

中国上市公司未来收益预测的实证研究

——贝叶斯动态模型及其预测

陆璇¹, 刘慧霞¹, 陈晓²

(1. 清华大学数学科学系, 北京, 100084; 2. 清华大学经济管理学院, 北京, 100084)

摘要: 预测上市公司的未来收益是投资者、证券商、债权人和管理层所关注的问题。国内外的一些实证研究结果表明: 上市公司定期公布的财务报告中包含关于公司未来收益变化的信息。基于此, 利用中国上市公司年报中的信息, 采用贝叶斯动态回归模型对公司未来收益的变化进行预测, 并将其结果与静态回归模型的预测结果进行比较, 认为贝叶斯动态回归模型的预测效果在一般情况下优于静态回归模型, 但要取得更好的预测效果还有待于更多的数据积累, 同时还应考虑宏观经济环境变化的影响。

关键词: 上市公司; 收益预测; 贝叶斯动态回归模型

中图分类号: F830.91

文献标识码: A

文章编号: 1672-3104(2003)02-0223-04

公司未来持续的赢利能力是投资者、证券商和债权人普遍关注的问题, 因为只有具备持续赢利能力的公司才能给投资者带来持续的红利回报。普通投资者所能得到的关于上市公司经营状况的最主要和直接的信息来自于上市公司定期公布的财务报告。因此, 站在普通投资者的立场, 以下问题是值得研究的: ①这些财务报告中是否包含关于公司未来赢利能力的信息? ②若回答是肯定的, 是否可以找到较好的方法来利用这些信息, 获得较好的预测效果?

国内外学者的研究结果对第一个问题做了肯定的回答。如 Ou 和 Penman 从美国上市公司的财务报告中选取了 68 个变量, 围绕公司下一年盈利增加的概率 Pr 建立了一个 Logit 回归模型。他们发现 Pr 与未来的盈利变化是正相关的^[1]。Abarbanell 和 Bushee 也对美国上市公司的财务报告做了类似研究。与 Ou 和 Penman 的做法不同, 他们从分析师常用的基本信号中选了 9 个作为预测变量, 用对横截面数据的线性回归模型来考察它们对于公司未来赢利的预测能力。分析结果亦表明: 这种预测能力是存在的^[2]。

在我国, 赵宇龙做了赢余数字在中国股市有用

性的研究^[3]。陈晓等人的研究结果也提供了类似证据^[4]。通常认为公司当前赢利的信息包含在股票的市价中, 但对于长期投资者(持股 1 年以上)来说, 重要的是预见未来。财务报告除了包含赢余数字外, 还有大量描述公司经营状况的数据。在有足够可信度的前提下, 这些数据应能为投资者提供关于公司未来赢利能力的信息。陆璇等^[5]分别采用 Ou、Penman 和 Abarbanell、Bushee 所用的方法对 A 股市场上上市公司的年报进行了分析, 得出了类似结论: 上市公司的财务报告所提供的信息可以用来对公司未来的赢利变化进行预测。但是否可以找到较好的方法来利用这些信息, 从而获得较好的预测结果, 即是本文要研究的问题。

经济领域常用的预测模型有 ARMA 模型、指数加权模型以及贝叶斯动态回归模型等。与西方发达国家成熟的股市相比较, 中国股市(以 A 股市场为代表)是一个新兴的、发展中的资本市场, 其主要特点是: 历史短、变化大(规模不断扩大, 政策不断变化)、披露信息量少(除年报外, 中报的历史很短, 且至今没有季报)、不够规范。基于以上原因, 我们认为 ARMA 模型和指数加权模型的预测效果不会太好, 因为这两种模型都要求有较长历史且较稳定的

数据。而采用贝叶斯动态回归模型有以下优点:首先,动态模型假定模型参数在随时间变化,这符合股票市场的实际情况。其次,基于贝叶斯观点的动态模型,其主要特征是根据过去所有的信息(即先验信息)和当前的数据对模型参数的分布进行不断修正,得到后验分布,再从后验分布出发对未来进行预测。由于这种修正是随时间推移不断进行的,因此可以更好地模拟股市的运动,从而有望得到更好的预测效果。

一、贝叶斯动态模型及有关理论

贝叶斯动态预测模型是英国统计学家 P J Harrison 和 C F Stevens 于 1971 年提出的一种预测方法,基本特点是:

- ① 模型参数是随机变量,有初始的先验分布;
- ② 在每一指定时刻,模型参数的先验分布是过去所有信息的综合;
- ③ 模型参数的后验分布是当前信息与先验分布的综合;
- ④ 对未来的预测基于当前的后验分布。

我们对贝叶斯动态线性回归模型做一些技术上的说明。这个模型是由两个方程和一个初始条件所确定的一个系统。其定义如下:

$$\text{观测方程: } y_t = F_t' \theta_t + v_t, \quad v_t \sim N[0, V_t]$$

$$\text{状态方程: } \theta_t = \theta_{t-1} + \omega_t, \quad \omega_t \sim N[0, W_t]$$

$$\text{初始先验: } (\theta_0 | D_0) \sim N[m_0, C_0]$$

在观测方程中, y_t 为被预测变量在 t 时刻的观测值向量, F_t 为预测变量在 t 时刻的观测值矩阵, θ_t 为线性回归模型在 t 时刻的参数向量, v_t 为正态零均值误差项, V_t 为方差。与传统的静态线性回归模型不同,在动态线性回归模型中,模型参数向量 θ_t 被看做是随机变量,它随时间的推移而变化。在状态方程中,假定 θ_t 对于 θ_{t-1} 有一个变化 w_t , 其分布为正态、零均值、方差 W_t 。假定观测和状态误差序列($\{v_t\}$ 和 $\{w_t\}$) 相互独立,且在每一序列内部对于不同的时刻 t 变量之间相互独立,当 $W_t = 0$ 时,状态方程退化为恒等式,动态线性回归模型退化为静态线性回归模型。

记任一时刻 t 时的信息集 $D_t = \{y_t, F_t, D_{t-1}\}$,

其中 D_{t-1} 是 $t-1$ 时刻的信息集, $\{y_t, F_t\}$ 是当前(t 时刻)的数据。首先求先验分布 $\theta_t | D_{t-1}$, 再求预测分布 $y_t | D_{t-1}$, 然后求后验分布 $\theta_t | D_t$, 最后得到预测分布 $y_{t+1} | D_t$ 。以上过程随时间推移不断进行递推修正。当观测方差 V_t 已知时,递推公式如下:

$$(I) \quad t-1 \text{ 时刻的后验分布: } (\theta_{t-1} | D_{t-1}) \sim N[m_{t-1}, C_{t-1}]$$

$$(II) \quad t \text{ 时刻的先验分布: } (\theta_t | D_{t-1}) \sim N[a_t, R_t]$$

$$(III) \quad t-1 \text{ 时刻的一步预测分布: } (y_t | D_{t-1}) \sim N[f_t, Q_t]$$

$$(IV) \quad t \text{ 时刻的后验分布: } (\theta_t | D_t) \sim N[m_t, C_t]$$

$$\text{其中: } a_t = m_{t-1}; R_t = C_{t-1} + W_t; f_t = F_t' a_t; Q_t = F_t' R_t F_t + V_t; e_t = y_t - f_t; A_t = F_t R_t / Q_t; m_t = m_{t-1} + A_t e_t; C_t = R_t - A_t A_t' Q_t$$

在上述递推公式中, W_t 是难于设定的,因而对 R_t 通常采用折扣法来确定: $R_t^{-1} = \delta C_{t-1}^{-1}$, 其中 δ 为折扣因子, $0 < \delta < 1$ 。即当前的先验精度等于折扣的上一次后验精度。通过选取不同的值,可以选择最佳的预测模型。当观测方差 V_t 未知时,对 V_t 也要假定一定的先验分布(χ^2 分布),并进行递推修正,此时模型参数的精确分布为 t 分布。

二、实证结果

(一) 预测模型 样本与变量

我们采用的贝叶斯动态回归预测模型如上节所述。其中 $y_t = (y_{1t}, \dots, y_{nt})'$, 下标 1 至 n 为公司编号, t 为年份; $y_{it} = \Delta EPS_{it+1} = EPS_{t+1} - EPS_t$, 为公司 i 在 $t+1$ 年度对于 t 年度的收益变化; F_t 为预测变量在 t 时刻的观测值矩阵(预测变量如表 1 所示); θ_t 为回归系数向量。我们不仅要直接考察贝叶斯动态模型的预测效果,还要将其与静态回归模型的预测效果进行比较。我们所使用的数据基本与文献 [5] 相同,均来自上海和深圳 A 股市场上市公司 1992~1999 年的年度财务报告。由于在预测 $t+1$ 年的收益时,要用到从 $t-2$ 到 $t+2$ 年的数据,所以我们利用这 8 年的财务报告数据,能做的预测周期只有 4 个,相应的样本数量见表 2。

表 1 预测变量

编号	变量	计算公式
1	存货	$\Delta \text{存货} = \Delta \text{主营业务收入}$
2	应收帐款	$\Delta \text{应收帐款净额} = \Delta \text{主营业务收入}$
3	毛利	$\Delta \text{主营业务收入} - \Delta \text{主营业务利润}$
4	消费与管理费用	$\Delta \text{消费与管理费用} = \Delta \text{营业收入}$
5	备抵坏帐	$\Delta \text{应收帐款} - \Delta \text{备抵坏帐}$
6	实际税率	当期税前收益* (去年的实际税率- 当期的实际税率)
7	总资产周转率	当期主营业务收入* 2/(当期总资产+ 上一年年报总资产)
8	资产负债率	(当期长期负债合计+ 当期流动负债合计)/当期总资产
9	固定资产周转率	当期主营业务收入* 2/(当期固定资产合计+ 上一年年报固定资产合计)
10	ΔPTE	税前收益变化* (1- 上一年的实际税率)
11	规模控制变量	$\log(\text{总资产})$

注: 1. Δ 代表变量与用 2 年平均模型计算出来的期望值的百分比变化。例如: $\Delta \text{存货} = [\text{存货}_t - E(\text{存货}_t)] / E(\text{存货}_t)$, 其中 $E(\text{存货}_t) = (\text{存货}_{t-1} + \text{存货}_{t-2}) / 2$
2. 实际税率= 当期税费/ 税前收益。

表 2 样本数量

年份	1995 预测 1996	1996 预测 1997	1997 预测 1998	1998 预测 1999
深市	130	219	321	371
沪市	143	225	337	384
合计	273	444	658	755

(二) 分析结果

用贝叶斯动态回归模型做预测时, 选择适当的折扣因子可以取得较好的预测效果。我们在拟合和预测中尝试了不同的 δ 值, 发现在多数场合下取 $\delta = 0.6$ 时的效果最好。我们用实际值落入预测区间的百分比来衡量预测效果的好坏, 预测区间的置信度分别取 85%、90% 和 95%。表 3 给出了在 $\delta = 0.6$ 时预测的准确率, 从中可以看出, 贝叶斯动态回归模型的预测效果相当好。也就是说, 实际预测的准确率在多数情况下高于置信度, 只在个别情况下与置信度相当。

表 3 贝叶斯动态回归模型预测准确率($\delta = 0.6$) / %

预测区间的 置信度	95 预测 96 的准确率	96 预测 97 的准确率	97 预测 98 的准确率	98 预测 99 的准确率
85	97.1	91.9	91.2	92.3
90	97.8	93.9	92.9	93.1
95	98.2	95.0	94.7	95.6

我们采用两种评价准则对静态回归模型和贝叶斯动态模型进行预测效果的比较: 平均绝对误差 $MAD = \sum_{i=1}^n |e_n| / n$ 和均方误差 $MSE = \sum_{i=1}^n |e_n| / n$, 其中 e_n 为预测值与实际观测值的差, 它们的值越小, 说明预测的效果越好。比较的结果见表 4。

表 4 静态回归模型和贝叶斯动态回归模型
预测效果比较($\delta = 0.6$)

年份	指标	静态回归模型	贝叶斯动态回归模型
1995 预测	MAD	0.184 2	0.184 2
1996	MSE	0.189 3	0.189 3
1996 预测	MAD	0.203 4	0.160 7
1997	MSE	0.519 5	0.073 6
1997 预测	MAD	0.214 9	0.180 0
1998	MSE	0.381 5	0.108 5
1998 预测	MAD	0.199 3	0.209 8
1999	MSE	0.133 5	0.192 7

从中可以看出, 在第 1 个预测年份(1995 预测 1996), 贝叶斯动态回归模型的预测效果与静态回归模型的预测效果相当。这是因为在开始阶段我们没有过去的先验信息, 所以贝叶斯动态回归模型不能提供更好的预测效果。在接下来的 2 个年份中(1996 预测 1997、1997 预测 1998), 贝叶斯动态回归模型的预测效果明显优于静态回归模型的预测效果。但是在由 1998 年的数据预测 1999 年的赢利时, 贝叶斯动态回归模型的预测效果还不如静态回归模型的预测效果。其原因可能是因为宏观经济的形势发生恶化(1997 年亚洲金融危机, 及 1998 年的自然灾害等) 造成上市公司的业绩与过去相比有较大的跳跃。

由于中国股市的历史很短, 积累的数据很少, 加之我们的研究经费和规模有限, 因此目前的研究结果虽然在一定程度上显示了贝叶斯动态回归模型在

预测公司未来赢利方面的优越性,但还不能充分显示其全部效能。

三、结语

随着我国计划经济向市场经济的转轨,中国证券市场在不到10年的时间里迅速发展,到2000年底沪深两个交易所已有上市公司近千家。证券市场已经在我国国民经济发展中产生了重要影响。对上市公司未来收益的预测既是对国民经济未来发展的预测,也是广大投资者所关注的问题。

我们利用A股市场上市公司的年度财务报告所提供的信息,在贝叶斯动态回归模型下对公司的未来盈利变化预测,取得了较为理想的预测效果。将其结果与静态回归模型的预测结果进行比较,以平均绝对误差和均方误差作为模型预测性能的评估准则,初步结论是:在正常情况下,贝叶斯动态模型的预测效果优于静态回归模型。但当宏观经济形势发

生大的变化时,用贝叶斯动态模型进行预测的效果反而不如静态回归模型。其原因并不在于贝叶斯动态模型本身,而是由于我们暂时未能将宏观经济形势的变化指标输入模型。

参考文献:

- [1] Ou J A, Penman S H. Accounting measurement, price-earnings ratio, and the information content of security prices[J]. *Journal of Accounting Research*, 1989, 27(1): 111-144.
- [2] Abarbanell J S, Bushee B J. Fundamental analysis, future earnings and stock prices[J]. *Journal of Accounting Research*, 1997, 35(1): 1-24.
- [3] 赵宇龙. 会计盈余披露的信息含量[J]. *经济研究*, 1998, 7(2): 41-49.
- [4] 陈晓, 陈小悦, 刘钊. A股盈余报告的有效性研究[J]. *经济研究*, 1999, 6(3): 21-28.
- [5] 陆璇, 陈小悦, 张岭松, 等. 中国上市公司财务基本信息对未来收益的预测能力[J]. *经济科学*, 2001, 6(2): 53-62.
- [6] West M, Harrison J. Bayesian forecasting and dynamic models[M]. New York: Springer-Verlag, 1996. 15-26.

On earnings forecast of Chinese listed companies ——Bayesian Dynamic model and forecasting

LU Xuan¹, LIU Huixia¹, CHEN Xiao²

(1. Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China;

(2. School of Economics and Management, Tsinghua University, Beijing 100084, China)

Abstract: Forecasting of listed companies' future earnings is an important issue for investors, stockbrokers, creditors and administrators. Studies by foreign and domestic researchers show that financial reports disclosed by the listed companies contain information about the earnings variability in the future. On the basis of these studies, the authors use Bayesian dynamic regression model to forecast the variation of the company's prospective earnings, using the data in the annual financial reports of Chinese listed companies, and compare the empirical analysis result under Bayesian dynamic regression model with the result under ordinary static regression model. The authors' conclusion is that Bayesian dynamic regression model generally offers better forecast than static regression model and that more empirical data and information about the macro-economic circumstances must be obtained in order to have more satisfactory forecast.

Key words: listed company; earnings' forecast; Bayesian dynamic linear model