Appendix: Oil Prices and Field Production: A Montecarlo Simulation Experiment

Johannes Mauritzen
Department of Business and Management Science
NHH Norwegian School of Economics
Helleveien 30, 5045
Bergen, Norway
johannes.mauritzen@nhh.no
jmaurit.github.io

April 8, 2014

1 Introduction

The data set used in this paper has a relatively complex structure. The dependent variable - field level oil production - has a non-linear autocorrelated structure. The production profiles are also correlated with each other - leading to a bell-shaped total production curve. In addition the exogenous variable of interest - oil prices - are autocorrelated and non-stationary. In such a setting, standard procedures for inference may not hold. In this appendix I instead present a Monte Carlo simulation experiment. Generating artificial data that is similar in character to the data used, I first show how failing to properly control for the production profile can lead to a heavily biased estimation of the effect of oil prices. I then show how using a generalized additive model can lead to correct inference.

In a simpler time-series setting the solution might be to take the first difference of the series. This is not possible in this setting since it is the actual levels of oil prices that are important. Intuitively, identification comes from comparison of fields of similar size and at similar points in their production lifetime but at different periods of oil production. This variation would be discarded by purely looking at first differences.

For illustration, I include code snippets in this document. The full code for the analysis can be found at `jmaurit.github.io\#oil_prices`

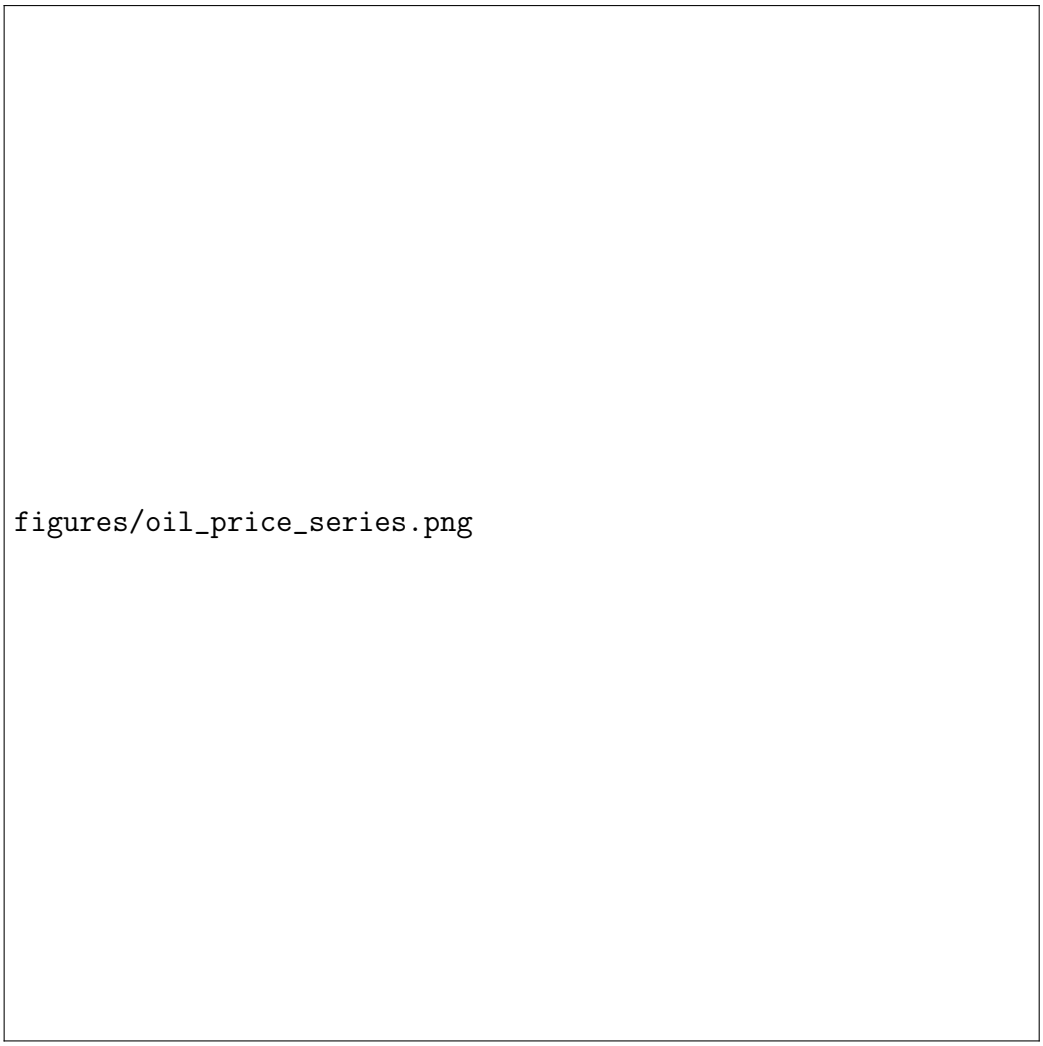

Figure 1: The Real Price of Brent Traded Oil from 1960 to 2013 in 2010 U.S. Dollars.

2 Generating Data

The first step is generate a set of fields from a size distribution. In the simulation, I generate 77 fields from a exponential normal distribution with mean 2.3 and standard deviation of 1.5. This is similar to the distribution of the actual 77 productive Norwegian fields.

```
field_size<-round(exp(rnorm(77,mean=2.3, sd=1.5)), digits=1)

gen_year<-function(size, maxsize){
#test
#size=15
#maxsize=1267
#

#let small fields be distributed uniformly from 1975 to 2013
if(size<10){
year<-trunc(runif(1, 1975, 2008))
} else{ #while large fields should be more common earlier on
range<-FALSE
while(range==FALSE){
year<-trunc(rnorm(1,mean=(1973+(maxsize+300)/(size+300)), sd=10))
ifelse(year>=1970 & year<=2013, range<-TRUE, range<-FALSE)
#print(paste("in while loop", year))
}
}
return(year)
}

max<-max(field_size)
init_year<-trunc(sapply(field_size, gen_year, maxsize=max))
```

3 Estimation of Coefficient