

## 附录 A 数据来源

将数据分为三类，原始输入数据，中间处理数据，输出结果。原始输入数据包括 2002 年，2007 年及 2012 年的中国国民经济投入产出表的中间投入部分，分别对应 2002-123-raw.csv,2007-129-raw.csv,2012-139-raw.csv 文件。以及 WIND 资讯中提取的证监会行业指数，包括自指数编制以来，每个交易日的开盘价，收盘价，涨幅和跌幅，对应文件是 Wind-证监会指数-完整版.csv。证监会的行业分类与投入产出表的分类存在一些差异，这些差异主要体现在三个方面。一：一部分的行业在证券市场中没有上市公司，例如公益服务业，政府机构。二：有一部分的行业指数，一个指数包含多个子行业。三：由于行业分类标准在 2011 年发生了变动，因此，2007 年及 2002 年的投入产出表中，有一部分指数对应的行业是当时没有细分的。为了克服差异造成统计上的不便，我们额外制作了三个表格，每个表格都建立了当年投入产出表中各行业与证监会行业指数的对应关系。接着，我们按照这些表格，将投入产出表各行业中的流量，转化为各行业指数对应实体经济的流量。具体的转换操作可见源程序 Raw-To-CC-Re.R，输出的流量矩阵对应的文件是 Flow12-Re.csv,Flow07-Re.csv,Flow02-Re.csv。接着，计算各个行业的收益率序列。我们先从月收益率序列开始，月收益率计算的方式是  $t$  月的指数价格，除于  $t-1$  月的指数的价格，再取对数。得到了月收益率序列后，将它存在文件 monthlyield-zoo-series.csv 中。转换完成后，则可以开始对行业的上游和下游产业作分析。对行业矩阵的列向量，取各列的前  $k$  个对应的行业，记为每个行业的上游产业的收益率序列。类似地，对流量矩阵的行向量，取各行前  $k$  个对应的行业，记为每个行业的下游产业的收益率序列。为了测试交叉预测性，对上游行业和下游行业的序列作一个单位的时延，将时延后得到的上下游收益率序列与行业收益率序列作组合，得到一个面板数据。再对这个面板数据作最小二乘回归，回归的程序保存在 SectorRegression-beta.R 文件中，回归的结果见附录 A-1 表。

表 A.1

<i>Dependent variable:</i>	
	yield.trimtail
i1	0.024 (0.021)
i2	-0.003 (0.016)
i3	0.008 (0.013)
i4	-0.010 (0.012)
o1	0.016 (0.016)
o2	0.003 (0.012)
o3	0.028** (0.011)
o4	-0.012 (0.008)
Observations	12,775
R <sup>2</sup>	0.001
Adjusted R <sup>2</sup>	0.001
F Statistic	1.618 (df = 8; 12592)

*Note:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01