第 5 章 随 机 波 动 率*

Eric Ghysels, Andrew C. Harvey and Eric Renault

1. 引言

随机波动率（SV）类模型产生于数理金融学和金融计量学。事实上，SV 模型的几种变形来源于对非常不同问题的研究，例如，Clark（1973）认为资产收益率是信息到达这一随机过程的函数，这种所谓的时间形变（time deformation）方法产生了资产收益率的时变（time-varying）波动率模型。Tauchen 和 Pitts（1983）对此做了改进，提出了资产收益率与信息到达短暂相关的混合分布模型。Hull 和 White（1987）并不直接关心资产收益率与信息到达的联系，而是对欧式期权的定价感兴趣，并假定标的资产服从连续时间 SV 模型。他们提出资产价格服从扩散过程，其波动率为正扩散过程。另一种方法源于 Taylor（1986）的研究，他建立了替代自回归条件异方差（ARCH）模型的离散时间 SV 模型。直到不久以前，估计 Taylor 模型或者任何其它 SV 模型几乎仍是不可能的，经济计量理论的最新发展使估计 SV 模型变得容易得多。从而，它们变成了一类很有吸引力的模型，并可替代其它类模型如 ARCH。

在数理金融学和经济计量学文献中均可以找到关于 SV 模型的研究。因此，我们面对着一系列广泛的主题。本文将很少涉及 ARCH 模型，因为最近已有几篇关于该主题的优秀综述，包括 Bera 和 Higgins（1995），Bollerslev、Chou 和 Kroner（1992），Bollerslev、Engle 和 Nelson（1994），以及 Dievold 和 Lopez（1995）的研究成果。而且，由于本文是为*《统计学手册》*写的，我们力图最小限度地涉及数理金融学的内容。不过，期权定价的内容显然是不必要的。事实上，在第 2 节定义波动率时已广泛涵盖 Black-Scholes 隐含波动率问题。第 2 节还概括了经验类型化事实（stylized facts），及波动率统计建模的结论。对统计学概念有浓厚兴趣的读者可跳过更偏重于金融学的第 2 节的前三小节，而从 2.4 节开始阅读；第 3 节讨论离散时间模型；第 4 节回顾了连续时间模型；第 5 节主要是 SV 模型的统计学推断；第 6 节是结论。

2. 金融市场的波动率

在对衍生证券定价时，波动率起了主要的作用。关于欧式期权定价的 Black-Scholes 模型是迄今为止应用最广泛的公式，即使其假设无法成立。因此，2.1 节将把 Black-Scholes 模型作为参考点，通过它来讨论波动率的一些概念；对有关波动率和期权价格的类型化事实的讨论在 2.2 节；这两小节为 2.3 节讨论的随机波动率的定义提供了正式的框架；最后，2.4 节介绍随机波动率的*统计*模型。

2.1. Black-Scholes 模型和隐含波动率

在 Louis Bachelier（1900）开创性研究后的半个多世纪，连续时间随机过程成为描述资产价格的标准工具，Black 和 Scholes（1973）以及 Merton（1990）的研究在这方面非常有影响。在 2.1.1 节中，我们回顾了利用扩散过程对资产价格建模时所做的一些假设，特别是要提到瞬时波动率（instantaneous volatility）的概念；2.1.2 节转向期权定价模型和隐含波动率（implied volatility）的各种概念。

2.1.1. 瞬时波动率的概念

考虑一种金融资产，比如股票，当前的（时间 t）市场价格记为 S_t，[2]令 I_t 表示 t 时已获得的

* Torben Anderson、David Bates、Frank Diebold、René Garcia、Eric Jacquier 和 Neil Shephard 对本文前期手稿提出了非常有用的建议。第一作者要感谢 FCAR（Québec）、SSHRC（Canada）的赞助，以及 CORE（Louvain-la-Neuve,Belgium）的热情与支持。第二作者要感谢 ESRC 的赞助。第三作者要感谢 Institut Universitaire de France、the Fédération Francaise des Cociétés 以及 CIRANO 和 C.R.D.E 的赞助。

[2] 此点以后，本文将集中讨论有关股票和汇率的期权文献，不涉及利率期限结构和相关衍生证券的大量文献。

信息，并且考虑给定I_t时，持有资产在期间$[t,t+h]$的收益S_{t+h}/S_t的条件分布。[3]贯穿全文的一个假设是：给定I_t，资产的收益具有有限的条件期望，或：

$$E_t(S_{t+h}/S_t)=S_t^{-1}E_tS_{t+h}<+\infty \tag{2.1.1}$$

并且给定I_t时，同样具有有限条件方差，即

$$V_t(S_{t+h}/S_t)=S_t^{-2}V_tS_{t+h}<+\infty \tag{2.1.2}$$

连续复利期望收益率由$h^{-1}\log E_t(S_{t+h}/S_t)$来表示，那么，第一个假设可以表述如下：

假设 2.1.1.A. 当$h>0$且趋近于零时，连续复利期望收益率几乎必定(almost surely)收敛于一个有限值$\mu_S(I_t)$。从这个假设可得$E_tS_{t+h}-S_t\sim h\mu_S(I_t)S_t$，或者用它的微分表达式：

$$\text{几乎必定}\quad \frac{d}{d\tau}E_t(S_\tau)\Big|_{\tau=t}=\mu_S(I_t)S_t \tag{2.1.3}$$

其中，对右边求导。等式（2.1.3）有时被近似地定义为：$E_t(dS_t)=\mu_S(I_t)S_tdt$。下一个假设是关于条件方差的，可以表述为：

假设 2.1.1.B. 当$h>0$且趋近于零时，收益$h^{-1}V_t(S_{t+h}/S_t)$几乎必定收敛于一个有限值$\sigma_S^2(I_t)$。此外，用它的微分表达是

$$\text{几乎必定}\quad \frac{d}{d\tau}Var_t(S_\tau)\Big|_{\tau=t}=\sigma_S^2(I_t)S_t^2 \tag{2.1.4}$$

它与表达式$V_t(dS_t)=\sigma_S^2(I_t)S_t^2dt$有近似关系。

通过如下形式的一个等式，假设 2.1.1.A 和 B 都推导出资产价格动态的一个表达式：

$$dS_t=\mu_S(I_t)S_tdt+\sigma_S(I_t)S_tdW_t \tag{2.1.5}$$

其中W_t是一个标准布朗运动。因此，每次资产价格过程表示为扩散方程，我们就自动定义为所谓的瞬时波动率过程$\sigma_S(I_t)$，上式的瞬时波动率过程也可写为：

$$\sigma_S(I_t)=\left[\lim_{h\downarrow 0}h^{-1}V_t(S_{t+h}/S_t)\right]^{1/2} \tag{2.1.6}$$

在转到下一节之前，先对假设 2.1.1.A 和 B 的某些基础做一个简要的讨论。Bachelier（1900）提出把布朗运动过程作为股票价格运动的一个模型。在现代术语学中，这相当于资产定价的随机游走理论。该理论认为，由于金融市场的信息有效性，资产收益应是不可预测的。因此，可以假设在连续规则样本期$[t+k,t+k+1]$，$k=0,2,...,h-1$，收益是独立（同）分布的。根据这个基准，把连续复利收益率$\log(S_{t+h}/S_t)$的期望值和方差看成与投资的到期期限h成比例就是合理的。

显然，我们不再用布朗运动作为资产价格的一个过程，但仍然值得注意的是假设 2.1.1.A 和 B 还意味着，在一个无限短的间隔$[t,t+h]$内，投资的期望收益率和相关联的平方风险（根据收益率的方差）是和h成比例的。通过“局部不可预测性”概念，Sims（1984）为两个假设提供了一些依据。

最后，简要地讨论（2.1.5）的一个特殊情形，它主要应用于理论发展，并且突出了本文包含的一个隐含约束。对所有t，$\mu_S(I_t)=\mu_S$和$\sigma_S(I_t)=\sigma_S$是常数时，资产价格是一个几何布朗运动。Black 和 Scholes（1973）应用该过程推导出著名的欧式期权定价公式。显然，由于$\sigma_S(I_t)$是一个常数，不再是一个瞬时波动率过程，因而只有一个参数σ_S——这无疑大大简化了很多东西，包括期权定价。必须强调的第二点是，假设 2.1.1.A 和 B 考虑到了资产价格过程中离散跳跃的可能性。这样的跳跃通常用 Poisson 过程来表示，并且自 Merton（1976）的研究以后，它在期权定价文献中占主导地位。然而，虽然假设在原则上考虑到了跳跃，但它们并没有在（2.1.5）中出现。事实上，本文中，始终维持样本路径连续性的假设，并且当单独讨论 SV 模型时，排除了跳跃的可能性。

---

[3] 2.3 节将更详细讨论信息集。应指出的是，由于 S_t 属于 I_t，本文在应用资产价格 S_{t+k} 和收益 S_{t+k}/S_t 的条件分布时不加区分。

2.1.2. 期权价格和隐含波动率

正如在引言中指出的，SV 模型部分源于期权定价文献。在过去二十年中期权和其它衍生证券市场得以空前地增长，这些市场有时表现为“波动率交易”的地方。这一节将提供有关这些论述的基本原理，并且研究所谓的期权的隐含波动率与标的资产收益过程的瞬时及平均波动率概念间的关系。

Black-Scholes 的期权定价模型是建立在标的资产价格可用对数正态或者几何布朗运动建模的基础上的：

$$dS_t = \mu_S S_t dt + \sigma_S S_t dW_t \tag{2.1.7}$$

其中 μ_S 和 σ_S 是固定参数。具有执行价 K 和到期日 $t+h$ 的欧式看涨期权可得收益

$$[S_{t+h}-K]^+ = \begin{cases} S_{t+h}-K, \text{如果} S_{t+h} \geq K \\ 0, \text{其他} \end{cases} \tag{2.1.8}$$

由于 Black 和 Scholes（1973）开创性的论文，现在已有了完整的文献资料，提出各种方法，能够导出这种合约的定价公式。显然，详细论述这些文献超出了本文的范围。[4]在此只是最小限度地介绍其内容，以便于讨论关于波动率的概念。

若假定连续的无成本交易是可行的，则可以在 Black-Scholes 框架中构造一个投资组合，该组合对标的股票只用一个看涨期权和一个卖空策略就可以消除所有的风险，这就是为什么仅仅使用套利理论——即通过包括看涨期权的无风险投资组合的市场收益率等于无风险利率——就可以确定期权价格，而且，这种以套利为基础的期权定价并不依赖于个人偏好。[5]

这就是为什么推导 Black-Scholes 期权定价公式最容易的方法是通过一个“风险中性世界”，其中资产价格过程是由*修正的概率测度*来规定的，指的是风险中性的概率测度，记为 Q（这在 4.2 节中有更清楚的讨论）。在这个虚拟环境中，概率一般与数据生成过程（DGP）是不一致的，这个虚拟环境只是用来导出在客观概率设置中仍然有效的期权价格。在风险中性世界中，有：

$$dS_t / S_t = r_t dt + \sigma_S dW_t \tag{2.1.9}$$

$$C_t = C(S_t, K, h, t) = B(t, t+h) E_t^Q (S_{t+h}-K)^+ \tag{2.1.10}$$

其中 E_t^Q 是在 Q 下的期望值，$B(t,t+h)$ 是一个在时间 $t+h$ 时（可以）得到一单位收益的纯贴现债券在 t 时的价格，并且

$$r_t = -\lim_{h \to 0} \frac{1}{h} Log B(t, t+h) \tag{2.1.11}$$

是无风险即时利率。[6]本文已隐含地假设在这个市场中利率是非随机的（W_t 是风险的唯一来源），这样：

$$B(t,t+h) = \exp\left[-\int_t^{t+h} r_\tau d\tau\right] \tag{2.1.12}$$

根据定义，在风险中性背景下没有风险溢价，因此 r_t 和股票的即时期望收益率是一致的，那么看涨期权价格 C_t 就是（2.1.10）中描述的其期末收益 $(S_{t+h}-K)^+$ 的贴现值。

给定 S_t 时，根据 S_{t+h} 的对数正态性可计算（2.1.10）中的期望值，从而得到 t 时的看涨期权公式：

$$C_t = S_t \phi(d_t) - KB(t,t+h)\phi(d_t - \sigma_S \sqrt{h}) \tag{2.1.13}$$

其中 ϕ 是累计标准正态分布函数，d_t 在随后定义。（2.1.13）就是所谓的 Black-Scholes 期权定价

[4] 然而，可参阅 Jarrow 和 Rudd（1983）、Cox 和 Rubinstein（1985）、Duffie（1989）、Duffie（1992）。在其他学者中，Hull（1993）或 Hull（1995）对期权和其它衍生证券作了更详尽论述。

[5] 它有时被称为*无偏好期权定价*，该术语可能有些误导，因为在股票市场和无风险债券定价中隐含考虑了个人偏好。然而，期权价格仅通过债券和股票的市场价格依赖于个人偏好。

[6]为了符号的方便，用一个同样的符号 W_t 表示在 P 下（(2.1.7）中）以及在 Q 下（(2.1.9）中）的一个布朗运动。实际上，Girsanov 定理建立了这两个过程之间的联系（例如，参见 Duffie（1992）和 4.2.1 节）。

公式，因此，期权价格 C_t 依赖于股票价格 S_t、执行价 K 和贴现因子 $B(t,t+h)$，这里定义：

$$x_t = Log\, S_t / KB(t,t+h) \qquad (2.1.14)$$

那么，有

$$C_t / S_t = \phi(d_t) - e^{-x_t}\phi(d_t - \sigma_S\sqrt{h}) \qquad (2.1.15)$$

其中 $d_t = (x_t / \sigma_S\sqrt{h}) + \sigma_S\sqrt{h}/2$，很容易看出数量 x_t 所起的关键作用，称为期权的货币性。

— 如果 $x_t = 0$，则当前股票价格 S_t 和执行价 K 的现值一致，换句话说，若不考虑在 t 和 $t+h$ 之间股票价格的随机变化，合约对每个人可能看起来是公平的，本文称在这种情况下有一个*两平期权*。

— 如果 $x_t > 0$（或者 $x_t < 0$），称期权是实值的（或者虚值的）。[7]

前面已经提到，Black-Scholes 公式在从业者中得到广泛运用，即使已知它的假设无法满足，特别是常数波动率 σ_S 的假设是不现实的（见 2.2 节的实证证据），这激发了 Hull 和 White（1987）提出具有随机波动率的期权定价模型，该模型假设波动率本身是一个独立于 W_t 的状态变量：[8]

$$\begin{cases} dS_t / S_t = r_t dt + \sigma_{St} dW_t \\ (\sigma_{St})_{t\in[0,T]}, (W_t)_{t\in[0,T]} \text{独立马尔可夫过程} \end{cases} \qquad (2.1.16)$$

值得注意的是（2.1.16）仍然被列入风险中性的背景下，因为 r_t 和股票的瞬时期望收益一致，另一方面，外生的波动率风险并不直接用于交易，这阻碍了对风险中性的概率测度的明确定义，在 4.2 节中对此有更详细的讨论。然而，给定 (S_t, σ_{St}) 时，如果期望值是按照马尔可夫过程 (S_t, σ_S) 的联合概率分布计算出来的，期权定价公式（2.1.10）仍是有效的。[9]那么，可把（2.1.10）写成如下形式：

$$\begin{aligned} C_t &= B(t,t+h)E_t(S_{t+h} - K)^+ \\ &= B(t,t+h)E_t\{E[(S_{t+h} - K)^+ \,|\, (\sigma_{St})_{t\le\tau\le t+h}]\} \end{aligned} \qquad (2.1.17)$$

其中方括号中的期望值是在给定 I_t 和一个波动率路径 σ_{St}， $t \le \tau \le t+h$ 时,根据 S_{t+h} 的条件概率分布计算的。然而，由于波动率过程 σ_{St} 是独立于 W_t 的，由（2.1.15）得

$$\begin{aligned} & B(t,t+h)E_t\{E[(S_{t+h} - K)^+ \,|\, (\sigma_{St})_{t\le\tau\le t+h}]\} \\ &= S_t E_t[\phi(d_{1t}) - e^{-x_t}\phi(d_{2t})] \end{aligned} \qquad (2.1.18)$$

其中的 d_{1t} 和 d_{2t} 定义如下：

$$\begin{cases} d_{1t} = (x_t / \gamma(t,t+h)\sqrt{h}) + \gamma(t,t+h)\sqrt{h}/2 \\ d_{2t} = d_1 - \gamma(t,t+h)\sqrt{h} \end{cases}$$

其中， $\gamma(t,t+h) > 0$，并且

$$\gamma^2(t,t+h) = \frac{1}{h}\int_t^{t+h} \sigma_{S\tau}^2 d\tau \qquad (2.1.19)$$

这就产生了所谓的 Hull 和 White 期权定价公式：

$$C_t = S_t E_t[\phi(d_{1t}) - e^{-x_t}\phi(d_{2t})] \qquad (2.1.20)$$

其中的期望值是在给定 σ_{St} 时，根据 $\gamma(t,t+h)$ 的条件概率分布（对于风险中性概率测度）计算的。

[7]这里用一个做过微小修改的术语代替通常的术语。事实上，当分别有 $S_t = K$ / $S_t > K$ / $S_t < K$ 时，它更普遍地被称为两平/实值/虚值期权。从经济学的观点来说，比较 S_t 和执行价 K 的现值更有意义。

[8]类似 Hull 和 White（1987）的其它随机波动率模型包括在 Johnson 与 Shanno（1987），Scott （1987），Wiggins（1987），Chesney 和 Scott（1989），Stein 和 Stein（1991）以及 Henston（1993）等等的论文中。

[9]这里隐含假定可得信息 I_t 包括过去的值 $(S_\tau, \sigma_\tau)_{\tau\le t}$，4.2 节将讨论这个假定。

[10]

在这一节的其余部分中，假设观察到的期权价格服从 Hull 和 White 的公式（2.1.20）。那么，期权价格将产生两种隐含波动率概念：（1）瞬时隐含波动率，和（2）平均隐含波动率。为了使其更精确，假设风险中性概率分布属于一个参数族，$P_\theta, \theta \in \Theta$，那么，Hull 和 White 期权定价公式产生出期权价格的一个函数表达式：

$$C_t = S_t F[\sigma_{St}, x_t, \theta_0] \tag{2.1.21}$$

其中θ_0是参数的真实未知值，（2.1.21）揭示了为什么人们经常称“期权市场可以被当成是交易波动率的市场”（参见 Stein（1989））。事实上，如果对任意给定的(x_t, θ)，$F(\cdot, x_t, \theta)$是一一对应的，那么将等式（2.1.21）求反函数，得到一个*隐含瞬时波动率*：[11]

$$\sigma_\tau^{imp}(\theta) = G[S_t, C_t, x_t, \theta] \tag{2.1.22}$$

通过证明期权价格和瞬时波动率之间一一对应关系的成立，Bajeux 和 Rochet（1992） 在事实上使期权市场的作用，正式成为一个规避波动率风险的适当工具。显然，只有知道真实未知值θ_0，或至少能够算出它足够精确的估计值时，隐含瞬时波动率（2.1.22）才能在实践中用于定价或作为套期保值的衍生工具。

然而，估计 SV 模型时所遇到的困难长久地阻碍了它们在实证中的广泛应用，这也是为什么从业者经常偏好另一个隐含波动率概念的原因，即由 Latane 和 Rendleman（1976）提出的所谓的 *Blace-Scholes 隐含波动率*。它是一个定义如下的$\omega^{imp}(t, t+h)$过程：

$$\begin{cases} C_t = S_t[\phi(d_{1t}) - e^{-x_t}\phi(d_{2t})] \\ d_{1t} = (x_t / \omega^{imp}(t,t+h)\sqrt{h}) + \omega^{imp}(t,t+h)\sqrt{h}/2 \\ d_{2t} = d_{1t} - \omega^{imp}(t,t+h)\sqrt{h} \end{cases} \tag{2.1.23}$$

其中C_t是观测到的期权价格。[12]

事实上 Hull 和 White 的期权定价模型可以看作是实践的理论基础；比较（2.1.23）和（2.1.20），可以把 Black-Scholes 隐含波动率$\omega^{imp}(t, t+h)$解释为隐含平均波动率，因为$\omega^{imp}(t, t+h)$可视为$\gamma(t, t+h)$的条件期望值（假设观测到的期权价格同 Hull 和 White 的定价公式相同）。更精确些，让我们考虑两平期权的最简单情形（一般情形将在 4.2 节中讨论），由于$x_t = 0$，则有$d_{2t} = -d_{1t}$，因此，$\phi(d_{1t}) - e^{-x_t}\phi(d_{2t}) = 2\phi(d_{1t}) - 1$，所以，$\omega_o^{imp}(t, t+h)$（加上标记$o$是为了清楚表明这里探讨的是两平期权）由下式定义：

$$\phi\left(\frac{\omega_o^{imp}(t,t+h)\sqrt{h}}{2}\right) = E_t \phi\left(\frac{\gamma(t,t+h)\sqrt{h}}{2}\right) \tag{2.1.24}$$

由于累计标准正态分布函数在零附近大致是线性的，如有（对短的到期期限h）：

$$\omega_o^{imp}(t,t+h)\sqrt{h} \approx E_t \gamma(t,t+h)$$

这就解释了为什么可以把 Black-Scholes 隐含波动率$\omega^{imp}(t, t+h)$看成一个隐含平均波动率：

$$\omega_o^{imp}(t,t+h)\sqrt{h} \approx E_t\left[\frac{1}{h}\int_t^{t+h} \sigma_{S\tau} d\tau\right]^{1/2} \tag{2.1.25}$$

2.2. *波动率的一些类型化事实(stylized facts)*

---

[10]条件是关于σ_t的，因为它概括了I_t的有关信息（过程σ假设为马尔可夫过程，并与W独立）。

[11] $F(\cdot, x_t, \theta)$是一一对应的，这一事实对某些正则性条件下的σ_{St}的任何扩散模型都成立，参见 Bajeux 和 Rochet（1992）。

[12]这里并不明确研究$\omega^{imp}(t, t+h)$和各种有关过程：C_t，S_t，x_t之间的依存关系，也为了简便的缘故，未把这种依存关系体现在符号$\omega^{imp}(t, t+h)$中。

探索模型的设定和选择总是以经验类型化事实做指导的。模型能够再现这些类型化事实是比较理想的，未能达到这一要求的，通常作为放弃某一设定的准则，虽然人们并不试图立刻用一个简单的模型去拟合或者解释所有可能的实证规律。在 ARCH 文献中详细阐述了许多关于波动率的类型化事实，例如 Bollerslev、Engle 和 Nelson（1994）。关于衍生证券和隐含波动率的实证规律也有很多，例如 Bates（1995a）。这一节中将总结经验类型化事实，补充和更新上述参考资料中包含的一些材料。

（a）肥尾

六十年代早期以来，人们开始注意到资产收益具有尖峰分布的性质，特别是 Mandelbrot（1963）、Fama（1963，1965）的发现。其结果是，大量的论文应用肥尾的独立同分布，如 Prantian 分布或者 Levy 分布，为资产收益建模。

（b）波动率群集 (clustering)

对金融时间序列的任何观测都表明了高或低波动率时段的聚集。事实上，波动率群集和资产收益肥尾是密切相关的，后者事实上是一个静态的解释，而 ARCH 模型的主要作用是给出了动态（条件）波动率行为和（无条件）肥尾间的正式联系。由 Engle（1982）提出，并且此后获得大量扩展的 ARCH 模型及 SV 模型，主要就是用于模拟波动率群集的。文献广泛讨论的还有 ARCH 效应随着时频归并(temporal aggregation)而消失，参见 Diebold(1988)以及 Drost 和 Nijman(1993)。

（c）杠杆效应

被 Black（1976）称为杠杆效应的现象指股票价格运动和波动率呈负相关。因为下跌的股票价格暗示公司财务杠杆提高，人们相信这意味着更多的不确定性及更高的波动率。然而，Black（1976），Christie（1982）及 Schwert（1989）的实证证据表明，杠杆效应自身作用太小，不足以解释股票价格中发现的不对称性。其他报告关于杠杆效应的实证证据的还包括 Nelson（1991），Gallant、Rossi 和 Tauchen（1992，1993），Campbell 和 Kyle（1993）以及 Engle 和 Ng（1993）。

（d）信息到达

资产收益通常是用固定频率的抽样观测资料进行度量和建模的，如每天、每周或每月的观测资料。Mandelbroty 和 Taylor（1967）以及 Clark（1973）等作者提出把资产收益明确地和信息到达流联系起来。事实上，人们已注意到 Clark 提出了 SV 模型的早期例子。随着时间变化，信息到达是不均匀的，并且经常是不能被直接观测到的。在概念上，人们可以把资产价格运动想象成是过程 $Y_t = Y^*_{Z_t}$ 的实现，其中 Z_t 是所谓的有向过程（directing process）。这个正的非递减随机过程 Z_t 可以被理解为和信息到达是相关的，这个时间形变或者从属随机过程的思想被 Mandelbrot 和 Taylor（1967）用来解释肥尾收益，被 Clark（1973）用来解释波动率。最近 Ghysels、Gouriéroux 和 Jasiak（1995a）对它做了改进和进一步研究。而且，Easley 和 O'Hara（1992）提出了一个涉及时间形变的微观结构模型。在实践中，它表明在与信息到达有关的许多其它现象中，同市场波动率有直接关系的包括（1）交易量，（2）报价到达，（3）可预测事件，例如股利通告或者宏观经济数据发布，（4）市场关闭。

关于交易量和波动率的关系，有几篇论文用类型化事实证明高交易量和市场波动率有显著的联系，例如参见 Karpoff（1987）或者 Gallant、Rossi 和 Tauchen（1992）。[13]被报价到达等所度量的波动率和市场活动的日内模式也很有名，并见诸文献中。Wood、McInish 和 Ord（1985）及 Harris（1986）研究了证券市场的这个现象，并且发现在市场开盘和收盘时波动率都很高，即具有 U-形模式。外汇市场的正点交易也产生独特的波动率模式，该波动率模式和市场活动的剧烈程度相联系，并产生了很强的周期性模式。许多学者分析了远期外汇市场的日内模式，例如 Műller 等（1990），Baillie 和 Bollerslev（1991），Harvey 和 Huang（1991），Dacorogna 等（1993），Bolerslev 和 Ghysels（1994），Addersen 和 Bolleslev（1995），以及 Ghysels、Gouriéroux 和 Jasiak（1995b）等。另一个有关的经验类型化事实是隔夜和周末市场的关闭以及它们对波动率的影响。Fama（1965）及 French 和 Roll（1986）发现，当 NYSE 和 AMEX 关闭时，信息的积累会变得更慢，从而导致周末及假日后的市场产生更高的波动率。Baillie 和 Bollerslev（1989）发表了关于远期市

[13]有许多理论的和实证的模型把交易量同资产收益联系起来，对此本文不作详细讨论。部分名单包括 Foster 和 Viswanathan（1993a,b），Ghysels 和 Jasiak（1994a,b），Hausman 和 Lo（1991），Huffman（1987），Lamoureux 和 Lastrapes（1990，1993），Wang（1993）以及 Andersen（1995）。

场的类似证据。最后，大量的论文证实，金融市场在股利通告（Cornell（1978），Patell 和 Wolfson（1979，1981））和宏观经济数据发布（Harvey 和 Huang（1991，1992），Ederington 和 Lee（1993））时波动率增大。

（e）长期记忆和持续性

一般来说，波动率是高度持续性的。特别是对于高频率数据，证据表明条件方差过程具有接近单位根的行为。在 ARCH 文献中，关于股票市场、商品、外汇和其它资产价格序列的 GARCH 模型的各种估计，是与 IGARCH 设定相一致的。同样，对随机波动率模型的估计显示了相似的持续性模式（参见 Jacquier、Polson 和 Rossi（1994））。这些发现导致了一场争论，即条件方差过程持久性的建模是通过单位根还是长期记忆过程。后一种过程还适用于 ARCH 模型和 SV 模型，参见 Baillie、Bollerslev 和 Mikkelsen（1993），Breidt 等（1993），Harvey（1993）以及 Comte 和 Renault（1995）。Ding、Granger 和 Engle（1993）研究了 c 为正值时 $[r(t,t+1)]^c$ 的序列相关性，其中 $r(t,t+1)$ 是投机性资产的一期收益，他们发现 $[r(t,t+1)]^c$ 对于长期滞后具有很高的自相关性，当 c 接近 1 时具有最强的时频依赖（temporal dependence）。最早在日 S&P500 收益序列中发现的这个结果，已证明它在其它股票市场指数、商品市场和外汇序列中也存在（参见 Granger 和 Ding（1994））。

（f）波动率的协同运动(comovement)

有大量的文献是讨论投机市场的跨国协同运动的。资本市场的全球化是否提高了价格的波动率和股票收益的相关性已经成为最近许多研究的主题，包括 von Fustenberg 和 Jean （1989），Hamao、Masulis 和 Ng（1990），King、Sentatna 和 Wadhwani（1994），Harvey、Ruiz 和 Sentana（1992）以及 Lin、Engle 和 Ito（1994）。人们通常运用因子模型来模拟国际波动率的共同性，比如 Diebold 和 Nerlove（1989），Harvey、Ruiz 和 Sentana（1992），Harvey，Ruiz 和 Shephard （1994），或者探索所谓的共同特征，如 Engle 和 Kozicki（1993），或共同趋势，如 Bollerslev 和 Engle（1993）。

（g）隐含波动率的相关性

类型化事实通常为无模型的实证观测所揭示。[14]隐含波动率显然是基于模型的，因为它们从一个特定模型的定价公式计算得到，即在 2.1.3 节中提到的 Black 和 Scholes 模型。由于它们是在每天数据的基础上计算出来的，且模型假定波动率为常数，因此明显存在内在的不一致性。然而，由于事实上许多期权价格是通过它们的隐含波动率来报价的，所以很自然地就要研究后者的时间序列行为。基于相同标的的资产，具有不同执行价和到期日的同期期权价格会产生不同的隐含波动率，所以经常要计算一个综合量度。综合量度通常是通过加权方案得到的，它给予接近实值的期权更多的权重，这种期权在有组织的市场中的交易量最大。[15]

从股票、股票指数和货币期权得到的隐含波动率的时间序列性质非常相似，它们体现了平稳性，并且可用一阶自回归模型很好地描述（参见 Merville 和 Pieptea（1989）以及 Sheikh（1993）关于股票期权的研究，Poterba 和 Summers（1986），Stein（1989），Harvey 和 Whaley（1992）以及 Diz 和 Finucane（1993）关于 S&P100 合约的研究，Taylor 和 Xu（1994），Campa 和 Chang（1995），以及 Jorion（1995）关于货币期权的研究）。等式（2.1.25）表明，人们认为隐含（平均）波动率包含了未来波动率的信息，因此可以预测未来波动率。通常对该假设的检验是用已实现波动率对过去的隐含波动率做回归。

关于隐含波动率可预测内容的实证结果多种多样。Lamoureux 和 Lastrapes（1993）对时间序列的研究考虑了无红利股票，并且比较了 GARCH、隐含波动率和历史波动率估计的预测能力，发现隐含波动率的预测优于其它波动率，尽管从式（2.1.25）可以看出预测偏差。明显不同的是，Canina 和 Figlewski（1993）研究交易活跃的 S&P100 指数看涨期权，他们发现隐含波动率在预测 S&P100 指数的未来实际波动率时几乎没有用。在设定以星期作为抽样间隔的 S&P100 期权合约和不同样本条件下，Day 和 Lewis（1992）发现隐含波动率不但具有预测内容，而且还是无偏的。对外币期权的研究，如 Jorion （1995），也发现隐含波动率能预测未来的实际值，并且还发现 GARCH 和历史波动率都不优于隐含波动率。

[14]这在一定程度上是假想的，即使对宏观经济数据也是如此，例如当这些数据是趋势消除或季节调整的。趋势消除和季节调整都以模型为基础。关于趋势消除对类型化事实的潜在严重影响，参见 Canova（1992）及 Harvey 和 Jaeger （1993），关于季节调整对实证规律的影响见 Ghysels 等人（1993）。

[15]有不同的加权方案被提出，参见 Latane 和 Rendleman （1976）， Chiras 和 Manaster （1978）， Beckers（1981），Whaley（1982），Day 和 Lewis（1988），Engle 和 Mustafa（1992）以及 Bates（1995b）。

（h）隐含波动率的期限结构

Black-Scholes 模型所预测波动率的期限结构是平缓的。事实上，当短期波动率很低的时候，实值期权的隐含波动率的期限结构是向上倾斜的，反之则向下倾斜（见 Stein（1989））。Taylor 和 Xu （1994）发现外汇期权隐含波动率的期限结构每几个月都要改变一次斜率方向。Stein（1989）也发现中短期隐含波动率的实际敏感度比预测期限结构得到的估计敏感度要更大一些，并且得出中期隐含波动率对信息具有过度反应的结论。Diz 和 Finucane（1993）运用不同的估计技术拒绝了过度反应假设，同时报告了反应不足的证据。

（i）微笑

如果市场中的期权价格满足 Black-Scholes 公式，则对应于相同资产的各种期权的所有 Black-Scholes 隐含波动率将和标的资产的波动率参数 σ 相一致。但事实并非如此。由（2.1.23）定义的 Black- Scholes 隐含波动率 $\omega^{imp}(t,t+h)$ 在很大程度上取决于日历时间 t、到期期限 h 和期权的货币性 $x_t = LogS_t / KB(t,t+h)$。当 BS 隐含波动率被用来评估具有不同执行价 K 和到期期限 h 的新期权时，这可能在期权定价和保值中产生偏差。这些对从业者来说是非常熟悉的价格失真在实证文献中通常被称为*微笑效应*，这里所谓的“微笑”指的是隐含波动率在不同执行价时产生的 U-形模式。说得更精确些，下列类型化事实被广泛讨论（参见 Rubinstein（1985），Clewlow 和 Xu（1993），Taylor 和 Xu（1993））：

- 作为 K（或者 $\log K$）的函数，$\omega^{imp}(t,t+h)$ 的 U-形模式当处于接近实值的期权时取其最小值（贴现的 K 接近于 S_t，即 x_t 接近于零）。
- 波动率的微笑作为 $\log K$（或者 x_t）的函数常是，但不总是对称的。当微笑是非对称的时候，偏度效应通常可以被描述为把一个单调曲线加到标准对称微笑上：如果一个下降的曲线被加上，则隐含波动率倾向于在执行价下降的时候比在它上升的时候增长得更快，并且隐含波动率曲线在处于虚值期权时取最小值。在相反的情形中（加上一个上升的曲线），隐含波动率在执行价上升时增长得更快，并且在处于实值期权时取最小值。
- 到期期限越短，微笑的幅度增长越快。事实上，对很短的到期期限，微笑效应是很显著的（同步期权价格的 BS 隐含波动率可能在 15％到 25％之间变化），而对很长的到期期限，它几乎完全消失了。

一般认为波动率的微笑效应必须由随机波动率模型来解释。这有几个理由：首先，应用随机时变波动率模型来表示随机时变 BS 隐含波动率是很自然的。其次，微笑下降的幅度是到期期限的函数，遵从象（2.1.25）这样的公式。确实，实际情况显示，当到期期限增加时，波动率的时频归并消除了条件异方差，从而减少微笑现象。最后，偏度本身也可以被归因于波动率过程的随机特征以及该过程与价格过程（所谓的杠杆效应）的整体相关性。事实上，这个效应对股票价格数据是很明显的，但是对利率和汇率序列却是很小的，这就是为什么微笑的偏度在以股票为标的期权时更常见。

然而，慎重对待一些诱人的关系是很重要的：随机隐含波动率和随机波动率；股票和微笑偏度的不对称性。这将在第 4 节中讨论，诸如此类的情况并不总具有严密的证明。而且，关于解释微笑及其偏度的其它论据（跳跃，交易成本，买卖差价，非同步交易，流动性问题，…）在理论上和实证上都应加以考虑。例如，实证证据表明最昂贵的期权（微笑曲线的上部）也是最小流动性的期权；因此偏度或许可归因于期权市场中流动性的特殊结构形式。

2.3. *信息集*

到目前为止，我们还未清楚地定义信息集。这样做的目的是集中讨论相关问题。在这节中，我们需要更正式的信息定义，以便能够阐明文献中介绍的各种 SV 模型之间以及 SV 和 ARCH 模型之间的一些遗漏的联系。我们知道，SV 模型是从一系列不同问题的研究中产生的。本节试图定义一个共同的思路和一般的统一框架。通过对信息集的认真分析来完成这个工作，并把它和 Granger 意义上的非因果性概念联系起来。这些因果性条件将允许我们在 2.4 节中刻画出 ARCH 和 SV 模型的独特性质。[16]

[16]这节中的分析——在关于运用信息集阐明 SV 和 ARCH 类模型之间的差别时——具有一些和 Andersen（1992）一样的特征。

2.3.1. 状态变量和信息集

Hull 和 White（1987）模型是衍生资产定价模型的一个简单例子，其中的股票价格动态是由一些不可观察的状态变量控制的，例如随机波动率。更一般地，可以很方便地假设一个多变量扩散过程(diffusion process)U_t，它在一定意义上概括了相关的状态变量：

$$\begin{cases} dS_t/S_t = \mu_t dt + \sigma_t dW_t \\ dU_t = \gamma_t dt + \delta_t dW_t^U \\ Cov(dW_t, dW_t^U) = \rho_t dt \end{cases} \quad (2.3.1)$$

其中随机过程 μ_t、σ_t、γ_t、δ_t 和 ρ_t 是满足（假设 2.3.1）$I_t^U = [U\tau, \tau \le t]$。这说明过程 U 概括了股票价格过程 S（它验证了术语“状态”变量）的全部动态特征。因为对于状态变量的一个样本路径 $(U\tau)_{0\le\tau\le T}$，连续收益 $S_{t_{k+1}}/S_{t_k}$，$0 \le t_1 < t_2 < ... < t_k \le T$ 是随机独立的并且服从对数正态分布（如在基准 BS 模型中的一样）。

2.1.2 节中的一些论证可以扩充到这里讨论的状态变量的框架中（参见 Garcia 和 Renault（1995））。实际上，Black 和 Scholes 模型通常通过期权价格的隐含波动率为它们提供一个报价的方法，上述的扩充为此提供了理论依据。[17]事实上，这是引入 BS 模型中被忽略的异质性的一个方法（参见 Renault（1995），这与引入劳动力市场和其他市场的微观经济计量模型的异质性相似）。

在连续时间模型中，t 时刻交易者（他们的信息决定期权价格）的可得信息被状态变量样本路径和股票价格过程样本路径这两类的连续时间观测结果所描述；即：

$$I_t = \sigma[U_\tau, S_\tau; \tau \le t] \quad 。 \quad (2.3.2)$$

2.3.2. 离散抽样和 Granger 非因果性

下一节将明确讨论离散时间模型。有必要建立类似等式（2.3.1）的离散时间模型。这里讨论的离散时间抽样和 Granger 非因果性条件将使我们进一步接近用离散时间数据建立统计模型的正式框架。很清楚，只要对过程 ε_t 加一些约束，类似等式（2.3.1）的离散时间形式便成为：

$$\log S_{t+1}/S_t = \mu(U_t) + \sigma(U_t)\varepsilon_{t+1} \quad 。 \quad (2.3.3)$$

要加上的约束必须足够灵活，以适应比如杠杆效应这样的情况，这种灵活性体现在：

假设 2.3.2.A. （2.3.3）中的过程 ε_t 是独立同分布的，并且不是状态变量过程 U_t 的 Granger 结果。

假设 2.3.2.B. （2.3.3）中的过程 ε_t 不是 U_t 的 Granger 原因。

假设 2.3.2.B 对 BS 隐含波动率的实际应用是很有帮助的，因为它是类似假设 2.3.1 的离散时间的情形。假设 2.3.1 描述的是，过程 U 的系数适合 I_t^U（详见 Garcia 和 Renault(1995)）。假设 2.3.2.A 对函数 $\mu(U_t)$ 和 $\sigma(U_t)$ 的统计解释很重要，$\mu(U_t)$ 和 $\sigma(U_t)$ 分别表示趋势和波动率系数，即，

$$\begin{aligned} &E[\log S_{t+1}/S_t | (S_\tau/S_{\tau-1}; \tau \le t)] \\ &\quad = E[E[\log S_{t+1}/S_t | (U_\tau, \varepsilon_\tau; \tau \le t)] | (S_\tau/S_{\tau-1}; \tau \le t)] \\ &\quad = E[\mu(U_t) | (S_t/S_{t-1}; \tau \le t)] \end{aligned} \quad (2.3.4)$$

由于 $E[\varepsilon_{t+1} | (U_\tau, \varepsilon_\tau; \tau \le t)] = E[\varepsilon_{t+1} | (\varepsilon_\tau; \tau \le t)] = 0$，该等式成立归因于假设 2.3.2.A 从 U_t 到 ε_t 的 Granger 非因果性。同样地，能够很容易得出

$$\begin{aligned} &Var[\log S_{t+1}/S_t - \mu(U_t) | (S_\tau/S_{\tau-1}; \tau \le t)] \\ &\quad = E[\sigma^2(U_t) | (S_\tau/S_{\tau-1}; \tau \le t)] \end{aligned} \quad (2.3.5)$$

（2.3.4）和（2.3.5）中已经隐含地引入了一个新的信息集，包括（2.3.2）中定义的 I_t，该信

[17] Garcia 和 Renault（1995）认为假设 2.3.1 实质上保证了关于价格对（股票价格，执行价格）期权价格(option prices with respect to the pair)的同质性，期权价格的同质性反过来保证了 BS 隐含波动率不依赖股票价格水平，而依赖货币性 S/K。Merton （1973）首次强调了这种同质性。

息集对进一步的分析同样是很有用的。事实上，人们经常限于用统计方法分析股票收益序列的离散时间抽样所传达的信息，该序列由如下信息集定义为：

$$I_t^R \equiv \sigma[S_\tau / S_{\tau-1} : \tau = 0,1,...,t-1,t] \tag{2.3.6}$$

其中上标 R 代表收益。通过扩展的 Andersen（1994），我们将采用单变量波动率模型的最一般框架，其假设包括 2.3.2.A，2.3.2.B 和：

假设 2.3.2.C. $\mu(U_t)$ 是 I_t^R 可测的。

因此，（2.3.4）和（2.3.5）事实上表明：

$$E[\log S_{t+1}/S_t \mid I_t^R] = \mu(U_t) \tag{2.3.7}$$

$$Var[(\log S_{t+1}/S_t) \mid I_t^R] = E[\sigma^2(U_t) \mid I_t^R] \tag{2.3.8}$$

2.4. 随机波动率的统计建模

金融时间序列是在离散时间间隔中观测到的，而多数理论模型是在连续时间中构建的。一般说来，有两种统计学方法可以解决这个问题。或者估计连续时间过程的统计离散时间模型，或者按连续时间设定统计模型，而后通过离散时间近似推断。本节将详细讨论前一种方法，后一种将在第 4 节中介绍。这里讨论一类一般的离散时间统计模型。在 2.4.1 节中介绍一些符号和术语。下一节讨论 Andersen（1994）提出的所谓的随机自回归波动率模型。该模型由 Andersen（1994）引入，作为一个非常一般的、灵活的半参数框架，包含文献中已有的随机波动率的各种表达式。它所要求的参数识别和约束在 2.4.3 节中讨论。

2.4.1. 符号和术语

在 2.3 节中，对趋势 $\mu(\cdot)$ 和波动率 $\sigma(\cdot)$ 所具有的函数形式未加以设定。事实上，在某种意义上本文建立了一个由 Lezan、Renault 和 de Vitry（1995）最近提出的非参数框架，他们提出这个框架是为了讨论*未知形式随机波动率*的符号。[18]这种非参数结构包含了标准参数模型（参见 2.4.2 节中更多正式的讨论)。为表述方便，先考虑两个极端的情形。为了简化，假设 $\mu(U_t) = 0$：（i）当 $\sigma(U_t) = \sigma_t$ 是一个独立于股票收益标准化新生（innovation）过程 ε 的随机过程时，可以得到 Hull 和 White 模型（2.1.16）的离散时间形式；（ii）σ_t 可以是过去新生的一个确定性函数 $h(\varepsilon_\tau, \tau \le t)$。后者是和（i）完全相反的，并且通过它可以得出 h 的参数化函数的许多种选择，从而得到 X-ARCH 模型（GARCH，EGARCH，QTARCH，周期性的 GARCH 等等）。

除了这两种极端的情形（其中假设 2.3.2.A 是通过不重要的退化方法实现的）之外，人们还可以提供杠杆效应。[19]特别是，U 中的新生和收益过程之间的同期相关性结构可以非零，因为 Granger 非因果性假设是处理瞬时(temporal)的因果联系，而不是同期的。例如，可以令 $\sigma(U_t) = \sigma_t$，且：

$$\log S_{t+1}/S_t = \sigma_t \varepsilon_{t+1} \tag{2.4.1}$$

$$Cov(\sigma_{t+1}, \varepsilon_{t+1} \mid I_t^R) \ne 0 \quad 。 \tag{2.4.2}$$

（2.4.2）中负的协方差是杠杆效应的一个标准情形，这并不违反非因果性假设 2.3.2.A 和 B。

有必要做一些结论性的观察来应付文献中迅速增长的各种术语。首先，本文没有考虑 Taylor（1994）提出的（2.4.1）的“滞后自回归随机方差模型”与“同期自回归随机方差模型”之间的区别，后者被定义为：

$$\log S_{t+1}/S_t = \sigma_{t+1} \varepsilon_{t+1} \quad 。 \tag{2.4.3}$$

---

[18] Lezan、Resault 与 Vitry（1995）详细讨论了怎样使这个框架适用于诸如波动率群集这样的现象。作为一种非参数的结构，在（稳健）估计方面也具有某些优点。例如，他们发展的模型能够作为一种特定参数模型（见第 5 节）的有效算法的第一步估计。

[19]假设 2.3.2 在情形（i）中发挥作用，但不能被用于 GARCH 情形（ii），当它在后一个情形中不成立时，它使得 GARCH 结构不太适合期权定价。

事实上，由于波动率过程σ_t是不可观测的，故只要（2.4.1）和（2.4.3）不是由精确的（非）因果性假设实现，它们的假定就观测而言就是等价的。例如：(i)（2.4.1）和假设 2.3.2.A 一起构成一个正确且非常一般化的 SV 模型定义，该模型可通过假设 2.3.2.B 实现期权定价，并通过(2.4.2)引入杠杆效应，(ii)（2.4.2）与（2.4.3）的结合将不是 SV 模型的一个正确定义，因为通常有这样的情形：$E[\log S_{t+1}/S_t \mid I_t^R] \neq 0$，并且模型将通过过程$\sigma$引入一个预测，该预测不但和波动率有关而且和期望收益也有关。

为了简化符号，第 3 节中使用的框架（2.4.3）的杠杆效应将由$Cov(\sigma_{t+1}, \varepsilon_t) \neq 0$代替$Cov(\sigma_{t+1}, \varepsilon_{t+1}) \neq 0$来实现。Amin 和 Ng（1993）提出期权定价的另一个术语，他们的“可预测的”和“不可预测的”波动率之间的差别很接近杠杆效应的概念，并且可以通过因果性的概念进行分析，这在 Carcia 和 Renault（1995）中有讨论。最后，不像相应的 ARCH 模型（参见 Drost 和 Nijman（1993）），它没有必要区分 SV 模型的弱的、半强的和强的概念。事实上，这里定义的 SV 模型族可以满足参数化。这种参数化在时频归并下是闭合的（关于时频归并见第 4.1 节）。

2.4.2. 随机自回归波动率

为简单起见，考虑如下的单变量波动率过程：

$$y_{t+1} = \mu_t + \sigma_t \varepsilon_{t+1} \tag{2.4.4}$$

其中μ_t是观测变量$y_t \in I_t^R, \tau \le t$的一个可测函数。这里将围绕（2.4.4）展开讨论，先讨论几个一般化的问题，而不局限于特定模型；第 3.5 节中会有更明确的扩展。根据（2.3.8）的结果，可得：

$$Var[y_{t+1} \mid I_t^R] = E[\sigma_t^2 \mid I_t^R] \tag{2.4.5}$$

它表明：(1）通过条件期望（2.4.5）中的自回归动态特征，可以反映波动率群集；(2）肥尾可以用三种方法中的任何一个得到，即（a）通过白噪声ε_t的分布的肥尾，(b）通过$E[\sigma_t^2 \mid I_t^R]$的随机特征，(c）通过波动率过程σ_t的特殊的随机性，该随机性使得它隐含起来，也即$\sigma_t \notin I_t^R$。[20]从（1）和（2）得出的波动率动态特性通常是关于σ_t的某个非线性函数的 AR（1）模型，因此，波动率过程被假设为平稳的并且是一阶马尔可夫过程，但并不要求σ_t本身是线性的 AR（1)。正是这激发了 Andersen（1994）提出随机自回归方差或者 SARV 类模型，其中σ_t（或者σ_t^2）是马尔可夫过程K_t的一个多项式函数$g(k_t)$，该马尔可夫过程具有如下动态设定：

$$K_t = w + \beta K_{t-1} + [\gamma + \alpha K_{t-1}]u_t \tag{2.4.6}$$

其中$\tilde{u}_t = u_t - 1$是具有单位方差的零均值白噪声。Andersen（1994）讨论了确保K_t满足平稳性和遍历性的充分正则性条件。在进入细节的讨论之前，要注意基本的非因果性假设 2.3.2.A 暗示（2.4.6）中的u_t过程并不是（2.4.4）中ε_t的 Granger 原因。非因果性条件实际上表明了对 Andersen（1994）定义的轻微修改。也就是，它暗示给定$\varepsilon_{t-j}, j \ge 0$时，假设$\varepsilon_{t+1}$独立于$u_{t-j}, j \ge 0$是对有条件的概率分布而言的，而不是对无条件的分布。这个修改并没有改变 Andersen 的 SARV 类模型是迄今为止有关波动率文献中最一般的参数统计模型这一事实。GARCH（1,1）模型是直接从（2.6.4)（通过令$K_t = \sigma_t^2$，$\gamma = 0$和$u_t = \varepsilon_t^2$）得到的。要注意（2.4.4）和（2.4.6）的随机分量之间的确定性关系$u_t = \varepsilon_t^2$强调了，在 GARCH 模型中，波动率过程不存在随机设定。由 Taylor（1986）推广的自回归随机方差模型也属于 SARV 类，这里：

$$\log \sigma_{t+1} = \xi + \phi \log \sigma_t + \eta_{t+1} \tag{2.4.7}$$

其中η_{t+1}是一个白噪声扰动，所以$Cov(\eta_{t+1}, \varepsilon_{t+1}) \neq 0$引入杠杆效应。这是具有$K_t = \log \sigma_t$、

[20] Kim 与 Shephard（1994）使用 S&P500 指数的周收益率数据发现，一个$t-$GARCH 模型具有同基于正态的 SV 模型几乎一样的似然值。这个例子表明σ_t的特殊随机性可以产生同白噪声ε的肥尾学生分布一样的边际峰度水平。

$\alpha = 0$ 和 $\eta_{t+1} = \gamma u_{t+1}$ 的一个 SARV 模型。[21]

2.4.3. 参数的识别

介绍一般类型的波动率过程，像前一节中讨论的 SARV 类型一样，提出有关识别的问题。再次假设：

$$\begin{aligned} y_{t+1} &= \sigma_t \varepsilon_{t+1} \\ \sigma_t^q &= g(K_t)，\quad q \in \{1,2\} \\ K_t &= w + \beta K_{t-1} + [\gamma + \sigma K_{t-1}]u_t \end{aligned} \quad (2.4.8)$$

Andersen（1994）指出，通过零均值白噪声过程 $\tilde{u}_t = u_t - 1$，能更好地解释模型：

$$K_t = (w + \gamma) + (\alpha + \beta)K_{t-1} + [\gamma + \alpha K_{t-1}]\tilde{u}_t \quad (2.4.9)$$

从后者可以很清楚地看出，经验上可能很难区分常数 w 和“随机”常数 γu_t，同样的，分别估计参数 α 和 β 也会有很多问题，因为 $(\alpha + \beta)$ 决定了对波动率冲击的持续性，这些估计问题通常是通过对两对参数（w，γ）和（α，β）施加（主观）约束来解决。

GARCH（1,1）和自回归随机方差设定分别假设 $\gamma = 0$ 和 $\alpha = 0$。没有这些约束的所有参数的识别一般要有附加的限制，例如通过 ε_{t+1} 和 u_t 分布的一些假设，把（2.4.6）的半参数结构限制在参数统计模型的范围内。

为了更严格地解决识别问题，Andersen（1994）认为考虑如下的再参数化是很有用的（为了符号使用上的方便，假设 $\alpha \neq 0$）：

$$\begin{cases} K = (w+\gamma)/(1-\alpha-\beta) \\ \rho = \alpha + \beta \\ \delta = \gamma/\alpha \end{cases} 。\quad (2.4.10)$$

因此方程（2.4.9）可写成：

$$K_t = K + \rho(K_{t-1} - K) + (\delta + K_{t-1})\tilde{U}_t$$

其中 $\tilde{U}_t = \alpha\tilde{u}_t$。

从（2.4.10）可以清楚看到只有原始参数 α，β，γ，w 的三个函数可以被识别，例如三个参数 K，ρ，δ 是由过程 K_t 的前三个无条件矩来识别的。

为了给这些识别结果以实证内容，必须知道：（1）怎样从可观测过程 Y_t 的矩转化到波动率过程 σ_t 的矩，（2）怎样从波动率过程 σ_t 的矩转化到潜在过程(latent process) K_t 的矩。第一点通过指定标准化的新生过程 ε 相应的矩很容易解决，例如，如果假设一个高斯概率分布，就可得到：

$$\begin{cases} E|y_t| = \sqrt{2/\pi}\,E\sigma_t \\ E|y_t||y_{t-j}| = 2/\pi\,E(\sigma_t\sigma_{t-j}) \\ E|y_t^2||y_{t-j}| = \sqrt{2/\pi}\,E(\sigma_t^2\sigma_{t-j}) \end{cases} \quad (2.4.11)$$

第二点的解决方法一般需要映射 g 的设定和（2.4.6）中 u_t 的概率分布的设定。对所谓的对数正态 SARV 模型，假设 $\alpha = 0$ 和 $K_t = \log\sigma_t$（Taylor 的自回归随机方差模型）以及 u_t 是正态分布的（波动率过程的对数正态性）,在这种情形下，很容易得出：

$$\begin{cases} E\sigma_t^n = \exp[nEK_t + n^2 Var\,K_t/2] \\ E(\sigma_t^m\sigma_{t-j}^n) = E\sigma_t^m E\sigma_{t-j}^n \exp[mnCov(K_t, K_{t-j})] \\ Cov(K_t, K_{t-j}) = \beta^j Var K_t \end{cases} \quad (2.4.12)$$

---

[21] Andersen（1994）还证明了 SARV 结构包含了另一类被认为是病态设定的随机方差模型，因为它合并了（2.4.2）和（2.4.3）。

对没有正态假设（即 QML，混合正态，学生分布……）的情形，这个模型将在第 3 节和第 5 节中从概率和统计的观点出发进行详细研究。而且，这是研究 SARV 类型模型的其它设定的一个样板，另外，各种设定将在第 4 节中考虑以代表连续时间模型。

3. 离散时间模型

本节的目的是运用简单的单变量情形，讨论离散时间模型的统计处理。先从与（2.4.7）中讨论的自回归随机方差模型相对应的最基本的 SV 模型的定义开始。第 3.2 节讨论它的统计特征。第 3.3 节给出与 ARCH 模型的比较。第 3.4 节研究过滤、预测和平滑。最后一节讨论各种扩展，包括多变量模型。第 5 节的最后部分讨论决定波动率过程的参数的估计。

3.1. 离散时间 SV 模型

离散时间 SV 模型可记为

$$y_t = \sigma_t \varepsilon_t \quad , \qquad t = 1,\dots,T \tag{3.1.1}$$

其中 y_t 代表去均值（demeaned）的收益过程 $y_t = \log(S_t/S_{t-1}) - \mu$，并且 $\log\sigma_t^2$ 服从 AR（1）过程。假设 ε_t 是独立同分布的随机扰动序列。通常 ε_t 被设定为具有标准分布，所以它的方差 σ_ε^2 是已知的。因而，对正态分布而言 σ_ε^2 是单位元素，而对具有 v 自由度的 $t-$分布而言 σ_ε^2 是 $v/(v-2)$，遵循文献中常采用的惯例，记 $h_t \equiv \log\sigma_t^2$：

$$y_t = \sigma\varepsilon_t e^{0.5h_t} \tag{3.1.2}$$

其中 σ 是一个尺度（scale）参数，它消除了平稳一阶自回归过程所需要的常数项，该过程为

$$h_{t+1} = \phi h_t + \eta_t, \eta_t \sim IID(0,\sigma_\eta^2) \,, \quad |\phi| < 1 \tag{3.1.3}$$

前面已经提到，如果允许 ε_t 和 η_t 相关，那么模型能够反映常在股价中发现的不对称行为。事实上，ε_t 和 η_t 之间的负相关性导致杠杆效应，如同 2.4.1 节中一样，对（3.1.3）中扰动的时间选择要确保观测值仍是鞅差分，用这种方法写出方程是为使它和状态空间的文献保持一致。

需要强调的是，上述模型只是第 2 节的连续时间模型在离散时间观测中的一个近似，Dassions（1995）运用 Edgeworth（埃奇渥斯）扩展（参见第 4.1 节和第 4.3 节中更进一步的讨论）检验了该近似的准确性。

3.2. 统计特征

即使 ε_t 和 η_t 是同期相关的，SV 模型仍具有如下性质。第一，y_t 是一个鞅差分；第二，h_t 的平稳性隐含了 y_t 的平稳性；第三，如果 η_t 服从正态分布，那么从对数正态分布的性质可得到 $E[\exp(ah_t)] = \exp(a^2\sigma_h^2/2)$，其中 a 是一个常数，σ_h^2 是 h_t 的方差，因此，如果 ε_t 具有有限的方差，那么 y_t 的方差如下

$$Var(y_t) = \sigma^2\sigma_\varepsilon^2\exp(\sigma_h^2/2) \tag{3.2.1}$$

同样，如果 ε_t 的四阶矩存在，则 y_t 的峰度是 $\kappa\exp(\sigma_h^2)$，其中 κ 是 ε_t 的峰度，这样 y_t 呈现出比 ε_t 更大的峰度，最后，所有的奇数阶矩都是零。

基于多种目的，需要考虑绝对值的幂的矩。再次，假设 η_t 是正态分布的，由于已假定 ε_t 具有标准正态分布，Harvey（1993）推导出了如下表达式：

$$E|y_t|^c = \sigma^c 2^{c/2}\frac{\Gamma(c/2+1/2)}{\Gamma(1/2)}\exp(\frac{c^2}{8}\sigma_h^2) \quad , \quad c > -1 \quad , \quad c \neq 0 \tag{3.2.2}$$

和

$$Var|y_t|^c = \sigma^{2c} 2^c \exp(\frac{c^2}{2}\sigma_h^2)\left\{\frac{\Gamma(c+1/2)}{\Gamma(1/2)} - \left[\frac{\Gamma(c/2+1/2)}{\Gamma(1/2)}\right]^2\right\} \quad , \quad c > -1 \quad , \quad c \neq 0$$

注意$\Gamma(1/2)=\sqrt{\pi}$，$\Gamma(1)=1$。对其它分布的也可以计算出ε_t相应的表达式，包括学生t－分布和一般误差分布（参见 Nelson（1991））。

最后，σ_t^2 变异系数的平方经常被用作 SV 过程相关强度的度量，即 $Var(\sigma_t^2)/[E(\sigma_t^2)]^2=\exp(\sigma_h^2)-1$。Jacquier、Polson 和 Rossi（1994）认为这比σ_η^2更容易解释。在实证研究中，他们指出该系数很少小于 0.1 或者大于 2。

3.2.1. 自相关函数

如果假设扰动ε_t和η_t是相互独立的，并且η_t是正态的，那么观测值的 ACF 的绝对值的c次方如下：

$$\rho_\tau^{(c)} = \frac{E(|y_t|^c|y_{t-\tau}|^c) - \{E(|y_t|^c)\}^2}{E(|y_t|^{2c}) - \{E(|y_t|^c)\}^2} = \frac{\exp(\frac{c^2}{4}\sigma_h^2\rho_{h,\tau})-1}{\kappa_c\exp(\frac{c^2}{4}\sigma_h^2)-1} \quad , \quad (3.2.3)$$

$$\tau \geq 1 \quad , \quad c > -0.5 \quad , \quad c \neq 0$$

其中κ_c是

$$\kappa_c = E(|y_t|^{2c})/\{E(|y_t|^c)\}^2 \quad (3.2.4)$$

且$\rho_{h,\tau}$，$\tau=0,1,2,...$表示h_t的 ACF，Taylor（1986）给出了c等于 1 和 2 且ε_t是正态分布时的该表达式。当$c=2$时，κ_c是峰度，对于标准正态分布它等于 3。更一般地，

$$\kappa_c = \Gamma(c+1/2)\Gamma(1/2)/\{\Gamma(c/2+1/2)\}^2 \quad , \quad c \neq 0 \quad 。$$

对于具有ν自由度的学生$t-$分布：

$$\kappa_c = \frac{\Gamma(c+1/2)\Gamma(-c+\nu/2)\Gamma(1/2)\Gamma(\nu/2)}{\{\Gamma(c/2+1/2)\Gamma(-c/2+\nu/2)\}^2} \quad , \quad (3.2.5)$$

$$|c| < \nu/2 \quad , \quad c \neq 0$$

注意，如果c是 2，ν必须至少是 5。

ACF，$\rho_\tau^{(c)}$具有如下特征：第一，与 Taylor（1986，p.74－5）比较，如果σ_h^2很小，并且/或者$\rho_{h,\tau}$接近于 1，

$$\rho_\tau^{(c)} \approx \rho_{h,\tau}\frac{\exp(\frac{c^2}{4}\sigma_h^2)-1}{(\kappa_c\exp(\frac{c^2}{4}\sigma_h^2)-1)} \quad , \quad \tau \geq 1 \quad (3.2.6)$$

则h_t的 ACF 的形状近似于$\rho_\tau^{(c)}$。除了乘上一个比例因子外，对于正的c，当κ_c大于 1 时，该比例因子必须小于 1；第二，对于$t-$分布，当ν趋于无穷时，κ_c下降。所以，对于正态分布，$\rho_\tau^{(c)}$是极大值。另一方面，具有比正态分布更小峰度的分布将会有更高的$\rho_\tau^{(c)}$值。

虽然（3.2.6）给出了$\rho_\tau^{(c)}$和c之间的一个明确关系，但仍不可能对某个c值下$\rho_\tau^{(c)}$的极大化作出概括性描述。事实上，不同的σ_h^2值将导致不同的c值以极大化$\rho_\tau^{(c)}$。如果σ_h^2的选择使得$\rho_\tau^{(c)}$和 Ding、Granger 和 Engle（1993）的对应项大小相同，那么对于稍微小于 1 的c值就可以得到该极大值。Harvey（1993）指出，与c相关的$\rho_\tau^{(c)}$曲线的形状同 Ding，Granger 和 Engle 报告的经

验关系很相似。

3.2.2. 对数变换

将（3.1.2）中的观测值进行平方并且取对数就得到

$$\log(y_t^2)=\log(\sigma^2)+h_t+\log\varepsilon_t^2 \qquad (3.2.7)$$

或

$$\log(y_t^2)=\omega+h_t+\xi_t \quad , \qquad (3.2.8)$$

其中$\omega=\log(\sigma^2)+E\log\varepsilon_t^2$，通过构造以使扰动$\xi_t$具有零均值。

根据 Abramovitz 和 Stegun（1970），当ε_t具有标准正态分布时，可知ε_t^2的均值和方差为-1.27和$\pi^2/2=4.93$。然而，$\log\varepsilon_t^2$的分布远非正态，而是严重偏斜的，且具有长尾。

更一般地，如果ε_t服从具有ν自由度的$t-$分布，它可以表示为：

$$\varepsilon_t=\zeta_t\kappa_t^{-0.5}$$

其中ζ_t是一个标准正态变量，κ_t是独立分布，从而使$\nu\kappa_t$服从具有ν自由度的卡方分布。这样

$$\log\varepsilon_t^2=\log\zeta_t^2-\log\kappa_t$$

再根据 Abramovitz 和 Stegun（1970）中的结果，得出结论$\log\varepsilon_t^2$的均值和方差分别是$-1.27-\psi(\nu/2)-\log(\nu/2)$和$4.93+\psi'(\nu/2)$，其中$\psi(\cdot)$是 digamma 函数*。注意，$\xi_t$的矩是存在的，即使模型的构造方式使$\varepsilon_t$服从 Cauchy 分布，即$\nu=1$。事实上，在这种情形中，$\xi_t$是对称的，具有超峰度为 2，而当$\varepsilon_t$是高斯分布时，$\xi_t$具有超峰度为 4。

由于$\log\varepsilon_t^2$序列独立，由此可直接算出当h_t服从任何平稳过程时$\log y_t^2$的 ACF：

$$\rho_t^{(0)}=\rho_{h,\tau}\Big/\{1+\sigma_\xi^2\big/\sigma_h^2\} \quad , \qquad \tau\geq 1 \qquad (3.2.9)$$

符号$\rho_t^{(0)}$表明观测值的绝对值的幂的 ACF 等同于 Box-Cox 变换，即$\{|y_t|^c-1\}\big/c$，因此任何（非零）幂乘方的绝对值的对数变换对应于$c=0$。（但要注意，在（3.2.3）中不能简单地规定$c=0$）。

根据 Harvey、Ruiz 和 Shephard（1994），即使ε_t和η_t不是相互独立的，如果ε_t和η_t的联合分布是对称的，即$f(\varepsilon_t,\eta_t)=\ \ f(-\varepsilon_t,-\eta_t)$，那么扰动$\eta_t$和$\xi_t$仍是不相关的，因此（3.2.9）中 ACF 的表达式仍是有效的。

3.3. 与 ARCH 模型的比较

GARCH（1,1）模型已被广泛运用于金融时间序列中。它假设（3.1.1）中的方差依赖于前一时间段的方差以及观测值的平方。这样

$$\sigma_t^2=\gamma+\alpha y_{t-1}^2+\beta\sigma_{t-1}^2 \quad , \qquad t=1,\ldots,T \qquad (3.3.1)$$

GARCH 模型是由 Bollerslev（1986）和 Taylor（1986）提出的，并且是 Engle（1982）提出的 ARCH 模型的推广。ARCH（1）模型是$\beta=0$时的特定 GARCH（1,1）模型。这样做的动机是来自预测的需要，在具有独立扰动的 AR（1）模型中，下一个观测值的最优预测是当前观测值的一小部分，并且在 ARCH（1）中它是当前观测值平方的一部分（加上一个常数）。原因是最优预测是在当前信息条件下做出的，在 ARCH 模型中下一时段的方差被假设为已知。一旦ε_t的分布被假定，该结构可直接导出模型的似然函数。这样，对σ_t^2所依赖的参数估计原则上是直接的。GARCH 模型引入的项类似于 ARMA 模型中的移动平均项，从而使得预测值成为过去观测值平方的分布滞后函数。

显然y_t是具有（无条件）方差$\gamma/(1-\alpha-\beta)$的鞅差分(martingale difference)。因此$\alpha+\beta<1$

* digamma 函数表示 gamma 函数的导数——译者注。

是协方差平稳的条件。正如 Bollerslev（1986）所述，高斯模型中使得四阶矩存在的条件是 $2\alpha^2+(\alpha+\beta)^2<1$，于是模型显示出超峰度，然而在现实中不可能总满足四阶矩条件。有趣的是，严平稳的条件反而弱得多，如 Nelson（1990）所指出的，甚至包括 $\alpha+\beta=1$ 的情形。

GARCH（1,1）的设定意味着

$$y_t^2=\gamma+\alpha y_{t-1}^2+\beta\sigma_{t-1}^2+\upsilon_t=\gamma+(\alpha+\beta)y_{t-1}^2+\upsilon_t-\beta\upsilon_{t-1}$$

其中 $\upsilon_t=y_t^2-\sigma_t^2$ 是一个鞅差分。从而，y_t^2 具有 ARMA（1,1）过程的形式，它的 ACF 可以用同样的方法计算数值。对应于 ARMA 模型的 ACF 反映了实际应用中在 y_t^2 的相关图里可能观察到的模式类型。

通过增加 σ_t^2 和 y_t^2 的更多滞后项可以扩展 GARCH 模型，而 GARCH（1,1）看起来是用得最广泛的。它显示了和 SV 模型相似的特征，特别是 ϕ 接近 1，这一点可以从具有 ARMA（1,1）过程模式的（3.2.6）清楚地看出。显然，ϕ 发挥了类似于 $\alpha+\beta$ 的作用。ACF 的主要不同似乎在滞后一期时最明显。Jacquier 等（1994，p.373）给出了纽约股票交易所中一个投资组合的周收益率平方的相关图，同时还给出拟合的 SV 模型和 GARCH （1,1）模型所隐含的 ACF。在这种情形下，SV 模型隐含的 ACF 接近于样本值。

因为 y_t 是混合分布，即使 ϕ 为零，SV 模型仍显示出超峰度。参数 σ_η^2 决定的混合程度独立于方差演化的平滑程度。GARCH 模型则不是这样，它的峰度是和方差方程的根相联系的，这些根在 GARCH（1,1）情形下是 α 和 β。因此，常常需要用非高斯的 GARCH 模型来捕捉金融时间序列中常见的高峰度。

基本的 GARCH 模型并不能反映 SV 模型在扰动同期相关下捕捉的不对称性，虽然 Engle 和 Ng（1993）提出它是能够修改的。Nelson（1991）提出的 EGARCH 模型通过使 $\log\sigma_t^2$ 成为过去观测值的平方以及绝对值的函数来处理不对称性问题。

3.4. 过滤，平滑和预测

为了期权定价的目的，需要估计和预测方差 σ_t^2，它当然是和 h_t 的指数成比例的。基于直到时间 t 的——并且可能包括时间 t 的——所有观测值的估计被称为*过滤*估计。另一方面，基于样本中所有观测值——包括时间 t 后的观测值——的估计被称为*平滑*估计。*预测*是估计未来值。出于对该方面演变历史的兴趣，我们可能希望通过考察平滑估计来研究方差随时间的演化。这些可以跟相应的期权价格所隐含的波动率做比较，如 2.1.2 节中讨论的那样。为了定价“两平”期权，可以只使用在样本尾端的过滤估计和对方差未来值的预测，这是由 Noh、Engle 和 Kane（1994）提出来的关于 ARCH 模型的方法。更一般地，价格的确定有必要基于方差未来值的完整分布。而这种分布可通过模拟技术得到。进一步的讨论见第 4.2 节。

人们可以考虑用一个简单但主观的方法来构造过滤和平滑估计，即采用变换后观测值的移动平均函数，这样：

$$\hat{\sigma}_t^2=g\left\{\sum_{j=t-1}^{r}w_{tj}f(y_{t-j})\right\}\quad,\quad t=1,\ldots,T \tag{3.4.1}$$

其中,对于过滤估计 $r=0$ 或 1，对于平滑估计 $r=t-T$。

既然我们已经构建一个随机波动率模型，自然就是用其作为过滤、平滑和预测的基础。对一个线性的高斯时间序列模型，状态空间形式能够用来作为最优过滤和平滑算法的基础。可惜的是，SV 模型是非线性的。这给我们留下三种可能性：

a. 基于线性状态空间模型的非有效估计；
b. 运用计算机的强化技术估计最优过滤直到一个满意的精确度；
c. 运用一个（非特定的）ARCH 模型来近似最优过滤。

下面分别对它们进行详细考察。

3.4.1. 线性状态空间形式

根据 Nelson（1988）和 Harvey、Ruiz 和 Shephard（1994），变换的观测值$\log y_t^2$能用于构造线性状态空间模型。度量(measurement)方程是（3.2.8），而（3.1.3）是过渡(transition)方程。状态h_t的初始条件是由它的无条件均值和方差给出的，分别是 0 和$\sigma_\eta^2/(1-\phi^2)$。

虽然假设η_t服从正态分布是合理的，但只有ε_t的绝对值服从对数正态分布，ξ_t才是正态分布的。而这是不太可能的。因此，运用 Kalman 滤子和相关联的平滑器产生状态h_t的估计量，它仅在基于$\log y_t^2$的线性组合的一类估计量中最优。而且，需要的不是h_t而是它们的指数。假定$h_{t|T}$代表基于线性空间的平滑估计量，那么$\exp(h_{t|T})$具有（3.4.1）与常数尺度σ^2相乘的形式。它可以被写成一个加权几何平均值。这就使得这些估计量对很小的观测值很敏感，同时也显示了该方法的局限性。

进行对数变换产生了一个重要的实际问题，即怎样处理零观测值。这是前述要点的一个反映，因为很显然包含零观测值的任何加权几何平均值都是零。更一般地，我们希望避免很小的观测值，一个可能的解决方法是去掉样本均值。另一个是由 Fuller 提出并由 Breidt 和 Carriquiry （1995）发展的更令人更满意的方法，基于 Taylor 级数展开式做如下变换：

$$\log y_t^2 \cong \log(y_t^2+cs_y^2)-cs_y^2\big/(y_t^2+cs_y^2)\ ,\quad t=1,\dots,T \tag{3.4.2}$$

其中s_y^2是y_t的样本方差，c是一个很小的数，其建议值为 0.02。这个变换的作用是通过切掉长尾来减少转变观测值的峰度，该长尾由通过取“内层”对数所得到的负值组成。换句话说，它是一种修剪（trimming）的形式。通过除以$\hat{\sigma}_t^2$的初始估计量来纠正观测值的异方差，然后再进行前面的程序，结果可能更令人满意。这样，$\log\hat{\sigma}_t^2$被加到了转变观测值中。$\hat{\sigma}_t^2$可以从第一轮中构造出来，或用完全不同的方法——也许是非参数方法——构造出来。

可对线性状态空间形式进行修改，以便于处理非对称模型。早就有人注意到，即使η_t和ε_t不是相互独立的，只要η_t和ε_t的联合分布是对称的，则状态空间形式中的扰动是不相关的。因此，前面的过滤和平滑操作仍是有效的，但会因为对观测值求平方而丢失一部分信息。Harvey 和 Shephard（1993）指出，通过给观测值的符号附加条件可以恢复该（丢失）信息，观测值的符号由s_t表示，当y_t是正的（负的）时s_t取值$+1$（-1）。当然，这些符号和ε_t的符号是一样的。令E_+（E_-）表示在ε_t是正的（负的）条件下的期望值，并对方差和协方差算子给予一个类似的说明。对ε_t的符号附加条件不影响ξ_t的分布，但是要记住$E(\eta_t|\varepsilon_t)$是ε_t的一个奇函数，

$$\mu^* = E_+(\eta_t)=E_+[E\eta_t|\varepsilon_t]=-E_-(\eta_t)\ ,$$

并且

$$\begin{aligned}\gamma^* &= Cov_+(\eta_t,\xi_t)=E_+(\eta_t\xi_t)-E_+(\eta_t)E(\xi_t)=E_+(\eta_t\xi_t)\\ &=-Cov_-(\eta_t,\varepsilon_t)\end{aligned}$$

因为ξ_t的期望值是零，并且

$$E_+(\eta_t\xi_t)=E_+[E(\eta_t|\varepsilon_t)\log\varepsilon_t]-\mu^*E(\log\varepsilon_t)=-E_-(\eta_t\xi_t)\ 。$$

最后

$$Var_+\eta_t=E_+(\eta_t^2)-[E_+(\eta_t)]^2=\sigma_\eta^2-\mu^{*2}\ 。$$

现在，线性状态空间形式是

$$\begin{aligned}&\log y_t^2=\omega+h_t+\xi_t\\ &h_{t+1}=\phi h_t+s_t\mu^*+\eta_t^*\ ,\\ &\begin{pmatrix}\xi_t\\ \eta_t^*\end{pmatrix}\Big|s_t\sim ID\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}\sigma_\xi^2 & \gamma^*s_t\\ \gamma^*s_t & \sigma_\eta^2-\mu^{*2}\end{pmatrix}\right)\ 。\end{aligned} \tag{3.4.3}$$

通过使h_0具有零均值和方差$\sigma_\eta^2 \big/ (1-\phi^2)$，Kalman 滤子仍然可以初始化。

（3.4.3）的参数化并不直接包括代表ε_t和η_t之间相关性的参数。μ^*和γ^*间的关系和模型中的初始参数只能通过假设关于ε_t及η_t分布来得到。当ε_t和η_t服从具有$Corr(\varepsilon_t,\eta_t)=\rho$，$E(\eta_t|\varepsilon_t)=\rho\sigma_\eta\varepsilon_t$的二元正态分布时，就有

$$\mu^* = E_+(\eta_t) = \rho\sigma_\eta E_+(\varepsilon_t) = \rho\sigma_\eta\sqrt{\pi/2} = 0.7979\rho\sigma_\eta \quad , \tag{3.4.4}$$

而且

$$\gamma^* = \rho\sigma_\eta E(|\varepsilon_t|\log\varepsilon_t^2) - 0.7979\rho\sigma_\eta E(\log\varepsilon_t^2) = 1.1061\rho\sigma_\eta \quad 。\tag{3.4.5}$$

当ε_t为$t-$分布时，它可以写成$\zeta_t\kappa_t^{-0.5}$，并且ζ_t和η_t能够被看作具有相关系数ρ的二元正态分布，而κ_t与两者都独立。μ^*和γ^*的数值计算与前面相同，除了初始条件是关于ζ_t而不是关于ε_t的，且必要的表达式和高斯情形下（的表达式）完全一样。

对数波动率h_t的过滤估计——记作$h_{t+1|t}$——具有形式：

$$h_{t+1|t} = \phi h_{t|t-1} + \frac{\phi(p_{t|t-1}+\gamma^* s_t)}{p_{t|t-1}+2\gamma^* s_t+\sigma_\xi^2}(\log y_t^2 - \omega - h_{t|t-1}) + s_t\mu^* \quad ,$$

其中$p_{t|t-1}$是$h_{t|t-1}$对应的均方误差(mean square error)。如果$\rho<0$，那么$\gamma^*<0$，且过滤估计将表现出和 Nelson（1991）估计的 EGARCH 模型一样的情形，即负观测值比相应的正观测值使被估计的对数波动率增加得更多。

3.4.2. 非线性滤子

原则上，可以为原始方程（3.1.2）和（3.1.3）写一个准确的滤子，且前者作为度量方程。计算这样一个滤子需要用数值方法近似一系列积分。Kitagawa（1987）提出了得到该滤子的一般方法，并由 Watanabe（1993）把它运用到 SV 模型中。可惜的是，它看起来非常耗时，在当前的计算机技术下不太适用。

通过将其作为整体模型的贝叶斯处理的一部分，Jacquier、Polson 和 Rossi（1994）说明了通过模拟得到波动率的平滑估计的可行性。需要的是以观测值为条件的波动率联合分布的均值向量。然而，因为模拟这样的联合分布在现实中是不可行的，他们把它分解成一系列单变量分布，每一个波动率都以所有其它的波动率为条件。这些分布可以记作$p(\sigma_t|\sigma_{-t},y)$，其中$\sigma_{-t}$表示除了$\sigma_t$之外的所有波动率，人们要做的是从这些分布中依次抽样，使得σ_{-t}集合中的元素等于它们最后的估计值，并且重复几千次。这就是 Gibbs 抽样法。然而存在一些困难。SV 模型的马尔可夫结构可以拓展成

$$p(\sigma_t|\sigma_{-t},y) = p(\sigma_t|\sigma_{t-1},\sigma_{t+1},y_t) \propto p(y_t|h_t)p(h_t|h_{t-1})(h_{t+1}|h_t)$$

虽然上述表达式的右边可明确写出，但密度却不是标准形式，并且没有正规化常数的解析式。Jacquier、Polson 和 Rossi 采用的解决方法是使用一连串的 Metropolis 接受/拒绝的独立链。

Kimyu 和 Shephard（1994）指出，如果ϕ接近于 1，并且/或者σ_η^2很小的话，因为σ_t变化很慢，Jacquier、Polson 和 Rossi 采用的单一移动算法（single mover）将会变慢。事实上，当它是一个常数时，运算将根本不收敛。基于线性状态空间形式的另一个方法是通过正态分布的混合获得度量方程中的非正态扰动项ξ_t。Watanabe（1993）提出了基于两矩的混合的近似方法。Kim 和 Shephard（1994）提出了基于线性状态空间形式的一个多移动抽样法（multimove sampler）。成批的h_t被抽取，而不是一次取一个。他们所采用的技术是混合适当数目的正态分布来达到（3.2.7）中所要求的近似扰动的精确程度。Mahieu 和 Schotman（1994a）通过在正态分布的混合中引进更多的自由度扩展了这种方法，其中的参数是估计的，而不是事先确定的。值得注意的是σ_t的分布能够从h_t的模拟分布得到。

Jacquier、Polson 和 Rossi（1994，p.416）指出，在 Kim 和 Shephard 的方法中，不论用到多

少混合成份，$\log\varepsilon_t^2$ 的尾部状态不可能被满意地近似。事实上，他们注意到给定 Kim 和 Shephard 状态空间的不连续性，在小数目的抽取中并不是所有的状态都能涉及到，也就是所谓的内层问题（也参见 3.4.1 节和 Nelson（1994））仍是存在的。

最后要注意，当超参数（hyperparameters）未知时，由贝叶斯方法生成的状态的模拟分布考虑到了它们的抽样变异性。

3.4.3. 作为近似滤子的 ARCH 模型

这里的目的是把注意力引向一个将在 4.3 节中更详细讨论的主题。在 ARCH 模型中，条件方差被假设成过去观测值的一个精确函数。Nelson 和 Foster（1994，p.32）指出，这个假设是经济学和统计学*特有*的。然而，由于 ARCH 模型相对而言易于估计，Nelson（1992）以及 Nelson 和 Foster（1994）主张的一个有用的策略，是把它们视为产生条件方差估计值的滤子，来估计条件方差。这样，即使我们相信有一个连续时间或者离散时间的 SV 模型，仍可以估计一个 GARCH（1,1）模型，并且把 σ_t^2 作为一个近似的滤子，如（3.4.1）中那样。于是，估计值就是过去观测值平方的加权平均，它给出了以时间 $t-1$ 的观测值为条件的 σ_t^2 分布的均值估计。此外，Taylor（1986）和 Schwert（1989）提出的模型也可以运用，该模型中的条件标准差构造成先前条件标准差和先前绝对值的线性组合，这对异常值来说更稳健，因为它是过去绝对值的线性组合。

Nelson 和 Foster 推导了一个 ARCH 模型，它给出*连续时间* SV 公式的最接近的近似（详见 4.3 节）。这并不对应某个标准模型，虽然它相当接近 EGARCH。对于*离散时间* SV 模型，过滤理论没有得到这样广阔的发展，事实上，Nelson 和 Foster 指出，从随机微分方程到*差分*方程，在极限定理和优化理论中都有相当大的差异。他们研究近似扩散的情形，作为例子来说明这些差异。

3.5. 模型的扩展

3.5.1. 持续性和周期性

最简单的非平稳 SV 模型是 h_t 服从随机游走。如果用观测值的对数变换 $\log y_t^2$ 来做研究，很容易得到这个模型的动态属性。所要做的全部工作只是求一阶差分以得到一个平稳过程。未变换的观测值是非平稳的，但是模型的动态结构将出现在 $[y_t/y_{t-1}]^c$ 的 ACF 中，只要 $c<0.5$。

该模型的可替代选择是 IGARCH，即 $\alpha+\beta=1$ 的（3.3.1）。IGARCH 模型使观测值的平方具有单整 ARMA（integrated ARMA）过程的一些特征，并且被认为呈现持续性；参见 Bollerslev 和 Engle（1993）。然而，它的性质并不是简单明了的。例如，它必须包含一个常数 γ，否则如 Nelson（1990）所指出的，σ_t^2 几乎必定收敛于零，且模型具有独特性质，即它是严平稳的而不是弱平稳的。另一方面，非平稳的 SV 模型能够在 h_t 是一阶标准积分过程的基础上进行分析。

因为 $\log y_t^2$ 只是一个随机游走加上噪声，过滤和平滑能够在线性状态空间框架中执行。初始条件用与非平稳结构时间序列模型相同的方法处理，即从第一个观测值开始，有效形成的状态适当优先，参见 Harvey（1989）。作为过去的 $\log y_t^2$ 线性组合的一类估计，h_t 的最优过滤估计 $h_{t|t-1}$，是一个常数加上过去 $\log y_t^2$ 的一个等权重移动平均（EWMA）。在 IGARCH 中，σ_t^2 由一个常数加上过去观测值平方的 EWMA 精确给出。

随机游走波动率能够被其它非平稳设定所替代。一种可能是双重积分随机游走，其中 $\Delta^2 h_t$ 是白噪声。在连续时间建模中，该模型等价于三次样条函数，应用于水平模型时，显示出相对平滑的趋势。在 SV 的框架下，如果目标是寻找满足平滑演化方差的加权函数，这将是很吸引人的。然而，对预测来说，它可能更不稳定。

其它非平稳成分可以很容易地引入 h_t。例如，周期性的或日间的成分能够包含在内；设定与 Harvey（1989）以及 Harvey 和 Koopman（1993）讨论的相应的水平模型完全一样。动态属性是通过运用于 $\log y_t^2$ 的一般变换直接给出的，且适当地变换绝对值并不困难。这样，如果波动率是由一个随机游走加上一个如 Harvey（1989，p.40－3）中的非平稳的周期性缓慢变化构成，那么

适当的变换则是$\Delta_s \log y_t^2$和$\left|y_t / y_{t-s}\right|^c$，其中$s$是周期数。状态空间的公式遵循相应的水平结构时间序列模型的方法。在 GARCH 框架中处理这样的效应并不容易。

通过运用将在稍后的小节中讨论的时间形变的思想，处理周期性的不同方法也可引入 SV 模型中。当处理周期性中的急剧变化时，这些方法可能特别重要，这些急剧变化似乎发生在高频数据中，如五分钟或者逐笔（tick－by－tick）的外汇交易数据。

3.5.2. *干预和其它确定性效应*

干预变量很容易引入 SV 模型。例如，波动率过程的一个突然的结构性变化能够通过如下假设来反映

$$\log \sigma_t^2 = \log \sigma^2 + h_t + \lambda w_t$$

其中w_t在突变前是 0，在突变后是 1，λ是一个未知参数。对数变换给出了（3.2.8），只是λw_t加到了右边。将该效应引入 ARCH 模型时应谨慎。例如，在 GARCH（1,1）中，反映意外的突变必须建立模型

$$\sigma_t^2 = \gamma + \lambda w_t - (\alpha + \beta)\lambda w_{t-1} + \alpha y_{y-1}^2 + \beta \sigma_{t-1}^2$$

对λ的约束使σ_t^2总是正的。

与干预虚拟变量相反，更一般的可观测解释变量可经方差引入模型。

3.5.3. *多变量模型*

对应于（3.1.2）的多变量模型，假设每一个序列都是由如下形式的模型生成的

$$y_{it} = \sigma_i \varepsilon_{it} e^{0.5 h_{it}} \quad , \qquad t = 1, \ldots, T \quad , \tag{3.5.1}$$

向量$\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$的协方差（相关）矩阵用$\Sigma_\varepsilon$表示。波动率向量$h_t$遵从 VAR（1）过程，也就是

$$h_{t+1} = \Phi h_t + \eta_t \quad ,$$

其中$\eta_t \sim IID(0, \Sigma_\eta)$。该设定允许波动率的变动通过$\Sigma_\eta$在不同序列间相关，相互作用通过$\Phi$的非对角元素体现。

观测值平方的对数变换导出了一个多变量线性状态空间模型，通过它可以计算波动率的估计值，如 3.4.1 节中那样。

一个简单的非平稳模型可通过假设波动率服从多变量随机游走得到，即$\Phi = I$。如果Σ_η是奇异的，具有秩$K < N$，在波动率中只有K个分量，也就是（3.5.1）中的每一个h_{it}都是$K < N$个共同趋势的线性组合，即

$$h_t = \Theta h_t^{\perp} + \bar{h} \tag{3.5.2}$$

其中$h_t^{\perp}$是共同随机游走波动率的$K \times 1$向量，$\bar{h}$是常数向量，Θ是因子载荷的$N \times K$矩阵。需要在Θ和$\bar{h}$上施加特定约束以确保可识别性，参见 Harvey、Ruiz 和 Shephard（1994）。在 Engle 和 Granger（1987）中，观测值平方的对数是“协整的”，这是由于有$N - K$个线性组合，这些线性组合是白噪声，因而是平稳的。这就意味着，例如，如果收益的两个序列表现出随机波动率特征，而当$\Theta' = (1,1)$时，波动率是相同的，那么序列的比率将不存在随机波动率特征。“共同持续性”相关概念的应用能够在 Bollerslev 和 Engle（1993）中找到。然而，如同在单变量情形中一样，对于是什么真正构成持久性仍存在疑问。

对于为何波动率中共同成分的思想不应扩展到平稳模型中，并无解释。公式（3.5.2）的应用不需要$\bar{h}$，而需要对$h_t^{\perp}$建模，例如通过 VAR（1）。

Bollerslev、Engle 和 Wooldridge（1988）指出，原则上一个多变量 GARCH 模型能够进行极大似然估计，但是由于涉及到大量的参数，除非施加约束，否则常会遇上计算问题。多变量 SV 模型比多变量 GARCH 的一般公式要简单得多。然而，它存在局限性，因为没有对变化的协方差建模。在这一意义上，它类似于 Bollerslev（1986）的有约束的多变量 GARCH 模型，在该模型中条件相关系数被假设为常数。

Harvey、Ruiz 和 Shephard（1994）对四种汇率运用非平稳模型，发现只有两个共同因子主导波动率。另一个运用在 Mahieu 和 Schotman（1994b）中。对汇率波动率建模的另一种完全不同的方法是 Diebold 和 Nerlove（1989）建立的潜在因子（latent factor）ARCH 模型。

3.5.4. 观测间隔、归并和时间形变

假设对一个 SV 模型每隔 δ 时段进行一次观测。在这种情形下，h_τ 仍是 AR（1），但参数变为 ϕ^δ，其中 τ 表示新的观测（抽样）间隔。扰动 η_t 的方差增大，但 σ_h^2 保持不变。SV 模型的这一性质使它很容易在不同抽样间隔中做比较；例如，为什么每天观测时 ϕ 大约是 0.98，而间隔一周（假定一周为 5 天）观测时 ϕ 值大约为 0.9，SV 模型的性质使这一原因变得很清楚。

如果观测平均很长一段时间才进行一次，比较就复杂得多。因为 h_τ 此时将服从一个 ARMA（1,1）过程。然而，AR 参数仍为 ϕ^δ。值得注意的是，要改变 ARCH 过程的观测间隔是很困难的，除非如 Drost 和 Nijman（1993）所述，结构将弱化，参见 4.4.1 节。

如同在 2.4 节中指出的，通常运用离散时间模型近似替代连续时间模型，这很容易通过运用如 Harvey （1989）所描述的线性状态空间模型来处理不规则时间间隔的观测值。事实上，最初由 Clark（1973）提出的以描述资产价格及其波动率的从属过程为基础的方法非常适合该框架。Stock（1988）指出，处理不规则时间间隔观测值的技术能够作为处理时间形变观测值的基础。Ghyseis 和 Jasiad（1994a,b）提出了一个 SV 模型，其中连续时间波动率方程的运作时间是由信息流决定的。这样的时间形变过程可能格外适合处理高频率数据。如果 $\tau = g(t)$ 是日历时间 τ 和运作时间 t 之间的映射，那么

$$dS_t = \mu S_t dt + \sigma(g(t)) S_t dW_{1t}$$

并且

$$d\log\sigma(\tau) = a(b - \log\sigma(\tau))d\tau + c dW_{2\tau}$$

其中 W_{1t} 和 $W_{2\tau}$ 是标准的独立 Wiener 过程。离散时间近似产生了（3.1.3），但它包含这样一项，该项在（3.1.2）中被包含在常数尺度因子 σ 中，因而有

$$h_{t+1} = [1 - e^{-a\Delta g(t)}]b + e^{-a\Delta g(t)} h_t + \eta_t$$

其中 $\Delta g(t)$ 是两个连续日历时间之间的运作时间变化，η_t 是具有均值为零和方差为 $c^2(1 - e^{-2a\Delta g(t)})/2a$ 的正态分布。明显地，如果 $\Delta g(t) = 1$，那么（3.1.3）中 $\phi = e^{-a}$。由于信息流，从而 $\Delta g(t)$，是不可观测的，必须详细设定日历时间到运作时间的映射以使模型可操作。Ghysels 和 Jasiak（1994a）讨论了一些设定，这些设定围绕比例指数函数考虑使 $g(t)$ 和可观测变量相关联，例如过去交易量和具有不对称杠杆效应的过去价格变化。该方法也被 Ghysels 和 Jasiak（1994b）用于模拟收益率－成交量的联合运动，以及被 Ghysels、Gourieroux 和 Jasiad（1995b）用于模拟日内高频数据，该数据具有很强的（3.5.1 节的）周期模式。

3.5.5. 长期记忆

Baillie、Bollerslev 和 Mikkelsen（1993）提出了扩展 GARCH 类模型的一种方式以解释长期记忆。他们称该模型为分数单整（Fractionally Integrated）GARCH（FIGARCH），其主要特征是在条件方差等式中的过去观测值平方的滞后结构中，加入分数差分(Fractional difference)算子 $(1-L)^d$，其中 L 是滞后算子。然而，当 $d = 0$ 时，这个模型只能是平稳的，并还原为 GARCH。在随后的一篇论文中，Bollerslev 和 Mikkelsen（1995）讨论了 Nelson（1991）EGARCH 模型的一个推广，其中 $\log\sigma_t^2$ 对包含分数差分算子的过去 ε_t 的分布滞后建模。如果 $|d| < 0.5$，该 FIEGARCH 模型是平稳和可逆的。

Breidt、Crato 和 de lima（1993）以及 Harvey（1993）提出了一个 SV 模型，其中 h_t 由分数噪声生成

$$h_t = \eta_t/(1-L)^d \quad , \quad \eta_t \sim NID(0, \sigma_\eta^2) \quad , \quad 0 \le d \le 1 \quad 。 \qquad (3.5.1)$$

如同（3.1.3）中的 AR（1）模型，这个过程还原为白噪声和一个在参数空间边界上的随机游

走，即分别为$d=0$和1。然而，只有当$d<0.5$时，它才是平稳的，因而，对 AR（1）模型来说，从平稳性到非平稳性的转变具有不同的途径。与 AR（1）的情形一样，限定（3.5.1）中的自相关为正是合理的。然而，一个负的d值却是相当合理的，实际上，当h_t是非平稳的时候，对其差分就可得到一个平稳的“中等记忆”过程，其中$-0.5\le d\le 0$。

具有长期记忆的 SV 模型的性质能够从 3.2 小节的公式中得到。在 Harvey（1993）中可以找到，当h_t服从$d=0.45$和$\sigma_h^2=2$的长期记忆过程时，以及当h_t是具有$\phi=0.99$的 AR（1）时，相应的 ACF 之间的比较。回顾长期记忆的一个特有属性是自相关的双曲线比率衰减，而非数据中观测到的（见 2.2e 节）指数比率衰减。长期记忆模型中，衰减缓慢是很明显的。事实上，对于$\tau=1000$，长期记忆的自相关系数仍然是0.14，而在 AR 情形中它只是0.000013。长期记忆的形状和 Ding、Granger 和 Engle（1993，p.86-8）中的形状非常相似。

通过令η_t为一个 ARMA 过程，并且/或者通过给波动率方程加上更多成分，模型可以扩展。

关于平滑和过滤，我们已经指出，由于涉及截断（truncation），状态空间方法是近似的，而且由于状态向量的长度，这个方法是比较麻烦的。在$\log y_t^2$的线性估计量中，平滑和过滤是最优的，如果已准备好构造并转置$\log y_t^2$的$T\times T$协方差矩阵，那么可以用直接的方法进行精确的平滑和过滤。

4. 连续时间模型

第 2 节的末尾给出了离散时间 SV 的统计建模框架，并且把整个第 3 节用于讨论离散时间 SV 模型的设定。为了讨论连续时间模型，首先研究微分方程和离散时间 SV 模型之间的所有精确关系（也没有近似误差）。在 4.1 节中通过一类在时频归并下是闭合的统计模型，进行（1）从高频的离散时间到较低频的离散时间，（2）从连续时间到离散时间的讨论。接下来，在 4.2 节中，用连续时间模型讨论期权定价和套期保值，并且详细阐述诸如微笑效应这样的特征。实际实行 SV 期权定价公式时，经常要求把离散时间的 SV 和/或者 ARCH 模型作为连续时间波动率过程的滤子和预测器。将在 4.3 节中涉及的滤子通常是连续时间 SV 模型的离散时间近似（并不是 4.1 节中那样的精确离散化）。4.4 节以基础模型的扩展结束。

4.1. 从离散到连续时间

这一节的目的是提供离散和连续时间 SV 模型之间关系的严格讨论。首先从对 SARV 类模型背景下时频归并的讨论开始，并着重讨论包括 GARCH 模型的特殊情形。这些内容涵盖了 4.1.1 节。接下来把注意力转向各类连续时间 SV 模型，以得出离散时间的表达式。这是 4.1.2 节的主题内容。

4.1.1. 离散时间模型的时频归并

2.4 节讨论 Andersen 的 SARV 类模型，以作为一般的离散时间参数 SV 统计模型。这里考虑零均值的情形，即：

$$y_{t+1}=\sigma_t\varepsilon_{t+1} \qquad (4.1.1)$$

当$q=1$或者2时，σ_t^q是马尔可夫过程K_t的一个多项式函数$g(K_t)$，K_t具有平稳的自回归表达式

$$K_t=\omega+\beta K_{t-1}+\upsilon_t \qquad (4.1.2)$$

其中$|\beta|<1$，并且

$$E[\varepsilon_{t+1}|\varepsilon_\tau,\upsilon \quad \tau\le t]=0 \qquad (4.1.3a)$$

$$E[\varepsilon_{t+1}^2|\varepsilon_\tau,\upsilon \quad \tau\le t]=1 \qquad (4.1.3b)$$

$$E[\upsilon_{t+1}|\varepsilon_\tau,\upsilon \quad \tau\le t]=0 \qquad (4.1.3c)$$

约束（4.1.3a-c）隐含υ是关于过滤$\Re_t=\sigma[\varepsilon_\tau,\upsilon_t,\tau\le t]$的鞅差分序列。[22]而且，（4.1.3a-c）中的条件矩条件也说明（4.1.1）中的ε在半强意义下是一个白噪声过程，也即$E[\varepsilon_{t+1}|\varepsilon_\tau,\tau\le t]=0$和$E[\varepsilon_{t+1}^2|\varepsilon_\tau,\tau\le t]=1$，并且不是由$\upsilon$ Granger 引起（Granger－caused）。[23]第 2 节一开始，就选择了在一个特定时间范围内的连续复利收益率作为讨论连续时间过程的起点。因而，令（4.1.1）中的y_{t+1}为资产价格过程S_t在$[t,t+1]$的连续复利收益率，因此：

$$y_{t+1}=\log S_{t+1}/S_t \tag{4.1.4}$$

由于抽样间隔的时间单位在很大程度上是主观的，因此要求由方程（4.1.1）到（4.1.3）所定义的 SV 模型（对于给定的q和函数g）在时频归并下是闭合的。由于收益率是流量变量，在时频归并下闭合意味着对任何整数m：

$$y_{tm}^{(m)}\equiv\log S_{tm}\Big/S_{tm-m}=\sum_{k=0}^{m-1}y_{tm-k}$$

它也适合（4.1.1）到（4.1.3）类型的模型，对于所包含的适当参数值，q和g的选择相同。这一节中的分析沿用 Meddahi 和 Renault(1995)，他们详细研究了 SV 模型的时频归并，尤其是对$\sigma_t^2=K_t$的情形，也即$q=2$并且g是恒等函数。它与 Drost 和 Werker(1994)的所谓的连续时间 GARCH 方法有关，因此，由（4.1.1）我们有：

$$\sigma_t^2=\omega+\beta\sigma_{t-1}^2+\upsilon_t \tag{4.1.5}$$

该模型具有条件矩约束（4.1.3a-c），在时频归并下是闭合的，例如，对于$m=2$：

$$y_{t+1}^{(2)}=y_{t+1}+y_t=\sigma_{t-1}^{(2)}\varepsilon_{t+1}^{(2)}$$

具有：

$$\left(\sigma_{t-1}^{(2)}\right)^2=w^{(2)}+\beta^{(2)}\left(\sigma_{t-3}^{(2)}\right)^2+\upsilon_{t-1}^{(2)}$$

其中：

$$\begin{cases}w^{(2)}=2\omega(1+\beta)\\ \beta(2)=\beta^2\\ \upsilon_{t-1}^{(2)}=(\beta+1)[\beta\upsilon_{t-2}+\upsilon_{t-1}]\end{cases}$$

而且，无论何时杠杆效应出现在归并水平中都是值得注意的，即：

$$Cov[v_{t-1}^{(2)},\varepsilon_{t-1}^{(2)}]\neq0$$

具有$\varepsilon_{t-1}^{(2)}=(y_{t-1}+y_{t-2})/\sigma_{t-3}^{(2)}$，它必须出现在非归并(disaggregate)水平中，即$Cov(\sigma_t,\varepsilon_t)\neq0$。

对于一般情形，Meddahi 和 Renault(1995)指出，模型（4.1.5）和条件矩约束（4.1.3a-c）合在一起就是在归并条件下闭合的一类过程。给定这一结果，同 Drost 和 Nijman(1993)关于 GARCH 时频归并的研究作比较是很有趣的。建立 Meddahi 和 Renault(1995)与 Drost 和 Nijman(1993)之间的联系时，也揭示了 GARCH 模型中杠杆性质的争论点。事实上，同人们通常相信的相反，发现了 GARCH 过程中的杠杆效应约束。而且，根据 Meddahi 和 Renault 的结果还发现弱 GARCH 过程包含某些 SV 模型。

为了找到在归并条件下闭合的一类 GARCH 过程，Drost 和 Nijman(1993)弱化了 GARCH 的定义，即对于一个正的平稳过程σ_t：

$$\sigma_t^2=w+ay_{t-1}^2+b\sigma_{t-1}^2 \tag{4.1.6}$$

其中$a+b<1$，他们定义：

—强 GARCH，如果y_{t+1}/σ_t是具有均值 0 和方差为 1 的独立同分布.。

—半强 GARCH，如果$E[y_{t+1}|y_\tau,\tau\le t]=0$，并且$E[y_{t+1}^2|y_\tau,\tau\le t]=\sigma_t^2$

[22]注意，这里不用（2.4.9）中出现的分解，即$\upsilon_t=[\gamma+aK_{t-1}]\widetilde{u}_t$。

[23]这里考虑的ε_t的 Granger 非因果性比假设 2.3.2A 要弱，因为它只运用了前两阶条件矩。

—弱 GARCH，如果 $EL[y_{t+1}^2 | y_\tau, y_\tau^2, \tau \le t] = \sigma_t^2$。[24]

Drost 和 Nijman 阐述了弱 GARCH 过程时频归并，而且给出了其系数的明确公式。在 2.4 节中提到，当波动率过程不存在特定随机性时，SARV 框架包含 GARCH 过程。这一性质将允许我们证明弱 GARCH 过程——如上面定义的——事实上包含更一般的、严格说来并不是 GARCH 的 SV 过程。按照 Meddahi 和 Renault(1995)的观点，要求根据 υ_t 和 y_t^2 之间的相关系数来对（4.1.3）和（4.1.5）所定义的模型进行分类，即：

（a） *具有完全相关性的模型*：第一种类型——以后就表示为 C_1——表现为在条件 $(\varepsilon_\tau, \upsilon_\tau, \tau < t)$ 下 υ_t 和 y_t^2 之间是一个线性关系，对（4.1.5）中的模型而言它就是 1 或者 -1。

（b） *不具有完全相关性的模型*：第二种类型——以后就表示为 C_2 ——如上述的条件相关性具有小于 1 的绝对值。

类型 C_1 包括所有的半强 GARCH 过程，事实上只要 C_1 中 $Var[y_t^2 | \varepsilon_\tau, \upsilon_\tau, \tau < t]$ 和 $Var[\upsilon_t | \varepsilon_\tau, \upsilon_\tau, \tau < t]$ 成比例，就有半强的 GARCH。因此，一个半强 GARCH 过程就是模型（4.1.5）满足：（1）约束（4.1.3），（2）如同 C_1 的完全条件相关性，（3）对条件峰度动态特性的约束。[25]

现在考虑如下假设：

假设 4.1.1. 如下两个条件期望为零：

$$E[\varepsilon_t \upsilon_t | \varepsilon_\tau, \upsilon_\tau, \tau < t] = 0 \tag{4.1.7a}$$

$$E[\varepsilon_t^3 | \varepsilon_\tau, \upsilon_\tau, \tau < t] = 0 \tag{4.1.7b}$$

这个假设相当于无杠杆效应，后者在条件协方差意义下定义，用于反映 2.4.1 节中讨论的瞬时因果关系的概念，这里它是在弱白噪声背景下运用。[26]还需要注意的是（4.1.7a）和（4.1.7b）一般不相等，除非在类型 C_1 的过程中。

类型 C_2 考虑了由于不完全相关性而产生的波动率过程的随机性。然而，尽管存在波动率特定的随机性，还是可以指出在*假设 4.1.1* 下 C_2 过程满足弱 GARCH 定义。不容置疑的是，任何满足（4.1.3a-c）、(4.1.5)、(4.7.1a-b) 以及*假设 4.1.1* 的 SV 模型，是弱 GARCH 过程。实际上，对称假设（4.7.1a-b）或 GARCH 中对杠杆的约束，使得 $EL[y_{t+1}^2 | y_\tau, y_\tau^2, \tau \le t] = \sigma_t^2$（连同条件矩约束（4.1.3a-c）），并产生了 Drost 和 Nijman(1993，例 2，p.915) 在所谓的对称弱 GARCH(1,1)中发现的时频归并的内在一致性。因此，这类弱 GARCH(1,1)过程能够被当作满足（4.1.3）和（4.1.5）过程的子类型。[27]

4.1.2. 连续时间模型的时频归并

为了方便讨论，将一般连续时间模型（2.3.1）的研究集中于零漂移的过程，即：

$$d\log S_t = \sigma_t dW_t \tag{4.1.8a}$$

$$d\sigma_t = \gamma_t dt + \delta_t dW_t^\sigma \tag{4.1.8b}$$

$$Cov(dW_t, dW_t^\sigma) = \rho_t dt \tag{4.1.8c}$$

其中随机过程 σ_t，γ_t，δ_t 和 ρ_t 满足 $I_t^\sigma = [\sigma_\tau; \tau \le t]$。为了保证 σ_t 是一个非负的过程，通常遵

[24]对于任何 L^2 的 Hilbert 空间 H，$EL[x_t | z, z \in H]$ 是 x_t 关于 1 和 $z \in H$ 的最佳线性预测值。需要注意的是，一个强 GARCH 过程，不容置疑地是一个本身也是弱 GARCH 过程的半强过程。

[25]事实上，Nelson 和 Foster(1994)发现，最常用的 ARCH 模型实际上假定了方差的方差(the variance of the variance)线性地增加到 σ_t^4，这是 ARCH 模型在近似连续时间 SV 模型（也见 4.3 节）时的主要缺点。

[26]条件期望（4.1.7b）可看成 ε_t 和 ε_t^2 之间的条件协方差。就是这个协方差——如果非零的话——产生 GARCH 中的杠杆效应。

[27]正如前面指出的，满足（4.1.3）和（4.1.5）的一类过程在时频归并下是闭合的，它包含了不满足*假设 4.1.1* 的杠杆效应的过程。

循如下两个策略中的一个：（1）考虑$\log\sigma_t^2$的扩散，或者（2）将σ_t^2作为 CEV 过程（或者服从 Cox(1975)以及 Cox 和 Ross(1976)的固定弹性方差过程）。[28]前者经常在期权定价文献中遇到（例如，参见 Wiggins(1987)），它明显与提出 EGARCH 的 Nelson(1991)以及 Taylor(1996)的对数正态 SV 模型有关。第二种类型的 CEV 过程可写作

$$d\sigma_t^2 = k(\theta - \sigma_t^2)dt + \gamma(\sigma_t^2)^\delta dW_t^\sigma \tag{4.1.9}$$

其中$\delta \le 1/2$保证了σ_t^2是一个取非负值的平稳过程。方程（4.1.9）可看作是 2.4 节中提到的离散时间 SARV 模型类型的连续时间情形。这一观测结果建立起同前面 4.1.1 节中讨论内容的联系，并且产生了连续时间 SV 模型精确的离散化结果。因此，就像前一节一样，主要是根据弱 GARCH 过程时频归并同 GARCH 类模型比较，特别是 Drost 和 Werker(1994)提出的扩散模型。

首先，应该注意（4.1.9）中的 CEV 过程隐含一个σ_t^2的离散时间自回归模型，即：

$$\sigma_{t+\Delta t}^2 = \theta(1-e^{-k\Delta t}) + e^{-k\Delta t}\sigma_t^2 + e^{-k\Delta t}\int_t^{t+\Delta t} e^{k(u-t)}\gamma(\sigma_u^2)^\delta dW_u^\sigma \tag{4.1.10}$$

Meddahi 和 Renault(1995)指出，只要（4.1.9）和它的离散化（4.1.10）决定波动率，离散时间过程$\log S_{t+(k+1)\Delta t}/S_{t+k\Delta t}$，$k \in Z$就是满足模型约束（4.1.3a-c）和（4.1.5）的 SV 过程。因此，如前一节所述，从扩散（4.1.9）可得到离散时间 SV 类模型，它在时频归并条件下是闭合的。更具体些，考虑例如$\Delta t = 1$，那么由（4.1.10），有：

$$\begin{aligned} y_{t+1} &= \log S_{t+1}/S_t = \sigma_t^{(1)}\varepsilon_{t+1} \\ \left(\sigma_t^{(1)}\right)^2 &= w + \beta\left(\sigma_t^{(1)}\right)^2 + \upsilon_t \end{aligned} \tag{4.1.11}$$

其中由（4.1.10）：

$$\begin{cases} \beta = e^{-k}, w = \theta(1-e^{-k}), \\ \upsilon_{t+1} = \left(\dfrac{1-e^{-k}}{k}\right)e^{-k}\displaystyle\int_t^{t+1} e^{k(u-t)}\gamma(\sigma_u^2)^\delta dW_u^\sigma. \end{cases} \tag{4.1.12}$$

值得注意的是（4.1.12），在连续时间中没有杠杆效应，即（4.1.8c）中$\rho_t = 0$，这意味着在低频率下没有该效应，且*假设 4.1.1* 的两个对称条件得到满足。这一系列的合理性还解释了 Drost 和 Werker(1994)的时频归并结果，但更具一般性的是，能够解释具有杠杆的连续时间 SV 模型，如何*精确*离散化为具有杠杆效应的离散时间 SV 模型。

4.2. 期权定价和套期保值

4.2.1 节用于研究具有 SV 的基本期权定价模型，即第 2 节的 Hull 和 White 模型。现在有了更好的方法来详细阐述它的理论基础，在 4.2.2 节讨论了实际意义，而 4.2.3 以一些基础模型的扩展结束。

4.2.1. 基本期权定价公式

回顾一下到期期限为$t+h=T$的欧式期权合约公式（2.1.10）。如在 2.1.2 节中提到的，假设交易是连续的和无摩擦的。并且，标的资产和无风险债券的交易中不存在套利机会，利率是非随机的，以使得由（2.1.12）定义的$B(t,T)$表示到期期限为T的一单位贴现债券在时间t的价格。现在，考虑概率空间$(\Omega, \Re, P)$，它是标的资产价格过程S的基础空间：

$$\begin{aligned} dS_t/S_t &= \mu(t,S_t,U_t)dt + \sigma_t dW_t^S \\ \sigma_t^2 &= f(U_t) \\ dU_t &= a(t,U_t)dt + b(t,U_t)dW_t^\sigma \end{aligned} \tag{4.2.1}$$

[28]偶尔会遇到不保证σ_t过程非负性的设定，为了计算的方便，如有些作者考虑σ_t和σ_t^2的 Ornstein- Uhlenbeck 过程（例如，参见 Stein 和 Stein(1991)）。

其中$W_t=(W_t^S,W_t^\sigma)$是定义在$(\Omega,\Re,P)$上的标准二维布朗运动（W_t^S和W_t^σ是独立的，具有零均值，方差为 1）。函数f被称为波动率函数，假设为一一对应。在该框架下（在适当的正则性条件下），*无免费午餐*假设相当于$(\Omega,\Re)$上有一个概率分布Q存在，它等于P，在此假设下贴现价格过程是鞅（参见 Harrison 和 Kreps(1979)）。这样的概率称为等鞅测度(an equivalent martingale measure)，当且仅当市场完全时它是唯一的（见 Harrison 和 Pliska(1981)）。[29]根据鞅表达式的积分形式（见 Karatzas 和 Shreve(1988)），任何等于P的概率测度Q的（正）密度过程可写成：

$$M_t=\exp\left[-\int_0^t\lambda_u^S dW_u^S-\frac{1}{2}\int_0^t(\lambda_u^S)^2du\right.$$
$$\left.-\int_0^t\lambda_u^\sigma dW_u^\sigma-\frac{1}{2}\int_0^t(\lambda_u^\sigma)^2du\right] \quad (4.2.2)$$

其中过程λ^S和λ^σ适合于自然过滤$\sigma_t=\sigma[W_\tau,\tau\le t]$，$t\ge 0$，并且满足可积性条件（几乎必定）：

$$\int_0^t(\lambda_u^S)^2du<+\infty \text{ 和 } \int_0^t(\lambda_u^\sigma)^2du<+\infty$$

根据 Girsanov 定理，过程$\widetilde{W}=(\widetilde{W}^S,\widetilde{W}^\sigma)'$定义为：

$$\widetilde{W}_t^S=W_t^S+\int_0^t\lambda_u^S du \text{ 和 } \widetilde{W}_t^\sigma=W_t^\sigma+\int_0^t\lambda_u^\sigma du \quad (4.2.3)$$

它们是Q条件下的二维布朗运动，Q条件下的标的资产价格的动态特征从（4.2.1）和（4.2.3）直接得到。而且，贴现资产价格过程$S_tB(0,t)$，$0\le t\le T$，是一个Q-鞅，当且仅当对于（2.1.11）中定义的r_t有：

$$\lambda_t^S=\frac{\mu(t,S_t,U_t)-r_t}{\sigma_t} \text{ 。} \quad (4.2.4)$$

由于S是唯一的交易资产，过程λ^σ不是固定的。（4.2.4）定义的λ^S被称为资产风险溢价。依此类推，任何满足所需的可积性条件的过程λ^σ都能够视为一个波动率风险溢价，并且对于任何λ^σ，由（4.2.2）密度过程M定义的概率$Q(\lambda^\sigma)$是一个等鞅测度。因此，给定风险溢价过程λ^σ：

$$C_t^{\lambda^\sigma}=B(t,T)E_t^{Q(\lambda^\sigma)}[Max[0,S_T-K]] \quad , \quad 0\le t\le T \quad (4.2.5)$$

是欧式看涨期权的一个可行的价格过程。[30]

Hull 和 White 的期权定价模型依赖于下面的假设，它对等鞅测度集作出约束：

假设 4.2.1. 波动率风险溢价λ_t^σ只依赖于波动率过程的当前值：$\lambda_t^\sigma=\lambda^\sigma(t,U_t)$，$\forall t\in[0,T]$。

该假设和跨期均衡模型是一致的，在该模型中行为者的偏好由时间可分的等弹性效用函数描述（参见 He(1993)以及 Pham 和 Touzi(1993)）。它保证了$\widetilde{W}^S$和$\widetilde{W}^\sigma$独立，使得以$\Re_t$和波动率路径$(\sigma_t,0\le t\le T)$为条件的$\log S_T/S_t$的$Q(\lambda^\sigma)$分布服从具有均值$\int_t^T r_u du-\frac{1}{2}\gamma^2(t,T)$和方差为$\gamma^2(t,T)=\int_t^T\sigma_u^2du$的正态分布。在*假设 4.2.1* 下，可以计算（4.2.5）中以波动率路径为条件的期望，最后得到：

$$C_t^{\lambda^\sigma}=S_tE_t^{Q(\lambda^\sigma)}[\phi(d_{1t})-e^{-x_t}\phi(d_{2t})] \quad (4.2.6)$$

它和（2.1.20）有相同的符号。最后，值得注意的是，文献中用的许多期权定价公式具有和（4.2.6）共同的特征，因为基于*波动率参数异质分布的期望值*（参见 Renault(1995)关于这个主题的详细讨论），它们能够被表示为 Black-Scholes 价格。

4.2.2. 用 Hull 和 White 模型进行定价和套期保值

[29]这里，（在考虑期权的市场价格之前）市场被看作是不完全的，所以必须描述一组等鞅测度的特征。

[30]这里以及其他地方，当价格的动态特征由Q决定时，$E_t^Q(\cdot)=E_t^Q(\cdot|\Re_t)$表示给定$\Re_t$时的条件期望算子。

过程(S,σ)的马尔可夫特征暗示期权价格（4.2.6）只依赖于标的资产及其波动率的同期值。而且，在适当的正则性条件下，这个函数是可微的。因此，在随机波动率背景下解决套期保值问题的一个很自然的方法，就是用Δ_t^*单位的标的资产和Σ_t^*单位的价格为C_t^2的其他任何期权，对价格为C_t^1的一个给定期权进行套期保值，其中套期保值比率的解为：

$$\begin{cases}\partial C_t^1 \big/ \partial S_t - \Delta_t^* - \Sigma_t^* \, \partial C_t^2 \big/ \partial S_t = 0 \\ \partial C_t^1 \big/ \partial \sigma_t - \Sigma_t^* \, \partial C_t^2 \big/ \partial \sigma_t = 0 \end{cases} \tag{4.2.7}$$

上述方法被称为 delta-sigma 套期保值策略，是由 Scott(1991)发展的。通过假设任何欧式期权在某一时点到期，即$\partial C_t^2 \big/ \partial \sigma_t \neq 0$，$0 \le t \le T$，Bajeux 和 Rochet(1992)显示充分证据证实 delta-sigma 套期保值问题（4.2.7）唯一解的存在性，还证实了前面几节中隐含的假设，即可获得的信息I_t包含过去值(S_τ,σ_τ)，$\tau \le t$。实际操作中，期权交易者经常把兴趣集中在标的资产价格变化引起的风险上，并且考虑不完全的套期保值策略$\Sigma_t = 0$和$\Delta_t = \partial C_t^1 \big/ \partial S_t$，那么，Hull 和 White 的期权定价公式（4.2.6）直接给出了Δ_t的理论值：

$$\Delta_t = \partial C_t^{\lambda^\sigma} \big/ \partial S_t = E_t^{Q(\lambda^\sigma)} \phi(d_{1t}) \tag{4.2.8}$$

这个理论值在实际中是很难运用的，因为：（1）给定I_t（被σ_t所概括）时，即使知道d_{1t}的$Q(\lambda^\sigma)$条件概率分布，期望（4.2.8）的求导在计算上仍然很费力，（2）给定σ_t时，该条件概率与$\gamma^2(t,T) = \int_t^T \sigma_u^2 du$的条件概率分布是直接相关的，进而可能非平凡地（nontrivially）包含潜在过程σ_t的参数。而且给定σ_t时，这些参数是风险中性概率$Q(\lambda^\sigma)$下，条件概率$\gamma^2(t,T)$的参数，它们通常与数据生成过程P不同。统计推断问题因此变得非常复杂，我们将在第 5 节讨论，只有基于模拟（Simulation-based），同时涉及资产*和*期权价格的推断方法（通过期权定价模型）可以提供满意的解。

不过，回避这些复杂性的一个实际做法是运用 Black-Scholes 期权定价模型，即使已知设定是错误的。事实上，期权交易者知道，使用 BS 公式，其波动率参数是标的资产价格时间序列的历史估计值，通常不能得到足够精确的期权价格和套期保值比率。然而，Black-Scholes 的隐含波动率（2.1.23）概念被视为改进了 BS 模型的定价和套期保值性质。这产生了两个问题：（1）同时运用（假设常数波动率的）BS 模型和隐含波动率明显随时间随机变化的 BS 模型，它们的内在一致性是怎样的；(2)怎样讨论期权定价误差的面板结构？[31]

关于第一个问题，在第 2 节中已指出，Hull 和 White 期权定价模型事实上可看作是这个定价实践的理论基础，套期保值问题和期权*定价误差*的面板结构在 Renault 和 Touzi (1992)中有详细研究。

4.2.3. 微笑和假笑

如同在 2.2 节中提到的，微笑效应现在已是被广泛研究的经验类型化事实。而且，*微笑*有时会变成*假笑*，因为它或多或少会向一方倾斜（即所谓的偏度效应），在第 2 节我们告诫说微笑/假笑效应的一些解释经常是基于推断，而不是严格的证明。

据我们所知，现有研究结果包括：（i）对 Hull 和 White 期权定价模型隐含一个对称微笑的最正式证明，是由 Renualt 和 Touzi(1992)给出的；（ii）微笑／假笑效应可由流动性问题替代解释（微笑曲线最上面的部分，即最贵的期权最少流动）的第一个最完整证明，是由 Platten 和 Schweizer(1994)运用微观结构模型做出的，（iii）没有正式证明标的资产价格过程（杠杆效应、非正态性）的概率分布的不对称性能够用来解释观测到的微笑形的偏度。Renualt(1995)给出了解释观测到的偏度的一种不同尝试。他证明，被用于推断 BS 隐含波动率的标的资产价格$\widetilde{S}_t$和期权交易者所考虑的股票价格S_t之间的差异，可能在微笑中产生一个经验上似乎合理的偏度。这种非同

[31]使 BS 公式同观测到的期权市场价格相等的σ值，主要依赖实际日期t、执行价格K、到期时间$T-t$，并且因此生成一个面板数据结构。

步的$\widetilde{S}_t$和S_t与各种问题相联系：买卖报价差额、两个市场间的非同步交易、基于杠杆效应的预测策略等等。

最后值得注意的是，由 Gouriéoux、Monfort、Tenreiro(1994)始创，并且被 Ait-Sahalia、Bichel、Stoker(1994)所继承的一种新方法，是运用某些观测的状态变量的非参数函数来解释 BS 的隐含波动率。例如，Gouriéroux、Monfort、Tenreiro(1995)得到的一个很好的非参数拟合，具有如下形式：

$$\sigma_t(S_t, K) = a(K) + b(K)(\log S_t / S_{t-1})^2$$

经典的微笑效应可以直接在截距$a(K)$上观察到，但是一个倒转的微笑效应的出现则由路径依赖的效应参数$b(K)$决定。对于美式期权，Broadie、Detemple、Ghysels 和 Torrès (1995)推行了一个不同的非参数方法，其中除了波动率外，还能得到期权合约的执行边界。[32]

4.3. 过滤和离散时间近似

在 3.4.3 节中提到 ARCH 类模型能够被视为从离散时间数据中选取（连续时间）条件方差过程的滤子，有些论文就致力于该主题的研究，如 Nelson（1990，1992，1995a,b）以及 Nelson 和 Foster(1994，1995)。把 ARCH 和连续时间 SV 联系起来，是 Nelson 的开创性贡献之一。在其 1990 年的论文中，Nelson 的第一个贡献是证明了把波动率模拟为过去收益（平方）的函数的 ARCH 模型弱收敛于一扩散过程，要么是$\log\sigma_t^2$的扩散，要么是如 4.2 节中描述的 CEV 过程。特别地，当具有条件方差参数$\omega_h = h\omega$，$\alpha_h = \alpha(h/2)^{1/2}$和$\beta_h = 1 - \alpha(h/2)^{1/2} - \theta h$以及条件均值$\mu_h = hc\sigma_t^2$的 GARCH(1,1)模型，在越来越密集的时间间隔$\Delta t = h$中被观察时，收敛于一个类似式（4.1.8a）与$\delta = 1$的（4.1.9）结合的扩散极限，即

$$d\log S_t = c\sigma_t^2 dt + \sigma_t dW_t$$
$$d\sigma_t^2 = (\omega - \theta\sigma_t^2)dt + \sigma_t^2 dW_t^\sigma$$

类似地，也可以证明，一个 AR(1)-EGARCH(1,1)模型序列将弱收敛于$\ln\sigma_t^2$的一个 Ornstein-Uhlenbeck 扩散：

$$d\ln\sigma_t^2 = \alpha(\beta - \ln\sigma_t^2)dt + dW_t^\sigma$$

因此，这些基本观点表明，作为 ARCH 模型扩散的极限出现的连续时间随机差分方程，不再是 ARCH 而是 SV 模型。而且，根据 Nelson(1992)，即使设定错误，ARCH 模型仍然保持连续时间波动率的理想性质。有争议的是，对于种类繁多的设定错误的 ARCH 模型，当抽样时间间隔的长度以一个适当的比率趋向于零时，ARCH 滤子的波动率估计和真实的潜在扩散波动率之间的差会依概率收敛于零。例如，具有如前面描述的ω_h，α_h和β_h的 GARCH(1,1)模型，$\hat{\sigma}_t^2$的估计如下：

$$\hat{\sigma}_t^2 = \omega_h(1-\beta_h)^{-1} + \sum_{i=0}^{\infty}\alpha_h\beta_h^i y_{t-h(i+1)}^2$$

其中$y_t = \log S_t / S_{t-h}$。该滤子可以被看作是式（3.4.1）的特例。通过过去收益平方在时间t附近的一个滞后多项式函数，GARCH(1,1)和很多其他模型有效地完成了σ_t的一致估计。

各种各样设定错误的 ARCH 模式始终如一地从高频率数据中选取σ_t，这就产生了关于滤子效率的问题。Nelson(1995a,b)以及 Nelson 和 Foster(1994,1995)给出了这个问题的答案，在 3.4 节中提到，线性状态空间 Kalman 滤子也能看成σ_t的一个（次优的）选取滤子。Nelson 和 Foster(1994)指出，渐近最优线性 Kalman 滤子具有标准化估计误差$h^{1/4}[\ln(\hat{\sigma}_t^2) - \ln\sigma_t^2]$的渐近方差等于$\lambda Y(1/2)^{1/2}$，其中$Y(x) = d[\ln\Gamma(x)]/dx$和$\lambda$是尺度因子。和如下形式的 EGARCH 密切相关的一个模型：

$$\ln(\hat{\sigma}_{t+h}^2) = \ln(\hat{\sigma}_t^2) + \rho\lambda(S_{t+h} - S_t)\hat{\sigma}_t^{-1}$$

[32]也见 Bossaerts 和 Hillion(1995)关于非参数套期保值方法和微笑效应的运用。

$$+\lambda(1-\rho^2)^{1/2}[\Gamma(1/2)^{1/2}\Gamma(3/2)^{1/2}|S_{t+h}-S_t|\hat{\sigma}_t^{-1}-2^{-1/2}]$$

产生了具有标准化估计误差的渐近方差等于 $\lambda[2(1-\rho^2)]^{1/2}$ 的渐近最优 ARCH 滤子，其中参数 ρ 度量杠杆效应。这些结果还显示了最有效的次优 Kalman 滤子和最优 ARCH 滤子之间的差别可能相当大。除了过滤之外，还必须处理平滑和预测。所有这些问题都在 3.4 节的离散时间 SV 模型中讨论过，Nelson 和 Foster(1995)对（设定错误的）ARCH 模型的预测性质进行了广泛的研究。Nelson(1995)通过研究平滑滤子使 ARCH 模型有了进一步的发展，即 ARCH 模型不仅包括了滞后收益的平方，还包括了将来的实现值，也就是等式（3.4.1）中 $r=t-T$ 。

4.4. 长期记忆

用连续时间 SV 模型中长期记忆的简要讨论来结束这一节。目的是建立连续时间的长期记忆随机波动率模型，该模型同高频率金融数据及长期期权定价有关。引发长期记忆模型付诸使用的原因在 2.2 节和 3.5.5 节中讨论过。*连续时间*长期记忆的优点，是其能够对主导短期和长期动态特征的参数提供结构性解释。第一小节定义分数布朗运动，然后把注意力转向分数 SV 模型，接着就是关于过滤和离散时间近似。

4.4.1. 关于分数布朗运动的随机积分

这一小节回顾了连续时间下分数和长期记忆过程的一些定义和性质，这些在（例如）Comte 和 Renault(1993)中有广泛的研究，考虑标量过程(scalar process)：

$$x_t=\int_0^t a(t-s)dW_s \tag{4.4.1}$$

这个过程在二次均值的意义上渐近等价于平稳过程：

$$y_t=\int_{-\infty}^t a(t-s)dW_s \tag{4.4.2}$$

只要 $\int_0^{+\infty} a^2(x)dx<+\infty$ 。如果 $a(x)=x^\alpha\tilde{a}(x)/\Gamma(1+\alpha)$ ，其中 $|\alpha|<1/2$ ，$\tilde{a}$ 在 $[0,T]$ 上连续可微，而 $\Gamma(1+\alpha)$ 是尺度因子，在 $[0,T]$ 上用于标准化分数导数算子，则该过程称为分数过程（fractional processes）。这些过程可有几种表示，特别它们能够写成：

$$x_t=\int_0^t c(t-s)dW_{\alpha s}\,,\quad W_{\alpha t}=\int_0^t \frac{(t-s)^\alpha}{\Gamma(1+\alpha)}dW_s \tag{4.4.3}$$

其中 W_α 是所谓的 α 阶分数布朗运动（fractional Brownian Motion of a ）（参见 Mandelbrot 和 Van Ness(1968)）。

函数 a 和 c 之间的关系是一一对应的。有人可能会指出,W_α *不*是一个半鞅(Semi-Martingale)（参见例如 Rogers (1995)），但是关于 W_α 的随机积分能够被正确定义。如果：

$$\lim_{x\to+\infty} x\tilde{a}(x)=a_\infty,\quad 0<\alpha<1/2 \text{ 并且 } 0<a_\infty<+\infty \tag{4.4.4}$$

过程 x_t 是长期记忆的，例如，

$$dx_t=-kx_t dt+\sigma dW_{\alpha t}\qquad x_t=0,\ k>0,\qquad 0<\alpha<1/2 \tag{4.4.5}$$

其解为：

$$x_t=\int_0^t (t-s)^\alpha(\Gamma(1+\alpha))^{-1}dx_t^{(\alpha)} \tag{4.4.6a}$$

$$x_t^{(\alpha)}=\int_0^t e^{-k(t-s)}\sigma dW_s \tag{4.4.6b}$$

注意，x_t 的 α 阶导数 $x_t^{(\alpha)}$ 是普通 SDE 的一个解：$dz_t=-kz_t dt+\sigma dW_t$ 。

4.4.2. 分数 SV 模型

为了方便 FIEGARCH 模型和 3.5.5 节中讨论的对数正态 SV 模型的分数扩展的比较，考虑如下分数 SV 模型（此后称为 FSV）：

$$dS_t/S_t = \sigma_t dW_t \qquad (4.4.7a)$$

$$d\log\sigma_t = -k\log\sigma_t dt + \gamma dW_{\alpha t} \qquad (4.4.7b)$$

其中 $k > 0$ 并且 $0 \le \alpha < 1/2$。分数的指数 α 如果是非零的，那么它将根据波动率过程的规则排列，给出一些自由度，即 α 越大，波动率过程的路径就越平滑。如果用 $r_\sigma(\cdot)$ 表示 σ 的自协方差函数，那么：

$$\alpha > 0 \Rightarrow (r_\sigma(h) - r_\sigma(0))/h \to 0 \quad 当 \quad h \to 0 \quad 时 \quad 。$$

这将被错误地解释为近似单整（near-integrated）行为，这在高频率数据中经常出现，例如当：

$$r_\sigma(h) - r_\sigma(0)/h = (\rho^h - 1)/h \to \log\rho \quad 当 \quad h \to 0 \quad 时 \quad ,$$

并且 σ_t 是连续时间 AR(1)，具有接近于 1 的相关性 ρ 时。

长期记忆的连续时间方法允许对具有如下特征的持续性建模：（1）波动率过程本身（不仅仅是它的对数形式）具有双曲线形衰减的相关图；（2）波动率冲击的持续性使收益产生尖峰态特征，它在缓慢的双曲线形衰减率下随着时频归并而消失。[33]事实上，对于 $[0,h]$ 上的收益率有：

$$\frac{E[\log S_{t+h}/S_t - E(\log S_{t+h}/S_t)]^4}{\left(E[\log S_{t+h}/S_t - E(\log S_{t+h}/S_t)]^2\right)^2} \to 3$$

如果 $\alpha \in [0,1/2]$，当 $h \to \infty$，其衰减速率为 $h^{2\alpha-1}$，如果 $\alpha = 0$，该衰减速率为 $\exp(-kh/2)$。

4.4.3. 过滤和离散时间近似

波动率过程的动态特征由 SDE(4.4.5)的解描述，即：

$$d\log\sigma_t = \int_0^t (t-s)^\alpha / \Gamma(1+\alpha) d\log\sigma_s^{(\alpha)} \qquad (4.4.6)$$

其中 $\log\sigma^{(\alpha)}$ 服从 O-U 过程：

$$d\log\sigma_t^{(\alpha)} = -k\log\sigma_t^{(\alpha)} dt + \gamma dW_t \qquad (4.4.7)$$

为了计算离散时间近似，必须只用过程 $\log\sigma^{(\alpha)}$ 在 $[0,t]$ 的离散分割（discrete partition）点 j/n，$j = 0,1,...,[nt]$ 的值计算积分（4.4.6）的数值。[34]一个很自然的方法是用阶梯函数，生成如下的替代过程：

$$\log\hat{\sigma}_t^n = \sum_{j=1}^{[nt]} (t-(j-1)/n)^\alpha / \Gamma(1+\alpha) \Delta\log\sigma_{j/n}^{(\alpha)} \qquad (4.4.8)$$

其中 $\Delta\log\sigma_{j/n}^{(\alpha)} = \log\sigma_{j/n}^{(\alpha)} - \log\sigma_{(j-1)/n}^{(\alpha)}$。Comte 和 Renault(1994)指出，对于 $n \to \infty$，$\log\hat{\sigma}_{nt}$ 在紧集(compact sets)上均匀地收敛到 $\log\sigma_t$。而且，通过重新整理（4.4.8）可以得到：

$$\log\hat{\sigma}_{j/n}^n = \left[\sum_{i=0}^{j-1} ([(i+1)^\alpha - i^\alpha]/n^\alpha \Gamma(1+\alpha)) L_n^i\right] \log\sigma_{j/n}^{(\alpha)} \qquad (4.4.9)$$

其中 L_n 是对应于抽样方案(sampling scheme) j/n 的滞后算子，即 $L_n Z_{j/n} = Z_{(j-1)/n}$。有了该抽样方案，$\log\sigma^{(\alpha)}$ 就是从连续时间过程推导出来的离散时间 AR(1)，该连续时间过程具有如下形式：

$$(1-\rho_n L_n)\log\sigma_{j/n}^{(\alpha)} = u_{j/n} \qquad (4.4.10)$$

其中 $\rho_n = \exp(-k/n)$，并且 $u_{j/n}$ 是相关联的新生过程。由于过程是平稳的，可以得到（假设对于 $j \le 0$ 有 $\log\sigma_{j/n}^{(\alpha)} = u_{j/n} = 0$）：

[33]对于一般的 GARCH 或者 SV 模型，它以指数比率消失（参见 Drost 和 Nijman(1993)以及 Drost 和 Werker(1994)关于短期记忆情形中这些问题的讨论）。

[34]当 $k \le z < k+1$ 时，$[z]$ 是整数 k。

$$\log\hat{\sigma}^{n}_{j/n}=\left[\sum_{i=0}^{j-1}\frac{(i+1)^{\alpha}-i^{\alpha}}{n^{\alpha}\Gamma(1+\alpha)}L_n^i\right](1-\rho_n L_n)^{-1}u_{j/n} \tag{4.4.11}$$

它在两个部分给出了波动率动态特征的参数化：（1）对应于滤子 $\sum_{i=0}^{+\infty}\mathrm{a}_i L_n^i/n^{\alpha}$ 的长期记忆部分，其中 $\mathrm{a}_i=[(i+1)^{\alpha}-i^{\alpha}]/\Gamma(1+\alpha)$；以及（2）表现为 AR(1)过程即 $(1-\rho_n L_n)^{-1}u_{j/n}$ 的短期记忆部分。事实上，可以证明，在两种类型的过程之间存在一个长期关系（协整关系）的意义上，该长期记忆滤子"长期等价于"通常的离散时间长期记忆滤子 $(1-L)^{-\alpha}$。然而，该长期记忆滤子和通常的离散时间长期记忆滤子 $(1-L)^{-\alpha}$ 之间的这个长期等价*并不表示*标准的参数化 FARIMA(1, α ,0)适合本文的框架。事实上，可以证明，通常的离散时间滤子 $(1-L)^{-\alpha}$ 引进了长期特征和短期特征之间的某种混合，而简约的连续时间模型却并不这样。[35]与离散时间模型及 GARCH 长期记忆模型相比，该特征无疑是连续时间 FSV 一个优势。

5. 统计推断

计算 ARCH 模型的似然函数是一个相对简单的工作。形成明显对照的是，对于 SV 模型，不可能得到似然函数的明确表达式。这是几乎所有的非线性潜变量模型共同具有的一般性特征。SV 模型估计方法的缺乏使其在很长一段时间里同 ARCH 相比，成为没有吸引力的一类模型。然而，最近几年，总体上关于非线性潜变量模型的估计已经有了显著的进步，特别是 SV 模型的估计，现在很多方法可用，随着 CPU 性能的提高，这些方法能够被计算机实现。对 SV 模型估计的早期尝试是用 GMM 方法，一个突出的例子是 Melino 和 Turnbull(1990)。5.1 节致力于 SV 模型的 GMM 估计。显然，GMM 不是用于处理连续时间扩散的，因为它要求离散时间过程来满足特定的正则性条件。由 Hansen 和 Scheinkman (1994) 提出的连续时间 GMM 方法包含了直接从连续时间过程的表达式推导出来的矩条件，该方法在 5.3 节中讨论。这两节之间，即 5.2 节中，讨论 QML 方法，这是 Harve、Ruiz 和 Shephard(1994)以及 Nelson(1988)提出的。事实表明，非线性（高斯）SV 模型能够转化为如第 3 节中的线性非高斯状态空间模型，并且根据这一点，可以计算高斯准似然(quasi- likelihood)。5.1 节到 5.3 节中包含的方法均没有涉及到模拟。然而，计算机能力的提高使基于模拟的估计方法日益流行。模拟矩方法，或者由 Duffie 和 Singleton(1993)提出的基于模拟的 GMM 方法，是第一个例子，它包含在 5.4 节中。接下来的第 5 节中，讨论 Gouriéroux、Monfort 和 Renault(1993)的间接推断方法，以及 Gallant 和 Tauchen(1994)的矩匹配方法（moment matching methods）。最后，5.6 节包括很大一类强化计算机马尔可夫链蒙特卡罗方法的估计量，这些方法由 Jacquier、Polson 和 Rossi(1994)以及 Kim 和 Shephard(1994)在 SV 模型的框架下运用，还包括在 Danielsson(1994)以及 Danielsson 和 Richard(1993)在论文中提出的基于 ML 估计的模拟。

在每一节中，将试图把焦点仅限于在 SV 模型框架下估计方法的运用，而回避有关经济计量理论的细节。完成本节有用的参考材料主要有：（1）Hansen(1992)，Gallant 和 White(1988)，Hall(1993)以及 Ogaki(1993)的 GMM 估计，（2）Gourieroux 和 Monfort(1993b)以及 Wooldridge(1994)的 QMLE，（3）Gouriéroux 和 Monfort(1995)以及 Tauchen(1995)包含间接推断和矩匹配的基于模拟的经济计量方法，还有（4）Geweke(1995)以及 Shephard (1995)的马尔可夫链蒙特卡罗法。

5.1. 广义矩方法

考虑（3.1.2）和（3.1.3）介绍的离散时间 SV 的简单形式，它具有新生过程 (ε_t,η_t) 的概率分布服从正态性这个附加假设。该对数正态 SV 模型至少已经是两个扩展的蒙特卡罗研究的主题，它们是关于 SV 模型的 GMM 估计，分别是 Andersen 和 Sørensen(1993)以及 Jacquier、Polson 和 Rossi(1994)做出的。主要思想是运用 SV 模型的平稳性和遍历性性质，这些性质使得样本矩收敛于它们的无条件期望。例如，二阶和四阶矩是 σ^2 和 σ_h^2 的简单表达式，即分别是 $\sigma^2\exp(\sigma_h^2/2)$

[35]即，$(1-L_n)^{\alpha}\log\hat{\sigma}^{n}_{j/n}$ 不是一个 AR(1)过程。

和$3\sigma^4\exp(2\sigma_h^2)$。如果要在样本中计算这些矩，$\sigma_h^2$的估计可直接来自样本峰度$\hat{\kappa}$，$\hat{\kappa}$是四阶矩与二阶矩的平方的比值，表达式为$\hat{\sigma}_h^2 = \log(\hat{\kappa}/3)$。参数$\sigma^2$的估计则可以通过在该估计中替换$\sigma_h^2$直接从二阶矩得到，还可以计算$y_t^2$的一阶自协方差，或者仅计算$y_t^2 y_{t-1}^2$的样本均值，其表达式为$\sigma^4\exp(\{1+\phi\}\sigma_h^2)$。给定$\sigma^2$和$\sigma_h^2$的估计，可以直接得到$\phi$的估计。

以上步骤是矩法运用的一个例子。笼统地说，计算了 m 个矩。对一个容量为T的样本，令$g_T(\beta)$表示用模型参数β表示的每一个样本矩和其理论表达式之间的差距的一个$m\times 1$向量。广义矩（GMM）估计量是通过最小化准则函数来构造的：

$$\hat{\beta} = Arg_{\beta}\min g_T(\beta)' W_T g_T(\beta)$$

其中W_T是一个$m\times m$的权重矩阵，它反映了赋予相应的每一个矩的重要性。当ε_t和η_t相互独立时，Jacquier、Polson 和 Rossi(1994)提出用 24 个矩，前面四个对于$c=1,2,3,4$由（3.2.2）给出，而其它的解析式是：

$$E\left[\left|y_t^c y_{t-1}^c\right|\right] = \left\{\sigma^{2c} 2^c \left[\Gamma\left(\frac{c}{2}+\frac{1}{2}\right)\right]^2 \Big/ \pi\right\} \exp\left(\frac{c^2}{4}\sigma_h^2[1+\phi^\tau]\right) ,$$

$$c=1,2 \quad , \quad \tau = 1,2,\ldots,10 \text{ 。}$$[36]

更一般情形下，当ε_t和η_t相关时，Melino 和 Turnbull(1990)引入$E\left[\left|y_t\right| y_{t-\tau}\right]$的估计，$\tau = 0,\pm1,\pm2,\ldots,10$。他们给出了$\tau=1$情形的一个明确表达式，并且指出它的符号完全由$\rho$决定。

GMM 方法还可以扩展，用来处理ε_t非正态分布的情形。需要的解析式能够从 3.2 节中得到。另一方面，2.4 节中提到的一般 SARV 模型的无条件矩的解析式可以在更一般的设定下给出 GMM 估计的基础（参见 Andersen(1994)）。

从一开始，我们就知道 GMM 估计量不是有效的，问题是需要忍受多大无效性以换取它的相对简单性。GMM 的一般设置方式没有矩条件的数目，除了识别所需要的最小数目以及矩的明确选择。而且，权重矩阵的计算也是一个问题，因为在实际中有很多选择。Andersen 和 Sørensen(1993)以及 Jaquier、Polson 和 Rossi(1994)的扩展的蒙特卡罗研究试图回答这些重要的问题。他们发现主要以 2.2 节中提到的类型化事实为基础的 GMM 大体上是一个相当无效的方法，（3.1.3）中的ϕ在许多实证结果中是接近于 1 的，因为波动率具有高度的持续性。对于ϕ接近于 1 的参数值收敛于无条件矩是相当缓慢的，这表明只有大样本能够解决这种情况。Andersen 和 Sørensen(1993)的蒙特卡罗研究对怎样控制无效性的程度提出了一些指导，尤其是把矩条件控制在一个很小的数目内。对于具有数据依赖带宽(data-dependent bandwidch)的权重矩阵估计量的选择，通过运用 Bartlett 核，他们也提出了一些具体建议。

5.2. 准极大似然估计(QML)

5.2.1. 基本模型

考虑 3.4.1 小节中描述的线性状态空间模型，其中（3.2.8）是度量方程，而（3.1.3）是过渡方程。参数ϕ，σ_η^2以及ξ_t的方差σ_ξ^2的 QML 估计，是把ξ_t和η_t视为正态的，并最大化由 Kalman 滤子获得的似然值的预测误差分解形式。Harvey、Ruiz 和 Shephard (1994)中指出，准极大似然（QML）估计量是渐近正态的，其协方差矩阵由 Dunsmuir(1979，p.520)中的理论给出。它假设η_t和ξ_t具有有限的四阶矩，并假设参数不是在参数空间的边界上。

参数ω能够与其它参数同时估计。换句话说，它可以作为$\log y_t^2$的均值被估计，因为在ϕ的绝对值小于 1 时，它们是渐近等价的。

QML 方法的运用不要求对ε_t的具体分布做假设。我们称之为无约束的 QML。然而，如果已

[36]导出这些矩条件的一个简单方法是通过一个类似于（2.4.8）和（2.4.9）或者（3.2.3）的两步法。

经做了分布假设，那么就没有必要估计σ_ξ^2，因为它是已知的，而尺度因子σ^2的估计能够从ω的估计中得到。换句话说，它能够象在 3.4.1 小节中提到的那样取得。

如果无约束的 QML 估计已经实现了，那么决定一类分布中某个特殊分布的参数值可以从ξ_t方差的估计值推导出来。例如，在学生t分布中，v可以根据ξ_t的方差的理论值$4.93+\psi'(v/2)$来决定（其中$\Psi(\cdot)$是 3.2.2 节中介绍的 digamma 函数）。

5.2.2. 不对称模型

在一个不对称模型中，QML 可以以（3.4.3）中修正的状态空间形式为基础。除了η_t和ξ_t的四阶矩存在以及ξ_t和η_t是联合对称的之外，在没有任何分布假设的条件下，参数σ_ξ^2、σ_η^2、ϕ、μ^*和γ^*能够通过 Kalman 滤子来估计。然而，如果想要得到ρ的估计，则需要作出关于扰动的分布假设，导出与（3.4.4）和（3.4.5）一样的公式。这些公式可以用于设定有关原始参数σ^2、σ_η^2、ϕ和ρ的优化。其优点是能够施行$|\rho|<1$的约束。注意，任何$t-$分布都给出参数之间的相同关系，所以在该类分布中不需要设定自由度。

运用原始扰动均符合高斯假设的 QML 方法，Harvey 和 Shephard(1993)估计了有关 1962 年 6 月 3 日到 1987 年 12 月 31 日 US 市场指数加权值的 CRSP 日收益的一个模型。Nelson(1991)在论文中使用这些数据阐述他的 EGARCH 模型，实证结果表明有一个非常高的负相关。

5.2.3. 频域（frequency domain）中的 QML

对一个长期记忆的 SV 模型，时域中的 QML 估计是缺乏吸引力的，因为只有当h_t表示成一个自回归或移动平均过程，并且截取适当高的滞后时，状态空间形式（SSF）才能使用。因此，该方法很麻烦，尽管原始状态的协方差矩阵很容易构造，而且截取并不影响估计量的渐进性质。如果自回归的近似，从而 SSF，没有使用，那么时域 QML 就要求重复构造$\log y_t^2$的$T\times T$协方差矩阵并求逆，参见 Sowell(1992)。另一方面，频域中的 QML 估计并不比 AR(1)情形中的 OML 估计更难得到。Cheung 和 iebold(1994)给出模拟证据，指出虽然时域估计在小样本中更有效，但是在必须估计均值时，其差异较不显著。

忽略常数项，频域（准）对数似然函数是：

$$\log L=-\frac{1}{2}\sum_{j=1}^{T-1}\log g_j-\pi\sum_{j=1}^{T-1}I(\lambda_j)/g_j \qquad (5.2.1)$$

其中$I(\lambda_j)$是$\log y_t^2$的样本谱(sample spectrum)，g_j是谱生成函数（SGF），对（3.5.1）该谱生成函数是：

$$g_j=\sigma_\eta^2[2(1-\cos\lambda_j)]^{-d}+\sigma_\xi^2$$

注意，（5.2.1）中的总和是从$j=1$开始而不是$j=0$，这是因为对于正的d，g_0无法计算。然而，省略零频数的确会除去均值。未知参数是σ_η^2、σ_ξ^2和d，但是通过σ_η^2由信号噪声比率$q=\sigma_\eta^2/\sigma_\xi^2$替换的再参数化，$\sigma_\xi^2$可通过似然函数的集中化消去(concentrated out)。另一方面，如果已假设了ε_t的分布，那么σ_ξ^2就是已知的。Breidt、Crato 和 de Lima(1993)阐述了 QML 估计量的一致性。

当d处于0.5和1之间时，h_t是不平稳的，但对$\log y_t^2$差分会产生一个零均值的平稳过程，其 SGF 是

$$g_j=\sigma_\eta^2[2(1-\cos\lambda_j)]^{1-d}+2(1-\cos\lambda_j)\sigma_\xi^2$$

长期记忆模型的一个吸引人的地方就是推断不受自回归产生的单位根问题的影响。因此，以似然函数为基础的假设检验$d=1$及其备择假设$d<1$，检验能够运用标准理论来构造，参见 Robinson(1993)。

5.2.4. GMM 和 QML 的比较

关于 GMM 和 QML 的有限样本行为的模拟证据能够在 Andersen 和 Sørensen(1993)，Ruiz (1994)，Jacquier、Polson 和 Rossi(1994)，Breidt 和 Carriquiry(1995)，Andersen 和 Sørensen(1996) 以及 Harvey 和 Shephard(1996)中找到。一般性结论似乎是，当波动率相对强时（反映为方差具有高系数），QML 使估计值具有更小的 MSE。这是因为度量方程（3.2.8）中正态分布的波动率分量和非正态的误差项相关。当方差系数较低时，GMM 较优。然而，在这种情形下，Jacquier、Polson 和 Rossi(1994，p.383)观察到"…QML 和 GMM 估计量的性能都迅速恶化"。换句话说，5.6 节中描述的强化计算机的方法表现更好。

其他方面都相同，一个接近于 1 的 AR 系数 ϕ 倾向于支持 QML，因为自相关性缓慢衰减，因此较难被 GMM 中运用的矩捕捉。同样的原因，用 GMM 估计长期记忆模型的结果可能相当差。

QML 的吸引人之处在于它很容易实现，并且很容易扩展到更一般的模型，例如非平稳的和多变量的模型。同时，它提供了状态的过滤、平滑估计以及预测。向前一步（one-step ahead）预测误差也可用于构造诊断，比如 Box-Ljung 统计量，虽然做这样的检验时必须记住观测值是非正态的。因此，即使超参数（hyperparameter）最终用其他方法估计，QML 在寻找适当的模型设定时仍可起到重要作用。

5.3. 连续时间 GMM

Hansen 和 Scheinkman(1995)提出应用 GMM 方法来估计连续时间扩散过程，GMM 方法专为该过程设定。在 5.1 节中，讨论了 SM 模型的估计，对于离散时间过程它可被明确公式化，或是连续时间扩散过程的离散化。在两种情形中，推断都是基于无条件矩与其等价样本矩之间差距的最小化。对于连续时间过程，Hansen 和 Scheinkman(1995)直接采用扩散而不是其离散化，来表示矩条件。为描述该方法的一般设置方式，他们建议考虑如下 n 个扩散方程组成的（多变量）系统：

$$dy_t = \mu(y_t;\theta)dt + \sigma(y_t;\theta)dW_t \qquad (5.3.1)$$

与第 2 节中的表述相比较即可发现该方法的某些局限性。首先，函数 $\mu_\theta(\cdot) \equiv \mu(\cdot;\theta)$ 和 $\sigma_\theta(\cdot) \equiv \sigma(\cdot;\theta)$ 的参数化只通过 y_t 进行，它将第 2 节中的状态变量过程 U_t 约束为 y_t 的同期值。（5.3.1）中的扩散过程包含一般的向量过程 y_t，因此 y_t 可包含适合 SV 模型的波动率过程，但是，向量 y_t 被假设为可观测的。暂时把这个问题放到一边，到本节的结尾再讨论。Hansen 和 Scheinkman(1995)考虑了极小算子 A，它是为如下的一类平方可积函数 $\varphi: R^n \to R$ 定义的：

$$A_\theta \varphi(y) = \frac{d\varphi(y)}{dy'}\mu_\theta(y) + \frac{1}{2}Tr\left(\sigma_\theta(y)\sigma'_\theta(y)\frac{d^2\varphi(y)}{dydy'}\right) \qquad (5.3.2)$$

由于该算子被定义为一个极限，即：

$$A_\theta \varphi(y) = \lim_{t\to 0} t^{-1}[E(\varphi(y_t)|y_0 = y) - y] \ ,$$

它不要求对所有平方可积函数 φ 都存在，而只是对某一限制域 D 必须存在。现在，从这类函数 $\varphi \in D$ 能够得到一组矩条件。事实上，例如正象 Revuz 和 Yor(1991)已经指出的，下列等式成立：

$$EA_\theta \varphi(y_t) = 0 \ , \qquad (5.3.3)$$

$$E[A_\theta \varphi(y_{t+1})\widetilde{\varphi}(y_t) - \varphi(y_{t+1})A^*_\theta \widetilde{\varphi}(y_t)] = 0 \qquad (5.3.4)$$

其中 A^*_θ 是 A_θ 与过程 y 的不变测度相关联的标量乘积的伴随极小算子。[37]通过选择一组适当的函数，Hansen 和 Scheinkman 利用矩条件（5.3.3）和（5.3.4）构造 θ 的 GMM 估计量。

函数 $\varphi \in D$ 以及 $\widetilde{\varphi} \in D^*$ 的选择决定了数据的什么矩被用来估计参数。显然，这就产生有关于选择函数以提高估计量有效性的问题，但是最先也是最重要的问题是通过条件（5.3.3）和（5.3.4）决定 θ 的可识别性。在本节一开头就提到，为了涵盖 SV 模型，多变量过程 y_t 必须以某种方式包含潜在的条件方差过程。Gourieroux 和 Monfort(1994，1995)指出，由于以 φ 和 $\widetilde{\varphi}$ 为基础的矩条件不能包括任何潜在过程，所以经常（但不总是）不能识别所有参数，尤其是那些决定潜在波动率

[37]请注意 A^*_θ 再一次同定义域 D^* 相关联，使得（5.3.4）中 $\varphi \in D$ 并且 $\widetilde{\varphi} \in D^*$。

过程的参数。一个可能的补救方法是，增加模型中不直接与潜在波动率过程相关联的观测值，旨在使其可观测。一个可能的备选方案是把证券价格以及通过标的资产的期权市场报价得到的 Black-Scholes 隐含波动率包含在 y_t 中。事实上，这个方法是 Pastorello、Renault 和 Touzi(1993)提出的，虽然不是在连续时间 GMM 的背景下，而是使用了间接推断方法，该推断方法将在 5.5 节中讨论。[38]另一个可能是依赖如 Conley 等人（1995）在连续时间 GMM 背景下讨论的 SV 模型的时间形变表示。

5.4. 模拟矩方法

到目前为止讨论的估计方法还没有涉及任何模拟技术。从现在起，我们将从模拟矩方法（SMM）估计量开始，讨论模拟和估计相结合的方法， 这在 Duffie 和 Singleton(1993)关于时间序列过程的研究中已有所涉及。[39]在 5.1 节中，提到 SV 模型的 GMM 估计是以一组选择的样本矩和无条件总体矩之间距离的最小化为基础，并表达为模型参数的解析函数。现在，假设这样的解析函数很难得到。当该表达式涉及潜在过程（如随机波动率过程）的边际化时，情况就是如此。那么，能否对特定参数值的模型进行数据模拟，用模拟数据的矩匹配样本矩作为替代？这个策略即为 SMM。事实上，模拟那些过程并且利用 SMM 方法经常是相当简单的，再考虑前一节（等式（5.3.1））中的（多变量）扩散作为参考，用离散化进行 H 个模拟 $i=1,...,H$：

$$\Delta\hat{y}_t^i(\theta)=\mu(\hat{y}_t^i(\theta);\theta)+\sigma(\hat{y}_t^i(\theta);\theta)\varepsilon_t \quad 且 \quad i=1,...,H \quad , \quad t=1,...,T$$

其中 $\hat{y}_t(\theta)$ 是给定参数 θ，而 ε_t 是独立同高斯分布的一个模拟，。[40]根据识别性和其它正则性条件，可以考虑

$$\hat{\theta}_T^H = Arg\min_{\theta}\left\| f(y_1,...,y_T)-\frac{1}{H}\sum_{i=1}^{H} f(\hat{y}_1^i(\theta),...,\hat{y}_T^i(\theta))\right\|$$

它具有适当的标准选择，也就是 GMM 中二次型的加权矩阵，以及数据的函数 f，即矩条件。其渐近分布理论和 GMM 中的非常相似，只是与 GMM 估计量相比，模拟行为引入了一个额外的随机误差源，影响了 SMM 估计量的有效性。有效性的损失能够由 H 的选择加以控制。[41]

5.5. 间接推断和矩匹配

Gouriéroux、Monfort 和 Renault(1993)对间接推断方法以及 Gallant 和 Tauchen(1994)对矩匹配方法的重要见解是，为了估计关注的模型，引入一个被向量（比方说 β）参数化的辅助模型。在这里，后者就是 SV 模型。[42]第一小节中，将描述一般原理，而第二小节集中于扩散的估计。

5.5.1. 原理

在第 5 节的开头提到，同 SV 模型相比，ARCH 类模型相对容易估计些。为此，ARCH 类模型可以作为辅助模型的候选对象。一个可供选择的策略是，通过 Gallant 和 Tauchen(1989)提出的 SNP 密度来尽量概括数据的特征，这种经验的 SNP 密度，或更具体地说是它的得分，也能够起到辅助模型的作用。其它可能性也可考虑。然后要运用辅助模型估计 β，使得：

$$\hat{\beta}_T = Arg\max_{\beta}\sum_{t=1}^{T}\log f^*(y_t|y_{t-1},\beta) \qquad (5.5.1)$$

这里为了便于阐述，把注意力限制在一个只有一期滞后的简单动态模型。当为了方便估计而有意

---

[38]在 2.1.3 节中提到，隐含波动率是有偏的。Pastorello、Renault 和 Touzi(1993)运用的间接推断方法能够处理这种偏差，这将在 5.5 节中解释。期权价格数据的使用会在 5.7 节中做进一步的讨论。

[39]最初提出 SMM 是为了横截面应用，参见 Pakes 和 Pollard(1989)以及 McFadden(1989)。也可参见 Gouriéroux 和 Monfort(1993a)。

[40]下一节将详细讨论模拟技术。事实上，为了控制离散化的偏倚，必须用一个好的抽样间隔来模拟。

[41] SMM 估计量的渐近方差通过因子 $(1+H^{-1})$ 依赖于 H，例如，参见 Gouriéroux 和 Monfort(1995)。

[42]值得注意的是，这里将描述的以推断方法为基础的模拟适用于许多其他类型的模型，如横截面、时间序列以及面板数据的模型。

使辅助模型设定错误时，（5.5.1）中的目标函数 f^* 可以是一个准极大似然函数。另一选择是 f^* 能够从 SNP 类密度中得到。[43]于是，Gouriéroux，Monfort 和 Renault 提出估计同样的参数向量 β，不采用真实的样本数据，而是使用在给定 θ 时从关注模型获得的 $i=1,\dots,H$ 次模拟样本 $\{\hat{y}_t^i(\theta)\}_{t=1}^T$，这就产生了 β 的一个新估计量，即：

$$\hat{\beta}_{HT}(\theta) = Arg\max_{\beta}(1/H)\sum_{i=1}^{H}\sum_{t=1}^{T}\log f^*(\hat{y}_t^i(\theta)\big|\hat{y}_{t-1}^i(\theta),\beta) \qquad (5.5.2)$$

下一步是基于 H 个重复模拟和 T 个观测值的样本，运用加权矩阵 W_t 来选择 θ 的一个间接估计量，使得二次距离最小化，即：

$$\hat{\theta}_{HT} = Arg\min_{\beta}\left(\hat{\beta}_T - \hat{\beta}_{HT}(\theta)\right)' W_T\left(\hat{\beta}_T - \hat{\beta}_{HT}(\theta)\right) \qquad (5.5.3)$$

Gallant 和 Tauchen(1994)的方法避免了通过计算 f^* 的得分函数来估计 $\hat{\beta}_{HT}(\theta)$ 这一步骤，并且不用最小化类似于（5.5.3）的二次距离，但是它要计算得分函数在 $\hat{\beta}_T$ 处的值，并且要用由关注模型生成的模拟序列来代替样本数据。在适当的正则性条件下，$\hat{\theta}_{HT}$ 是根 T 一致（root T consistent）且渐近正态的估计量。如同 GMM 和 SMM，还要有一个最优加权矩阵，最后的渐进协方差矩阵依赖于模拟的次数，就象 SMM 估计量依赖于 H 一样。

Gouriéroux、Monfort 和 Renault(1993)用简单的例子阐明了间接推断估计量的运用，这里对该例做简要讨论。典型的 AR 模型很容易估计，而 MA 模型则需要更多的详细步骤。假设关注的模型是具有参数 θ 的一阶移动平均模型，他们提出估计一个包含参数向量 β 的 AR(p)模型，而不是直接从数据估计 MA 参数。下面的步骤包括运用 MA 模型模拟数据，并接着进行上面的步骤。[44]他们发现 $\hat{\theta}_{HT}$ 的间接推断估计量看起来比 MA 参数更传统的极大似然估计量有更好的有限样本性质。事实上，间接推断估计量呈现出与 Andrews(1993)提出的中值无偏估计量（median unbiased Estimator）相似的特征。Gouriéroux、Renault 和 Touzi(1994)证实并阐明了这些性质，他们研究了间接推断估计量的二阶渐近扩展及其降低有限样本偏倚的能力。

5.5.2. 扩散估计

下面考虑一个同 5.3 节中一样的扩散过程，该扩散方程处理连续时间 GMM，即：

$$dy_t = \mu(y_t;\theta)dt + \sigma(y_t;\theta)dW_t \qquad (5.5.4)$$

在 5.3 节中提到，上述方程在特定约束下成立，例如约束 y_t 为函数 μ 和 σ 的自变量。这些约束适用于 5.3 节的设定，而不适用于这里讨论的估计方法。事实上，方程（5.5.4）只是用于作为一个解释性的例子。那么，可以通过确切的离散化或者通过某类近似的离散化（例如 Euler 或者 Mil'shtein，参见 Pardoux 和 Talay(1985)或者 Kloeden 和 Platten(1992)中更详细的讨论）来模拟扩散，更精确的 $y_t^{(\delta)}$ 定义为：

$$y_{(k+1)\delta}^{(\delta)} = y_{k\delta}^{(\delta)} + \mu(y_{k\delta}^{(\delta)};\theta)\delta + \sigma(y_{k\delta}^{(\delta)};\theta)\delta^{1/2}\varepsilon_{(k+1)\delta}^{(\delta)} \qquad (5.5.5)$$

在适当的正则性条件（例如，参见 Strook 和 Varandhan(1979)）下，已知扩散（依分布）容许有一个唯一解，并且当 δ 趋于零时，过程 $y_t^{(\delta)}$ 收敛于 y_t。因此，当 δ 足够小时，可以期望非常精确地模拟 y_t，选择 $\delta=1$，辅助模型可能是(5.5.4)的离散化。所以，令 $\delta=1$，可以在（5.5.5）非线性 AR 模型基础上建立一个 ML 估计量。例如，为了控制离散化的偏倚，可以在 $\delta=1/10$ 或者 $\delta=1/20$ 时模拟潜在扩散，且归并模拟的数据以符合 DGP 的抽样频率。Borze、Scaillet 和 Zakoïan(1994)讨论了模拟步幅（step size）对渐近分布的影响。

---

[43]这个讨论不应该留下如下印象，即辅助模型只能通过 ML-类型的估计量来估计。任何根 T (root T consistent)的一致渐近正态估计方法都可以运用。

[44]这里，遵循 Gallant 和 Tauchen(1994)，可以再次运用得分原则。事实上，在一个线性高斯设定中，拟合由 MA(1)模型生成数据的 SNP 方法可以估计 AR(p)模型。Ghysels、Khalaf 和 Vodounou(1994)更详细地讨论了 MA 模型基于得分的和间接推断的估计量，以及它们同更标准估计量的关系。

当扩散包含潜在的过程，基于模拟推断方法的运用变得格外适合和富有吸引力，例如 SV 模型的情况。Gouriéroux 和 Monfort(1994，1995)讨论了几个例子，并且通过蒙特卡罗模拟研究它们的性能。需要注意的是，在一个粗糙的离散化下估计扩散并不是辅助模型唯一可能的选择。事实上，Pastorrello、Renault 和 Touzi(1993)，Engle 和 Lee(1994)以及 Gallant 和 Tauchen(1994)提出使用 ARCH 类模型。

这些方法在金融时间序列中已经有一些成功的运用。它们包括 Broze 等（1995），Engle 和 Lee(1994)，Gallant、Hsieh 和 Tauchen(1994)，Gallant 和 Tauchen(1994，1995)，Ghysels、Gouriéroux 和 Jasiak(1995b)，Ghysels 和 Jasiak(1994a,b)，Pastorello 等(1993)等等。

5.6. *基于似然的方法与贝叶斯方法*

在高斯线性状态空间模型中，似然函数通过向前一步预测误差构造。似然值的这一预测误差分解形式被用来作为 QML 中的准则函数，当然，在这种情形中它不是精确的似然值。原则上，Watanabe(1993)提出的精确滤子将产生精确似然值，然而，正如 3.4.2 节中提到的，由于该滤子运用了数值积分，则需要长时间的计算，并且如果还需要计算关于超参数的数值最优化，那就会变得不切实际。

Kim 和 Shephard(1994)研究了用于 QML 的线性状态空间形式，除了用混合正态分布近似测量误差的 $\log(\chi^2)$ 分布。对于这些正态分布中的每一个，预测误差分解的似然函数都可以计算。一种模拟的 EM 算法被用于寻找最好的混合，并以此计算超参数的近似 ML 估计。

精确的似然函数也能被构造为以波动率为条件的观测值分布的混合，即

$$L(y;\phi,\sigma_\eta^2,\sigma^2)=\int p(y|h)p(h)dh$$

其中 y 和 h 分别包含 T 个元素 y_t 和 h_t。这个表达式能够根据 σ_t^2 写出，而不是根据它们的对数形式 h_t，但是随后会有小的差异。当然，问题是上面的似然函数没有闭合的形式，所以它必须用某种模拟方法计算。在 Shephard(1995)以及 Jacquier、Polson 和 Rossi(1994)中有精彩的讨论，包括评论。在概念上，最简单的方法是运用蒙特卡罗积分获得在给定参数 $(\phi,\sigma_\eta^2,\sigma^2)$ 值时 h 的无条件分布，估计似然值作为 $p(y|h)$ 的平均值，然后重复进行，对 ϕ、σ_η^2 进行搜索，直到找到模拟似然函数的极大值。按照现在的情况，这个方法不很令人满意，但是运用重点抽样的思想可以改善。Danielsson 和 Rechand (1993)以及 Danielsson(1994)在 SV 模型的 ML 估计中已经采用了该方法，然而，随着样本容量的增加，这个方法变得越来越困难。

通过模拟技术处理似然估计的一个更有前途的方法是运用马尔可夫链蒙特卡罗（MCMC）得到以观测值为条件的波动率的分布。能够做到这一点的方法在关于非线性滤子和平滑器的 3.4.2 小节中已经有概述。Kim 和 Shephard(1994)提出了计算 ML 估计量的一个方法，就是在模拟的 EM 算法中赋予它们多移动算法(multimove algorithin)。Jacquier、Polson 和 Rossi(1994)采用了贝叶斯方法，这个方法中模型的设定具有一个分级的结构，其中超参数 $\varphi=(\sigma_\eta,\phi,\ \sigma)'$ 的先验分布连接着条件分布 $y|h$ 和 $h|\varphi$。（实际上，是用 σ_t 而不是用 h_t）。h 和 φ 的联合后验与三个分布的乘积成比例，即 $p(h,\varphi|y)\propto p(y|h)p(h|\varphi)p(\varphi)$。$h$ 的引入使得统计处理易于进行，它是数据增扩（augmentation）的一个例子；见 Tanner 和 Wong(1987)。考虑到超参数中样本的变易性，边际的 $p(y|h)$ 能够根据联合后验 $p(h,\varphi|y)$ 解决未观测到的波动率的平滑问题。以 h 为条件的 φ 的联合后验 $p(\varphi|h,y)$ 很容易从线性模型的标准贝叶斯处理计算得到。如果可以以很低的成本从 $p(h|\varphi,y)$ 中直接抽取样本，那么通过来回地轮流提取 $p(\varphi|h,y)$ 和 $p(h|\varphi,y)$，构造一个马尔可夫链就变得很容易了，这将产生一个循环链，它是 Gibbs 抽样法的一个特殊情形。然而，正如在 3.4.2 小节中提到的，Polson 和 Rossi(1994)指出，把 $p(h|\varphi,y)$ 分解成一系列单变量分布更好，其中每一个 h_t（而不是 σ_t）都以所有其它的为条件。

ω 的先验分布，即 JPR(1994)中波动率过程的参数，是线性模型的标准共轭先验，一个（截取的）正态—伽玛。这些先验可以高度扩散，而其余部分不扩散。JPR 进行广泛的抽样试验，以

记录这个方法以及更传统方法的表现。通过模拟随机波动率序列，他们把后验均值的抽样性能同 QML 和 GMM 点估计做了比较。MCMC 后验均值展示出*均方根误差*是在 GMM 和 QML 点估计均方根误差的大约一半与大约四分之一之间。更为显著的是波动率平滑表现的结果。由贝叶斯滤子生成的 h_t 的后验均值的均方根误差比点估计小 10%，该点估计由有*真实*参数支持的近似 Kalman 滤子生成。

Shephard 和 Kim 在对 JPR(1994)的评论中指出，对于很高的 ϕ 和很小的 σ_η，JPR 算法的收敛比率将变缓。那么，将需要更多的抽取以得到同样数量的信息。他们提出用正态分布的离散混合来近似波动率扰动，这个方法的好处是向量 h 的抽取成为可能，并快于每个 h_t 的 T 个抽取。然而，这样做的代价是，由于离散化的影响，抽取将在更高维数的空间中进行。同样，基于离散混合的链的收敛对组合成分的数目以及赋于它们的概率权重很敏感，通过让数据产生离散状态空间特征的估计（概率，均值和方差）值，Mahieu 和 Schotman(1994)使 Shephard 和 Kim 的思想更具有一般性。

JPR 算法的最初实行被限制在随机波动率的一个非常基本的模型 AR(1)上，它具有不相关的均值和波动率扰动。在一个单变量的设定中，相关的扰动对股票收益可能是很重要的，即所谓的杠杆效应。Gallant、Rossi 和 Tauchen(1994)中的证据也涉及具有偏度和峰度的非正态条件误差。Jacquier、Polson 和 Rossi(1994a)展示了分层的框架怎样允许 MCMC 算法方便地扩展到更一般的模型。也就是，他们估计了单变量随机波动率模型，该模型具有相关扰动以及偏斜和肥尾的方差扰动，同时还考虑了多变量模型。换句话说，MCMC 算法可以扩展到因子结构，这些因子显示了随机波动率，可以是可观测的或不可观测的。

5.7. 推断和期权价格数据

当前在文献中可以找到的一些连续时间 SV 模型，是为解决关于衍生证券的定价问题而发展起来的。给出衍生产品和 SV 扩散之间相当清楚的联系，有点令人惊讶的是，人们对运用期权价格数据估计连续时间扩散的关注相对较少。事实上，Melino(1994)在他的调查中提到：“*很清楚，关于资产价格随机波动率的信息包含在历史的资产价格和以其为标的资产的期权中。当前把这两个信息源（包括隐含估计）结合在一起的策略特别不能令人满意。从统计学来说，需要模拟期权定价中预测误差的来源，并且需要把这些误差的分布同股票价格过程联系起来*”。例如，隐含的估计（如 BS 隐含波动率的计算），从统计学的观点来说的确是特别不能令人满意。一般说来，在同定价模型进行比较的时候，每个观测的期权价格都引进了预测误差的一个来源，通过许多未观测到的状态变量对期权和资产价格的*联合*非退化概率分布建模是一个挑战。最近许多论文都在从事该方法的研究，包括 Christensen(1992)，Renault 和 Touzi(1992)，Pastorello 等(1993)，Duan(1994)以及 Renault(1995)。

Christensen(1992)考虑了一个定价模型，其中的 n 个资产是状态向量 x_t 的函数，x_t 是 $l+n$ 维的，被分成 l 维观测到的(z_t)和 n 维未观测到的(ω_t)部分。令 p_t 表示该 n 个资产的价格向量，那么：

$$p_t = m(z_t, \omega_t, \theta) \tag{5.7.1}$$

对于给定的 z_t 和 θ，(5.7.1) 提供了 n 个潜在状态变量 ω_t 和 n 个观测到的价格 p_t 之间的一一对应关系。从金融角度看，它意味着，如果假设观测到的 n 个状态变量 z_t 已经被其他（原始）资产的价格动态所模拟，那么这 n 个资产就是完整描述市场的适当工具。而且，从统计的角度看，在给定观测到的价格的对数似然函数可以很容易地从 x_t 的一个统计模型推导出时，就可以得到完全结构的极大似然估计。例如，在马尔可夫设定中，以 x_0 为条件，$x_1^T = (x_t)_{1\le t\le T}$ 的联合分布具有如下密度：

$$f_x(x_1^T | x_0, \theta) = \prod_{t=1}^{T} f(z_t, \omega_t | z_{t-1}, \omega_{t-1}, \theta) \tag{5.7.2}$$

给定 $D_0 = (p_0, z_0)$，则数据 $D_1^T = (p_t, z_t)_{1\le t\le T}$ 的条件分布可从常用的雅克比公式(Jacobian formula)得到：

$$f_D(D_1^T|D_0,\theta)=\prod_{t=1}^{T}f[z_t,m_\theta^{-1}(z_t,p_t)|z_{t-1},m_\theta^{-1}(z_{t-1},p_{t-1}),\theta]x$$

$$\times\left|\nabla_\omega m(z_t,m_\theta^{-1}(z_t,p_t),\theta)\right|^{-1} \quad (5.7.3)$$

其中$m_\theta^{-1}(z_t,\cdot)$是$m(z,.,\theta)$的$\omega-$逆，其正式定义是$m_\theta^{-1}(z,m(z,\omega,\theta))=\omega$，而$\nabla_\omega m(\cdot)$代表与$\omega$的雅克比矩阵相对应的列。利用了衍生（证券）价格数据的这一 MLE 是由 Christensen(1992)和 Duan(1994)独立提出的。Renault 和 Touzi(1992)对 Hull 和 White 的期权定价公式特别有兴趣，该公式具有：$z_t=S_t$个观测到的标的资产价格，以及$\omega_t=\sigma_t$个未观测到的随机波动率过程。于是，由于联合过程$x_t=(S_t,\sigma_t)$是马尔可夫链的，就有了买权（看涨期权）价格的形式：

$$C_t=m(x_t,\theta,K)$$

其中$\theta=(\alpha',\gamma')$包含两类参数：（1）描述联合过程$x_t=(S_t,\sigma_t)$的动态特征的参数向量α，在等鞅测度下，它能够计算关于$\gamma^2(t,t+h)$的（风险中性）条件概率分布在给定σ_t时的期望；（2）描述风险溢价特征的参数向量γ，该风险溢价决定了x过程的风险中性概率分布和数据生成过程之间的关系。

结构化的 MLE 通常很难执行。这就激发 Renault 和 Touzi(1992)以及 Pastorello、Renault 和 Touzi(1993)考虑有较低效率，但更简单且更稳健的方法，该方法涉及构造似然函数（5.7.3）的一些替代。

为了阐明这些方法，先考虑连续时间的标准对数正态 SV 模型：

$$d\log\sigma_t=k(a-\log\sigma_t)dt+cdW_t^\sigma \quad (5.7.4)$$

标准的期权定价理论允许忽略对标的资产价格过程漂移的错误设定，因此，朝简化和稳健迈出的第一步是分离波动率动态特征的似然函数，即：

$$\prod_{i=1}^{n}(2\pi c^2)^{-1/2}\exp\left[-(2c^2)^{-1}(\log\sigma_{t_i}-e^{-k\Delta t}\log\sigma_{t_{i-1}}-a(1-e^{-k\Delta t}))\right]^2 \quad (5.7.5)$$

和样本σ_{t_i}联合在一起，$i=1,...n$且$t_i-t_{i-1}=\Delta t$。为了求出该表达式的近似值，可以考虑*直接方*法，如 Renault 和 Touzi(1992)，或者间接方法，如 Pastorello 等(1993)。前者包括根据 Hull 和 White 模型计算隐含波动率，以生成参数为k、a和c的伪样本(pseudo sample)σ_{t_i}，还包括计算（5.7.5）关于那三个参数的最大值。[45]Pastorello 等人(1993)提出了在（5.7.5）背景下的几种间接推断方法，这在 5.5 节中有描述。例如，他们提出运用一个间接的推断策略，以包含从标的资产得到的 GARCH(1,1)波动率估计（也由 Engle 和 Lee(1994)独立提出）。这产生了渐近无偏但相当无效的估计量。Pastorello 等人果然发现，Renault 和 Touzi 涉及期权价格的直接方法的一个间接推断简化将有效得多，与适当的统计方法联系在一起的期权价格数据的运用，将大大提高波动率扩散参数估计的精确度，这是很清楚的直观阐述。

5.8. 具有随机波动率的回归模型

扰动项中具有随机波动率的一个简单的回归模型可以写成

$$y_t=x_t'\beta+u_t \quad , \quad t=1,...,T \quad , \quad (5.8.1)$$

其中y_t表示第t个观测值，x_t是解释变量的$k\times1$向量，β是系数的$k\times1$向量，且如第 3 节中的讨论，$u_t=\sigma\varepsilon_t\exp(0.5h_t)$。作为一个特殊情形，观测值可以简化为具有零均值，使得$x_t'\beta=\mu\ \forall t$。

[45]运用 BS 隐含波动率的（5.7.5）的直接最大化也已提出，例如 Heynen、Kemna 和 Vorst(1994)。显然，由于 BS 模型的假设，对 BS 隐含波动率的运用导致了一个错误设定偏倚。

由于u_t是平稳的，故y_t对x_t的 OLS 回归生成β的一致估计量，但不是有效的。

对于给定的 SV 参数ϕ和σ_η^2的值，h_t的一个平滑估计量$h_{t|T}$可以用 3.4 节中概述的方法中的一种来计算。（5.8.1）两边乘以$\exp(-.5h_{t|T})$得到

$$\tilde{y}_t = \tilde{x}_t'\beta + \tilde{u}_t \quad , \quad t = 1,...,T \tag{5.8.2}$$

其中，$\tilde{u}_t$可以看作是校正异方差的扰动。Harvey 和 Shephard(1993)指出，这些扰动具有零均值、常数方差，并且是序列不相关的，因此提出了构造可行的 GLS 估计量的结构

$$\tilde{\beta} = \left[\sum_{t=1}^{T} e^{-h_{t|T}} x_t x_t'\right]^{-1} \sum_{t=1}^{T} e^{-h_{t|T}} x_t y_t \tag{5.8.3}$$

在古典的异方差回归模型中，h_t是确定的，并且依赖于一个固定数目的未知参数。由于这些参数能够被一致地估计，所以可行的 GLS 估计量具有同 GLS 估计量一样的渐近分布。这里h_t是随机的，其估计量的 MSE 是$o(1)$的。因此，情况就有些不同，Harvey 和 Shephard (1993)指出，在x_t序列的正则性条件下，$\tilde{\beta}$是渐近正态的，具有均值β和一个协方差矩阵，该协方差矩阵可以被一致地估计为

$$a\,\mathrm{v\tilde{a}r}(\tilde{\beta}) = \left[\sum_{t=1}^{T} e^{-h_{t|T}} x_t x_t'\right]^{-1} \sum_{t=1}^{T} (y_t - x_t'\tilde{\beta})^2 e^{-2h_{t|T}} x_t x_t' \left[\sum_{t=1}^{T} e^{-h_{t|T}} x_t x_t'\right]^{-1} \tag{5.8.4}$$

当$h_{t|T}$是由线性状态空间给出的平滑估计量时，Harvey 和 Shephard(1993)的分析表明，渐近地，可行的 GLS 估计量几乎和 GLS 估计量一样有效，并且比 OLS 估计量有效得多。用从 3.4 节所描述方法中的一种方法计算，得到更好的估计量替换$\exp(h_{t|T})$将是可能的，但这样做，对β的可行 GLS 估计量的有效性没有很大影响。

当h_t非平稳或接近非平稳，Hansen (1995) 证明，可以构建可行的适应性强的最小二乘估计量，它渐近地等价于 GLS。

6. 结论

不存在永远完美的研究。我们期待两个领域将在未来的几年中繁荣起来，但本文未能将它们包括进来。第一个是市场微观结构领域，Goodhart 和 O'Hara(1995)在最近的一篇回顾文章中对其做了很好的综述。随着高频数据序列可得性的不断提高，我们期待有更多的研究涉及博弈论模型。由于经济计量方法的最新发展，现在这些都能估计，就象能够实现对扩散的估计一样。希望出现令人感兴趣的研究的另一个领域，是运用非参数方法估计 SV 连续时间和衍生证券模型，最近的论文包括 Ait-Sahalia(1994)，Ait-Sahalia 等(1994)，Bossaerts、Hafner 和 Härdle(1995)，Broadie 等(1995)，Conley 等(1995)，Elsheimer 等(1995)，Gouriéroux、Monfor 和 Tenreiro(1994)，Gouriéroux 和 Scaillet(1995)，Hutchinson、Lo 和 Poggio(1994)，Lezan 等(1995)，Lo(1995)，Pagan 和 Schwert(1992)。

随机波动率模型的计量经济学研究相对较新。本文的回顾已经证明，近几年来，对最新统计技术的应用有了空前的增长。至于它与 ARCH 的关系，本文的观点是，SV 和 ARCH 未必是直接竞争对手，在某些方面，它们是互补的。最近的一些研究——例如 ARCH 模型作为滤子的应用，GARCH 和时频归并的弱化，以及引入非参数模型拟合条件方差——都说明对波动率建模的统一策略需要同时利用 ARCH 和 SV。

参考文献

Abramowitz, M. and N. C. Stegun (1970). Handbookof Mathematical Functions. Dover Publications Inc., New York.

Ait-Sahalia, Y. (1994). Nonparametric pricing of interest rata derivative securities. Discussion Paper, Graduate School of Business, Univcrsity of Chicago.

Ait-Sahalia, Y. S. J. Bickel and T. M. Stoker (I994). Goodness-of-Fit tests for regression using kernel methods. Discussion Paper, University of Chicago.
Amin, K. L. and V. Ng (1993). Equilibrium option valuation with systematic stochastic volatility. J. Finance 48, 881-910. .
Andersen, T. G. (1992). Volatility. Discussion paper, Northwestern University.
Andersen, T. G. (1994). Stochastic autoregresaive volatility: A framework for volatility modeling. Math. Finance 4, 75-102.
Andersen, T. G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. J. Finance, to appear.
Andersen, T. G. and T. Bollerslev (1995). Intraday seasonality and volatility persistence in financial Markets. J. Emp. Finance, to appear.
Andersen, T. G. and B. Sørensen (1993). GMM estimation of a stochastic volatility model: A Monte Carlo study. J. Business Econom. Statist. to appear.
Andersen, T. G. and B. Sørensen (1996). GMM and QML asymptotic standard deviations in stochastic volatility models: A response to Ruiz (1994). J. Econometrics, to appear.
Andrews, D. W. K. (1993). Exactly median-unbiased estimation of first order autoregresaive unit root models. Econometrica 61, 139--165.
Bachelier, L. (1900), Théorie de la spéculation. Ann. Sci. Ecole Norm. Sup. 17, 21-86, [On the Random Character of Stock Market Prices (Paul H. Cootner, ed.) The MIT Press, Cambridge, Mass. 1964].
Baillie, R. T. and T. Bollerslev (1989). The message in daily exchange rates: A conditional variance tale. J. Business Econom. Statist. 7, 297-305.
Baillie, R. T. and T. Bollerslev (1991), Intraday and Interday volatility in foreign exchange rates. Rev. Eeonom. Stud. 58, 565-585.
Baillie, R. T., T. Bollerslev and H. O. Mikkelsen (1993). Fractionally integrated generalized autoregressive conditional hcteroskedasticity. J. Econometrics, to appear.
Bajeux, I. and J. C. Rochet (1992). Dynamic spanning: Are options an appropriate instrument? Math. Finance, to appear.
Bates, D. S. (1995a). Testing option pricing models. In: G. S. Maddala ed., Handbook af Statistics, Vol. 14, Statistical Methods in Finance. North Holland, Amsterdam, in this volume.
Bates, D. S. (1995b). Jumps and stochastic volatility: Exchange rate processes implicit in PHLX Deutschemark options. Rev. Financ. Stud., to appear.
Beckers, S. (1981). Standard deviations implied in option prices as predictors of future stock price variability. J. Banking Finance 5, 363-381,
Bera, A. K. and M. L. Higgins (1995). On ARCH models: Properties, estimation and testing. In: L. Exley, D.A. R. George, C. J.Roberts and S. Sawyer eds., Surveys in Econometrics. Basil Blackwell: Oxford, Reprinted from J. Econom. Surveys.
Black, F. (1976). Studies in stock price volatility changes. Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section, Amer. Statist. Assoc. 177-181.
Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. J. Politic. Econom. 81, 637-654.
Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. J. Econometrics 31, 307-327.
Bollerslev, T., Y. C. Chou and K. Kroner (1992). ARCH modelling in finance: A selective review of the theory and empirical evidence. J. Econometrics 52, 201-224.
Bollerslev, T. and R. Engle (1993). Common persistence in conditional variances. Econometrica 61, 166-187.
Bollerslev, T., R. Engle and D. Nelson (1994). ARCH models. In: R. F. Engle and D. McFadden eds., Handbook of Econometrics, Volume IV. North-Holland, Amsterdam.
Bollerslev, T., R. Engle and J. Wooldridge (1988). A capital asset pricing model with time varying covariances. J. Politic. Econom. 96, 116-131.
Bollerslev, T. and E. Ghysels (1994). On periodic autoregression conditional heteroskedasticity. J. Business Econom. Statist., to appear.
Bollerslev, T. and H. O. Mikkelsen (1995). Modeling and pricing long-memory in stock market volatility. J. Econometrics, to appear.
Bossaerta, P., C. Hafner and W. Härdle (1995). Foreign exchange rates have surprising volatility. Discussion Paper, CentER, University of Tilburg.
Bossaerts, P. and P. Hillion (1995). Local parametric analysis of hedging in discrete time. J. Econo-

metrics, to appear.
Breidt, F. J., N. Crato and P. de Lima (1993). Modeling long-memory stochastic volatility. Discussion paper, Iowa State University.
Breidt, F. J. and A. L. Carriquiry (1995). Improved quasi-maximum likelihood estimation for stochastic volatility models. Mimeo, Department of Statistics, University of Iowa.
Broadie, M., J. Detemple, E. Ghysels and O, Torrès (1995). American options with stochastic volatility: A nonparametric approach. Discussion Paper, CIRANO.
Broze, L., O. Scaillet and J. M, Zakoian (1994). Quasi indirect inference for diffusion processes. Discussion Paper CORE.
Broze, L., O. Scaillet and J. M. Zakoian (1995). Testing for continuous time models of the short term interest rate. J. Emp. Finance, 199-223.
Campa, J. M. and P. H. K. Chang (1995). Testing the expectations hypothesis on the term structure of implied volatilities in foreign exchange options. J. Finance 50, to appear.
Campbell, J. Y. and A. S. Kyle (1993). Smart money, noise trading and stock price behaviour. Rev. Econom. Stud. 60, 1-34.
Canine, L. and S. Figlewski (1993). The informational content of implied volatility. Rev. Financ. Stud. 6, 659-682.
Canova, F. (1992). Detrending and Business Cycle Facts. Discussion Paper, European University Institute, Florence.
Chesney, M. and L. Scott (1989). Pricing European currency option: A comparison of the modified Black-Scholes model and a random variance model. J. Financ. Quant. Anal. 24, 267-284.
Cheung, Y.-W. and F. X. Diebold (1994). On maximum likelihood estimation of the differencing parameter of fractionally-integrated noise with unknown mean. J. Econometrics 62, 301-316.
Chiras, D. P. and S. Manaster (1978). The information content of option prices and a test of market efficiency. J. Financ. Econom. 6, 213-234.
Christensen, B. J. (1992). Asset prices and the empirical martingale model. Discussion Paper, New York University.
Christie, A. A. (1982). The stochastic behavior of common stock variances: Value, leverage, and interest rate effects. J. Financ. Econom. 10, 407-432.
Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. Econometrica 41, 135-156.
Clewlow, L and X. Xu (1993). The dynamics of stochastic volatility. Discussion Paper, University of Warwick.
Comte, F. and E. Renault (1993). Long memory continuous time models. J. Econometrics, to appear.
Comte, F. and E. Renault (1995). Long memory continuous time stochastic volatility models. Paper presented at the HFDF-I Conference, Zürich.
Conley, T., L. P. Hansen, E. Luttmer and J. Scheinkman (1995). Estimating subordinated diffusions from discrete time data. Discussion paper, University of Chicago.
Cornell, B. (1978). Using the options pricing model to measure the uncertainty producing effect of major announcements. Finane. Mgmt. 7, 54-59.
Cox, J. C. (1975). Notes on option pricing I: Constant elasticity of variance diffusions. Discussion Paper, Stanford University.
Cox, J. C. and S. Ross (1976). The valuation of options for alternative stochastic processes. J. Financ. Econom. 3, 145-166.
Cox, J. C. and M. Rubinstein (1985). Options Markets. Englewood Cliffs, Prentice-Hall, New Jersey.
Dacorogna, M. M., U. A. Müller, R. J. Nagler, R. B. Olsen and O. V. Pictet 0993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. J. Internat. Money Finance 12, 413-438.
Danielsson, J. (1994). Stochastic volatility in asset prices: Estimation with simulated maximum likelihood. J. Econometrics 61,375-400.
Danielsson, J. and J. F. Richard (1993). Accelerated Ganssian importance sampler with application to dynamic latent variable models. J. Appl. Econometrics 3, S153-S174.
Dassios, A. (1995). Asymptotic expressions for approximations to stochastic variance models. Mimeo, London School of Economics.
Day, T. E. and C. M. Lewis (1988). The behavior of the volatility implicit in the prices of stock index options. J. Financ. Econom. 22, 103-122.
Day, T. E. and C. M. Lewis (1992). Stock market volatility and the information content of stock index

options. J. Econometrics 52, 267-287.
Diebold, F. X. (1988). Empirical Modeling of Exchange Rate Dynamics. Springer Verlag, New York.
Diebold, F. X. and J. A. Lopez (1995). Modeling Volatility Dynamics. In: K. Hoover ed., Macro-economics: Developments, Tensions and Prospects.
Diebold, F. X. and M. Nerlove (1989). The dynamics of exchange rate volatility: A multivariate latent factor ARCH Model. J. Appl. Econometrics 4, 1-22.
Ding, Z., C. W. J. Granger and R. F. Engle (1993). A long memory property of stock market returns and a new model. J. Emp. Finance 1, 83-108.
Diz, F. and T. J. Finucane (1993). Do the options markets really overreact? J. Futures Markets 13, 298-312.
Drost, F. C. and T. E. Nijman (1993). Temporal aggregation of GARCH processes. Econometrica 61, 909-927.
Drost, F. C. and B. J. M. Werker (1994). Closing the GARCH gap: Continuous time GARCH modelling. Discussion Paper CentER, University of Tilburg.
Duan, J. C. (1994). Maximum likelihood estimation using price data of the derivative contract. Math. Finance 4, 155-167.
Duan, J. C. (1995). The GARCH option pricing model. Math. Finance 5, 13-32.
Duffie, D. (1989). Futures Markets. Prentice-Hall International Editions.
Duffie, D. (1992). Dynamic Asset Pricing Theory. Princeton University Press.
Dnitic, D. and K. J. Singleton (1993). Simulated moments estimation of Markov models of asset prices. Econometrica 61, 929--952.
Dunsmuir, W. (1979). A central limit theorem for parameter estimation in stationary vector time series and its applications to models for a signal observed with noise. Ann. Statist. 7, 490-506.
Easley, D. and M. O'Hara (1992). Time and the process of security price adjustment. J. Finance, 47, 577-605.
Ederington L. H. and J. H. Lee (1993). How markets process information: News releases and volatility. J. Finance 48, 1161-1192.
Elsheimer, B., M. Fisher, D. Nychka and D. Zirvos (1995). Smoothing splines estimates of the discount function based on US bond Prices. Discussion Paper Federal Reserve, Washington, D.C.
Engle, R. F. (1982). Autorcgressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987-1007.
Eagle, R. F. and C. W. J. Granger (1987). Cointegration and error correction: Representation, estimation and testing Econometrica 55, 251-576.
Engle, R. F. and S. Kozicki (1993). Testing for common features. J. Business Econom. Statist. 11, 369-379.
Engle, R. F. and G. G. J. Lee (1994). Estimating diffusion models of stochastic volatility. Discussion Paper, University of California at San Diego.
Engle, R. F. and C. Mustafa (1992). Implied ARCH models from option prices. J. Econometrics 52, 289-311.
Engle, R. F. and V. K. Ng (1993). Measuring and testing the impact of news on volatility. J. Finance 48, 1749-1801.
Fama, E. F. (1963). Mandelbrot and the stable Paretian distribution. J. Business 36, 420-429.
Fama, E. F. (1965). The behavior of stock market prices. J. Business 38, 34-105.
Foster, D. and S. Viswanathan (1993a). The effect of public information and competition on trading volume and price volatility. Rev. Financ. Stud. 6, 23-56.
Foster, D. and S. Viswanathan (1993b). Can speculative trading explain the volume volatility relation. Discussion Paper, Fuqua School of Business, Duke University.
French, K. and R. Roll (1986). Stock return variances: The arrival of information and the reaction of traders. J. Financ. Econom. 17, 5-26.
Gallant, A. R., D. A. Hsieh and G. Tauchen (1994). Estimation of stochastic volatility models with suggestive diagnostics. Discussion Paper, Duke University.
Gallant, A. R., P. E. Rossi and G. Tauchen (1992). Stock prices and volume. Rev. Financ. Stud. 5, 199-242.
Gallant, A. R., P. E. Rossi and G. Tauchen (1993). Nonlinear dynamic structures. Econometrica 61, 871-907.
Gallant, A. R. and G. Tauchen (1989). Semiparametric estimation of conditionally constrained het-erogeneous processes: Asset pricing applications. Econometrica 57, 1091-1120.

Gallant, A. R. and G. Tanchen (1992). A nonparametric approach to nonlinear time series analysis: Estimation and simulation. In: E. Parzen, D. Brillinger, M. Rosenblatt, M. Taqqu, J. Geweke and P. Caines eds., New Dimensions in Time Series Analysis. Springer-Verlag, New York.

Gallant, A. R. and G. Tauchen (1994). Which moments to match. Econometric Theory, to appear.

Gallant, A. R. and G. Tauchen (1995). Estimation of continuous time models for stock returns and interest rates. Discussion Paper, Duke University.

Gallant, A. R. and H. White (1988). A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models. Basil Blackwell, Oxford.

Garcia, R and E. Renault (1995). Risk aversion, intcrtemporal substitution and option pricing. Discussion Paper CIRANO.

Geweke, J. (1994). Comment on Jacquier, Polson and Rossi. J. Business Econom. Statist. 12, 397-399.

Geweke, J. (1995). Monte Carlo simulation and numerical integration. In: H. Amman, D. Kendrik, and J. Rust eds., Handbook of Computational Economics. North Holland.

Ghysels, E., C. Gouriéroux and J. Jasiak (1995a). Market time and asset price movements: Theory and estimation. Discussion paper CIRANO and C.R.D.E., Univeristé de Montréal.

Ghyseis, E., C. Gouriéroux and J. Jasiak (1995b). Trading patterns, time deformation and stochastic volatility in foreign exchange markets. Paper presented at the HFDF Conference, Zürich.

Ghysels, E. and J. Jasiak (1994a). Comments on Bayesian analysis of stochastic volatility models. J. Business Econom. Statist. 12, 399--401.

Ghysels, E. and J. Jasiak (1994b). Stochastic volatility and time deformation an application of trading volume and leverage effects. Paper presented at the Western Finance Association Meetings, Santa Fe.

Ghysels, E., L. Khalaf and C. Vodounou (1994). Simulation based inference in moving average models. Discussion Paper, CIRANO and C.R.D.E.

Ghysels, E., H. S. Lee and P. Siklos (1993). On the (mis)specification of seasonality and its consequences: An empirical investigation with U.S. Data. Empirical Econom. 18, 747-760.

Goodhart, C. A. E. and M. O'Hara (1995). High frequency data in financial markets: Issues and applications. Paper presented at HFDF Conference, Zürich.

Gouriéroux, C. and A. Monfort (1993a). Simulation based Inference: A survey with special reference to panel data models. J. Econometrics 59, 5-33.

Gouriéroux, C. and A. Monfort (1993b). Pseudo-likelihood methods in Maddala et al. ed., Handbook of Statistics Vol. I I, North Holland, Amsterdam.

Gouriéroux, C. and A. Monfort (1994). Indirect inference for stochastic differential equations. Diseussion Paper CREST, Paris.

Gouriéroux, C. and A. Monfort (1995). Simulation-Based Econometric Methods. CORE Lecture Series, Louvain-la-Neuve.

Gouriéroux, C., A. Monfort and E. Renault (1993). Indirect inference, J. Appl Econometrics 8, S85-Sll8.

Gouriéroux, C., A. Monfort and C. Tenreiro (1994). Kernel M-estimators: Nonparametric diagnostics for structural models. Discussion Paper, CEPREMAP.

Gouriéroux, C., A. Monfort and C. Tenreiro (1995). Kernel M-estimators and functional residual plots. Discussion Paper CREST - ENSAE, Paris.

Gouriéroux, C., E. Renault and N. Touzi (1994). Calibration by simulation for small sample bias correction. Discussion Paper CREST.

Gouriéroux, C. and O. Scaillet (1994). Estimation of the term structure from bond data. J. Emp. Finance, to appear.

Granger, C. W. J. and Z. Ding (1994). Stylized facts on the temporal and distributional properties of daily data for speculative markets. Discussion Paper, University of California, San Diego.

Hall, A. R. (1993). Some aspects of generalized method of moments estimation in Maddala et al. ed., Handbook of Statistics Vol. 11, North Holland, Amsterdam.

Hamao, Y., R. W. Masulis and V. K. Ng (1990). Correlations in price changes and volatility across international stock markets. Rev. Financ. Stud. 3, 281-307.

Hansen, B. E. (1995). Regression with nonstationary volatility. Econometrica 63, 1113-1132.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. Econometrica 50, 1029-1054.

Hausen, L. P. and J. A. Scheinkman (1995). Back to the future: Generating moment implications for continuous-time Markov processes. Econometrica 63, 767-804.

Harris, L. (1986). A transaction data study of weekly and intradaily patterns in stock returns. J. Financ. Econom. 16, 99-117.
Harrison, M. and D. Kreps (1979). Martingale and arbitrage in multiperiod securities markets. J. Econom. Theory 20, 381-408.
Harrison, J. M. and S. Pliska (1981). Martingales and stochastic integrals in the theory of continuous trading. Stochastic Processes and Their Applications 11, 215-260.
Harrison, P. J. and C. F. Stevens (1976).-Bayesian forecasting (with discussion). J. Roy. Statis. Soc.. Ser. B, 38, 205-247.
Harvey, A. C. (1989). Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press.
Harvey, A. C. and A. Jaeger (1993). Detrending, stylized facts and the business cycle. J. Appl. Econometrics 8, 231-247.
Harvey, A. C. (1993). Long memory in stochastic volatility. Discussion Paper, London School of Economics.
Harvey, A. C. and S. J. Koopman (1993). Forecasting hourly electricity demand using time-varying splines. J. Amer. Statist. Assoc. 88, 1228-1236.
Harvey, A. C., E. Ruiz and E. Sentana (1992). Unobserved component time series models with ARCH Disturbances, J. Econometrics 52, 129-158.
Harvey, A. C., E. Rniz and N. Shepherd (1994). Multivariate stochastic variance models. Rev. Econom. Stud. 61, 247-264.
Harvey, A. C. and N. Shepherd (1993). Estimation and testing of stochastic variance models, STI-CERD Econometrics. Discussion paper, EM93/268, London School of Economics.
Harvey, A. C. and N. Shephard (1996). Estimation of an asymmetric stochastic volatility model for asset returns. J. Business Econom. Statist. to appear.
Harvey, C. R. and R. D. Huang (1991). Volatility in the foreign currency futures market. Rev. Financ. Stud. 4, 543-569.
Harvey, C. R. and R. D. Huang (1992). Information trading and fixed income volatility. Discussion Paper, Duke University.
Harvey, C. R. and R. E. Whaley (1992). Market volatility prediction and the efficiency of the S&P 100 index option market. J. Financ. Econom. 31, 43-74.
Hausman, J. A. and A. W. Lo (1991). An ordered probit analysis of transaction stock prices. Discussion paper, Wharton School, University of Pennsylvania.
He, H. (1993). Option prices with stochastic volatilities: An equilibrium analysis. Discussion Paper, University of California, Berkeley.
Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev. Financ. Stud. 6, 327-343.
Heynen, R., A. Kemna and T. Vorst (1994). Analysis of the term structure of implied volatility. J. Financ. Quant. Anal.
Hull, J. (1993). Options, futures and other derivative securities. 2nd ed. Prentice-Hall International Editions, New Jersey.
Hull, J. (1995). Introduction to Futures and Options Markets. 2nd ed. Prentice-Hall, Englewood Cliffs, New Jersey.
Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. J. Finance 42, 281-300.
Huffman, G. W. (1987). A dynamic equilibrium model of asset prices and transactions volume. J. Politic. Econom. 95, 138-159.
Hutchinson, J. M., A. W. Lo and T. Poggio (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. J. Finance 49, 851-890.
Jacquier, E., N. G. Polson and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models (with discussion). J. Business Econom. Statist. 12, 371--417.
Jacquier, E., N. G. Polson and P. E. Ross (1995a). Multivariate and prior distributions for stochastic volatility models. Discussion paper CIRANO.
Jacqnier, E., N. G. Poison and P. E. Rossi (1995b). Stochastic volatility: Univariate and multivariate extensions. Rodney White center for financial research. Working Paper 19-95, The Wharton School, University of Pennsylvania.
Jacquier, E., N. G. Polson and P. E. Rossi (1995c). Efficient option pricing under stochastic volatility. Manuscript, The Wharton School, University of Pennsylvania.

Jarrow, R. and Rudd (1983). Option pricing. Irwin, Homewood III.
Johnson, H. and D. Shanno (1987). Option pricing when the variance is changing. J. Financ. Quant. Anal. 22, 143-152.
Jorion, P. (1995). Predicting volatility in the foreign exchange market. J. Finance 50, to appear.
Karatzas, I. and S. E. Shreve (1988). Brownian Motion and Stochastic Calculus. Springer-Verlag: New York, NY.
Karpoff, J. (1987). The relation between price changes and trading volume: A survey. J. Financ. Quant. Anal. 22, 109-126.
Kim, S. and N. Shephard (1994). Stochastic volatility: Optimal likelihood inference and comparison with ARCH Model. Discussion Paper, Nuffield College, Oxford.
King, M., E. Sentana and S. Wadhwani (1994). Volatility and links between national stock markets. Econometrica 62, 901-934.
Kitagawa, G. (1987). Non-Ganssian state space modeling of nonstationary time series (with discussion). J..Oner. Statist. Assoc. 79, 378-389.
Kloeden, P. E. and E. Platten (1992). Numerical Solutions of Stochastic Differential Equations. Springer-Verlag, Heidelberg.
Lamoureux, C. and W. Lastrapes (1990). Heteroskedasticity in stock return data: Volume versus GARCH effect. J. Finance 45, 221-229.
Lamoureux, C. and W. Lastrapes (1993). Forecasting stock-return variance: Towards an understanding of stochastic implied volatilities. Rev. Financ. Stud. 6, 293-326.
Latane, H. and R. Jr. Rendleman (1976). Standard deviations of stock price ratios implied in option prices. J. Finance 31, 369-381.
Lezan, G., E. Renault and T. deVitry (1995) Forecasting foreign exchange risk. Paper presented at 7th World Congres of the Econometric Society, Tokyo.
Lin, W. L., R. F. Engle and T. Ito (1994). Do bulls and bears move across borders? International transmission of stock returns and volatility as the world turns. Rev. Financ. Stud., to appear.
Lo, A. W. (1995). Statistical inference for technical analysis via nonparametric estimation. Discnssion Paper, MIT.
Mahieu, R. and P. Schotman (1994a). Stochastic volatility and the distribution of exchange rate news. Discussion Paper, University of Limburg.
Mahieu R. and P. Schotman (1994b). Neglected common factors in exchange rate volatility. J. Emp. Finance 1, 279-311.
Mandelbrot, B. B. (1963). The variation of certain speculative prices. J. Business 36, 394-416.
Mandelbrot, B. and H. Taylor (1967). On the distribution of stock prices differences. Oper..Res. 15, 1057-1062.
Mandelbrot, B. B. and J.W. Van Ness (1968). Fractal Brownian motions, fractional noises and applications. SIAM Rev. 10, 422-437.
McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. Econometrics 57, 1027-1057.
Meddahi, N. and E. Renault (1995). Aggregations and marginalisatious of GARCH and stochastic volatility models. Discussion Paper, GREMAQ.
Melino, A. and M. Turnbull (1990). Pricing foreign currency options with stochastic volatility. J. Econometrics 45, 239-265.
Metino, A. (1994). Estimation of continuous time models in finance. In: C.A. Sims ed., Advances in Econometrics (Cambridge University Press).
Merton, R. C. (1973). Rational theory of option pricing. Bell J. Econom. Mgrnt. Sci 4, 141-183.
Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. J. Financ. Econom. 3, 125--144.
Merton, R. C. (1990). Continuous Time Finance. Basil Blackwell, Oxford.
Merville, L. J. and D. R. Pieptea (1989). Stock-price volatility, mean-reverting diffusion, and noise. J. Financ. Econom. 242, 193--214.
Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller (1954). Equation of state calculations by fast computing machines. J. Chem. Physics 2l, 1087-1092.
Müller, U. A., M. M. Dacorogna, R. B. Olsen, W. V. Pictet, M. Schwarz and C. Morgenegg (1990). Statistical study of foreign exchange rates. Empirical evidence of a price change scaling law and intraday analysis. J. Banking Finance 14, 1189-1208.
Nelson, D.B. (1988). Time series behavior of stock market volatility and returns. Ph.D. dissertation,

Nelson, D. B. (1990). ARCH models as diffusion approximations. J. Econometrics 45, 7-39.
Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59, 347-370.
Nelson, D. B. (1992). Filtering and forecasting with misspecified ARCH Models I: Getting the right variance with the wrong model. J. Econometrics 25, 61-90.
Nelson, D. B. (1994). Comment on Jacquier, Polson and Rossi. J. Business Econom. Statist. 12, 403-406.
Nelson, D. B. (1995a). Asymptotic smoothing theory for ARCH Models. Econometrica, to appear.
Nelson, D. B. (1995b). Asymptotic filtering theory for multivariate ARCH models. J. Econometrics, to appear.
Nelson, D. B. and D. P. Foster (1994). Asymptotic filtering theory for univariate ARCH models. Econometrica 62, 1-41.
Nelson, D. B. and D. P. Foster (1995). Filtering and forecasting with misspecified ARCH models II: Making the right forecast with the wrong model. J. Econometrics, to appear.
Nob, J., R. F. Engle and A. Kane (1994). Forecasting volatility and option pricing of the S&P 500 index. J. Derivatives, 17-30.
Ogaki, M. (1993). Generalized method of moments: Econometric applications. In: Maddala et al. ed., Handbook of Statistics Vol. 11, North Holland, Amsterdam.
Pagan, A. R. and G. W. Schwert (1990). Alternative models for conditional stock volatility, J. Econometrics 45, 267-290.
Pakes, A. and D. Pollard (1989). Simulation and the asymptotics of optimization estimators. Econometrica 57, 995-1026.
Pardoux, E. and D. Talay (1985). Discretization and simulation of stochastic differential equations. Acta Appl. Math~ 3, 23-47.
Pastorello, S., E. Renault and N. Touzi (1993). Statistical inference for random variance option pricing. Discussion Paper, CREST.
Patell, J. M. and M. A. Wolfson (1981). The ex-ante and ex-post price effects of quarterly earnings announcement reflected in option and stock price. J. Account. Res. 19, 434-458.
Patell, J. M. and M. A. Wolfson (1979). Anticipated information releases reflected in call option prices. J. Account. Econom. 1, 117-140.
Pham, H. and N. Touzi (1993). Intertemporal equilibrium risk premia in a stochastic volatility model. Math. Finance, to appear.
Platten, E. and Schweizer (1995). On smile and skewness. Discussion Paper, Australian National University, Canberra.
Poterba, J. and L. Summers (1986). The persistence of volatility and stock market fluctuations, Amer. Econom. Rev. 76, 1142-1151.
Renault, E. (1995). Econometric models of option pricing errors. Invited Lecture presented at 7th W.C.E.S., Tokyo, August.
Renault, E. and N. Touzi (1992). Option hedging and implicit volatility. Math. Finance, to appear.
Revuz, A. and M. Yor (1991). Continuous Martingales and Brownian Motion. Springer-Verlag, Berlin.
Robinson, P. (1993). Efficient tests of nonstationary hypotheses. Mimeo, London School of Economics.
Rogers, L. C. G. (1995). Arbitrage with fractional Brownian motion. University of Bath, Discussion paper.
Rubinstein, M. (1985). Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23, 1976 through August 31, 1978. J. Finance 40, 455-480.
Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic volatility models. J. Econometrics 63, 289-306.
Schwert, G. W. (1989). Business cycles, financial crises, and stock volatility. Carnegie-Rochester Conference Series on Public Policy 39, 83-126.
Scott, L. O. (1987). Option pricing when the variance changes randomly: Theory, estimation and an application. J. Financ. Quant. Anal 22, 419-438.
Scott, L. (1991). Random variance option pricing. Advances in Futures and Options Research, Vol. 5, 113-135.
Sheikh, A. M- (1993). The behavior of volatility expectations and their effects on expected returns. J. Business 66, 93-116.

Shephard, N. (1995). Statistical aspect of ARCH and stochastic volatility. Discussion Paper 1994, Nuffield College, Oxford University.
Sims, A. (1984). Martingale-like behavior of prices. University of Minnesota.
Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. J. Econometrics 53, 165--188.
Stein, J. (1989): Overreaetions in the options market. J. Finance 44, 1011-1023.
Stein, E. M. and J. Stein (1991). Stock price distributions with stochastic volatility: An analytic approach. Rev. Ftnanc. Stud. 4, 727-752.
Stock, J. H. (1988). Estimating continuous time processes subject to time deformation. J. Amer. Statist. Assoc. 83, 77-84.
Strook, D. W. and S. R. S. Varadhan (1979). Multi-dimensional Diffusion Processes. Springer-Verlag, Heidelberg.
Tanner, T. and W. Wong (1987). The calculation of posterior distributions by data augmentation. J. Amer. Statist. Assoc. 82, 528-549.
Tauchen, G. (1995). New minimum chi-square methods in empirical finance, invited Paper presented at the 7th World Congress of the Econometric Society, Tokyo.
Tauchen, G. and M. Pitts (983). The price variability-volume relationship on speculative markets. Econometrica 51, 485-505.
Taylor, S. J. (1986). Modeling Financial Time Series. John Wiley: Chichester.
Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. Math. Finance 4, 183-204.
Taylor, S. J. and X. Xu (1994). The term structure of volatility implied by foreign exchange options. J. Financ. Quant Anal. 29, 57-74.
Taylor, S. J. and X. Xu (1993). The magnitude of implied volatility smiles: Theory and empirical evidence for exchange rates. Discussion Paper, University of Warwick.
Von Furstenberg, G. M. and B. Nam Jeon (1989). International stock price movements: Links and messages. Brookings Papers on Economic Activity I,125-180.
Wang, J. (1993). A model of competitive stock trading volume. Discussion Paper, MIT.
Watanabe, T. (1993). The time series properties of returns, volatility and trading volume in financial markets. Ph.D. Thesis, Department of Economics, Yale University.
West, M. and J. Harrison (1990). Bayesian Forecasting and Dynamic Models. Springer-Verlag, Berlin.
Whaley, R. E. (1982). Valuation of American call options on dividend-paying stocks. J. Financ. Econom. 10, 29-58.
Wiggins, J. B. (1987). Option values under stochastic volatility: Theory and empirical estimates. J. Financ. Econom. 19, 351-372.
Wood, R. T. McInish and J. K. Ord (1985). An investigation of transaction data for NYSE Stocks. J. Finance 40, 723-739.
Wooldridge, J. M. (1994). Estimation and inference for dependent processes. In: R.F. Engle and D. McFadden eds., Handbook of Econometrics Vol. 4. North Holland, Amsterdam.