SAS® GLOBAL FORUM 2020 #SASGF

Paper 4870-2020

Is it possible to reduce smoking by increasing taxes? An analysis on Danish data using SAS® software

Anders Milhøj, University of Copenhagen

ABSTRACT

In Denmark all ideas to reduce smoking are discussed. It is commonly accepted that increasing prices is an efficient tool to reach this goal. In a Danish context taxes are frequently used for such purposes as we already have a heavy taxation of all tobacco products as we also tax alcohol, fuel and even for some years ago and extra taxation on the content of fat in goods like meat, cheese, chocolate etc.

The question is however whether it is true that smoking could be reduced by increasing taxes. This paper presents several statistical time series analyses by several procedures - VARMAX, X12, ARIMA and UCM - in the SAS® ETS software package to answer this question. The data is series for monthly and yearly sales of cigarettes combined with series for the price of cigarettes and the overall consumer price index together with detailed information on previous changes in the consumer tax for cigarettes.

The results indicate a price elasticity of at least 0.6, but the chock effect of a sudden price increase by an increasing tax fades out rater fast.

SALES OF CIGARETTES IN DEMARK

Statistics Denmark publishes data for the sales of cigarettes. The basis for this series is the detailed information upon the tax of cigarettes which is available each month. Moreover, consumer price indices for the price of cigarettes and also the general consumer price index, also published by Statistics Denmark, are used.

The Danish taxation on cigarettes is heavy. Roughly speaking the tax for a single cigarette at present is around one Danish Krone which corresponds to say 15 cents. The politicians discuss how much this tax should be increased in order to drastically reduce smoking in order to ensure public health. It is remarkable that the taxation on cigarettes in some European countries is even higher than in Denmark. Experiences from e.g. Norway are often referred to as the price of cigarettes in Norway are taxed much higher than in Denmark.

The yearly sales and price series for cigarettes are plotted in Figure 1. The general picture is of course the price has increased due to inflation etc. The sales decreased in the eighties, then increased in the period up to 2005 and afterwards the sales decreased rapidly - almost by a third.

Figure 1. Yearly sales and price for cigarettes

The price as such is of course not so relevant as inflation, exchange rates etc affect the price in (fixed) Danish Kroner. Instead the relative price index defined as

$$\text{relative price index} = \frac{\text{price of cigarettes}}{\text{consumer price index}}$$

is used in the analyses. Figure 2 displays this relative price series and the sales series. Both series are log-transformed as a starting point for the analyses to come.

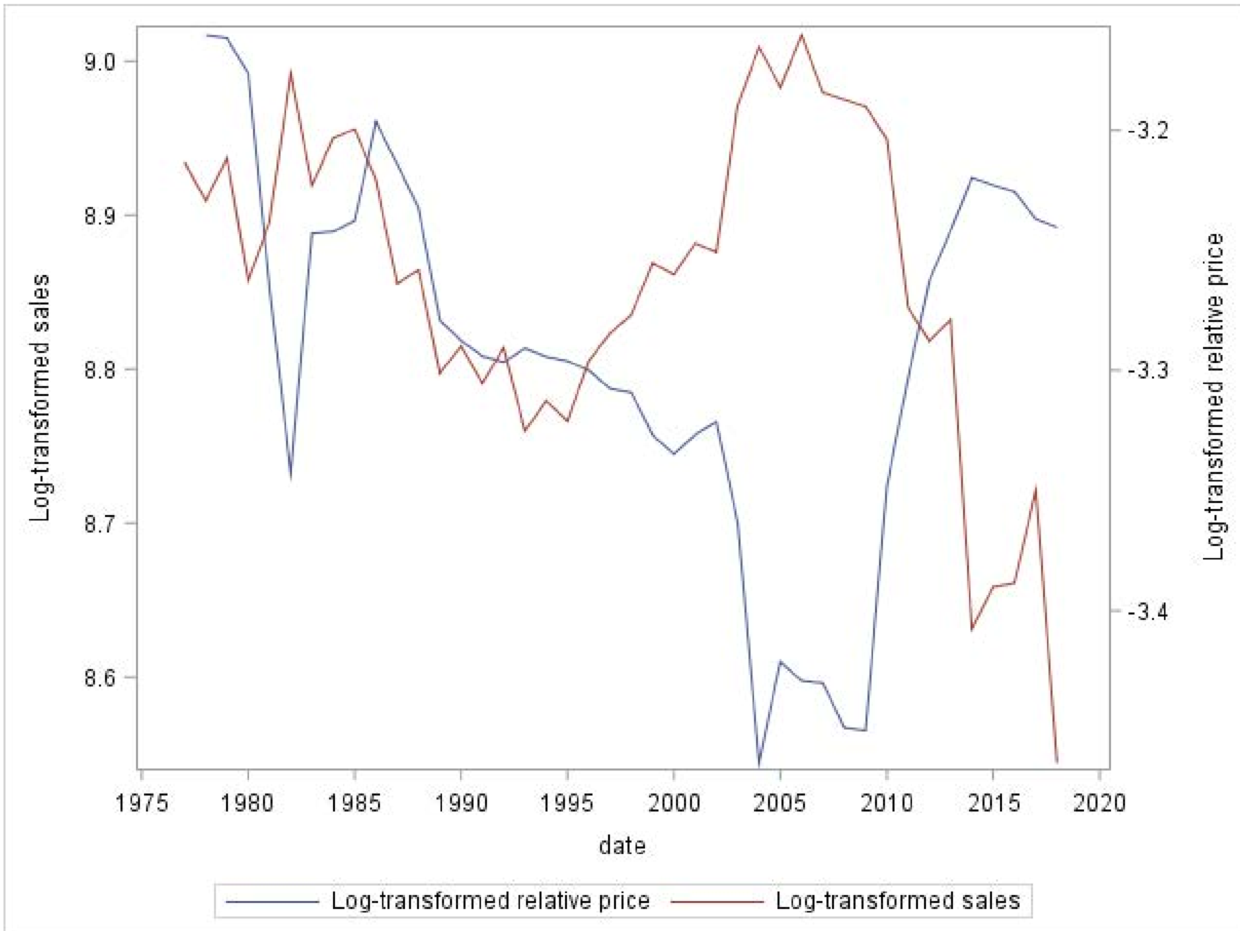

Figure 2. Log-transformed yearly sales and relative price

TIME SERIES MODELS FOR THE YEARLY SERIES USING PROC VARMAX

It is hard to see from Figure 1 and Figure 2 whether the price series affects the sales series. A thorough analysis following the principles by Box & Jenkins ends up with a rather simple model which tells that the year to year number of sold cigarettes is dependent upon the year to year relative price series with a lag one moving average term included in order to explain for some autocorrelation. This model could be estimated by several SAS® ETS procedures. Here is the code for an analysis using the VARMAX procedure

```
proc varmax data=c print=all plots=all;
  model Log_sales=Log_relative_price/
    dif=(Log_sales(1) Log_relative_price(1)) method=ml q=1 p=0 noint;
run;
```

The estimated parameters of the resulting model are seen in Output 1.

Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
Log_sales	XL0_1_1	-0.65697	0.20672	-3.18	0.0029	Log_relative_price(t)
	MA1_1_1	0.26786	0.14011	1.91	0.0633	e1(t-1)

Output 1. Model Parameter Estimates

which gives the ARIMA(0,1,1) model

$$(1 - B)S_t = -0.66(1 - B)P_t + \varepsilon_t - 0.27\varepsilon_{t-1}$$

The regression parameter, - 0.66, is clearly different from zero, but it is not significantly different from minus one. This means that the hypothesis of the price elasticity for sales of cigarettes equal one is accepted. Or in other words, that the budget for purchase of cigarettes on a macro level is fixed to a specific amount. This model passes the usual checks for model fit. However, two outliers for year 2014 and 2018 are present as seen from the Model Plot, Figure 3.

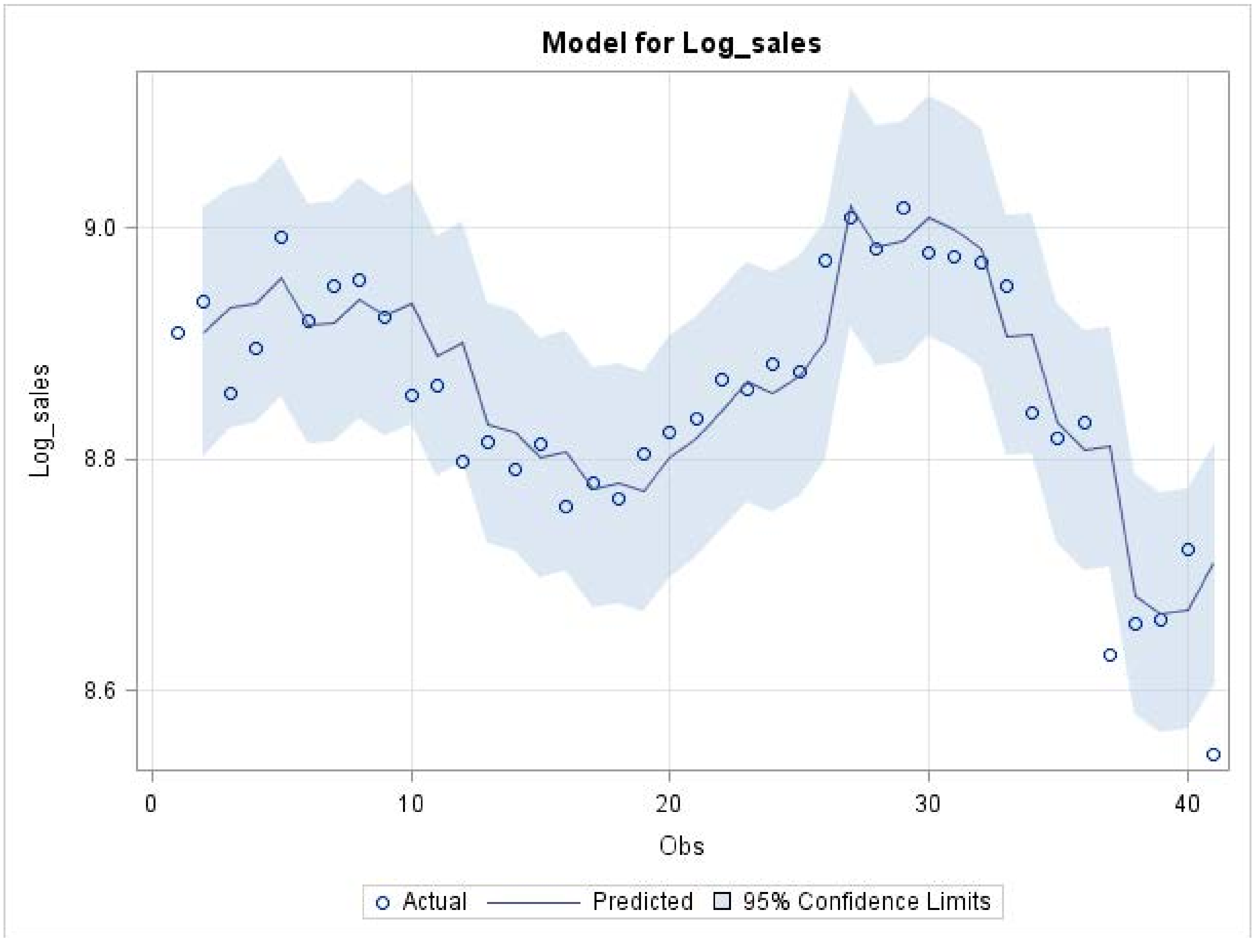

Figure 3. Model plot from the VARMAX procedure

TIME SERIES MODELS FOR THE YEARLY SERIES USING PROC UCM

Another approach to analyze these yearly series is to apply unobserved component models by the UCM procedure. In this setup the sales series is decomposed into a level component, a trend component which both could be time varying and an irregular component which could include some autocorrelation parameters. it also possible to include an explanatory variable and even allow the regression coefficient to vary over time.

The actual series, the log-sales series, has no visible trend, so a trend component is excluded. The level seems however to vary over time so a simple unobserved component models is estimated by this code

```
PROC UCM data=c plot=all;
  id date interval=year;
  model Log_sales;
  level;
  irregular q=1;
  estimate;
run;
```

The results indicate no serious autocorrelation problems. Figure 4 gives the Model Plot.

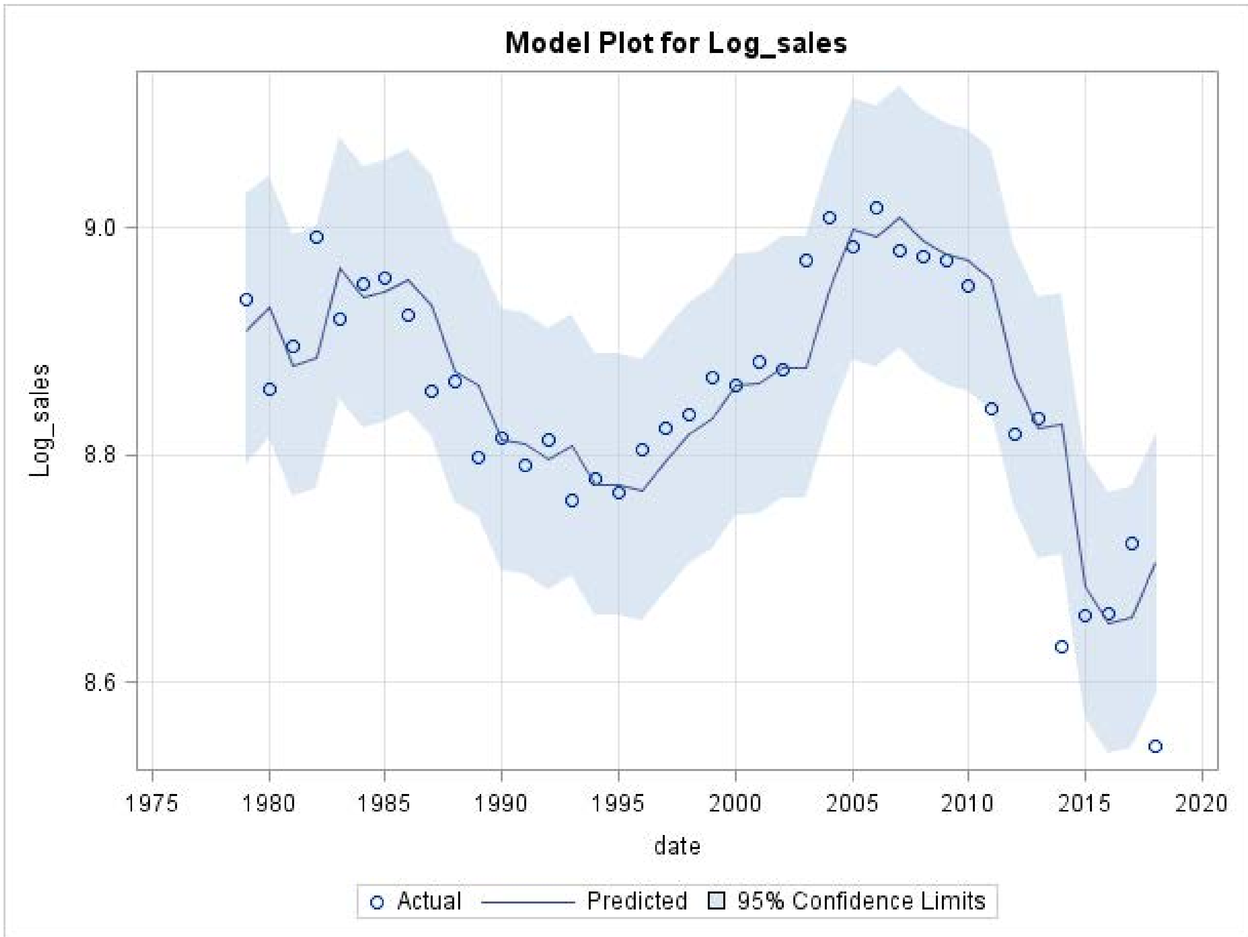

Figure 4. Model plot from PROC UCM

This model could be extended by the relative price series as an exogenous variable with a time varying regression coefficient which is possible by a randomreg statement in the UCM procedure:

```
PROC UCM data=c plots=all;
 id date interval=year;
  model Log_sales;
  level;
```

```
    randomreg Log_relative_price;
    irregular p=1;
    estimate;
  run;
```

The resulting parameters are presented as Output 2.

Component	Parameter	Estimate	Approx Std Error	t Value	Approx Pr > |t|
Irregular	Error Variance	0.00016568	0.0002981	0.56	0.5783
Irregular	AR_1	-0.72486	0.41766	-1.74	0.0826
Level	Error Variance	0.00206	0.0007642	2.70	0.0069
Log_relative_price	Error Variance	4.51976E-10	4.4797E-7	0.00	0.9992

Output 2. Estimates of the Free Parameters

The results from Output 2 prove that the error variance of the regression component is zero and hence the regression coefficient is constant, the randomreg plot, that is the plot of the time varying regression coefficient is a horizontal line at $\beta = -0.63$.

A regression with a fixed regression coefficient is fitted by the UCM procedure by the code

```
  PROC UCM data=c plots=all;
    id date interval=year;
    model Log_sales=Log_relative_price;
    level;
    irregular p=1;
    estimate;
  run;
```

The result, Output 3, again gives a regression coefficient corresponding to the price elasticity 0.63.

Component	Parameter	Estimate	Approx Std Error	t Value	Approx Pr > |t|
Irregular	Error Variance	0.00016568	0.0002981	0.56	0.5783
Irregular	AR_1	-0.72486	0.41762	-1.74	0.0826
Level	Error Variance	0.00206	0.0007642	2.70	0.0069
Log_relative_price	Coefficient	-0.63139	0.20406	-3.09	0.0020

Output 3. Estimates of the Free Parameters

The model is accepted and it has no autocorrelation problems.

THE MONTHLY SERIES

It is possible to give more precise answers by use of monthly data as many changes in taxation were effective from say April 1st. and not at the beginning of a new year. Monthly sales data is however only available from January 1. 2001, so it is impossible to go as far back in time as in the analysis by yearly data.

Figure 5 presents the monthly log-transformed relative price index of cigarettes. The red vertical lines represent changes in the taxation on cigarettes.

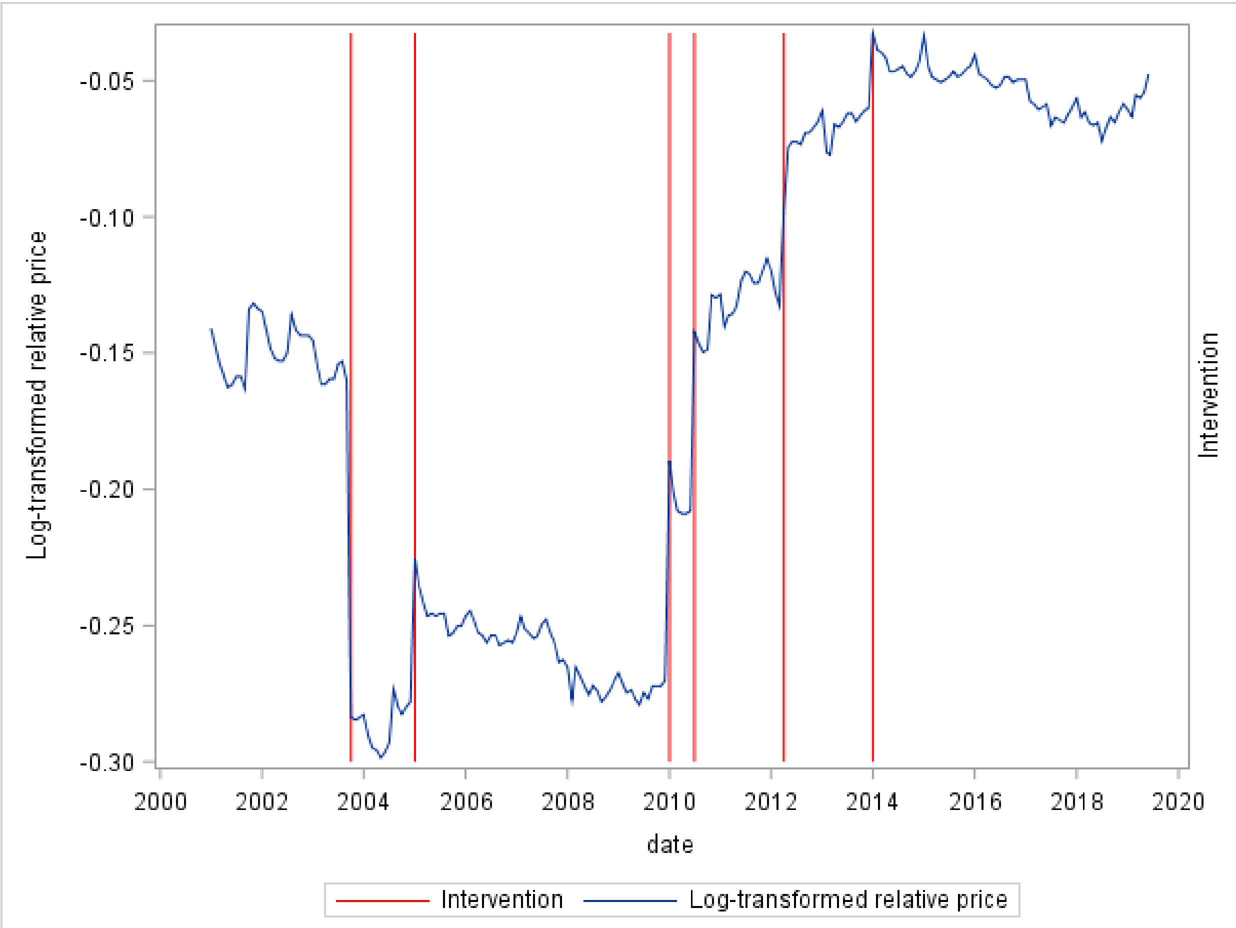

Figure 5. Log-transformed relative price with marked tax changes

Note that the change in taxation by October 1. 2003 gave a price reduction as the special tax on cigarettes was reduced in order to reduce border trade as limits on private import from abroad were deleted.

SEASONAL STRUCTURE OF THE MONTHLY SERIES

The monthly sales series include some seasonal variation which could be analyzed by the Census Seasonal Adjustment method, which is possible in SAS® by one of the procedures X11, X12 or X13.

The code is

```
PROC X12 data=i date=date;
  var log_sales;
  outlier type=ao;
  x11 mode=add trendma=23;
  regression predefined=(td Easter(4));
```

```
    automdl;
    output out=out a1 c17 d10 d11 d12 d13 c17;
  run;
```

The sales series is decomposed in output dataset into a trend cyclic, a seasonal component and an irregular component. Figure 6 presents two years of the seasonal component. In this plot the seasonal structure is seen in detail. Most remarkable are the small sales figure in January and high sales in October and December. One reason for this is the holiday season in December.

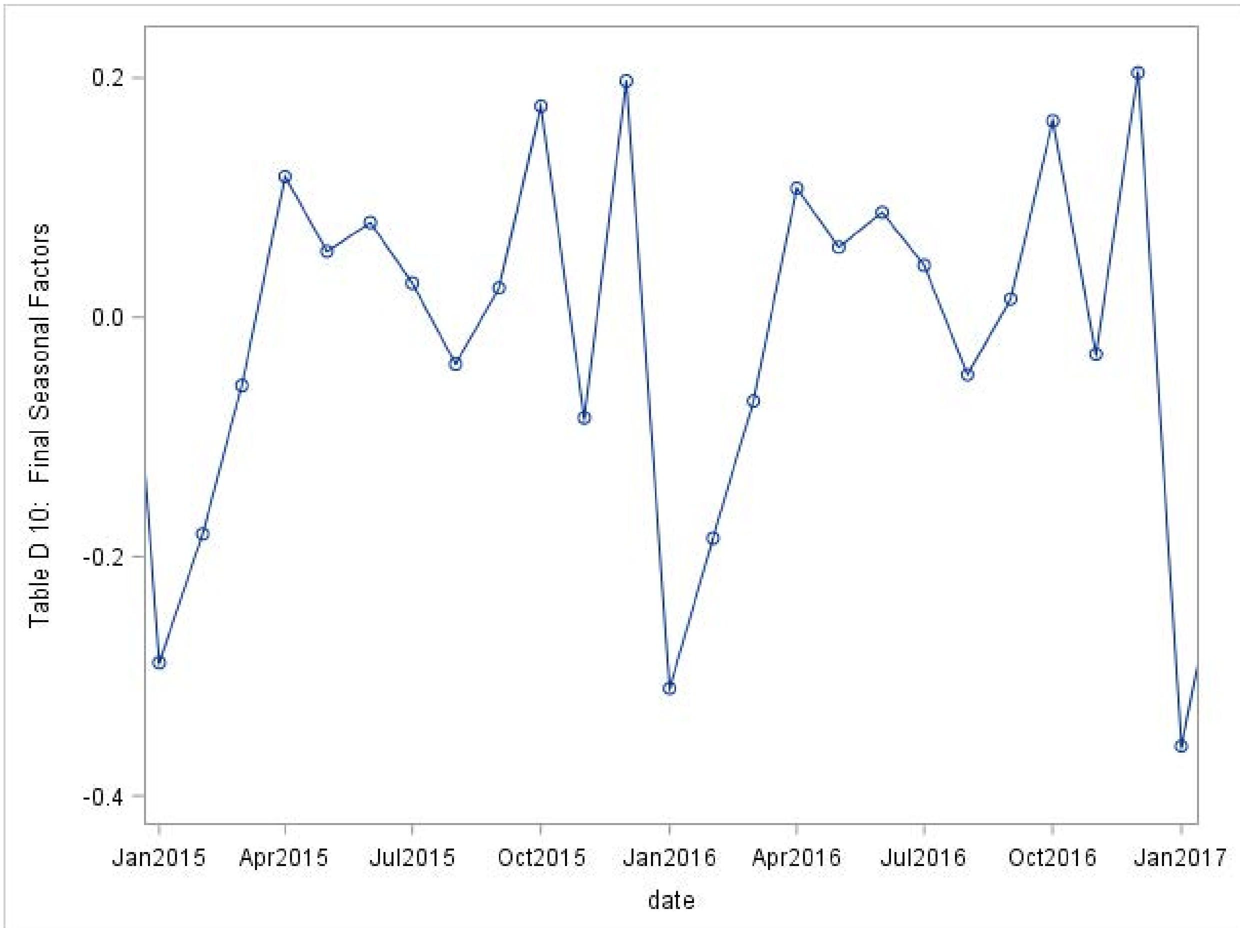

Figure 6. Seasonal factors for 2015 and 2016

The next Figure 7 gives the estimated trend cyclic component compared to the changes in taxation. Note that the trend cyclic component by the option `trendma=23` in the code is smoothed by a 23 terms Henderson average and not by the default 13 terms average. It is clear that the changes in taxation seem to have some effect on the sales series.

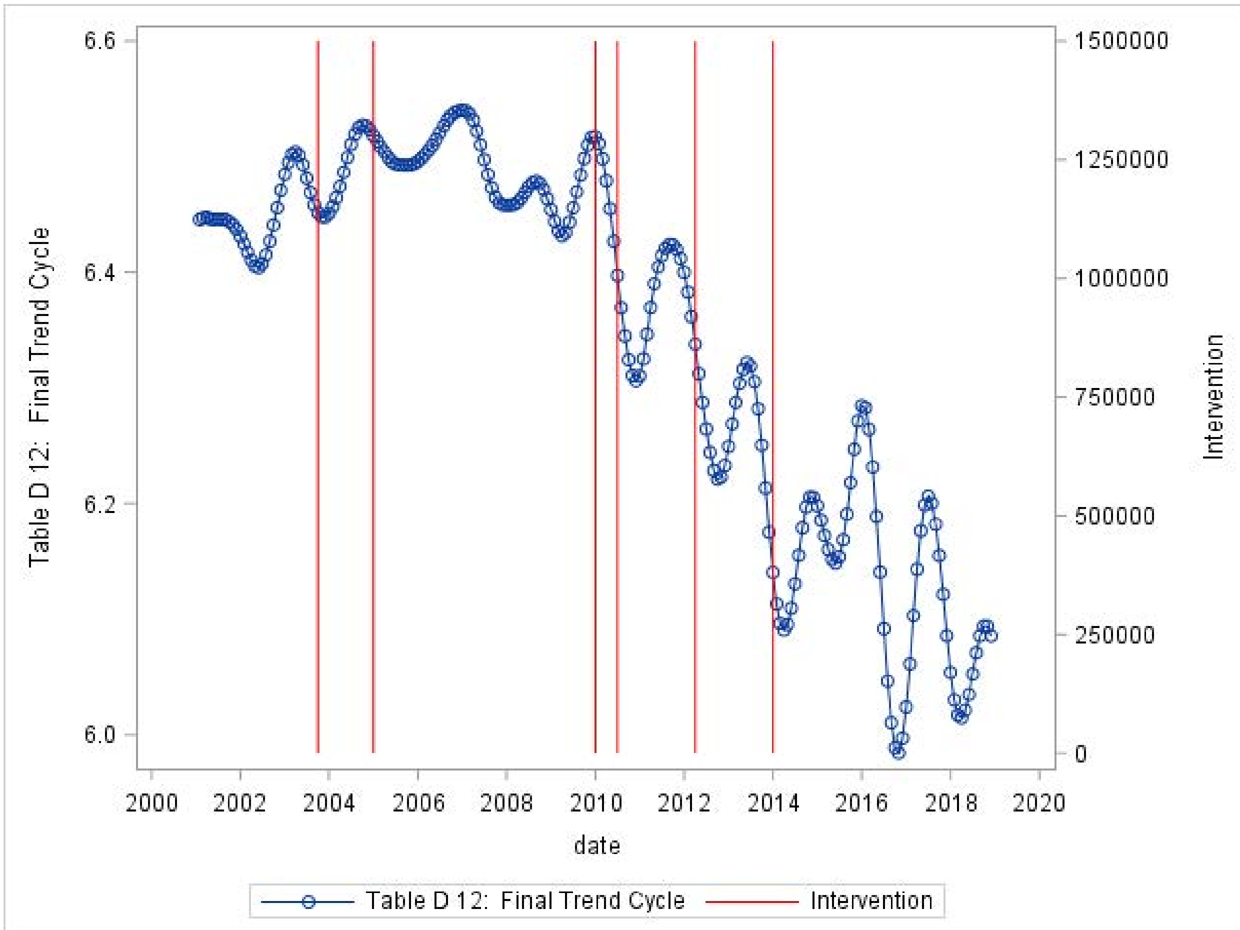

Figure 7. The trend-cyclic component with marked tax changes

The irregular component compared to the changes in taxation, Figure 8, also seems to be affected by changes in the taxation. But large values of the irregular component are also seen for other months.

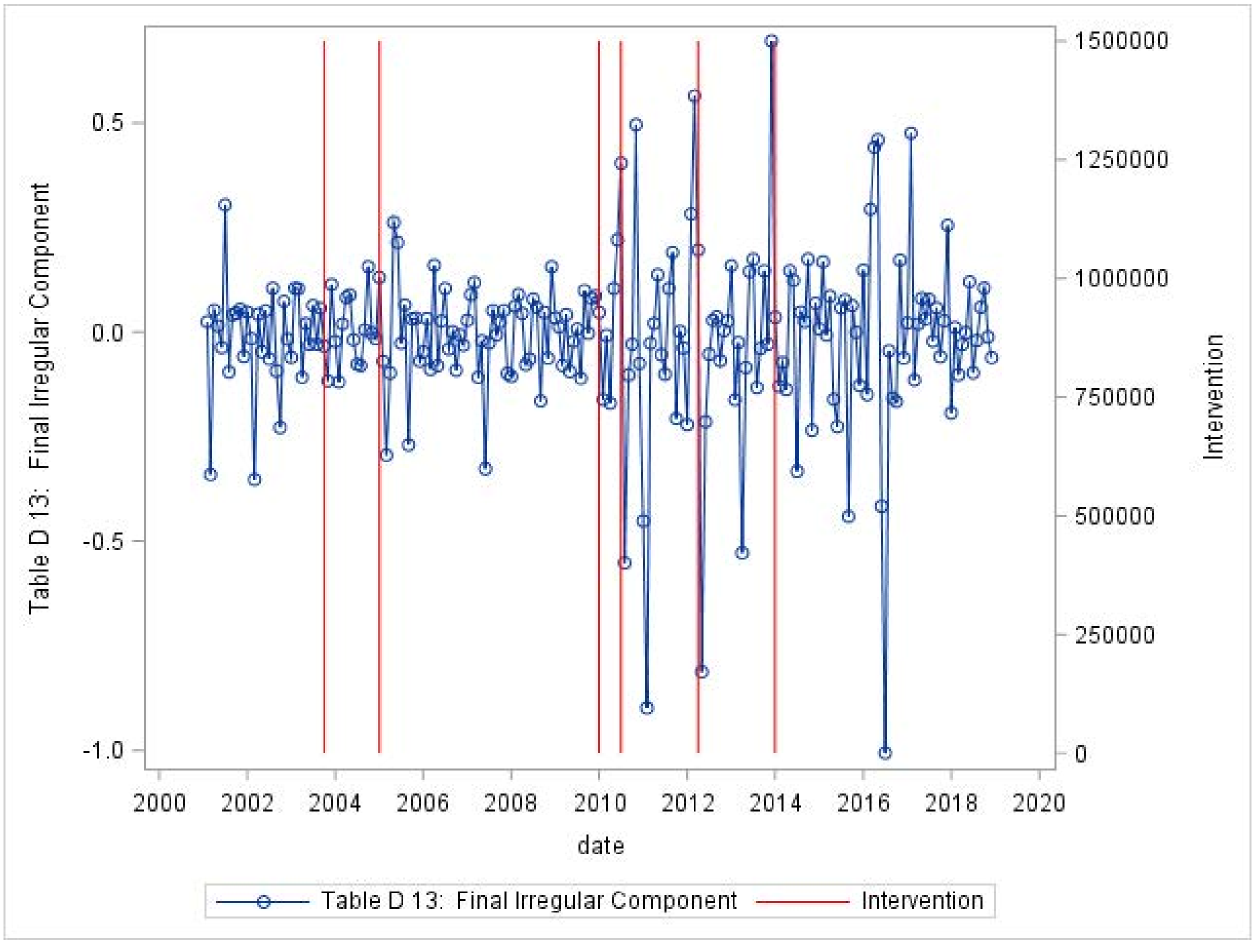

Figure 8. The irregular component with marked tax changes

PROC ARIMA APPLIED TO THE MONTHLY SERIES

By usual BOX & Jenkins(1976) model identification techniques we end up with an ARIMA(1,0,0)×ARIMA12(0,1,1) model

$$(1 - 0.25)(1 - B^{12})S_t = (1 - 0.72B^{12})\varepsilon_t \text{ , } var(\varepsilon_t) = 0.2584^2$$

This model is extended by the log-transformed relative price series as an input variable by the code

```
PROC ARIMA data=i plots=all;
  identify var=Log_sales(12);
  estimate p=1 p=(12) noint method=ml;
run;
```

The estimation results are given by Output 4.

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MA1,1	0.88115	0.06384	13.80	<.0001	12	Log_sales	0
AR1,1	0.15248	0.06710	2.27	0.0231	1	Log_sales	0
NUM1	-1.44333	0.22174	-6.51	<.0001	0	Log_relative_price	0

Output 4. Maximum Likelihood Estimation

The estimated elasticity 1.44 which is hardly significantly different from one confirms that sales of cigarettes could be controlled by the price of cigarettes.

This model could be extended by several intervention components following the principles suggested by Box & Jenkins(1976).

Every change in the taxation were announced before the change was actually put in action. This means that it was possible to buy extra cigarettes to the low price before the tax increase. This point is relevant as the stores of cigarettes in the shops are already taxed, so the sales considered in this analyses is from the wholesaler part to the retail stores. In the analyses this means that an intervention is relevant also the month before the actual new tax takes effect. This intervention is formally defined as

$B_t = 0$ for all t unless $B_t = 1$ for t = the month before the change in taxation.

Also of course the new taxation could lead to an immediate, but temporary, reaction in the sales which also could be due to delayed sales figure in case of a tax reduction or reduced sales in case of a tax increase due to increased sales prior to the actual date of the taxation change. This intervention is formally defined as

(1) $I_t = 0$ for all t unless $I_t = 1$ for t = the first month after the change in taxation.

A consumer reaction the first month after the change in taxation could also indicate an everlasting change in smoking behavior due to the change in the price of cigarettes - at least this is often an argument for the change in taxation. The question is whether this reaction is a lasting effect. This intervention is formally defined as

(2) $L_t = 0$ for all t, but $L_t = 1$ for t = all months after the change in taxation.

Often the effect of an intervention is however declining - perhaps smokers get used to the new price levels and find the money for increasing prices of cigarettes elsewhere in their personal budget. The idea is that the effect is slowly decaying by a factor δ as

(3) $X_t = \omega I_t + \omega\delta I_{t-1} + .. + \omega\delta^r I_{t-r} = \frac{\omega}{1-\delta B} I_t$

where I_t is as defined in (1). Here the effect of the intervention is given as steps in the following way

Time t_0 : The effect is ω

Time t_0+1 : The effect is ωδ

..

Time t_0+r : The effect is $\omega\delta^r$

... etc.

The value δ = 0 gives the simple intervention (1). But the situation of a positive parameter 0 < δ < 1 gives a gradual decrease towards the previous level of the series which is never reached but in practice the first steps are the largest. The parameter δ is therefore restricted to the interval] −1, 1[. For δ close to +1 takes the form of the step function (2). These kinds of models are discussed in detail by Box and Jenkins(1976).

Such models are easily fitted by PROC ARIMA. The next code gives an application of only one of the six changes in taxation in the period of analysis. In this first application only the change in taxation by January 1. 2012 is included. This intervention is modeled by an effect the month before and an exponentially declining effect afterwards. It however turns out that the effect for January 2012 was very small, while the effect form February 2012 and onwards was significant, so the shift by one month as coded by `1$` tells that the effect began February.

In this model the effect of the increasing tax in January 2012 fades out by the damping factor δ = 0.25 which in fact is insignificant meaning that the smokers reacted with some kind of chock but rapidly got used to the new price level.

```
PROC ARIMA data=i plots=all;
  identify var=log_sales(12)
    crosscor=(int_2012(12) lead_int_2012(12));
  estimate p=(1) Q=(12) method=ml noint
    input=( (1)/(1) int_2012  lead_int_2012);
run;
```

The results are given by Output 5.

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MA1,1	0.69342	0.05851	11.85	<.0001	12	Log_sales	0
AR1,1	0.22567	0.06965	3.24	0.0012	1	Log_sales	0
NUM1	0.29560	0.23127	1.28	0.2012	0	int_2012	0
NUM1,1	0.91273	0.23510	3.88	0.0001	1	int_2012	0
DEN1,1	0.25523	0.25766	0.99	0.3219	1	int_2012	0
NUM2	0.53688	0.22736	2.36	0.0182	0	lead_int_2012	0

Output 5. Maximum Likelihood Estimation

When this model is extended by the log-transformed relative price series the code becomes

```
PROC ARIMA data=i plots=all;
 identify var=Log_sales(12)
   crosscor=(Log_relative_price(12) int_2012(12) Lead_int_2012(12));
 estimate p=(1) q=(12) method=ml noint
    input=(Log_relative_price 1 $/(1) int_2012  Lead_int_2012);
run;
```

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MA1,1	0.84909	0.05866	14.47	<.0001	12	Log_sales	0
AR1,1	0.14641	0.06907	2.12	0.0340	1	Log_sales	0
NUM1	-1.34565	0.22566	-5.96	<.0001	0	Log_relative_price	0
NUM2	-0.77058	0.22030	-3.50	0.0005	0	int_2012	1
DEN1,1	0.12638	0.28309	0.45	0.6553	1	int_2012	1
NUM3	0.53645	0.21805	2.46	0.0139	0	lead_int_2012	0

Output 6. Maximum Likelihood Estimation

CONCLUSION

In this paper the relation between the price of cigarettes and the sales of cigarettes in Denmark corresponds to a price elasticity which is significantly positive. The elasticity is around 0.6. According this result the answer to the question as stated in the title is clear:

YES: It is possible to reduce smoking by increasing taxes.

But of course there are many reservations to this conclusion, like the amount of border trade (cigarettes are cheap in Poland), that have to be kept in mind!

REFERENCES

Box, G.E.P. and Jenkins, G.M. (1976), *Time Series Analysis: Forecasting and Control*, Revised Edition, San Francisco, CA: Holden−Day.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Anders Milhøj
Department of Economics
University of Copenhagen
anders.milhoj@econ.ku.dk