量和波动率的动态关系进行了实证分析。得到如下结论:

1. 我国沪深300股指期货市场具有显著的"杠杆效应"。该效应在大多国外证券市场大多成立,但在我国,研究者的结论不一致,如陈浪南(2002)、楼迎军(2003)、运怀立(2006),这些学者分别做了不同的解释,总的来说是新兴市场出现的异象。而我国股指期货市场经历了10年的酝酿、开市前数年的仿真交易以及在规则上的严格控制,这些对市场的理性和效率都有极大的帮助,因此出现了不同于我国股票市场显著的"杠杆效应"。

2. 作为量价关系的主流理论——MDH假说是成立的,而该假说强调信息是交易量和资产波动率变化的原因,暗含着市场效率的命题。市场效率高的市场信息是决定交易量和资产波动率二者变化的唯一原因,而低效率的市场二者的变化应该由其他原因解释。因此本文认为我国沪深300股指期货市场的效率优于我国其他的证券市场。MDH假说的成立,对于市场的参与者和管理者理解价格的形成机制、市场的微结构都有积极意义。

3. 非预期交易量是比原始交易量或取势交易量更好的信息代理变量。当把交易量分解得到的非预期交易量代入波动率模型中,我们能得到波动的持续性降低且不显著的结论,同时该变量与波动率没有相互的Granger因果关系,得到了支持MDH成立的强有力证据,但用取势交易量是得不到上述证据的。这说明原始交易量包含着大量噪音,这是新兴的我国股指期货市场必然的情况,对监管机构如何把握市场的运行规律,制定合理有效的监管政策,以引导市场的良性运行提出了严峻考验。

【注】 本文受国家社会科学基金项目"马尔科夫链抽样方法的金融市场随机波动的联动性及预警机制研究"(批准号:10BJL020)资助。

主要参考文献

1. Karpoffm J M. The relation between price changes and trading volume: a survey. Journal of Financial and Quantitative Analysis,1987;22

2. Schwert G. W. Why Does Stock Market Volatility Change Over Time?,The Journal ofFinance,1989;5

3. Lamoureux. C. G,Lastrapes W. D. Heteroskedasticity In Stock Return Data:Volume Versus GARCH Effects,The Journal of Finance,1990;1

4. Bessembinder H,Seguin P. J. Price Volatility,Trading Volume and Market Depth:Evidence from Futures Markets,The Journal of Financial and Quantitative Analysis,1993;1

5. Sharma J.L,M. Mougoue,R. Kamath. Heteroskedasticity in Stock Market Indicator Return Data: Volume versus GARCH Effects Applied Financial Economics,1996;6

6. Arag V,Nieto L. Heteroskedasticity in the Returns of The Main World Stock Exchange Indices: Volume versus GARCH Effects,Journal of International Financial Markets, Institutions and Money,2005;3

7. Berna Okan. Volume and Volatility: A Case of ISE-30 Index Futures ,International Research Journal of Finance and Economics,2009

8. Wiley M K,Daigler R T . A Bivariate GARCH approach to the futures volume-volatility issue . Presented at the Eastern Finance Association Meeting s,Miami Beach,Florida, 1999;4

9. Clark,P. K. A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices,Econometrica,1973;1

10. Andersen T.G. Return volatility and trading volume: an information flow interpretation of stochastic volatility. Journal of Finance,1996;51