

林晓明 执业证书编号: S0570516010001  
研究员 0755-82080134  
linxiaoming@htsc.com

# A 股收入和利润的定量预测模型

## 宏观多因子模型初探

### 相关研究

- 1 《行业指数频谱分析及配置模型》2016.05
- 2 《市场的频率》2016.05
- 3 《从周期调整市盈率(CAPE)看,目前 A 股市场估值较低,处于底部区域》2012.08

### 上证指数和中小板指数的收入、利润增长率均与上期增长率有较强正相关性

在 A 股市场收入和利润的定量预测模型中,自回归因子始终是比较重要的正向因子,且最优滞后数均为 1,即上期收入与利润增长率对本期收入与利润增长率有比较重要的正向影响。

### 上证指数收入增长率主要受自相关性影响,油价、商品价格亦有正向作用

上证指数的收入增长率与上期增长率、油价、商品价格具有显著的相关性,在动态带权的线性回归模型中,上述三个因子的最优滞后期数分别为 1、2、0,最近 12 个月回归系数的平均值分别为 0.24、0.011、0.016,表明上证指数收入增长率主要受上期增长率影响,其次油价和商品价格均有微弱正向作用。

### CPI 对上证指数利润增长率有较强负向作用

上证指数的利润增长率与上期增长率、CPI、商品价格具有显著的相关性,在动态带权的线性回归模型中,上述三个因子的最优滞后期数分别为 1、2、0,最近 12 个月回归系数的平均值分别为 0.32、-0.14、0.035,表明上证指数利润增长率主要受上期增长率和 CPI 影响,其中 CPI 具有明显的负向作用。

### 美元指数对中小板指数收入增长率有较强负向作用

中小板指数的收入增长率与上期增长率、美元指数、油价有显著的相关性,在动态带权的线性回归模型中,上述三个因子的最优滞后期数分别为 1、1、2,最近 12 个月回归系数的平均值分别为 0.26、-0.19、-0.034,表明中小板指数收入增长率主要受上期增长率和美元指数影响,其中美元指数具有明显的负向作用。

### GDP 对中小板指数利润增长率有较强正向作用

中小板指数的利润增长率与上期增长率、GDP、美元指数、油价有显著的相关性,在动态带权的线性回归模型中,上述四个因子的最优滞后期数分别为 1、1、1、2,最近 12 个月回归系数的平均值分别为 0.30、0.54、-0.064、-0.037,表明中小板指数利润增长率主要受上期增长率和 GDP 影响,其中 GDP 具有明显的正向作用。

### 下月预测:上证指数的收入和利润持平微降,中小板指数的收入和利润持平微升

根据报告中收入和利润的定量预测模型,上证指数的收入和利润下月预测增长率分别为-0.25%和-0.19%,中小板指数的收入和利润下月预测增长率分别为 0.76%和 0.38%。

风险提示:报告根据历史数据进行规律总结,历史规律有可能失效。

## 正文目录

宏观多因子模型介绍 .....	4
多因子模型概述 .....	4
宏观多因子模型的研究历史 and 意义 .....	5
研究框架 .....	5
模型介绍及数据处理 .....	6
模型介绍 .....	6
数据选取及平滑处理 .....	6
计算方法及参数优化 .....	8
计算方法介绍 .....	8
误差分析与参数优化 .....	8
主板市场的实证研究 .....	9
研究框架 .....	9
收入增长率与宏观因子的定量关系 .....	9
利润增长率与宏观因子的定量关系 .....	12
收入与利润预测 .....	13
中小板市场的实证研究 .....	14
收入增长率与宏观因子的定量关系 .....	14
利润增长率与宏观因子的定量关系 .....	15
收入与利润预测 .....	15
总结 .....	16

## 图表目录

图 1: 上证指数的利润 (EBITDA) TTM 数据及其环比增长率 .....	7
图 2: 上证指数的利润 (EBITDA) 经滤波处理后 TTM 数据及其环比增长率 .....	7
图 3: 上证指数回归拟合收入增长率与误差 .....	10
图 4: 上证指数回归拟合收入增长率 $\pm 2$ 倍标准差曲线 .....	10
图 5: 上证指数回归预测收入增长率与误差 .....	11
图 6: 上证指数收入动态回归中美元指数 $p$ 值 .....	12
图 7: 上证指数收入动态回归中油价和商品价格的 $p$ 值 .....	12
图 8: 上证指数回归预测收入增长率与误差 (优化后模型) .....	12
图 9: 上证指数回归预测利润增长率与误差 (优化后模型) .....	13
图 10: 中小板指数回归预测收入增长率与误差 (优化后模型) .....	14
图 11: 中小板指数回归预测利润增长率与误差 (优化后模型) .....	15
表格 1: 宏观因子的彭博代码及具体含义 .....	6

表格 2: 上证指数收入增长率与宏观因子回归分析表 .....	10
表格 3: 上证指数收入与宏观因子动态回归系数在最后一个面板期的值 .....	11
表格 4: 上证指数收入与宏观因子动态回归最优滞后数与最近 12 个月系数均值 .....	12
表格 5: 上证指数利润与宏观因子动态回归最优滞后数与最近 12 个月系数均值 .....	13
表格 6: 上证指数收入 2016 年 6 月预测表 .....	13
表格 7: 上证指数利润 2016 年 6 月预测表 .....	14
表格 8: 中小板指数收入与宏观因子动态回归最优滞后数与最近 12 个月系数均值 .....	14
表格 9: 中小板指数收入与宏观因子动态回归最优滞后数与最近 12 个月系数均值 .....	15
表格 10: 中小板指数收入 2016 年 6 月预测表 .....	16
表格 11: 中小板指数利润 2016 年 6 月预测表 .....	16

## 宏观多因子模型介绍

### 多因子模型概述

多因子模型是量化领域最经典的模型之一，其基本思想是股票的收益率可以被一组共同的因子和一个仅与该股票有关的特异因子解释。多因子模型并不蕴含任何意义下的因果关系，只是代表解释风险的维度。

多因子模型可以分三类：

1. 对外部变化的响应模型；
2. 资产属性的横截面因子模型；
3. 纯粹的内在或统计因子模型。

第一类模型的基本想法是：股票市场和外部经济力量之间应该存在某种可被证实的关联。通常的外部响应因子：GDP、CPI、债券市场收益率、油价变动、汇率变动、工业产量变动等等。这些因子通常被称为宏观类因子。宏观因子主要从两个方面影响股市，一是对上市公司的业绩产生影响，二是改变投资者的投资行为。经济形势好，上市公司业绩蒸蒸日上，投资者对前景乐观，投资积极；经济形势不好，上市公司业绩差强人意，投资者对前景悲观，投资则变得格外谨慎。宏观因子对股票收益率的解释力可能会非常强，但是也具有一些公认的缺陷：一是变量误差问题，最常见的研究宏观因子的方法是通过回归分析或类似技术来估计宏观因子的影响系数，假设每次回归需要利用 1000 只股票过去 60 个月的数据信息，这就不可避免地会产生估计误差；二是影响系数可能是非静态的，即使能够在统计意义下精确获取过去的信息，这些数值可能也无法精确描述现在的情况或预测未来的变化；三是宏观数据通常质量不佳，因为大部分宏观数据是由政府采集的而不能在市场中直接被观测到，所以数据质量和可信度都存在疑问，并且宏观数据还经常存在延迟及频率过低等情况。其中，缺陷三是很难被克服的，前两点还存在一定改进空间。本文将着眼 A 股市场，基于“动态”回归的方法，建立利润和收入的定量预测模型，主要克服了缺陷二，也对缺陷一有一定减轻效应。

第二类模型主要比较分析股票的各种属性，与宏观因子没有直接联系。横截面因子一般可以归为两类：基本面类和市场类。基本面类因子包括：市盈率、市净率等估值因子，资产周转率、毛利率等财务因子，净利润增长率、营业收入增长率等成长因子，资产负债率、现金比率等杠杆因子，以及未来每股盈利预测等分析师情绪因子。市场类因子包括：动量因子，波动率因子，换手率因子，macd、dea 等技术因子。这些因子某种程度上也具有变量误差、非静态性等缺陷，但与宏观因子的解释有所不同，横截面因子并非用来预测股票市场收益率是会持续还是反转，而只是用来发现过去一段时间因子值相近的股票在未来的表现会有共同之处。根据经验，横截面因子是非常有效的因子。后续我们将推出多因子系列研究报告，即主要针对横截面因子，依赖数学工具和投资直觉，发掘各因子与股票收益率之间的关系，建立量化策略，有效获得超额收益。

第三类模型是通过收集大量的股票收益率和因子信息，利用“因子生成器”，得出统计规律。能被用作因子生成器的统计工具很多，常见的有主成分分析（principal component analysis）、最大似然分析（maximum likelihood analysis）和预期最大化分析（expectations maximization analysis）等。统计因子虽然效果可能十分强大，但因为它们的直观含义通常难以理解，并且在估计过程中很容易受到“伪相关性”的影响，故在 A 股市场中并不是很通用。再者，统计工具依赖于每只资产对每个因子的暴露度在估计时段内都是恒定的，因此也不能捕获暴露度随时间变化的因子，例如非常有效的动量类因子。综上所述，这一类模型暂时不在我们的研究计划中。

## 宏观多因子模型的研究历史和意义

本文的主要研究目标是上一小节中描述的第一类模型，即宏观多因子模型。

由于宏观因子逻辑简单、易于理解、对股票市场收益率存在重要影响，国内外有许多学者及金融从业者致力于宏观因子对股票市场收益影响的研究，以期发现不同类型的宏观因子对股票市场收益影响的程度、途径和传导机制。

在国外早期的研究学者中，Fama 和 Schwert 的研究最具代表性。Fama 和 Schwert (1977) 研究发现，普通股票的收益与预期通货膨胀率呈显著负相关，与未预期通货膨胀率和预期通货膨胀率的变化也是负相关关系，但这种负相关性并不能对股票的收益波动做出解释。在这个结论的基础上，Fama (1981) (1990) 进一步实证认为，使用预期的现金流和预期的收益波动率相关性分析，股票价格波动被解释的部分为 30%；而股票价格与实际经济增长存在着正相关关系，可使用实际经济增长率来预测期望现金流的波动，股票价格波动被解释的部分为 48%。Schwert (1989) 检验了宏观经济因素对股票市场收益率的影响，发现宏观经济水平能够解释一部分的股票价格变动。Chen (1986) 在套利定价理论框架下进行研究并得出相似的结论，认为宏观经济变量与股票价格存在长期的均衡关系，经济变量会对股市收益率产生系统的影响。

不同地域的股票市场与宏观因子之间的关系不尽相同，亦有很多学者对此方面展开实证研究。Kaneko 和 Lee (1995) 对美国与日本市场进行研究，发现在美国股票市场中，时间溢价、风险溢价和工业产出的增长率都被显著定价；而在日本股票市场中，国际因素更为重要，石油价格的变化、贸易条件的变化和汇率对日本股票市场价格波动有更为显著的影响。Kwon 和 Shin (1999) 发现韩国股票市场指数和产出指数、货币供给量、汇率、贸易收支等宏观经济变量之间存在着长期的协整关系。Hassapis 和 Kalyvitis (2002) 基于七个亚洲的工业化国家的历史数据，运用向量自回归的检验方法，发现这七个国家的股票价格变化和产出增加率之间高度相关。Morelli (2002) 则在检验英国股市波动性和宏观经济条件波动性的关系时，认为宏观变量的波动并不能对股市的条件波动给出解释。Wong 和 KHAN (2006) 发现新加坡的股票价格普遍表现与利率、货币供应量有长期均衡关系。Arnold (2008) 的研究表明，从 1969 到 1996 这些年间，美国股市波动与专业的经济预测有着显著的相关性，但是该相关性从 1997 年开始消失。

我国股票市场起步晚，相对于其他发达国家的资本市场而言还不够完善，相关研究不如国外多。许多学者及金融从业者对我国股票市场与宏观因子之间的关系进行过实证研究，由于宏观因子的选取不尽相同、覆盖时间长度和计量方法也不一致，所得出的结论亦并不相同，但总体而言，大部分实证观点均认为我国股票市场收益和宏观经济运行情况之间存在一定关联性。

虽然已有不少关于 A 股市场与宏观经济间关系的研究报告，但能够利用宏观因子较为精确地预测 A 股市场利润和收入的定量模型却不多见。在本报告中，我们将利用动态回归的方法，寻找影响 A 股市场的关键宏观因子，并预测 A 股主板和中小板的利润和收入增长率，帮助读者更为清楚地认识 A 股市场和宏观经济之间的逻辑关系。

## 研究框架

大致分为四个部分：

1. 模型介绍及数据处理；
2. 计算方法及参数优化；
3. 主板市场的实证研究；
4. 中小板市场的实证研究。



## 模型介绍及数据处理

### 模型介绍

市场的利润和收入增速是衡量市场表现的一个重要指标。本报告旨在对利润和收入增长率进行建模,使用一系列宏观变量对其进行解释、拟合和预测。设  $t$  时刻的收入或利润为  $Z_t$ , 则我们令

$$\frac{Z_{t+1} - Z_t}{Z_t} = \beta_{0,t} + \beta_{1,t}x_{1,t} + \cdots + \beta_{m,t}x_{m,t} + \varepsilon_t$$

$x_{i,t}$  是宏观变量的同比增长率。需要注意的是: 由于利润或收入采用连续 12 个月的价值 (TTM), 所以  $Z_{t+1} - Z_t$  表示同比之差。因此与之对应, 宏观变量我们统一采用了其同比增长率。

我们的模型采用了某种“动态”的回归, 即回归系数  $\beta_{i,t}$  是随  $t$  变化的, 我们认为这样的回归能够更好地刻画运动的趋势。

### 数据选取及平滑处理

**被预测变量:** 市场的利润和收入的增长率;

**自变量:** 和经济发展相关的宏观变量, 比如美元指数、商品价格指数、油价、CPI、GDP 等;

**数据选取:** 针对我国宏观数据统计情况, 自变量的选取见表 1 (来源为彭博), 报告里使用的都是它们的同比增长率 (已经是同比增长率的除外), 因变量则选取上证指数 (代表主板市场)、中小板指数 (代表中小板市场) 的每股息税折旧及摊销前利润 (EBITDA) 和每股营业收入的最近 12 个月之和 (即 TTM 数据) 的环比增长率, 实际上, 第  $i$  个月 TTM 数据与第  $i-1$  月 TTM 数据之差即为第  $i$  个月的单月数据减去上一年度第  $i$  个月的单月数据, 因此 TTM 数据的环比增长率与单月数据的同比增长率是一个近似值, 只是两个表达式的分母不同;

**样本频率:** 由于国内市场数据不多, 按季选取的话数据非常少, 不利于回归, 所以我们的数据均为按月选取, 对于一些不能按月提取的宏观变量, 通过线性插值得到月度数据。

**表格1: 宏观因子的彭博代码及具体含义**

彭博代码	含义
CNGDPYOY Index	实际 GDP 同比增长率
WRCOPIA Index	油价三个基准平均值
CNCPIYOY Index	CPI 同比增长率
CRY Index	汤森路透核心商品价格指数
USTWBROA Index	贸易加权美元指数

资料来源: 华泰证券研究所

我们先查看一下需要拟合的因变量: 上证指数 2000 年以后的利润 TTM 数据及其月度环比增长率 (见图 1)。从图中我们明显可以看到数据非常“不平滑”, 相邻两月之间的变化非常剧烈, 故可以认为数据存在噪音。因此, 我们首先对利润数据进行去除噪声的处理, 我们把这样的操作定义为  $F$ 。去噪可以有不同的方式, 比如简单移动平均、指数移动平均等, 但是简单的移动平均等均会产生滞后, 这里我们采用了如下形式的滤波:

$$(Fz)_t = \left( \alpha - \frac{\alpha^2}{4} \right) z_t + \frac{\alpha^2}{2} z_{t-1} - \left( \alpha - \frac{3\alpha^2}{4} \right) z_{t-2} + 2(1-\alpha)(Fz)_{t-1} - (1-\alpha)^2 (Fz)_{t-2}$$

我们选取  $\alpha = 0.3$ ，进行滤波操作的结果如图 2 所示。

我们对被预测变量（利润、收入）均采用了这样的滤波操作，我们预测的实际上是滤波过后的增长率，从图中可以看出，滤波操作不改变被预测变量的变化趋势，只起到一个去噪的作用。此时，我们的模型变为：

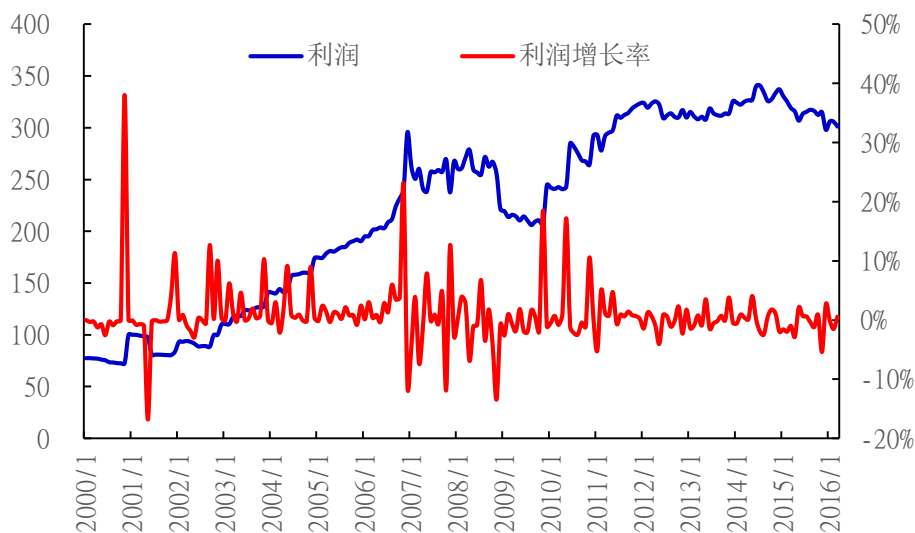
$$\frac{(Fz)_{t+1} - (Fz)_t}{(Fz)_t} = \beta_{0,t} + \beta_{1,t}x_{1,t} + \cdots + \beta_{m,t}x_{m,t} + \varepsilon_t$$

这里一个值得关注的地方在于  $x_{i,t}$  并不一定只宏观变量的当期同比增长率，也可能是滞后若干期的同比增长率。为了方便我们有时将上面的模型写为：

$$y_t = X_t \beta_t + \varepsilon_t$$

$X_t$ ， $\beta_t$  分别是行向量和列向量。

图1： 上证指数的利润（EBITDA）TTM 数据及其环比增长率



资料来源：华泰证券研究所

图2： 上证指数的利润（EBITDA）经滤波处理后 TTM 数据及其环比增长率



资料来源：华泰证券研究所

## 计算方法及参数优化

### 计算方法介绍

不同于传统的普通最小二乘法（OLS）线性回归，我们采用一种“动态”的线性回归模型，这种做法利于捕获股票市场的利润、收入与宏观因子的时间序列的“动态”相关关系。在普通最小二乘法的基础上，当回归模型由第  $t$  个月进入第  $t+1$  个月时，模型增加了  $(X_{t+1}, y_{t+1})$  这一组属于第  $t+1$  个月的数据，我们想要得到回归系数  $\beta_{t+1}$ ，只需把新的数据纳入到回归方程里即可。在把新的数据纳入进来的过程中，我们自然会认为离当前时间比较近的数据更有效，离当前时间远的数据效用不是那么大，于是我们采用了带权线性回归的思想。

所谓带权线性回归，指的是对每个点在回归时赋予不同的权重。具体来说，在求  $\beta_{t+1}$  时，我们求解下列问题：

$$\beta_{t+1} = \arg \min_{\beta} \sum_{i=1}^{t+1} D\left(\frac{t+1-i}{\lambda}\right) (y_i - X_i \beta)^2$$

$D(t)$  表示权重函数， $\lambda$  表示某种宽度的变量。我们提供了如下若干种形式的权重函数：

1. 示性函数：

$$D(t) = \begin{cases} 1, & |t| < 1 \\ 0, & \text{otherwise} \end{cases}$$

2. Tri-cube 函数：

$$D(t) = \begin{cases} (1 - |t|^3)^3, & |t| < 1 \\ 0, & \text{otherwise} \end{cases}$$

3. 高斯函数：

$$D(t) = \exp(-t^2)$$

4. 柯西函数：

$$D(t) = \frac{1}{1+t^2}$$

5. Epanechnikov 函数：

$$D(t) = \begin{cases} \frac{3}{4}(1-t^2), & |t| < 1 \\ 0, & \text{otherwise} \end{cases}$$

如果采用示性函数，我们实际上设置了某种滚动窗口，每次把  $[t-\lambda+1, t+1]$  的数据拿出来做传统的线性回归。而高斯函数函数和柯西函数，则是把之前所有数都纳入到回归的范围，不过离  $t+1$  近的权重大。其他函数即设置了滚动窗口，又给窗口内的数据不同权重。

### 误差分析与参数优化

在带权线性回归中，我们关心模型与算法的有效性，如何评价这样的有效性呢？我们认为有两种观点，一种是站在拟合的角度，即使用我们估计的系数  $\beta_{i,t}$ ，计算出相应的  $y_t$ ，我们记做  $y_t^{fit}$ （即  $y_t^{fit} = X_{i,t} \beta_{i,t}$ ），考察拟合误差：

$$E_f = \sqrt{\sum_t (y_t - y_t^{fit})^2}$$



需要注意的一点是系数 $\beta_{i,t}$ 是使用 $[1, t]$ 的数据计算出来的。

另一种是站在预测的角度：我们预测 $t$ 时刻的数据时，由于 $y_t$ 未知，我们不能通过上面的方式估计出 $\beta_{i,t}$ ；我们只有 $\beta_{i,t-1}$ ，借此，我们可以得到预测数据：

$$y_t^{\text{predict}} = X_{i,t} \beta_{i,t-1}$$

所谓预测误差就是预测值与准确值的差，定义为：

$$E_p = \sqrt{\sum_i (y_i - y_i^{\text{predict}})^2}$$

这里定义的拟合误差与预测误差类似于机器学习中的训练误差与测试误差。

注意到在计算中，我们有若干待定参数：变量的滞后数、带权线性回归中的 $\lambda$ 等，我们在优化这些参数时，采用最小化预测误差的方式。为了使问题简单，我假定参数只有滞后数，其他固定（ $\lambda$ 固定为 15）。在优化中，这样的整数规划是 N-P 难问题，我们采用贪心法。固定所有的滞后数范围为 0-3 个月（即最多滞后一个季度），首先对第一个变量优化，其余固定，之后对第二个，依次类推，可以得到一个局部最优值。我们调整变量的优化次序得到不同的最有结果，取其中的最优为最后的结果。

## 主板市场的实证研究

### 研究框架

在整个计算中，我们采用这样的框架：

1. 对所有数据进行普通最小二乘法线性回归（OLS），看是否所有宏观变量都显著，如果不显著，我们尝试调整该宏观变量的滞后数，看能否使其变得显著，如果还不显著，剔除该变量；
2. 对于剩下的变量重新进行带权线性回归，分析系数序列和 p 值序列；
3. 通过第 2 步，我们找出真正有用的变量，该变量在拟合中持续显著或者在较大的时间区域内显著；
4. 使用第 3 步中筛出的变量进行调整优化参数过程，得到最优参数。

我们的自变量 $X_{i,t}$ 除了包含宏观因子之外，我们还将滞后几期的因变量纳入进来，即还会包含 $y_{t-k}$ 。再次说明，这里 $y_t$ 的指利润或收入的 TTM 数据的月度环比增长率，即：

$$y_t = \frac{\text{PROF}_{t+1} - \text{PROF}_t}{\text{PROF}_t} \quad \text{或} \quad y_t = \frac{\text{SALE}_{t+1} - \text{SALE}_t}{\text{SALE}_t}$$

这里的 PROF 和 SALE 分别指主板市场的利润和收入，在第二章中曾提及，我们使用上证指数的每股息税折旧及摊销前利润（EBITDA）和每股营业收入代表主板市场的利润和收入，它们在进行回归之前都需经过滤波处理。宏观变量使用的是其同比增长率（已经是同比增长率的除外）。

### 收入增长率与宏观因子的定量关系

先考察被预测变量是每股收入的情形。首先我们对所有数据进行普通最小二乘法线性回归（OLS），我们发现 CPI 和美元指数同比增长这两个宏观因子并不显著，通过调整滞后数也不能使其显著，我们剔除这两个因变量后得到所有变量均显著，OLS 回归结果如下表。

表格2: 上证指数收入增长率与宏观因子回归分析表

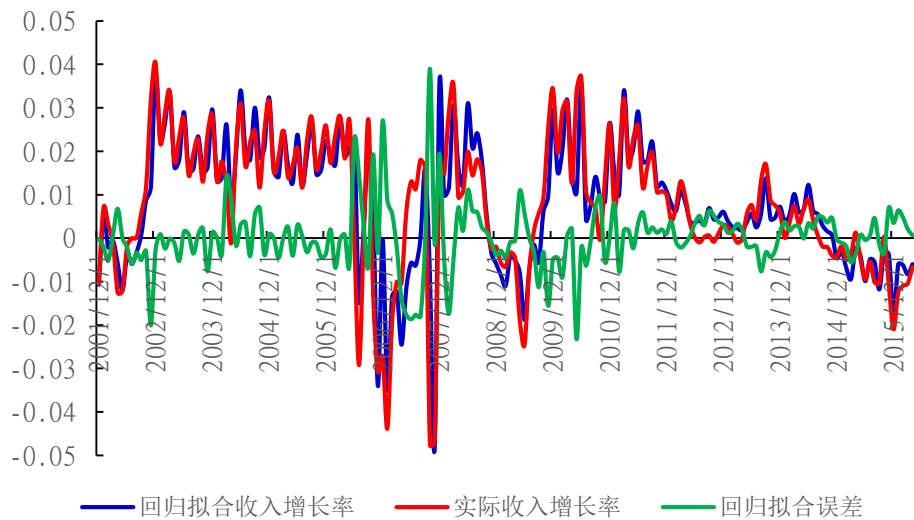
变量	滞后	回归系数	p 值
每股收入	1	2.23e-1	0
GDP	1	-1.34e-1	1.72e-3
油价	2	1.38e-2	4.23e-5
商品价格	0	2.69e-2	1.54e-07
常数		1.63e-2	6.31e-5

资料来源: 华泰证券研究所

下面我们采用带权线性回归法进行计算(每股收入、GDP、油价、商品价格的初始滞后数分别取 1,1,2,0)。

在计算中, 我们权重函数取柯西函数, 回归结果与拟合误差见图 3。

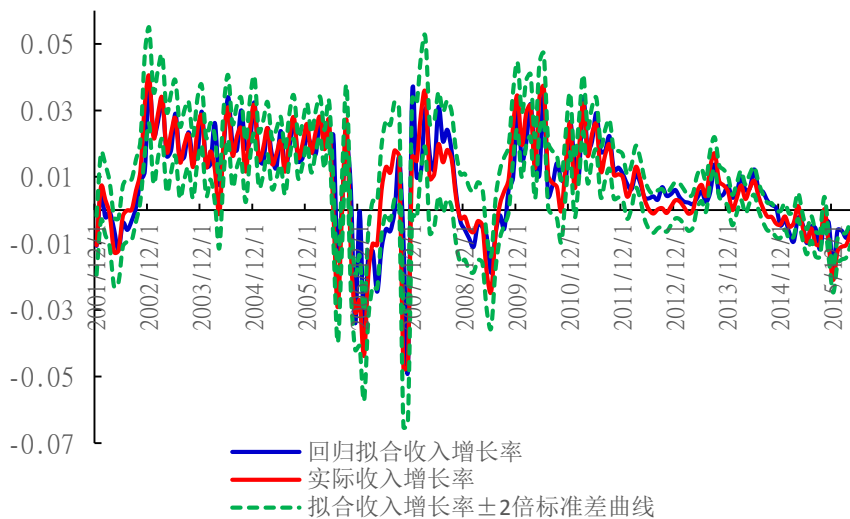
图3: 上证指数回归拟合收入增长率与误差



资料来源: 华泰证券研究所

根据回归结果, 还可以作出拟合收入增长率 $\pm 2$  倍标准差曲线, 见图 4。

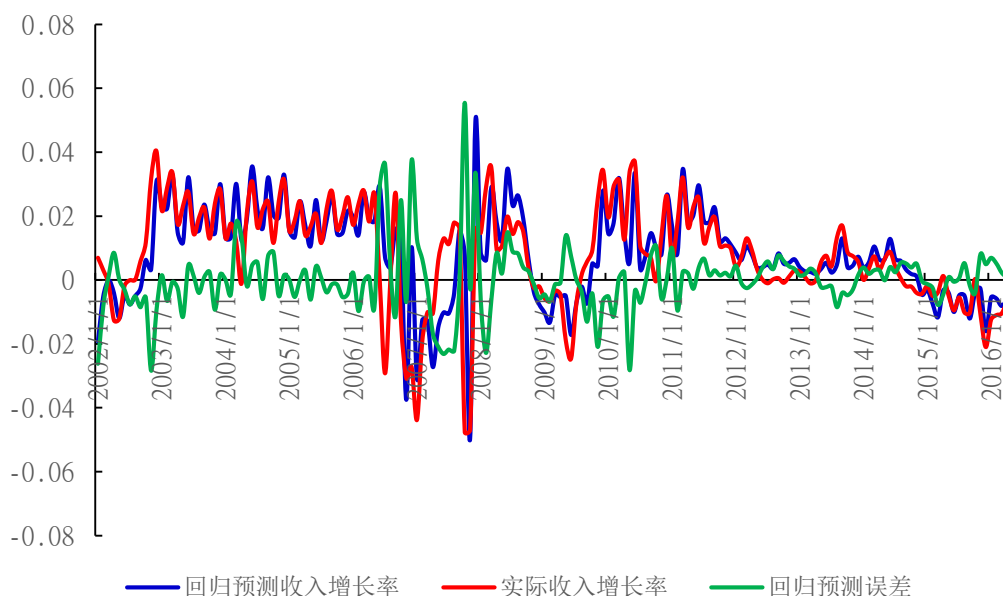
图4: 上证指数回归拟合收入增长率 $\pm 2$  倍标准差曲线



资料来源: 华泰证券研究所

观察以上两图可知，回归拟合效果相当不错。下面将使用回归所得的因子系数作为下期预测值，得到预测曲线如图 5 所示。

**图5： 上证指数回归预测收入增长率与误差**



资料来源：华泰证券研究所

根据第三章中提及的计算公式，图 3 中拟合误差  $E_f = 0.0984$ ，图 5 中预测误差  $E_p = 0.1373$ ，这说明拟合的效果是优于预测的。在我们的带权线性回归中，回归系数是动态的，最后一次的回归系数如下表所示：

**表格3： 上证指数收入与宏观因子动态回归系数在最后一个面板期的值**

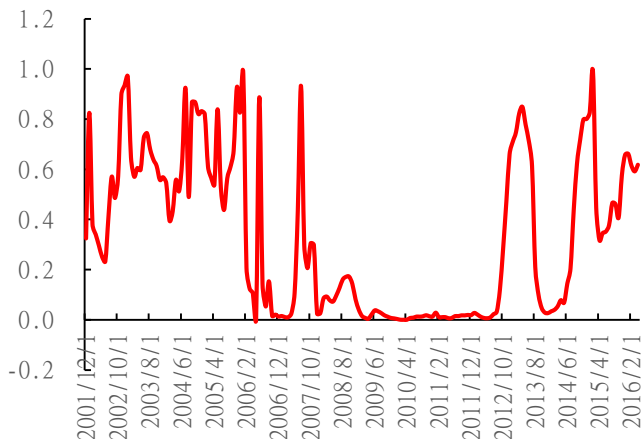
变量	回归系数	p 值
每股收入	2.36e-1	0
GDP	2.12e-2	6.18
油价	1.21e-2	9.04e-5
商品价格	1.61e-2	1.21e-3
常数	1.58e-3	6.60e-1

资料来源：华泰证券研究所

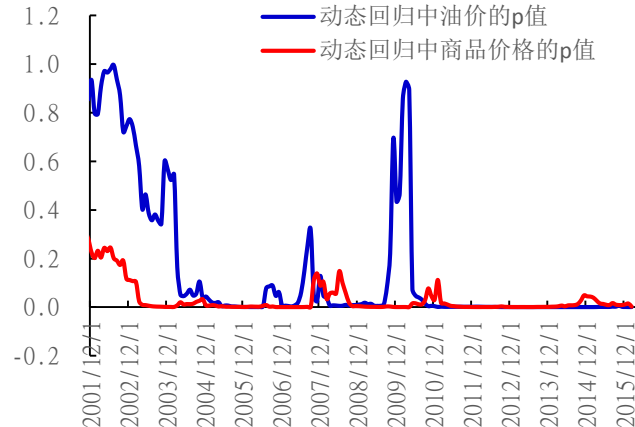
我们发现最后一期的结果中 GDP 竟然不显著！这和我们普通最小二乘法线性回归的结果大相径庭。我们现在一探原因，由于我们是动态线性回归，所以系数随着时间变化，相应的，每一次的 p 值也随时间变化。美元指数 p 值随时间的变化曲线如图 6 所示。

我们可以看出，GDP 在使用普通最小二乘法线性回归中显著，然而在动态线性回归中，该变量仅在一小段时间内显著，在绝大部分时间上不显著。对于油价和商品价格的 p 值见图 7，我们发现这两个变量几乎在整个测试区间上显著。至此，我们提取这两个变量作为最后宏观变量。

删去 GDP 这个因变量后，模型的拟合误差为  $E_f = 0.1032$ ，较之前变大，这是因为变量减少自然会使得拟合误差增大，然而预测误差  $E_p = 0.1326$  却有所减小，这说明新的变量选取降低了过拟合程度。我们认为能用尽可能少的变量刻画收入（或利润）比用非常多的变量进行刻画更有意义。

**图6： 上证指数收入动态回归中美元指数 p 值**

资料来源：华泰证券研究所

**图7： 上证指数收入动态回归中油价和商品价格的 p 值**

资料来源：华泰证券研究所

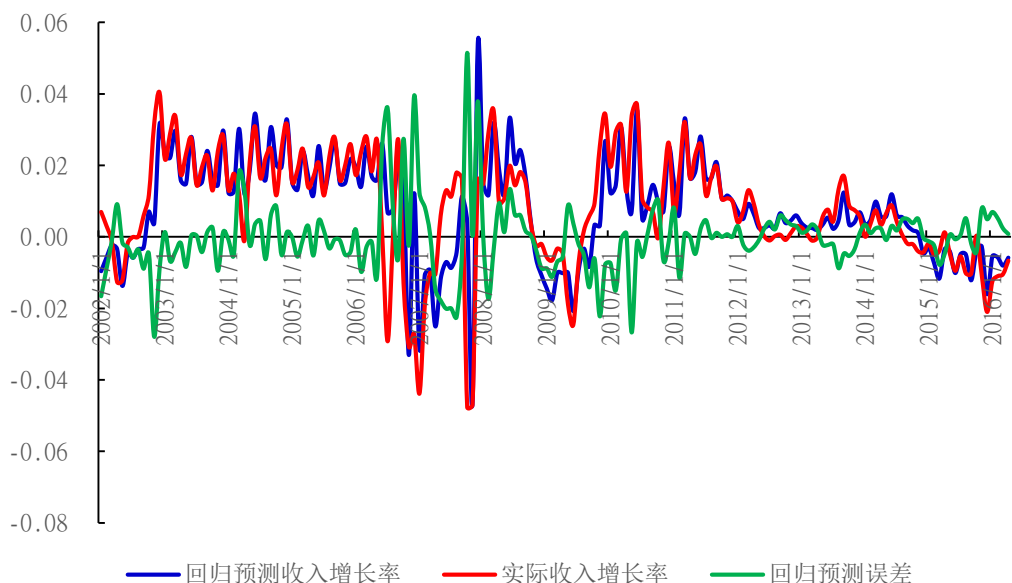
我们现在进行参数优化，我们优化的参数有收入、油价和商品价格的滞后数，按照如前的方式优化后，发现之前的系数即为最优结果。由于系数是随时间变化的，我们列出最近12个月系数的平均值如下：

**表格4： 上证指数收入与宏观因子动态回归最优滞后数与最近12个月系数均值**

变量	滞后数	最近12个月平均系数
每股收入	1	2.39e-1
油价	2	1.05e-2
商品价格	0	1.56e-2
常数		3.8e-3

资料来源：华泰证券研究所

上述模型的预测结果与真实结果的误差如下图所示：

**图8： 上证指数回归预测收入增长率与误差（优化后模型）**

资料来源：华泰证券研究所

### 利润增长率与宏观因子的定量关系

对于利润增长率的动态回归模型，我们复制上一小节的研究框架，首先筛选出CPI、商品价格、油价做为宏观变量（当然还要加上滞后的利润作为变量），不过油价与利润的相关

性比油价与收入的相关性低，最后经过计算删去油价这一因变量。经过参数优化，得出各因子的最优滞后数与最近 12 个月动态系数均值：

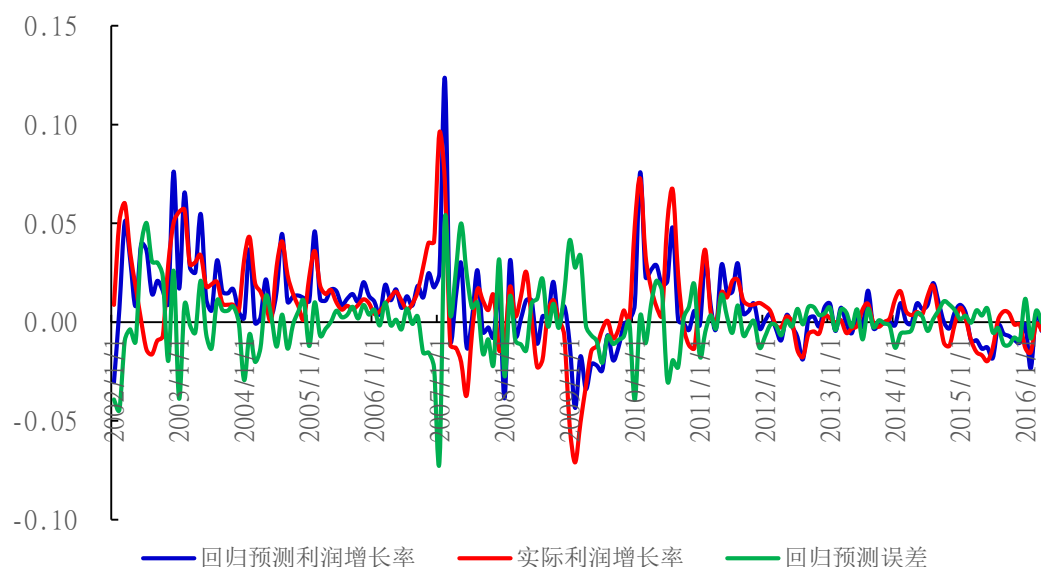
**表格5： 上证指数利润与宏观因子动态回归最优滞后数与最近 12 个月系数均值**

变量	滞后数	最近 12 个月平均系数
每股利润	1	3.15e-1
CPI	2	-1.39e-1
商品价格	0	3.46e-2
常数		6.1e-2

资料来源：华泰证券研究所

从表 5 可见，上证指数利润增长率与 CPI 是负相关关系。上述模型的预测结果与真实结果的误差如下图所示：

**图9： 上证指数回归预测利润增长率与误差（优化后模型）**



资料来源：华泰证券研究所

## 收入与利润预测

根据主板市场收入、利润增长率和宏观因子的定量关系研究，我们对下期（即 2016 年 6 月）主板市场的收入、利润增长率作出预测，见表 6 和表 7。

**表格6： 上证指数收入 2016 年 6 月预测表**

变量	滞后数	因子值	预测系数		收入增长率	收入
每股收入	1	-6.66e-3	2.37e-1	当前值	-	2308.0
油价	2	-2.93e-1	1.27e-2	预测值	-0.25%	2302.2
商品价格	0	-1.66e-1	3.35e-3	预测上限	0.10%	2310.3
常数		1	3.35e-3	预测下限	-0.60%	2294.1

资料来源：华泰证券研究所



表格7: 上证指数利润 2016 年 6 月预测表

变量	滞后数	因子值	预测系数		利润增长率	利润
每股利润	1	1.26e-3	2.90e-1	当前值	-	303.2
CPI	2	2.30e-2	-1.47e-1	预测值	-0.19%	302.7
商品价格	0	-1.66e-1	3.37e-2	预测上限	0.41%	304.5
常数		1	6.69e-3	预测下限	-0.79%	300.8

资料来源: 华泰证券研究所

我们的预测结果显示, 下个月主板市场的收入、利润水平均基本与上月持平, 可能微降。

## 中小板市场的实证研究

研究思路与主板市场完全相同, 这里我们使用中小板指数的每股息税折旧及摊销前利润 (EBITDA) 和每股营业收入代表中小板市场的利润和收入, 它们在进行回归之前都需经过滤波处理。宏观变量使用的是其同比增长率 (已经是同比增长率的除外)。

## 收入增长率与宏观因子的定量关系

经过宏观因子筛选和参数优化后, 得出各因子的最优滞后数与最近 12 个月动态系数均值为:

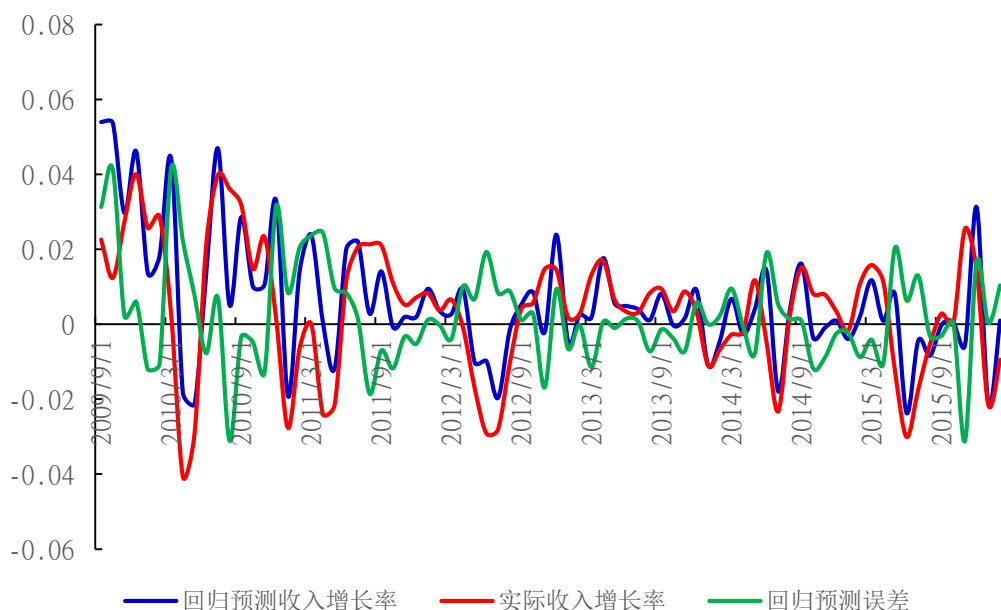
表格8: 中小板指数收入与宏观因子动态回归最优滞后数与最近 12 个月系数均值

变量	滞后数	最近 12 个月平均系数
每股收入	1	2.61e-1
美元指数	1	-1.91e-1
油价	2	-3.4e-2
常数		5.40e-2

资料来源: 华泰证券研究所

可以看出, 中小板指数收入增长率与美元指数有较强的负相关关系。上述模型的预测结果与真实结果的误差如下图所示:

图10: 中小板指数回归预测收入增长率与误差 (优化后模型)



资料来源: 华泰证券研究所

## 利润增长率与宏观因子的定量关系

经过宏观因子筛选和参数优化后，得出各因子的最优滞后数与最近 12 个月动态系数均值为：

**表格9： 中小板指数收入与宏观因子动态回归最优滞后数与最近 12 个月系数均值**

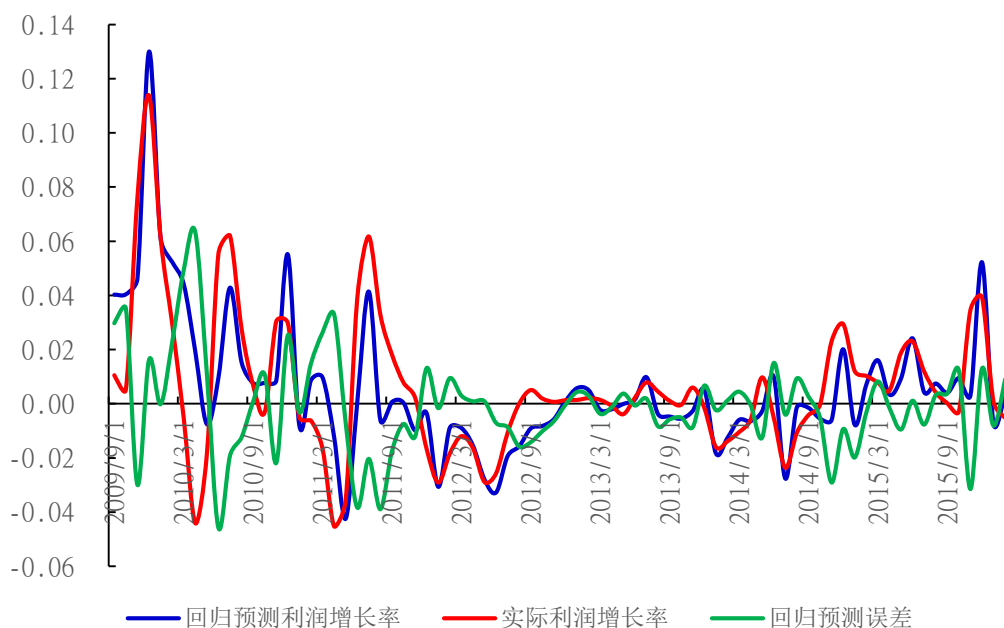
变量	滞后数	最近 12 个月平均系数
每股利润	1	2.95e-1
GDP	1	5.37e-1
美元指数	1	-6.40e-2
油价	2	-3.74e-2
常数		-3.96e-2

资料来源：华泰证券研究所

可见不同于收入，GDP 对利润的增长有促进作用，美元指数同样是呈负相关关系，不过作用没有收入那么显著，预测结果见图 11。

我们注意到中小板和主板的一个重要区别是：主板收入的回归系数主要体现了自回归效应，主板利润和 CPI 负相关；而对于中小板，美元指数对收入具有一个很强的负作用（其他变量的系数大小与自回归系数和美元指数的系数大小都不在一个量级上），而 GDP 对盈利有非常强的正作用。

**图11： 中小板指数回归预测利润增长率与误差（优化后模型）**



资料来源：华泰证券研究所

## 收入与利润预测

根据中小板市场收入、利润增长率和宏观因子的定量关系研究，我们对下期（即 2016 年 6 月）中小板市场的收入、利润增长率作出预测，见表 10 和表 11。

**表格10: 中小板指数收入 2016 年 6 月预测表**

变量	滞后数	因子值	预测系数		收入增长率	收入
每股收入	1	4.82e-3	2.32e-1	当前值	-	2021.1
美元指数	1	3.87e-2	-1.48e-1	预测值	0.76%	2036.5
油价	2	-2.93e-1	-2.62e-2	预测上限	1.74%	2056.3
常数		1	4.53e-3	预测下限	-0.22%	2016.7

资料来源: 华泰证券研究所

**表格11: 中小板指数利润 2016 年 6 月预测表**

变量	滞后数	因子值	预测系数		利润增长率	利润
每股利润	1	5.61e-3	2.53e-1	当前值	-	301.0
GDP	1	6.70e-2	6.59e-1	预测值	0.38%	304.4
美元指数	1	3.87e-2	-2.97e-2	预测上限	1.39%	307.5
油价	2	-2.93e-1	-3.32e-2	预测下限	-0.63%	301.3
常数		1	-5.03e-2			

资料来源: 华泰证券研究所

我们的预测结果显示, 下个月中小板市场的收入、利润水平均基本与上月持平, 可能微升。

## 总结

本报告对主板市场和中小板市场的收入与利润增长率进行了建模与计算。由于数据存在噪音, 我们首先对收入与利润进行滤波法去噪处理, 然后使用带权线性回归的方式对收入与利润增长率进行拟合、预测, 并通过最小化预测误差的方式进行参数优化, 测试结果显示了我们方法的可行性。我们最终的结果显示, 美元指数、商品价格指数、油价等均可影响收入与利润增长率, 自回归因子始终是比较重要的正向因子, 且最优滞后数均为 1, 即上期收入与利润增长率对本期收入与利润增长率有比较重要的正向影响。对于主板市场, CPI 对利润具有负向作用; 对于中小板市场, 美元指数对收入具有负向作用, GDP 则对利润具有正向作用。

基于定量模型, 我们预测 2016 年 6 月主板、中小板市场的收入、利润基本与上期持平, 主板收入、利润可能微降, 中小板收入、利润可能微升。

## 免责声明

本报告仅供华泰证券股份有限公司（以下简称“本公司”）客户使用。本公司不因接收人收到本报告而视其为客户。

本报告基于本公司认为可靠的、已公开的信息编制，但本公司对该等信息的准确性及完整性不作任何保证。本报告所载的意见、评估及预测仅反映报告发布当日的观点和判断。在不同时期，本公司可能会发出与本报告所载意见、评估及预测不一致的研究报告。同时，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。本公司不保证本报告所含信息保持在最新状态。本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司力求报告内容客观、公正，但本报告所载的观点、结论和建议仅供参考，不构成所述证券的买卖出价或征价。该等观点、建议并未考虑到个别投资者的具体投资目的、财务状况以及特定需求，在任何时候均不构成对客户私人投资建议。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。对依据或者使用本报告所造成的一切后果，本公司及作者均不承担任何法律责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

本公司及作者在自身所知情的范围内，与本报告所指的证券或投资标的不存在法律禁止的利害关系。在法律许可的情况下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券头寸并进行交易，也可能为之提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本报告版权仅为本公司所有。未经本公司书面许可，任何机构或个人不得以翻版、复制、发表、引用或再次分发他人等任何形式侵犯本公司版权。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“华泰证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。本公司保留追究相关责任的权力。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

本公司具有中国证监会核准的“证券投资咨询”业务资格，经营许可证编号为：Z23032000。

© 版权所有 2016 年华泰证券股份有限公司

## 评级说明

### 行业评级体系

- 报告发布日后的 6 个月内的行业涨跌幅相对同期的沪深 300 指数的涨跌幅为基准；

- 投资建议的评级标准

增持行业股票指数超越基准

中性行业股票指数基本与基准持平

减持行业股票指数明显弱于基准

### 公司评级体系

- 报告发布日后的 6 个月内的公司涨跌幅相对同期的沪深 300 指数的涨跌幅为基准；

- 投资建议的评级标准

买入股价超越基准 20%以上

增持股价超越基准 5%-20%

中性股价相对基准波动在 -5%~5%之间

减持股价弱于基准 5%-20%

卖出股价弱于基准 20%以上

## 华泰证券研究

### 南京

南京市建邺区江东中路 228 号华泰证券广场 1 号楼/邮政编码: 210019

电话: 86 25 83389999 / 传真: 86 25 83387521

电子邮件: ht-rd@htsc.com

### 深圳

深圳市福田区深南大道 4011 号香港中旅大厦 24 层/邮政编码: 518048

电话: 86 755 82493932 / 传真: 86 755 82492062

电子邮件: ht-rd@htsc.com

### 北京

北京市西城区太平桥大街丰盛胡同 28 号太平洋保险大厦 A 座 18 层  
 邮政编码: 100032

电话: 86 10 63211166 / 传真: 86 10 63211275

电子邮件: ht-rd@htsc.com

### 上海

上海市浦东新区东方路 18 号保利广场 E 栋 23 楼/邮政编码: 200120

电话: 86 21 28972098 / 传真: 86 21 28972068

电子邮件: ht-rd@htsc.com