译文

关于资本市场反应不足、动量交易和过度反应的统一理论（HS）

一、导言

过去几年里，大量的实证工作已经证明了有许多方法可以利用获得的公开信息来预测资产收益。尽管不同的研究使用的预测变量不同，但我们认为大多数结论都属于针对两类现象的概括范畴：一种现象是收益表现出中短期连续性；另一种是收益在长期内有逆转或根本性反转的趋势。

Barberis, Shleifer 和 Vishny (BSV,1998), 及 Daniel, Hirshleifer 和 Subrahmanyam (DHS, 1998)都假定价格由一组单独的有代表性的作用者形成，而且这些代表性作用者可能有一些认知的偏差，并研究了这种偏差足以同时导致短期连续性和长期反转的程度。

文中，我们继续 BSV 和 DHS 的同一目标，建立了统一行为模型。但我们选择了一种在根本上不同的方法。我们没有试图过多讨论代表性作用者的心理，我们的重点放在不同作用者的相互作用上。宽泛地，我们的模型中投资者行为较少是来自于个别投资者的认知偏差，而更多地来自于他们的相互作用机制。

更明确地说，我们的模型的特点在于假设了两组投资者：“消息观察者”和“动量交易者”。两组不是通常意义上的完全理性者，而是有限理性的，具有形成一种简单形式的有限性：每个组都只能“处理”所有可获得的公开信息中的一个子集。消息观察者基于他们私自观测到的关于未来基本情况的信号来做出预测，局限性是他们不能根据当前和过去价格的信息进行预测。动量交易者正好相反，他们可以根据过去价格变化做出预测，但预测是过去价格的简单（单变量）函数。

除了对投资者施加信息处理能力的限制外，我们进一步假设，这个假设在本质上更为保守：私人信息在信息观测者之中逐步扩散。这样，我们所有的结论都将从这三个关键的假设中得出。一开始，模型中只有消息观察者活跃在市场，价格慢慢调整到新的信息——这就会反应不足，而决不是反应过度。正如后面讨论的，在消息观察者不从价格中发现信息的假设下，这个结果随着消息的逐渐扩散自然得出。

接着，我们引入动量交易者。由于动量交易者能依赖过去的价格做判断，所以很容易推出他们会对消息观察者的反应不足套利；若有足够的风险承受，人们可能预期动量交易者能推动市场使市场更接近有效。但是，如果动量交易者仅限于采取简单策略，结果是这种直觉是不完全的，例如，假设一个动量交易者在**t**时刻可能仅依靠过去一些靠前的时间间隔，如从**t-2**到**t-1**的价格变化作为交易判断的基础。我们指出在这种情况下，动量交易者的套利努力会导致错误的结果：最初价格沿基本消息方向的反应实际上被加速，结果是制造了最后的对任何消息的过度反应，这即使当动量交易者是风险中性时也成立。

尽管我们作出了两种截然不同的有限理性的假设，在下述意义上，我们的模型仍称得上是统一了反应不足和过度反应。一开始，对一组投资者，模型设立的价格趋势是对私人信息的反应不足。接着我们提出当第二组投资者试图利用这种反应不足进行简单策略的套利时，他们只部分地抵消了反应不足，并在这个过程中制造了过度的价格冲量，从而不可避免地积累成反应过度。因而，通过使动量交易者进入市场有利可图，反应不足为过度反应埋下了祸

根。或者换个说法，统一性在于我们的模型把反应不足和过度反应都归结为关于基本信息的逐渐扩散。而不包括其他的对投资者情感刺激和流动性交易的需要。

下面，我们将建立一个简单的无期限模型来表达这个理念。

二、模型

（一）只有消息观察者时的价格形成

如前所述，我们的模型中有两类投资者，消息观察者和动量交易者。我们从描述只有消息观察者存在时的模型开始。每个**t**时刻，消息观察者对一风险资产进行投资。这个资产将在**T**时刻支付清算股利。最终，清算股利的价值为：$D_T = D_0 + \sum_{j=0}^{T} e_j$。$e$ 服从独立分布，零均值，方差为 s^2 的正态分布。T趋于无穷大，以使交易策略呈稳态，不用考虑离到期日有多久。

为了表示消息如何在消息观察者中扩散，我们把消息观察者分为z个规模一样的组。再假设每个股息的变化 e_j 可被分解为z个独立的子变化，每个子变化都有相同的方差 $\frac{s^2}{z}$：$e_j = e_j^1 + e_j^2 + ... + e_j^z$。消息扩散的时间选择如下。t时刻 e_{t+z-1} 开始传播，组1观察到 e_{t+z-1}^1，组2观察到 e_{t+z-1}^2，……，组z观察到 e_{t+z-1}^z。这样，在t时刻，每个 e_{t+z-1} 的子变化都被 $1/z$ 的人观察到。

接着，在t+1时刻，观察者开始旋转，组1观察到 e_{t+z-1}^2，组2观察到 e_{t+z-1}^3，直到组z，观察到 e_{t+z-1}^1。这样，每个 e_{t+z-1} 的子变化可以被 $2/z$ 的人观察到。这个过程一直到t+z-1时刻，z个组的每个人都直接观察到了 e_{t+z-1} 的所有子变化。所以t+z-1时刻，e_{t+z-1} 被完全公开。尽管这个过程很烦琐，但这个旋转的特征很有用。因为它表示了信息的逐渐扩散直至每个人都得知。在这部分中，参数z代表信息流动率，z值越大，信息扩散的越慢。我们马上将看到，这个对称使价格的解变得清晰简单。

所有消息观察者的效用都是风险厌恶系数相同的恒定风险厌恶（CARA）的函数，（CARA：$U = -e^{-(2g^w)}$，其中，g 是绝对风险厌恶系数，w是最终财富），直到到期日T所有的消息观察者都存在。无风险利率为0。资产供给不变，固定为Q。到此为止，所以这些假设都是完全传统的假设。接着我们提出两个较不传统的假设。首先，每个t时刻，消息观察者对资产的需求是基于静态最优化作出的决定，并一直持有资产到T时刻。第二，更为关键，消息观察者可以依赖前面描述的消息的子集，但他们不利用当前或过去的价格反映的信息。也就是说，我们的均衡是瓦尔拉均衡。与完全理性下的预期均衡不同，瓦尔拉均衡实质上是一种实物均衡而不是货币均衡。

正如介绍中提到的，这两个无条件的假设都基于有限理性的简单形式。人们可以把消息观察者看成只计算 e 的意义对最终红利 D_T 的影响，而不用当前的和过去市场的价格对 D_T 进行更复杂的预测，也不对未来价格的变动进行预测，因而无法执行动态的策略。

在这些假设下，结合前面描述的消息扩散的对称结构，基本信息的条件方差对所以消息观察者是相同的，则t时刻的股价可以表示为：

$$P_t = D_T + \{(z-1)e_{t+1} + (z-2)e_{t+2} + ... + e_{t+z-1}\}/z - qQ \qquad (1)$$

其中，q 表示消息观察者的风险厌恶系数与 e 的方差 s_e^2 的函数。标准化 q 以简化模型：q =1。方程（1）说明了新消息在z期内对价格的影响是线性的。这暗示在z期（短期）内，收益呈连续正相关。同时可以看出，此时的价格不会超过长期价值，即任何时间都不存在负相关。

即使提出显著的似乎公平的假设，即私人信息在消息观察者中逐渐扩散，等式（1）中价格逐渐调整的结果是以进一步假设消息观察者不依靠价格的假设为依据。如果他们确实如此——只要Q是非随机的——Grossman(1976)的逻辑将意味着一个完全有启迪作用的均衡，均衡价格 P_t^* 遵循随机游走（对 $q=1$）：$P_t^* = D_{t+z-1} - Q$ （2）

因此我们强调等式（1）表达的观察到的反应不足的结果只不过是个开始。它提出了一个显而易见的问题：即使消息观察者忙于处理基本数据而没有时间考虑用价格来预期，那到没有其他的投资者能专注于以价格为基础进行预测，并由此而产生近似于等式（2）中的理性预期均衡的结论吗？为了解决这个重要的问题，我们引入了动量交易者。

（二）引入动量交易者

动量交易者的效用也是恒定风险厌恶（CARA）的函数。t时刻，动量交易者进入，并持有股票j期（有限期），直到t+j时，j为外生变量。

动量交易者通过市场指令与消息观察者交易，他们提交指令但不知道这些指令将以什么价格执行。这个价格将由消息观察者的竞争决定。这时消息观察者成为造市商。为了确定指令的量，动量交易者必须预测（$P_{t+j} - P_t$）,他们的价格预测建立在过去价格的变动上。我们假定这些预期的形式很简单，唯一的条件变量是价格在长期的累积变化值：$P_{t-1} - P_{t-k-1}$。

k并不重要，可设k=1，则t时刻要预期的变量为 $P_{t-1} - P_{t-2} = \Delta P_{t-1}$。更重要的是，假定动量交易者所作的单变量的预期是建立在过去价格变化上的。因为如果用过去滞后n期的价格变化来预期，且赋予每个滞后期一定的权数，则当n充分大时，我们的结论就不成立了。此外，动量交易者是有限理性人，他们不具有计算复杂的多元回归的能力。

对k=1，从第t时刻进入市场的动量交易者递交的指令 F_1 表示：

$$F_1 = A + f\Delta P_{t-1} \qquad (3)$$

A为常数，f 为弹性，都有动量交易者最大化效用决定。这个交易指令必须被消息观察者吸收。假设消息观察者没有认识到这种供给的增长对价格的影响，这正与之前我们所假设的消息观察者不依赖价格变化进行预测一致。如果交易指令是过去价格变化的线性函数，则若允许消息观察者从中吸收信息，就相当于允许他们间接依赖价格做判断。

如果有j批动量交易者进入市场，则总的被消息观察者吸收的供给 S_t 表示为：

$$S_t = Q - \sum_{i=1}^{j} F_{t+1-i} = Q - jA - \sum_{i=1}^{j} f\Delta P_{t-i} \qquad (4)$$

继续假设t时刻，消息观察者买入并持有股票至清算股利T时刻，即价格为（1）中所示，用 S_t 代替Q，从而有：

$$P_t = D_t + \{(z-1)e_{t+1} + (z-2)e_{t+2} + ... + e_{t+z-1}\}/z - Q + jA + \sum_{i=1}^{j} f\Delta P_{t-i} \qquad (5)$$

大多数分析中常数Q和A不起作用，所以可以不予考虑以简化分析。

我们注意到，消息观察者的行为前后不一致或者说不连续。t时刻，他们的需求是建立在不再交易的基础上，而之后他们又打破了这一基础。我们利用了这种不一致性，因为它能简化分析。否则，如果消息观察者在t时刻不仅依赖他们对红利 D_T 的预测，还预期未来整个股价的走势的话，我们将面对一个复杂的动态的规划问题。

（三）均衡的本质

在前述的假设下我们来解模型。唯一的工作是计算 f 的均衡值。常数可以不予考虑，动量交易者的效用最大化意味着：

$$f\Delta P_{t-1} = g\,E_M(P_{t+j}-P_t)\big/\mathrm{var}(P_{t+j}-P_t) \qquad (6)$$

其中g 是动量交易者总的风险厌恶系数。E_M 和var_M 是均值和方差的表示。（6）又可以表示为：

$$f = g\,\mathrm{cov}(P_{t+j}-P_t, \Delta P_{t-1})\big/\{\,\mathrm{var}(\Delta P)\,\mathrm{var}_M(P_{t+j}-P_t)\} \qquad (7)$$

f 满足等式（7）时就达到均衡的固定点，这时价格动态地满足（5），。我们只限于研究稳态协方差均衡，那么均衡的必要条件就是$|f|<1$，附录中有证明。这样的均衡可能在任意参数值下不成立，但我们也不能一般地排除多重均衡的可能性。但附录中证明了，只要g 足够小，就可以保证$|f|<1$。一般说来，很难在封闭的环境中解这个模型，我们不得不求助于运算法则来解这个固定点。对于一系列任意的参数值，我们总是从j=1开始求固定点。有了这个限制（j=1)，我们可以得出，$|f|<1$是稳态均衡的充分必要条件。我们也可以从一个小的风险厌恶程度出发，并在一开始假设f 为0。这个过程确保在存在多重稳态均衡时，我们也可以选出f 的最小值。

即使不做任何计算，我们也能观察到一些均衡的性质。首先，我们得到

辅助定理1，在任何稳态均衡中，f >0。即，动量交易者必须表现为趋势追逐者。

现在我们对动态的价格进行定性描述。首先，考虑价格对突发的新闻的冲量反应。t时刻，有一个单位的积极的消息e_{t+z-1}开始在消息观察者中传播，从这个时点起，没有新的消息出现。价格变化的走势将如何？

为了得到答案，把在任何时间的价格分解成两部分：一部分归于消息观察者，一部分归于动量交易者，另一部分归于动量交易者。t到t+z-1时期，消息观察者完全把消息反应到对D_T的预期中，消息观察者对D_T值的总的估计上升。从而在t-z+1时刻，价格正好是动量交易者的交易指令流入市场的价格。但f >0，任何积极的消息都会产生最初的对动量交易者交易指令的推动力，使动量交易者进入市场。交易指令累积增加至t+j，使得t+j+1时刻已没有任何一个动量交易者会受已传播开的消息的刺激而进入市场。这种推理可引入以下结论：

命题1， 在任何一个稳态协方差均衡中，设e_{t+z-1}在t时刻开始在消息观察者中扩散。

一）总是存在过度反应，即价格累计的冲量反应的峰值大大高于均值1。

二）若j>z-1,累积的冲量在t+j时刻达到峰值，然后开始下降，最终回到1。

三）若j<z-1，累计的冲量未达到t+j时的值，但最终会集中在1。

除了冲量反应函数，考虑价格在不同时间的自协方差也很有意思。当动量交易者的风险厌恶系数g 趋于无穷大时。（7）意味着，均衡中必须有$\mathrm{cov}(\Delta P_{t+j}-P_t, \Delta P_{t-1})=0$。扩展为该表达式：$\mathrm{cov}(\Delta P_{t+1}, \Delta P_{t-1})+\mathrm{cov}(\Delta P_{t+2}, \Delta P_{t-1})+\ldots+\mathrm{cov}(\Delta P_{t+j}, \Delta P_{t-1})=0$。 （8）

方程（8）可推出：

命题2，在任何稳态协方差均衡中，如果价格变化在短期内是正相关的（即$\mathrm{cov}(P_{t+j}-P_t, \Delta P_{t-1})>0$），那么，对风险中性的动量交易者而言，在j+1内价格为负相关，即i≤j，$\mathrm{cov}(\Delta P_{t+i}, \Delta P_{t-1})<0$。

研究命题1和命题2在细节上的区别很有用，因为乍一看它们似乎是有些矛盾的。一方面，命题1认为，由于对好消息的反应，至少在比较期内价格连续上升，且可能更长时期内（如果j〈z-1)。另一方面，命题 2 认为价格在j＋１期内反转，且可能早于j＋１。

两个命题可以统一。前者是有限制条件下的定性描述，即它谈及的价格变化轨迹是从t

时刻向前的，条件是t时刻存在一个突发的新闻。从而命题 1 暗示投资者不知何故知道t时刻有新闻发生，通过t时刻买入并持有到t＋j时刻，赚取预期收益。这种策略可被定义为“在动量周期的早期中买入”，即在新闻一出现时就买入。但是显然我们的模型中动量交易者不能采取这种策略，因为他们不能直接依赖消息e 。

相反，命题 2 是个绝对的无条件的关于价格自协方差的描述。如果一个交易者在t时刻响应t-1时刻价格无条件的上升的影响买入，并一直持有到t+j时刻，平均来说他的利润为0，当动量交易者为风险中性时，这种0利润要求必须有效，因为他们可采取这个无条件的策略。

为什么命题2中的无条件限制地追逐价格的任何上升策略，与命题1中有限制地只追逐由于好消息引起的价格上升策略不同？原因很简单。并非所有的价格上升都是新闻驱使的。特别是基于t时刻的价格上升而买入股票的交易者导致了后来的风险。动量周期的后期即已没有任何好消息时，若最后的好消息是在t-i时刻发生的，t时刻上升的价格应归于最后一批动量交易者的买入。而那些受t-i时将发生的新闻吸引而较早买入交易会在不久后的将来（如，t-i+j+1时）使最后一批交易者在他们的交易到期前遭受损失。

（四）赢者和输者

消息观察者或动量交易者的有限理性是否导致他们发生系统性亏损？通常，只要资产净供给量Q为正，两类投资者都能获得正的预期收益。首先考虑Q=0时，这时可以得出，只要动量交易者的风险承受程度有限，就能赚取正的收益。因为Q=0为零和博弈，因此消息观察者必然亏损。

当Q>0时，为正和博弈，两个组按风险比例分配收益，即使消息观察者在与动量交易者的交易中亏损，这个亏损也可以通过在按风险分配收益中抵消，且他们还能获得净收益。在动量交易者是风险中性的限制下，两个组会扯平，与Q=0的逻辑相似，因为风险中性的动量交易者放弃了按风险分配的收益，重新回到零和博弈。

三、总结

我们认为任何关于资产定价的新的“行为”理论应该根据三个标准来评判：（一）假设投资者的行为似乎合理或与偶然的观察一致；（二）用简化和统一的方式解释现存的迹象；（三）作出大量的远期预测，这种预测能用“样本外推”检验，并最终被证明是有效的。

从这三点来看我们做得怎么样。首先，我们相信我们对有限理性的特定的解释——以无偏的方式处理可获得的信息的子集的能力——是似乎合理和凭直觉的。而且，在我们的研究框架中，这种有限理性意味着套利者对简单动量策略的普遍信赖。正如我们所讨论的，这一点与实际中观察到的高度一致。

从简化的标准来看，模型中一切东西都是由一个简单的刺激来驱动：关于未来基本面的消息的逐渐扩散。没有其他的对投资者情感刺激和流动性交易的需要。我们主要的概念上的贡献是，提出了如果存在对一部分投资者的短期内对消息的反应不足，那么假设套利策略很简单）最终在长期内必然存在反应过度。

Hong, Harrison, and Jeremy Stein (1999), “A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets", Journal of Finance 54, 2143-2184.
张莉　翻译