

数据收集

数据来源:

萝卜投研

用到的指标/变量:

1. GDP(当季值)
2. 沪深300收盘价(日度)
3. 宏观经济景气指数(先行值)
4. 宏观经济景气指数(一致值)

一致指数是反映当前经济的基本走势，由工业生产、就业、社会需求（投资、消费、外贸）、社会收入（国家税收、企业利润、居民收入）等4个方面合成；先行指数是由一组领先于一致指数的先行指标合成，用于对经济未来的走势进行预测

5. 活期存款利率(月度)
6. CPI(当月同比)
7. 货币供应量M2(同比)
8. 定期存款利率(1年整存整取)
9. 国房景气指数

全国房地产开发景气指数（简称“国房景气指数”）遵循经济周期波动的理论，以景气循环理论与景气循环分析方法为依据，运用时间序列、多元统计、计量经济分析方法，以房地产开发投资为基准指标

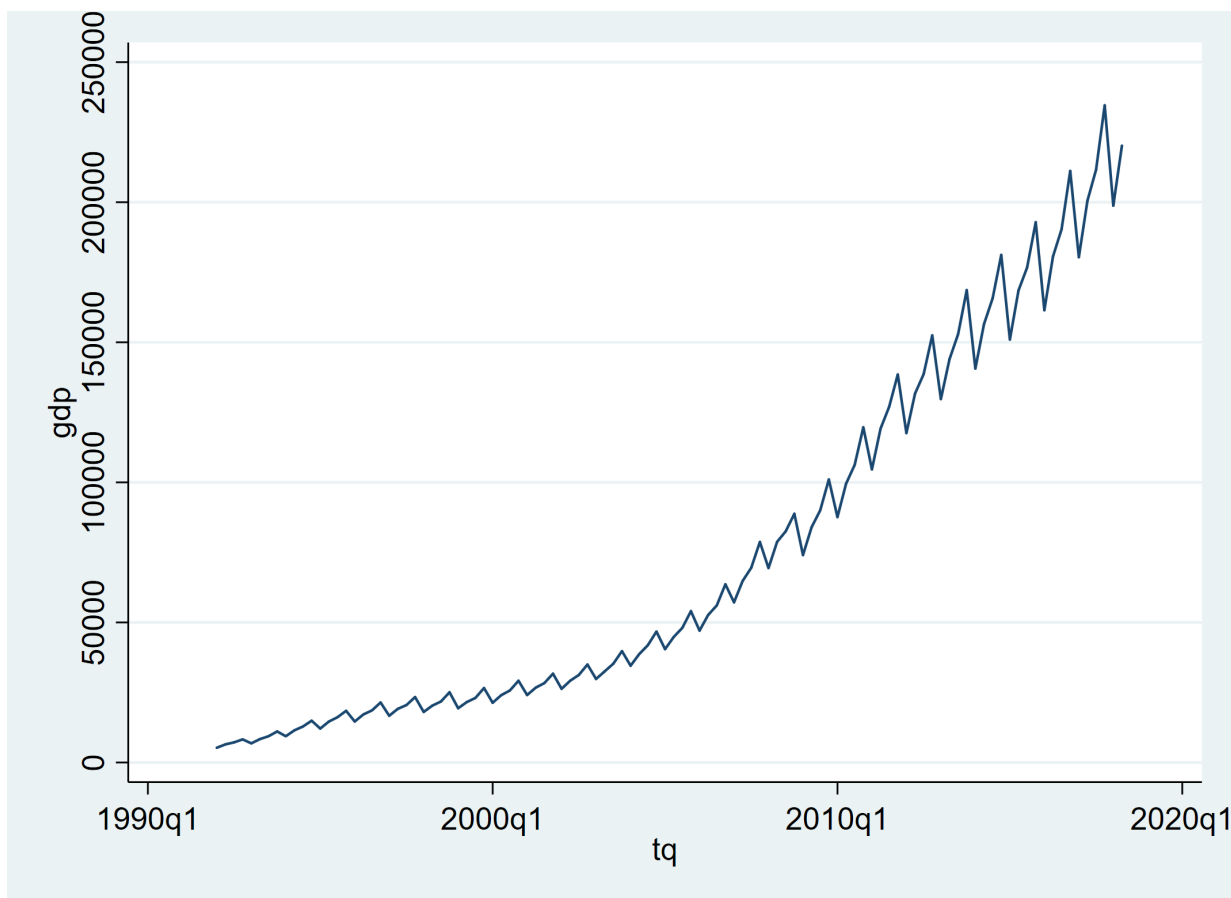
国房景气指数100点是最合适的景气水平，95至105点之间为适度景气水平，95以下为较低景气水平，105以上为偏高景气水平

数据处理

数据频率转化

GDP

1. 首先，GDP的季度值存在季节趋势



为了消除这种季节趋势，我们需要对GDP季度值数据进行**季节调整**

季节调整

季节调整通常通过估计“**季节因子(seasonal factor)**”来进行。根据季节因子起作用的方法，主要分为“**加法季节因子(additive seasonal factor)**”，与“**乘法季节因子(multiplicative seasonal factor)**”。

加法因子意味着对所有的第一月(或第一季)加上相同的季节因子，以此类推。

$$Y_t = TC_t + S_t + I_t$$

其中

Y_t 为原序列， TC_t 为趋势循环要素， S_t 为季节要素，而 I_t 为不规则要素

乘法模型的数学表达式如下：

$$Y_t = TC_t \times S_t \times I_t$$

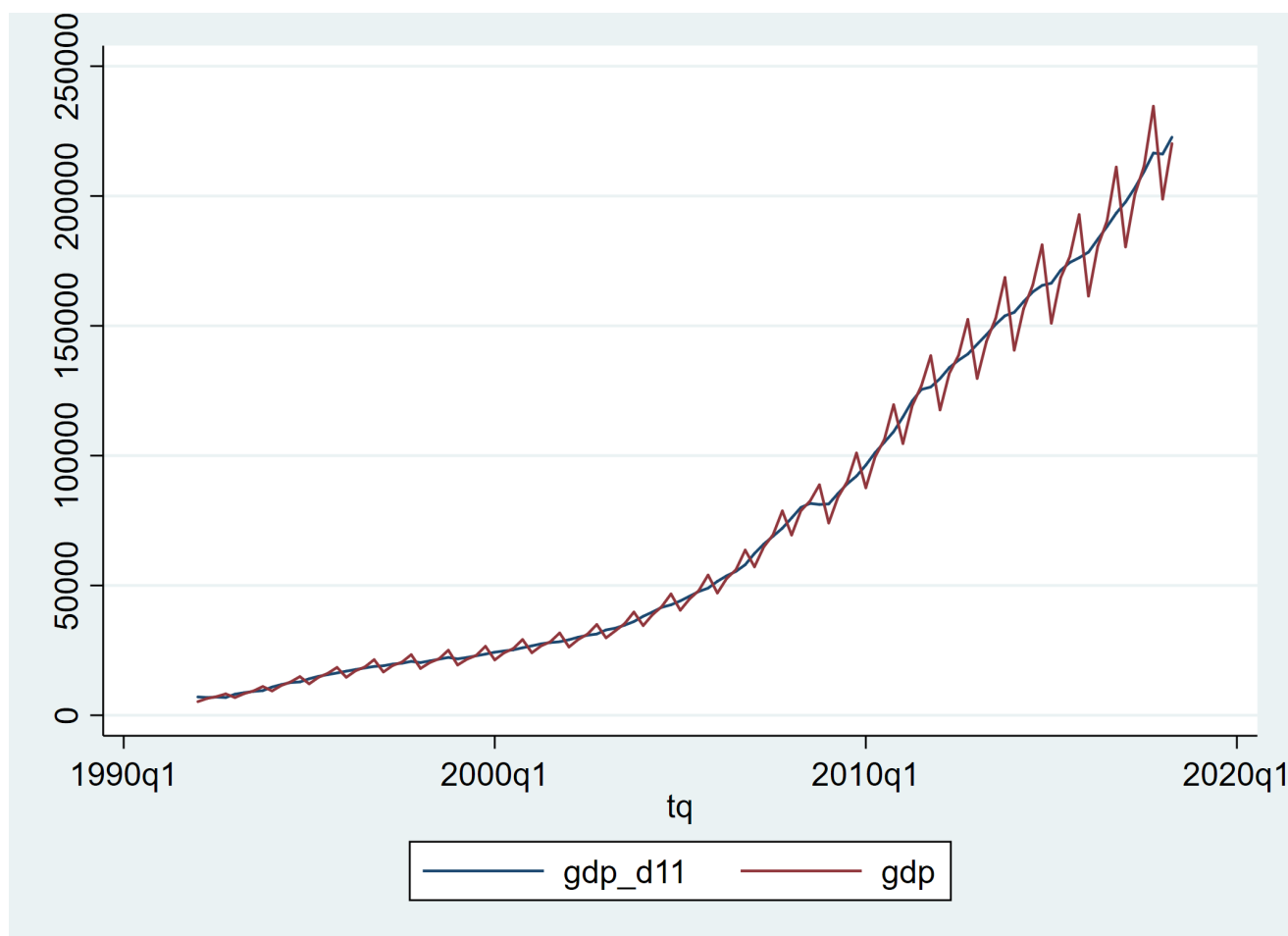
或者以对数形式

$$\ln Y_t = \ln TC_t + \ln S_t + \ln I_t$$

以上为季节调整的主要两种模型。季节调整的具体方法有多种，不同方法有不一样的结果，所以这是一项有主观性的工作。陈强老师编著的《*高级计量经济学及Stata应用(第二版)*》中介绍了季节调整的三种方法：**回归法**，**移动平均率法**与**X12方法**。其中X12方法是美国商务部人口普查局开发的用于在计算机上运行的季节调整程序。我们采用了X12方法。

关于X12-ARIMA方法的详细内容见PDF附件：《时间序列X - 12 - ARIMA季节调整——原理与方法（中国人民银行调查统计司）》

经过季节调整以后的GDP stata tsline绘图如下：其中gdp_d11是调整过后的序列，gdp是原序列



2. GDP是最重要的变量，但是我国GDP统计为一个季度一次，因此需要将GDP转化为**月度频率数据**，这是数据**频率由低到高的转化**，存在一些不可避的问题。如果我们学过**随机过程**那么可以用**MIDAS模型**，这也是目前最热门的方法。因为知识受限，我们采用**Eviews**直接转化数据频率的方法

我们选择Quadratic方法

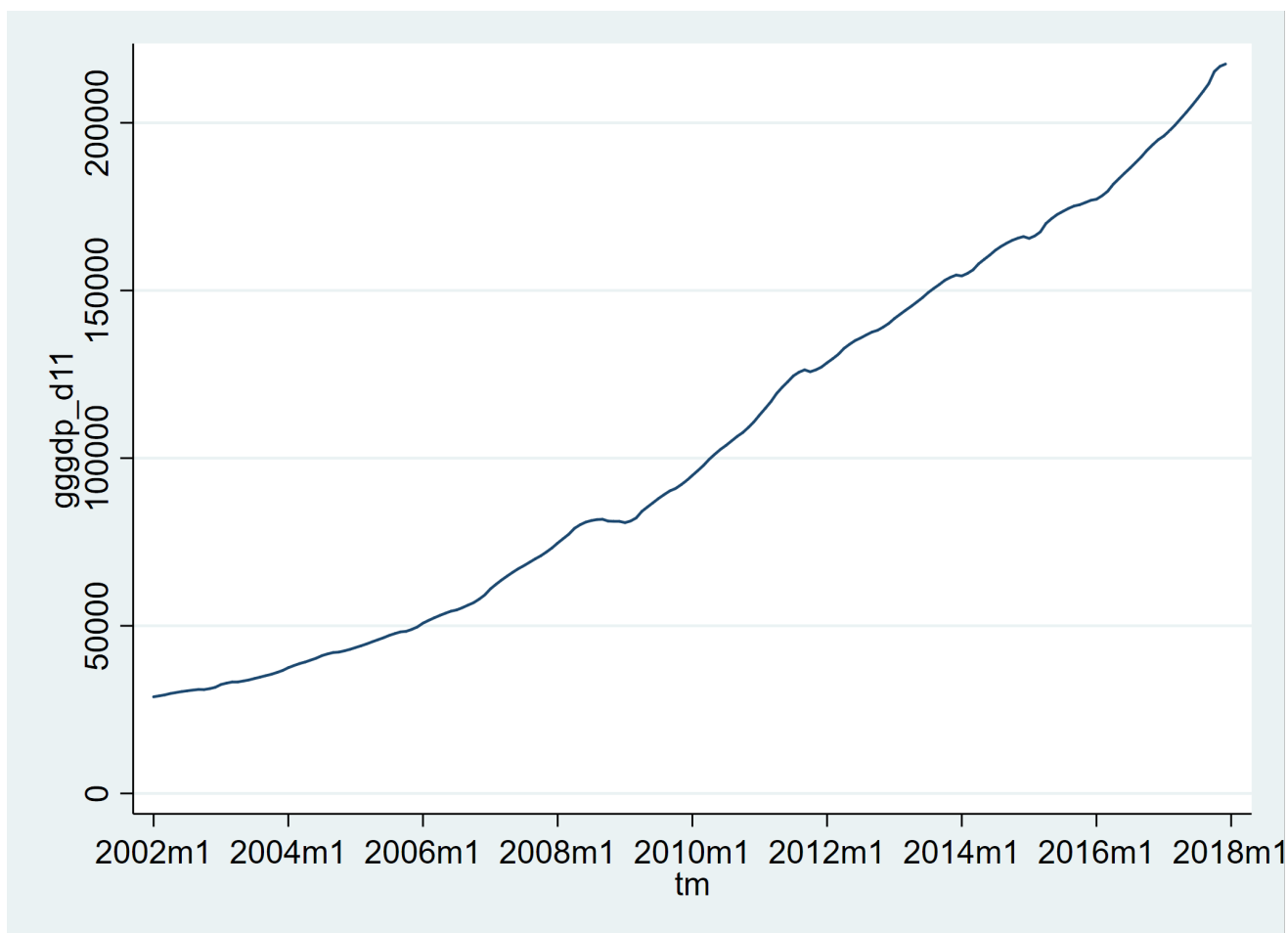
(Eviews的这几种转化方法还没全部学习，之后会详细学习，如果这里有比Quadratic更适合的方法随时可以更换使用)

Eviews关于Quadratic的描述:

Fits a local quadratic polynomial for each observation of the low frequency series, then use this polynomial to fill in all observations of the high frequency series associated with the period. The quadratic polynomial is formed by taking sets of three adjacent points from the source series and fitting a quadratic so that either the average or the sum of the high frequency points matches the low frequency data actually observed. For most points, one point before and one point after the period currently being interpolated are used to provide the three points. For end points, the two periods are both taken from the one side where data are available.

This method is a purely local method. The resulting interpolation curves are not constrained to be continuous at the boundaries between adjacent periods. Because of this, the method is better suited to situations where relatively few data points are being

经过频率转化以后的gdp走势图如下(这里截取了从2002年开始到现在的观测值):



沪深300收盘价

沪深300收盘价是日度数据，每周一到周五有观测者。我们将沪深300收盘价同样利用Eviews转化数据频率，这次是高频数据转低频。

Paste Special

Paste hs300 as

Pattern: *

Name: hs300

Paste as

☐ Series (by Value)

☒ Link

Merge by

☒ Date with frequency conversion

☐ General match merge criteria

Frequency conversion options

High to low frequency method

Average observations

Specified in source

Average observations

Sum observations

First observation

Last observations

Max observation

Min observation

No down conversions

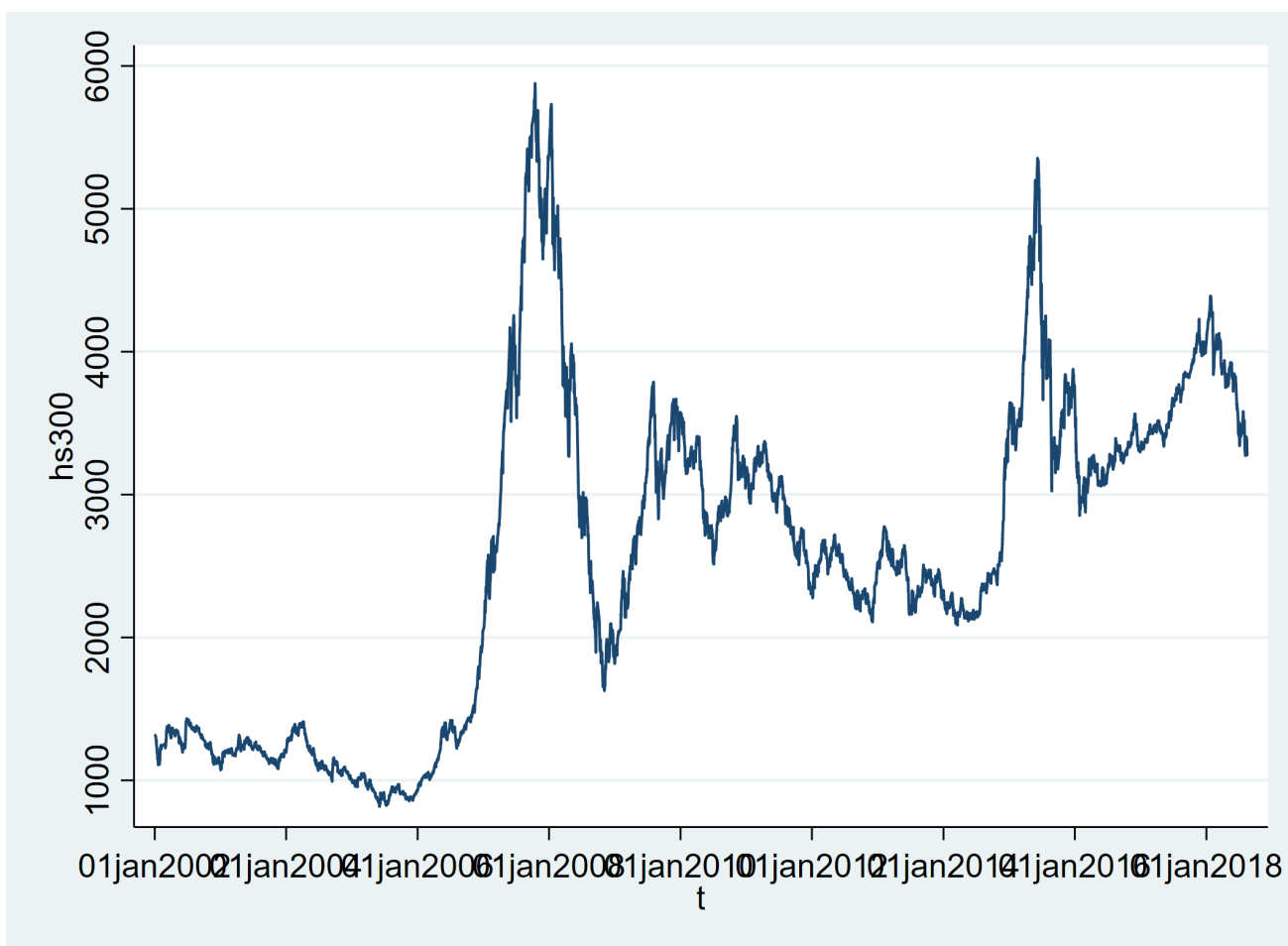
OK

OK to All

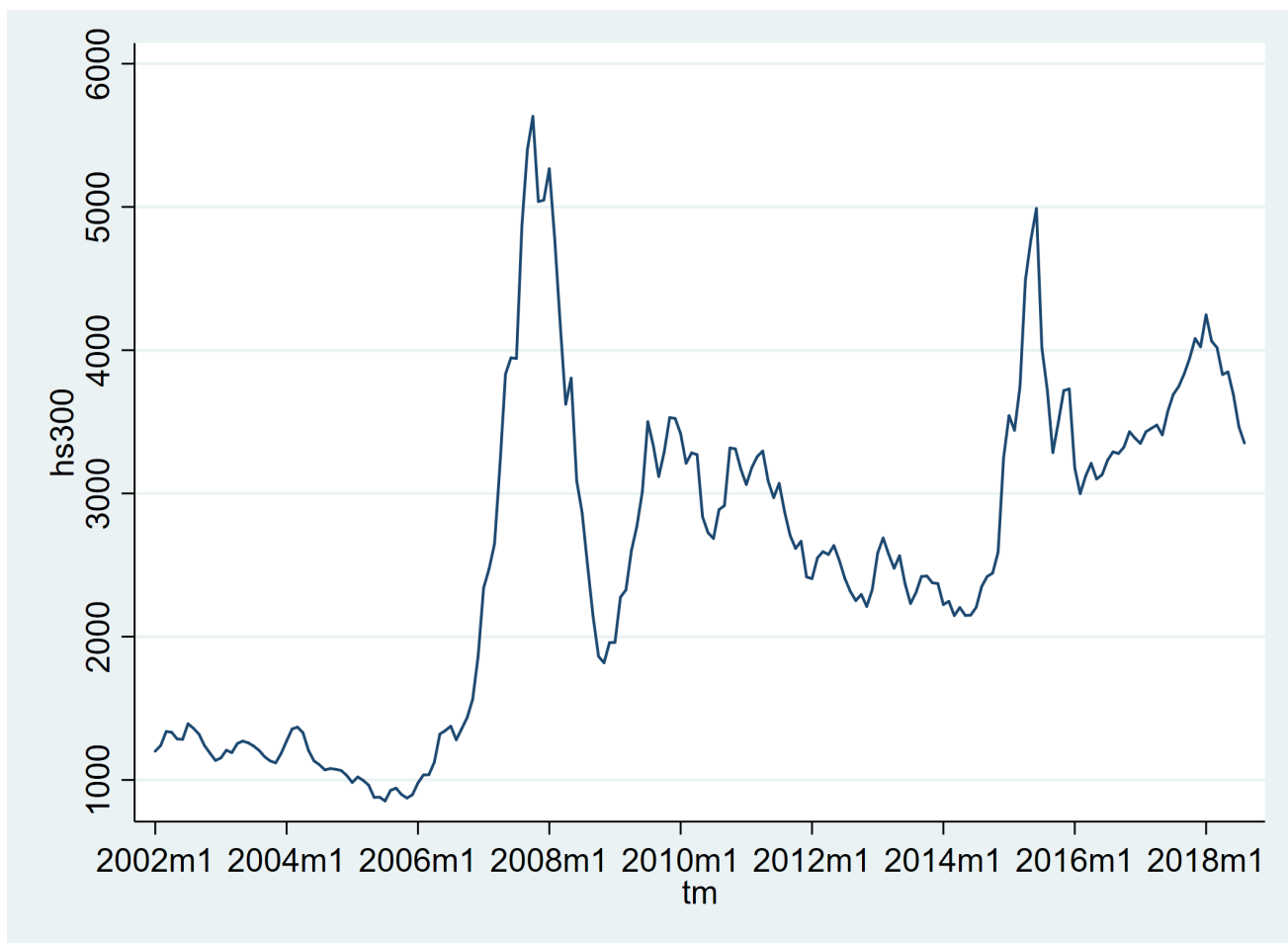
Cancel

Cancel All

频率转化前的HS300收盘价:



频率转化过后的HS300收盘价：



数据整合

我们在处理GDP和HS300时候是在一个只有时间变量和GDP/HS300的dta文件下操作，同时我们还有其他频率已经是monthly的指标，所以我们要将他们合并到同一个dta中。这里用到stata的merge命令

```
merge 1:1 tm using gdp.dta , assert(1)
merge 1:1 tm using hs300.dta , assert(1)
```

最后的dta文件见群文件

模型建立

详细的操作记录见群文件的stata log文件

我采用了VAR模型

首先varsoc选取滞后阶数，最后确定二阶

之后建立var模型，检验var稳定性，发现结果为不稳定