

股票代码的数字偏好

四川大学 赵绍阳 王 坤*

1. 数字偏好

数字偏好与厌恶根植于每一个国家的文化之中,人们认为特定的数字或数字组合能为他们带来好运或者厄运,这是一种典型的非理性心理行为。比如,由于宗教原因,西方国家一般比较忌讳数字“13”。在中国,受到传统文化的影响,数字“4”由于读音接近“死”而被人们所回避。在沿海一些城市,尾号为“4”的手机号码的销售价格会相对便宜,尾号为“4”的车牌号人们也很少选择;相反,数字“8”由于谐音“发”而备受青睐。2006 年时,温州车牌号“浙 C88888”拍出了 166 万元的高价,北京奥运会开幕式选在 2008 年 8 月 8 日晚 8 时举办,也非常直观地体现出中国人对数字“8”的喜爱。此外,代表着“六六大顺”的数字“6”,个位数中最大的阳数、象征着至高无上的数字“9”,都受到人们格外的偏爱。那么这种广泛存在于生活中的数字偏好现象是否也会出现在金融市场中呢?

2. 金融市场中的数字效应

国外学者对外国金融市场的研究中,数字效应相关研究主要集中于股票价格的聚类效应。Smith 和 Kendall (1939) 对数字样本的分布情况进行检验,发现样本中的数字并没有完全服从随机分布。Osborne (1962) 通过对纽约证券交易所的股票收盘价数据的分析发现,股票收盘价为整数的概率最大。Goodhart 和 Curcio (1991) 则研究了外汇市场上的价格聚类效应,发现在外汇报价尾数中,0 出现的概率最大,5 为其次。Grossman 等 (1997) 在对伦敦股票市场进行研究的过程中发现,股票市场报价 0 和 5 出现次数最多。Kandel 等 (2001) 得出结论,即:在日常生活中,人们使用偶数的频率要比使用奇数的频率要大。

研究人员对我国股票市场进行分析的过程中,也发现了股票价格的尾数偏好现象。Brown 等 (2002) 对香港股票市场股票收盘价进行研究,发现香港股票收盘价尾数有明显的数字偏好及厌恶现象,数字“8”出现的频率最高,数字“4”出现的频率最低。Brown 和 Mitchell (2008) 对大陆股票市场的收盘价进行研究,也发现了明显的数字偏好现象,“8”多“4”少。刘凤元 (2008) 研究了中国股票市场的价格集群现象,选取沪深两市 20 只股票在 2005 年 1 月 3 日至 2006 年 12 月 29 日间的每笔成交价共计 5762544 个数据为样本,实证结果明显拒绝了关于 0-9 这十个数字是均匀分布的假设,得出结论国内股票市场存在价格集群效应,0、5、8 出现的次

* 赵绍阳获西南财经大学经济管理研究院经济学博士学位,现为四川大学经济学院讲师;他的电子邮箱是 zhaoshaoyang@scu.edu.cn。王坤为四川大学经济学院在读本科生;她的电子邮箱是 404121481@qq.com。

数最多。

3. 股票代码的尾数效应

我国的股票代码为六位编码制,其中沪市前三位和深市前两位标志着股票代码类型,沪市后三位和深市后四位为股票的顺序编码区,而尾数则是顺序编码区里最有代表意义的数字,也是最能引起数字偏好的一位数字。那么,在这里是否也存在明显的吉利数字偏好现象呢?

赵静梅和吴风云(2009)选取1995年1月至2005年12月在上交所上市的569家A股非金融股票的市盈率为样本,分析表明:对数字“8”的崇拜现象抬高了股票的相对价格,即“发财代码价格贵”。研究还发现,“8”的崇拜并不能抬高股票的收益率,尽管代码尾数为“8”的股票组合一场收益率在上市首日是上升最高的,但在随后的12个月中的下降幅度也是较大的。叶成超(2010)选择我国制造业A股股票代码尾数为8和4且在2007年收盘价在20元以内的企业为样本,对样本进行了最小二乘法回归检验,得出结论,即:股票代码尾数为“8”对股票的交易量有显著正面的影响,且能引起股价的上涨,表明中国文化中对于“8”的数字偏好对股价涨跌有显著影响。盛卫锋、张兵和谢世宏(2011)以上海证券交易所A股市场2007年6月10日至2008年11月4日之间的周收益率为样本,使用CAPM模型对样本进行检验发现:在牛市中,投资者对于股票代码尾数为“4”的股票规避现象十分显著,但熊市时并不显著;另一方面,投资者对以吉利数字为代码尾数的股票并没有显示出明显的偏好。张顺明、唐唯(2015)以2004年以后非主板市场上的股票作为研究对象,通过对不同尾数组合市盈率及收益率的分析,得出中国证券市场曾经出现过显著的代码效应,但在近几年不复显著的结论,认为这在一定程度上反映了数字迷信的逐步破除和投资者素质的不断增加,即中国证券市场的有效性在不断增强。

接下来,我们将使用1990年2月至2014年12月之间于沪、深二市A股及创业板上市的全部公司(共2,100家,其中上证A股1,108家,深圳A股508家,创业板484家)的相关数据,对其进行检验,数据取自CSMAR(国泰安)数据库。

3.1 “数字偏好”现象的检验

为了避免各市场之间存在差异,我们在对全部上市公司股票数据进行统计的同时,也对沪市上市公司股票和深市上市公司股票以及创业板上市公司股票数据分别进行了统计。

首先,我们将所有上市公司股票以股票代码尾数0-9进行分组,统计了股票代码尾数0-9出现的公司数目,并计算出了各组公司数占总数的频率,绘制出表1。

表1 上市公司股票代码尾数分布情况

尾数	所有上市公司		上证A股上市公司		深证A股上市公司		创业板上市公司	
	公司数	频率%	公司数	频率%	公司数	频率%	公司数	频率%
0	209	9.95	106	9.57	56	11.02	47	9.71
1	205	9.76	106	9.57	51	10.04	48	9.92
2	193	9.19	93	8.39	51	10.04	49	10.12
3	206	9.81	103	9.3	54	10.63	46	10.12
4	106	5.05	46	4.15	14	2.76	49	9.50

续表

尾数	所有上市公司		上证 A 股上市公司		深证 A 股上市公司		创业板上市公司	
	公司数	频率%	公司数	频率%	公司数	频率%	公司数	频率%
5	206	9.81	104	9.39	53	10.43	49	10.12
6	232	11.05	126	11.37	57	11.22	49	10.12
7	205	9.76	106	9.57	50	9.84	49	10.12
8	290	13.81	177	15.97	64	12.6	49	10.12
9	248	11.81	141	12.73	58	11.42	49	10.12

可以看出,无论是对全部上市公司,还是对沪市、深市分别进行统计,股票代码尾数为 4 的上市公司都最少。而以数字 8 为代码尾数的上市公司最多,尾数为 6 和 9 的上市公司数量紧随其后。具体来看,在全部上市公司中,股票代码尾数为 4 的上市公司频率仅为股票代码尾数为 8 的上市公司频率的 36.57%;在沪市,代码尾数为 4 上市公司比例几乎是尾数为 8 的上市公司的 1/4;而深市则更为明显,代码尾数为 4 的公司数量只有全部公司数量的 2.76%。据此不难得出结论,即:主板市场的股票代码尾数分布呈现出极为明显的“吉利数字偏好”现象,也就是说,上市公司普遍偏好吉利数字 6、8、9 作为代码尾数,而回避忌讳数字 4。

为了对股票代码尾数分布状况的年变动趋势有所把握,接下来我们按照年份对 1990—2014 期间沪市 A 股新上市公司的代码尾数进行了描述性统计,发现:自 2004 年之后,沪市股票代码尾数中“吉利数字”和“忌讳数字”的频率分布更是发展成为两个极端:忌讳数字 4 已完全消失,而吉利数字开始占据绝对优势,单单数字“8”的出现频率就已接近全部数字的 1/3,6、8、9 的吉利数字组占比更是高达 63%。

那么,上证股票代码中同样属于顺序编码区的第二位、第三位数字是否会有同样的“数字偏好”现象呢?将沪市 2004 年以后上市公司的股票代码倒数第二、第三位也做了同样的描述性统计,结果如表 2 所示。由统计图表可以看到,上述问题的答案是肯定的:沪市股票代码中位于顺序编码区的三位数字都未能免遭影响,忌讳数字“4”的出现频率近乎为零,数字“6”“8”“9”的出现频率却明显高于本应差别不大的临近数字“5”“7”(0、1 的高频率应该是数字序号靠前的结果)。据此,可以进一步证实前文所提出的“股票市场存在数字偏好现象”的猜想。

表 2 沪市 2004 年以后上市的公司股票代码后三位数字分布情况

数字	倒数第一位		倒数第二位		倒数第三位	
	公司数	频率%	公司数	频率%	公司数	频率%
0	17	6.23	45	16.48	55	20.15
1	20	7.33	52	19.05	39	14.29
2	10	3.66	28	10.26	19	6.96
3	16	5.86	19	6.96	28	10.26
4	0	0	1	0.37	2	0.73
5	15	5.49	14	5.13	21	7.69
6	34	12.45	34	12.45	30	10.99
7	23	8.42	14	5.13	17	6.23
8	87	31.87	31	11.36	25	9.16
9	51	18.68	35	12.82	37	13.55

3.2 不同尾数组组合的市盈率比较

我们参考赵静梅、吴风云(2009)的方法,选取静态市盈率(股票价格/股票在上市前一年的每股收益)作为市盈率的计算指标,并在剔除负值、异常值后采取简单算术平均法计算平均市盈率。在此基础上,计算任意指定数字尾数的股票与其他股票的平均市盈率差值。

表3是所有尾数组组合与其对应组合上市一年后的平均市盈率及其差值。可以看到,“吉祥数字”6组合、8组合、9组合都未呈现出明显的代码溢价效应,8组合甚至出现了折价现象;而单看“忌讳数字”4组合时,的确存在最为明显地组合市盈率远低于对照组的情况。

表3 上市一年所有尾数组组合与其对应组合的平均市盈率及其差值

	0	1	2	3	4	5	6	7	8	9
M	47.78	41.36	48.29	59.10	32.05	46.36	50.59	41.81	47.37	51.73
\hat{M}	47.64	48.36	47.58	46.33	48.34	47.81	47.29	48.30	47.69	47.14
差值	0.14	-7.01	0.71	12.78	-16.29	-1.45	3.30	-6.48	-0.32	4.59

3.3 股票代码尾数的影响

经统计,上市公司股票代码选择时的“数字偏好”现象已经被证实的确存在,那么该现象会对股票本身带来什么影响呢?拥有吉利数字作为股票代码尾数的公司是否真的能够收获他们所期待的高收益率呢?下面我们开始进行检验。根据上文中的统计结果,沪、深二市的“数字偏好”现象明显较为明显,适合用来检验股票代码尾数的影响,但由于2004年之后沪市不再出现尾数为4的股票代码,为了保证被检验数据的统一性,避免将大多于2004年以前上市的代码尾数为4的股票与上市时间分布均匀的其他代码尾数进行比较,可能导致因上市时间分布的差别而造成的数据本身的差异,本文仅选择使用沪、深二市A股(中小板除外)2004年以前上市的公司的相关数据作为主要研究对象。此外,选择创业板自上市以来的数据作为参照对比。

这里我们选择1990—2004年期间上市的1492支股票作为样本,并将其按照股票代码尾数的不同分为3组,即“忌讳数字组”(包括尾数为4的所有样本,以下简称为“4组”),“普通数字组”(包括尾数为1、2、3、5、7、0的所有样本,以下简称为“5组”),“吉利数字组”(包括尾数为6、8、9的所有样本,以下简称为“6组”),通过剔除异常值后取组内均值的方法,分别追踪了每组股票上市十年内的收益率、交易量、换手率表现。为了更加准确地对三组股票的表现做出总体判断,本文又选择股票代码选择方面几乎不存在“数字偏好”现象的创业板作为对照,但由于创业板上市时间较短,数据相对较少,因此使用了2014年12月31日之前上市的全部406支股票,并采用同样的处理方法,追踪了其自上市至今的表现。

3.3.1 股票代码尾数对收益率的影响

从图1不难看出,沪、深二市,“4”“5”“6”三组股票收益率变化方向大致相同,波动幅度也基本相似,且从长期来看,三组股票变化趋势亦相同。不过,“4”组的收益率较之于“5”“6”组的收益率存在周期性的差别:上市后的前6年,“4”组平均收益率要高于“5”“6”组,但是之后“4”组平均收益率要低于“5”“6”组。在样本期间,这三个组在上市十年之内全部月收益率的

算数平均值分别是 1.80%、1.65%、1.64%。如果考虑每组收益率的波动,则三组股票的平均收益率没有显著的差别。另外需要注意的是,“4”组收益率波动更大可能是沪、深二市“4”组样本数量较少的原因。经统计,图 1 中“5”组包括 926 只股票,“6”组包括 496 支,而“4”组只有 70 支,这样在作图前的均值处理时,“4”组均值更容易受到极端值的影响,从而使得总体变化幅度较大。图 2 中创业板股票“4”“5”“6”三组股票收益率变化则基本相同。

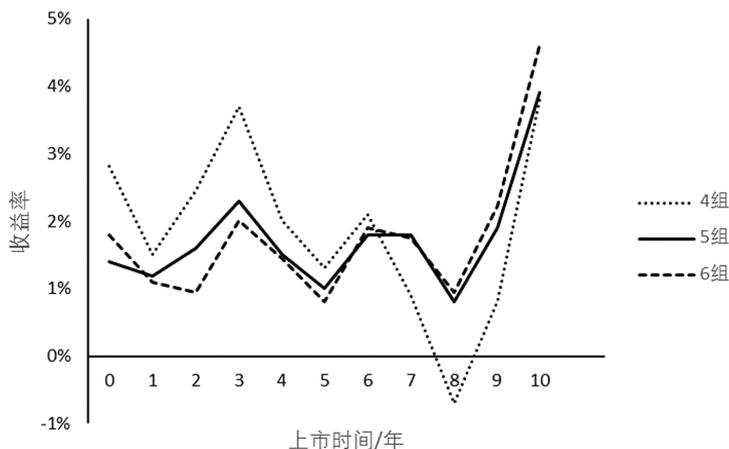


图 1 沪、深二市“4”“5”“6”组合上市十年收益率变化

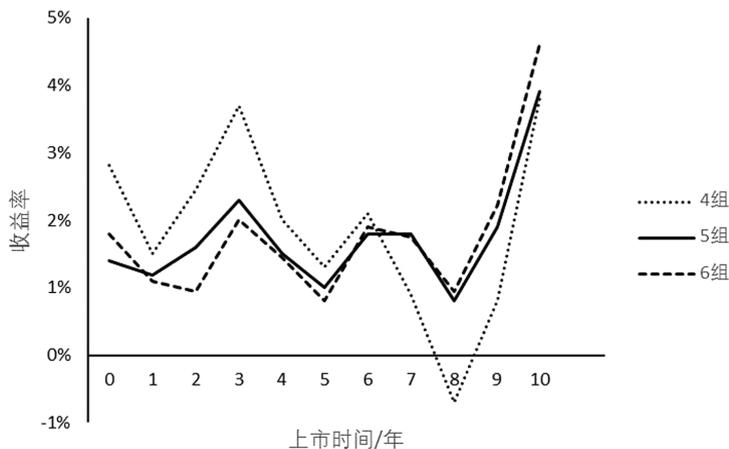


图 2 创业板“4”“5”“6”组合上市至今收益率变化

3.3.2 股票代码尾数对换手率的影响

换手率(流通市值/交易金额)是反映股票流动性的重要指标,因此我们按照同样的方法对该指标进行分组求均值处理,并做出相应图形如下图 3 和图 4。

结果显示,沪、深二市“4”组的换手率较之于“5”“6”组的换手率存在周期性的差别:上市后的前 6 年,“4”组平均换手率要高于“5”“6”组,但是之后“4”组平均换手率要低于“5”“6”组。但如果整体来看(上市十年内),则三组换手率没有显著差异。创业板的换手率整体要比

主板的换手率要高,“4”组的换手率在长期来看,要比“5”“6”组的换手率稍低。

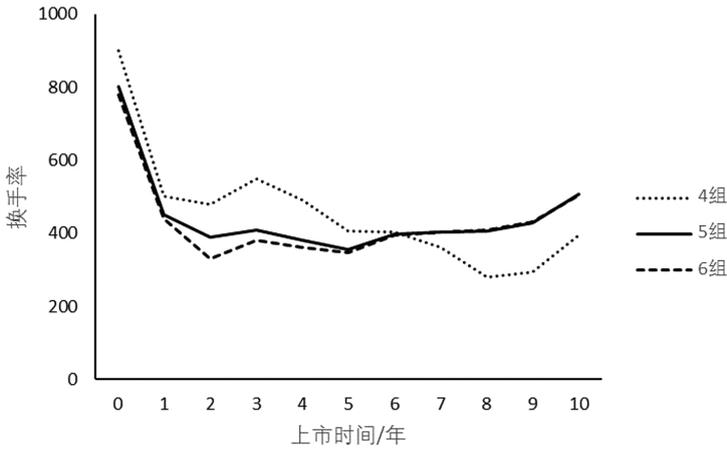


图3 沪、深二市“4”“5”“6”组合上市十年换手率变化

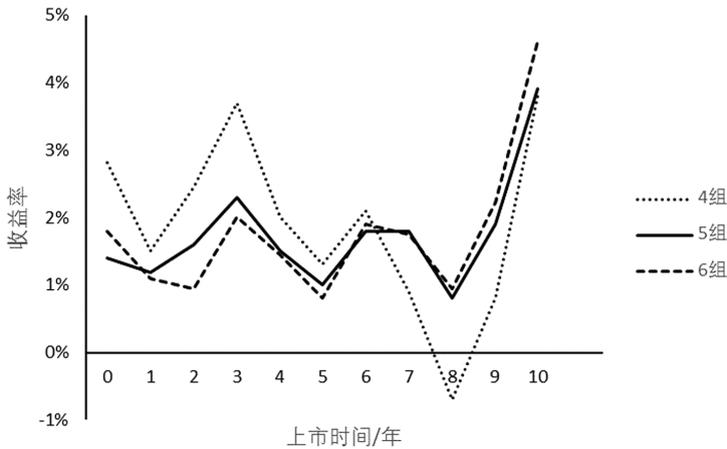


图4 创业板“4”“5”“6”组合上市至今换手率变化

4. 结果解释

通过以上分析可以看到,上证、深证A股的上市公司在选择股票代码时出现了明显的“数字偏好”现象,即股票代码为“忌讳数字”4的公司数量远少于平均值,“吉祥数字”

则相对较多,这可能是因为证券交易所确定股票代码的方式给了上市公司选择吉祥数字的空间。

根据我国股票代码分配规则,证券代码采用六位阿拉伯数字编码,取值范围为000000~999999。六位代码的前三位为证券种类标识区,其中第一位为证券产品标识,第二位至第三位为证券业务标识,六位代码的后三位为顺序编码区。对于顺序编码球的数字,上海证券交易所目前采用申请选择制,即上市公司可以自主申请喜欢的股票代码;而深圳证券交易所采取批文

顺序制,即按照上市批文的时间顺序来确定上市公司的股票代码,但上市公司仍可以通过改变上市时间来避开不喜欢的尾数或等待喜欢的代码,如华润三九医药股份有限公司成功选择了和其公司名称相配的代码:000999。两种确定上市公司股票代码的方式虽有所不同,但都给予了上市公司认为选择股票代码的可能,于是也在市场中表现出显著的“数字偏好”效应。相比之下,在创业板上市公司中就没有出现很明显的“数字偏好”现象,各数字出现的频率几乎相等。

这种现象的出现既有可能是上市公司管理层数字偏好的反映,也可能是公司对投资者行为预测后的应对策略。具体而言,上市公司预期潜在的投资者具有一定程度的吉利数字偏好,进而预期具有相应吉祥数字作为尾数的代码会受到投资者追捧,这些股票的需求增加,股价上升。因此,上市公司为了募集到更多资金,也会表现出与预测中投资者一致的数字偏好,而这种偏好在公司实力越强大时越容易得到满足。相较之下,在创业板上市的公司大多数规模较小或创立时间不长,没有足够的实力来自主选择上市的股票代码,只能由证券交易所按各公司上市时间来分配股票代码,并没有出现数字偏好的现象。但从股票的实际表现来看,我国投资者并非如预期中一样迷信吉利数字,说明投资群体素质正不断提高,投资行为日趋理性,因而上市公司并没有过分在意代码尾数的必要。

参考文献:

- Brown, P. and Mitchell, J. 2008, Culture and Stock Price Clustering: Evidence from PRC, *Pacific-Basin Finance* 1, 95-120.
- Goodhart, C. and Curcio, R., 1991, The Clustering of Bid-ask Prices and Spread in the Foreign Exchange Market, *LSE Financial Markets Group Discussion Paper Series* 110.
- Grossman, J., Miller M. and Cone K., 1997, Clustering and Competition in Asset Markets, *Journal and Economics* 40, 23-60.
- Kandel, S., Sarig, O. and Wohl, A., 2001, Do Investors Prefer Round Stock Price? Evidence from Israeli IPO Auctions, *Journal of Banking and Finance* 25, 1543-1551.
- Kendall, M. and Smith, B., 1938, Randomness and Random Sampling Numbers, *Journal of the Royal Statistical Society* 1, 147-166.
- Osborne, M., 1962, Periodic Structure in the Brownian Motion of Stock Prices, *Operations Research* 10, 345-375.
- 刘凤元,2008,社会文化、数字偏好与股票报价:中国股市的价格群集现象研究,《中国软科学》,第6期,28-33。
- 盛卫锋、张兵、谢世宏,2011,中国投资者对股票代码的数字偏好及其影响——来自上海证券交易所的经验证据,《西部论坛》,第4期,59-65。
- 叶成超,2010,数字偏好与股票市场反应——来自中国上市公司的经验证据,《财会通讯》,第6期,95-97。
- 张顺明、唐唯,2015,数字迷信与金融资产价格异象——中国证券市场“代码折溢价效应”研究,《投资研究》,第3期,107-118。
- 赵静梅、吴风云,2009,数字崇拜下的金融资产价格异象,《经济研究》,第6期,129-141。