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

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


Price index
250
200
150
100
8000
7000
6000
5000
Sales in millions
1975
1980
1985
1990
1995
2000
2005
2010
2015
2020
date
Price index
Sales in millions

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

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


9.0
8.9
8.8
8.7
8.6
Log-transformed sales
-3.2
-3.3
-3.4
Log-transformed relative price
1975
1980
1985
1990
1995
2000
2005
2010
2015
2020
date
Log-transformed relative price
Log-transformed sales

```
proc varmax data=tt.cigarettes_year_2020 print=all plots=all;
model l_sales=log_rel_pris/
dif=(l_sales(1) log_rel_pris(1)) method=ml q=1 noint;
where year(date)>=1977 and year(date)<2020;
run;
```

Model Parameter Estimates

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

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 one.

UCM for yearly data

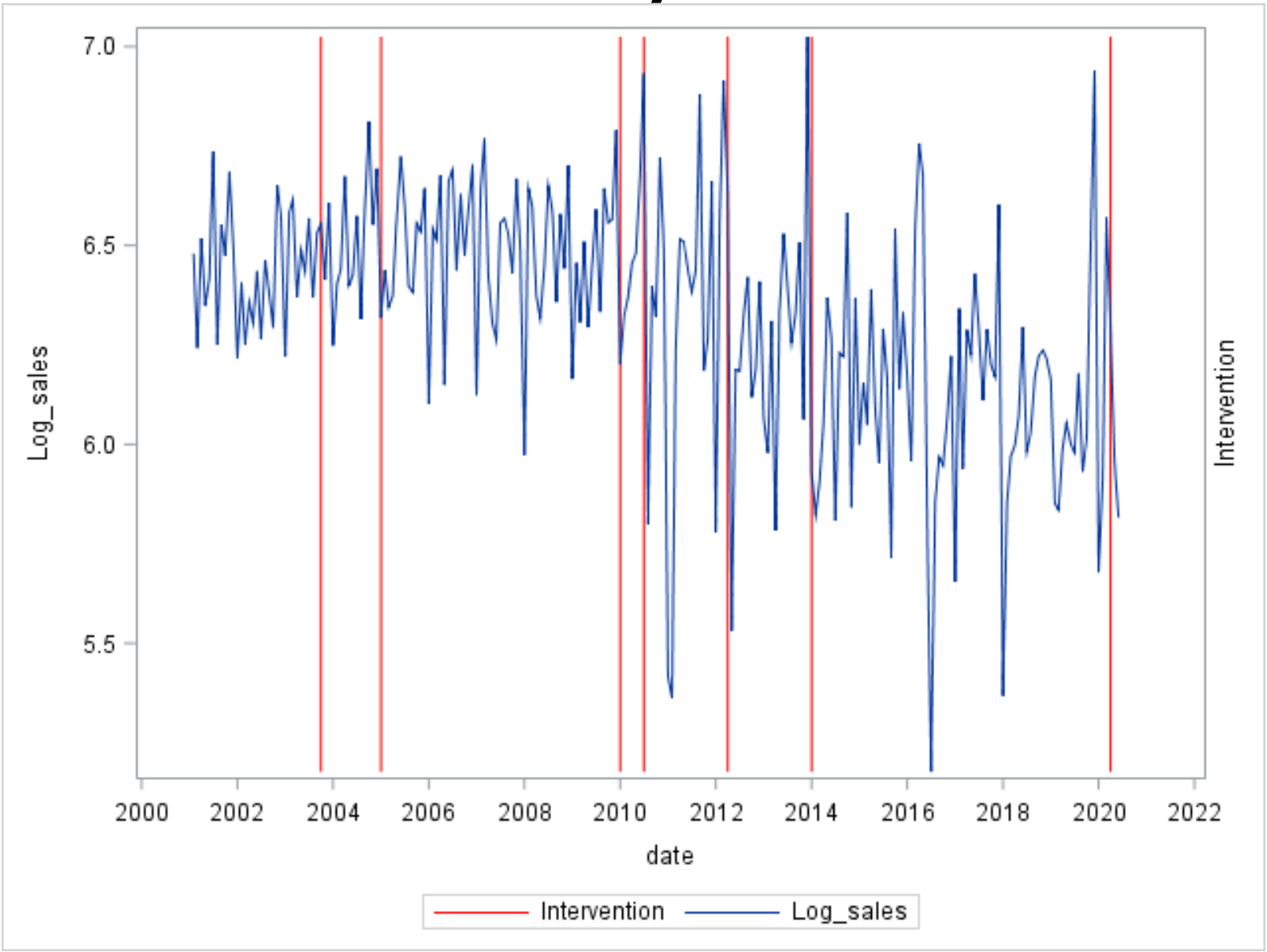
```
PROC UCM data=tt.cigarettes_year_2020 plots=all;
id date interval=year;
model l_sales;
level ;
randomreg log_rel_pris/ print=smooth;
irregular p=1;
estimate extradiffuse=2;
where year(date)>1977 and year(date)<2020;
run;
```

Final Estimates of the Free Parameters

Component	Parameter	Estimate	Approx Std Error	t Value	Approx Pr > |t|
Irregular	**Error Variance**	0.00019758	0.0004217	0.47	0.6394
Irregular	**AR_1**	-0.72733	0.52160	-1.39	0.1632
Level	**Error Variance**	0.00190	0.0008246	2.31	0.0209
log_rel_pris	**Error Variance**	2.62951E-10	3.3793E-7	0.00	0.9994

The results of this table 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 β = -0.6268.

Monthly data




0.2
0.1
0.0
-0.1
-0.2
-0.3
Log-transformed relative price
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
date
Intervention
Intervention
Log-transformed relative price

```
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;
```

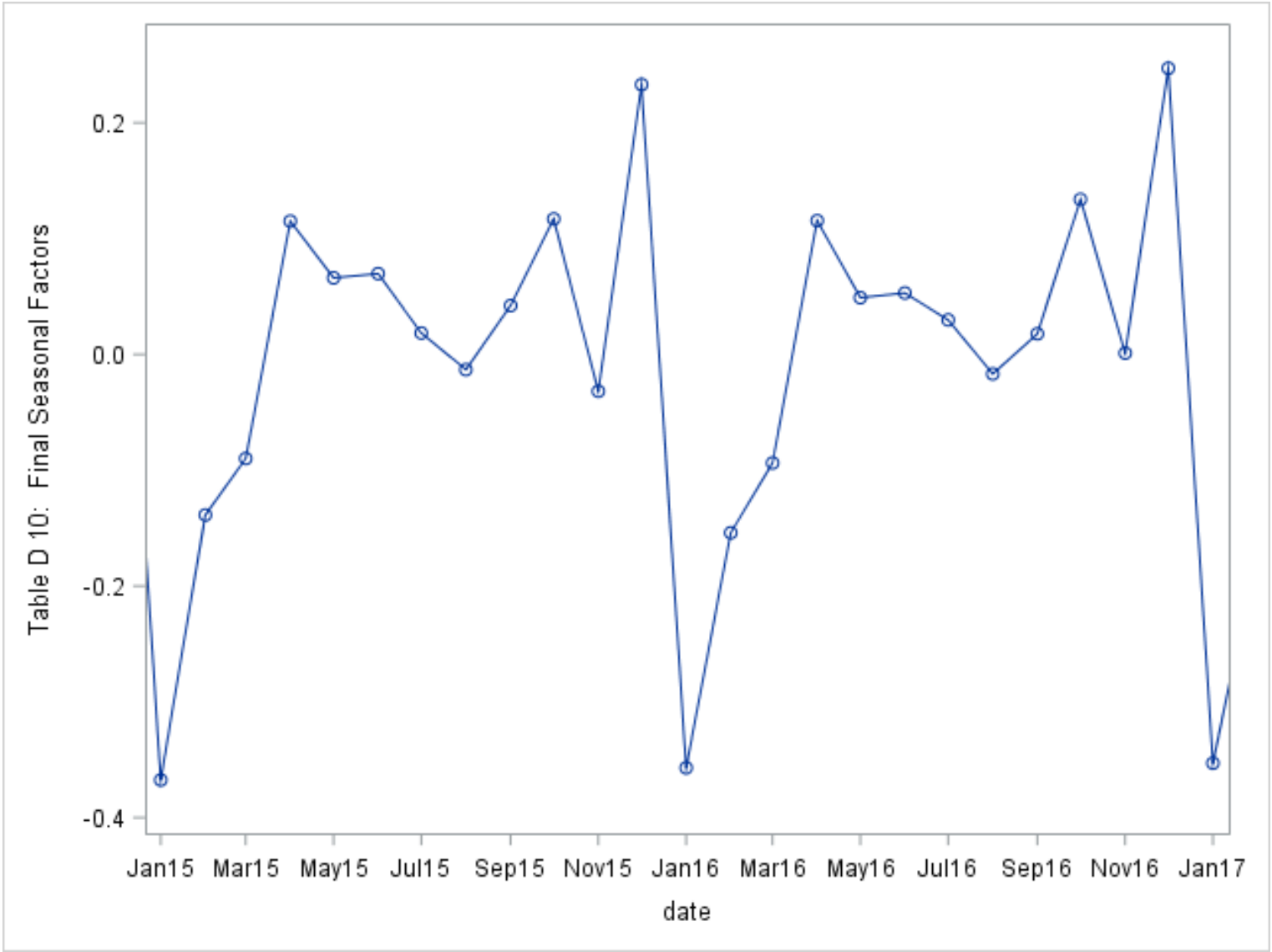


Table D 10: Final Seasonal Factors
0.2
0.0
-0.2
-0.4
Jan15
Mar15
May15
Jul15
Sep15
Nov15
Jan16
Mar16
May16
Jul16
Sep16
Nov16
Jan17
date

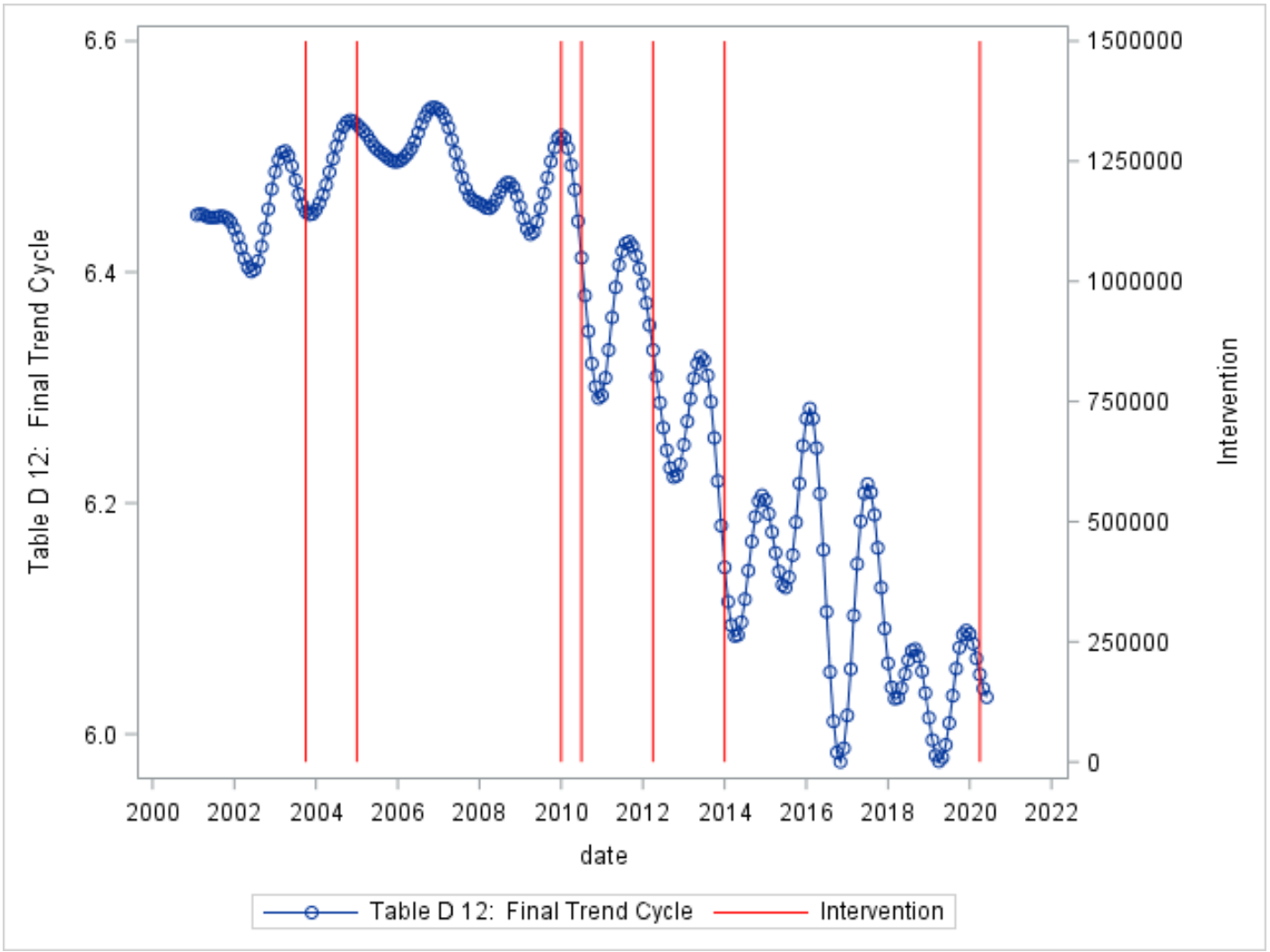


Table D 12: Final Trend Cycle
Intervention
6.6
6.4
6.2
6.0
1500000
1250000
1000000
750000
500000
250000
0
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
date
Table D 12: Final Trend Cycle
Intervention

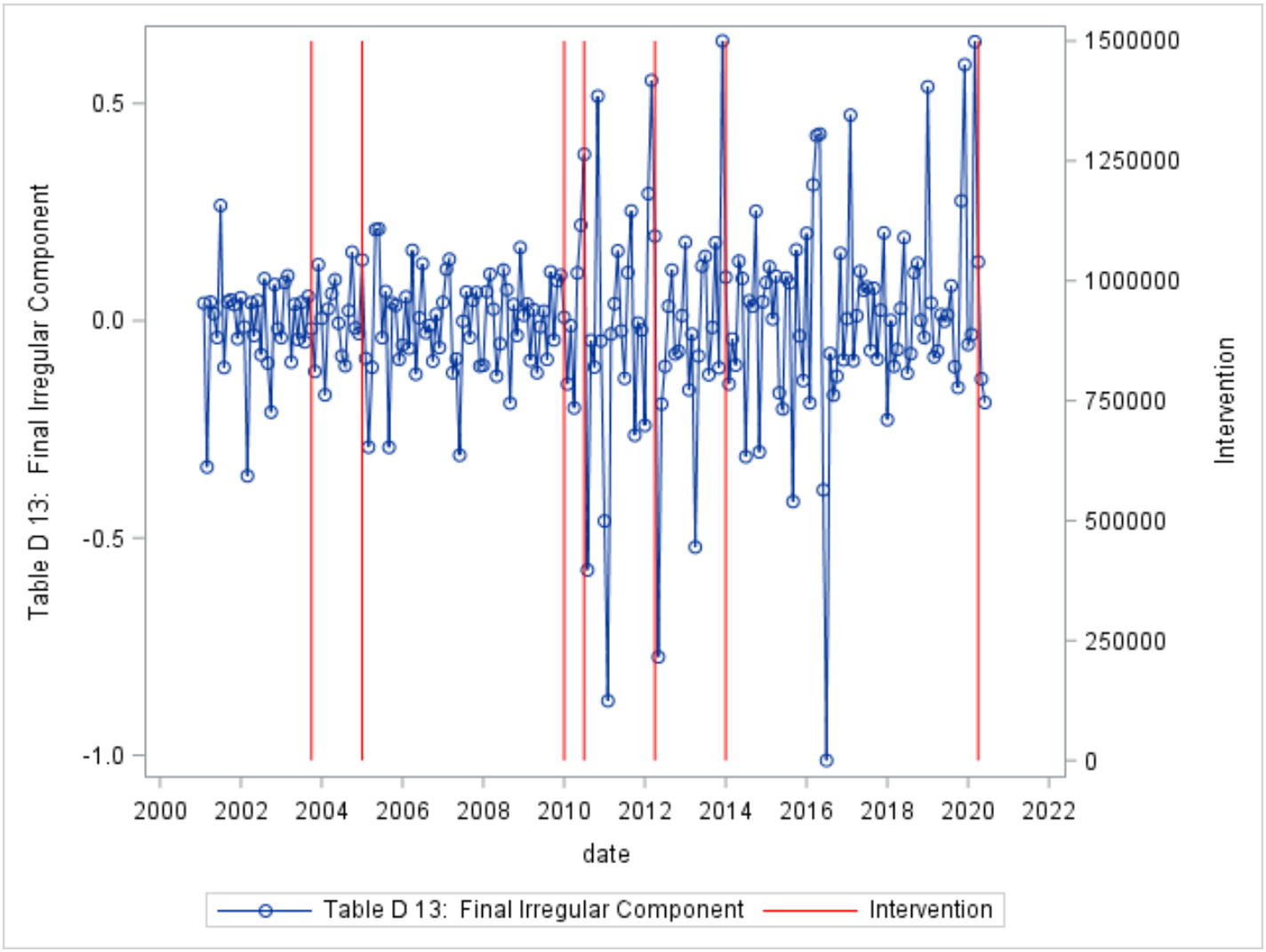


Table D 13: Final Irregular Component
0.5
0.0
-0.5
-1.0
1500000
1250000
1000000
750000
500000
250000
0
Intervention
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
date
Table D 13: Final Irregular Component
Intervention

Estimated seasonal ARIMA as identified from Proc X12

Exact ARMA Maximum Likelihood Estimation

For Variable Log_sales

Parameter	Lag	Estimate	Standard Error	t Value	Pr > |t|
Nonseasonal AR	**1**	0.07077	0.06630	1.07	0.2870
	2	-0.13698	0.06561	-2.09	0.0380
Nonseasonal MA	**1**	0.93565	0.02666	35.10	<.0001
Seasonal MA	**12**	0.96299	0.03654	26.36	<.0001

```
regression  uservar=log_relative_price;
```

Regression Model Parameter Estimates

For Variable Log_sales

Type	Parameter	NoEst	Estimate	Standard Error	t Value	Pr > |t|
Constant	**Constant**	Est	-0.01379	0.00243	-5.68	<.0001
User Defined	**Log_relative_price**	Est	-0.96206	0.14144	-6.80	<.0001

Exact ARMA Maximum Likelihood Estimation

For Variable Log_sales

Parameter	Lag	Estimate	Standard Error	t Value	Pr > |t|
Nonseasonal AR	**1**	0.92191	0.11847	7.78	<.0001
	2	-0.13201	0.08668	-1.52	0.1293
	3	-0.00126	0.07041	-0.02	0.9857
Nonseasonal MA	**1**	0.88067	0.09995	8.81	<.0001
Seasonal MA	**12**	0.99910	0.03945	25.33	<.0001

```
PROC ARIMA data=i plots=all;
identify var=Log_sales(12)
  crosscor=(Log_relative_price(12)) ;
estimate q=(12) method=ml
  input= (Log_relative_price) ;
forecast id=date interval=month back=5;
RUN;
quit;
```

Maximum Likelihood Estimation

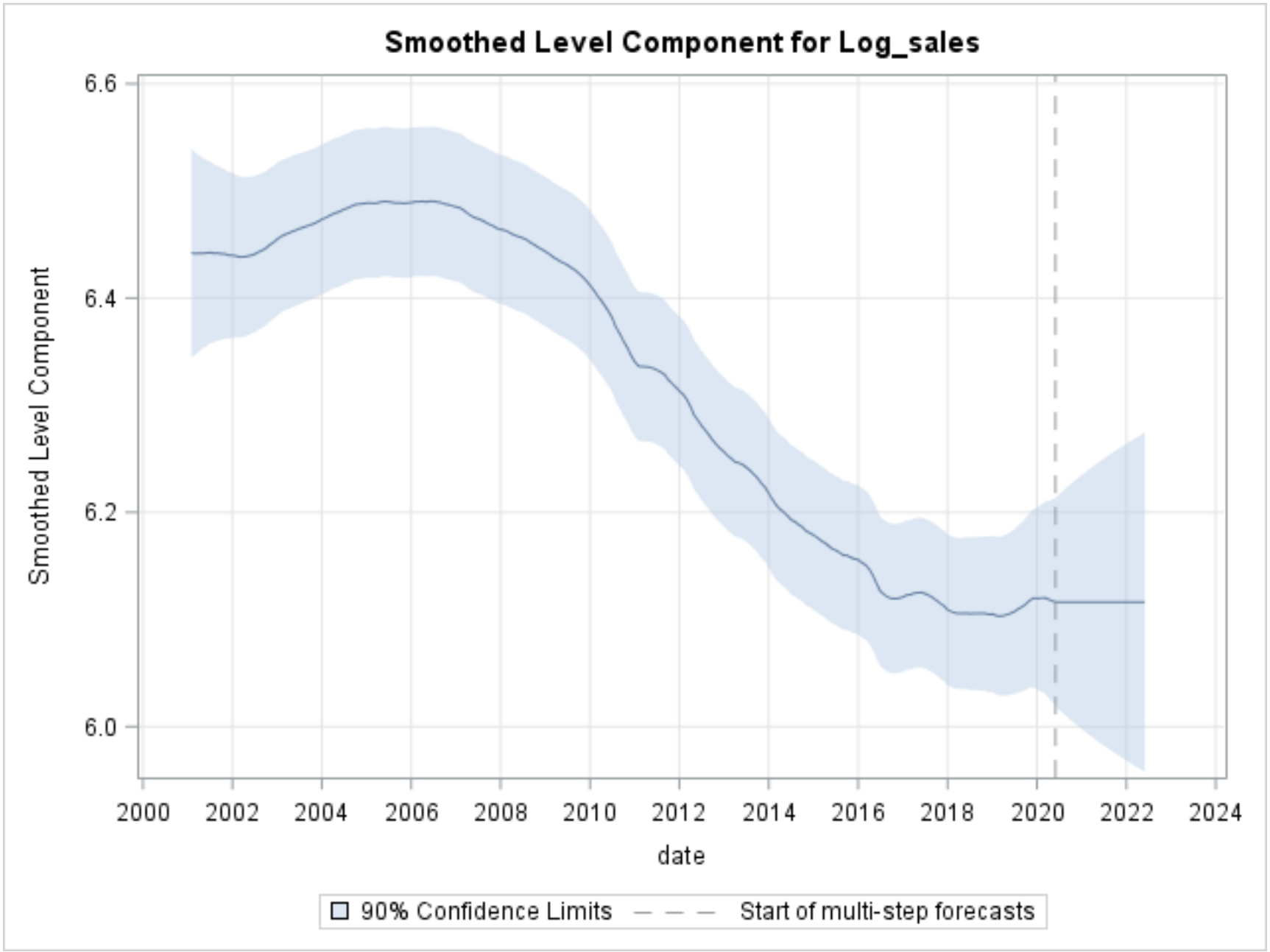
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	-0.01341	0.0038562	-3.48	0.0005	0	Log_sales	0
MA1,1	0.94484	0.09675	9.77	<.0001	12	Log_sales	0
NUM1	-1.00320	0.23298	-4.31	<.0001	0	Log_relative_price	0

UCM for monthly data

```
PROC UCM data=i plots=all;
id date interval=month;
model log_sales;
irregular p=2;
level ;
season length=12 print=smooth var=0 noest;
estimate extradiffuse=3;
forecast lead=24 plot=all alpha=0.1;
run;
```

Final Estimates of the Free Parameters

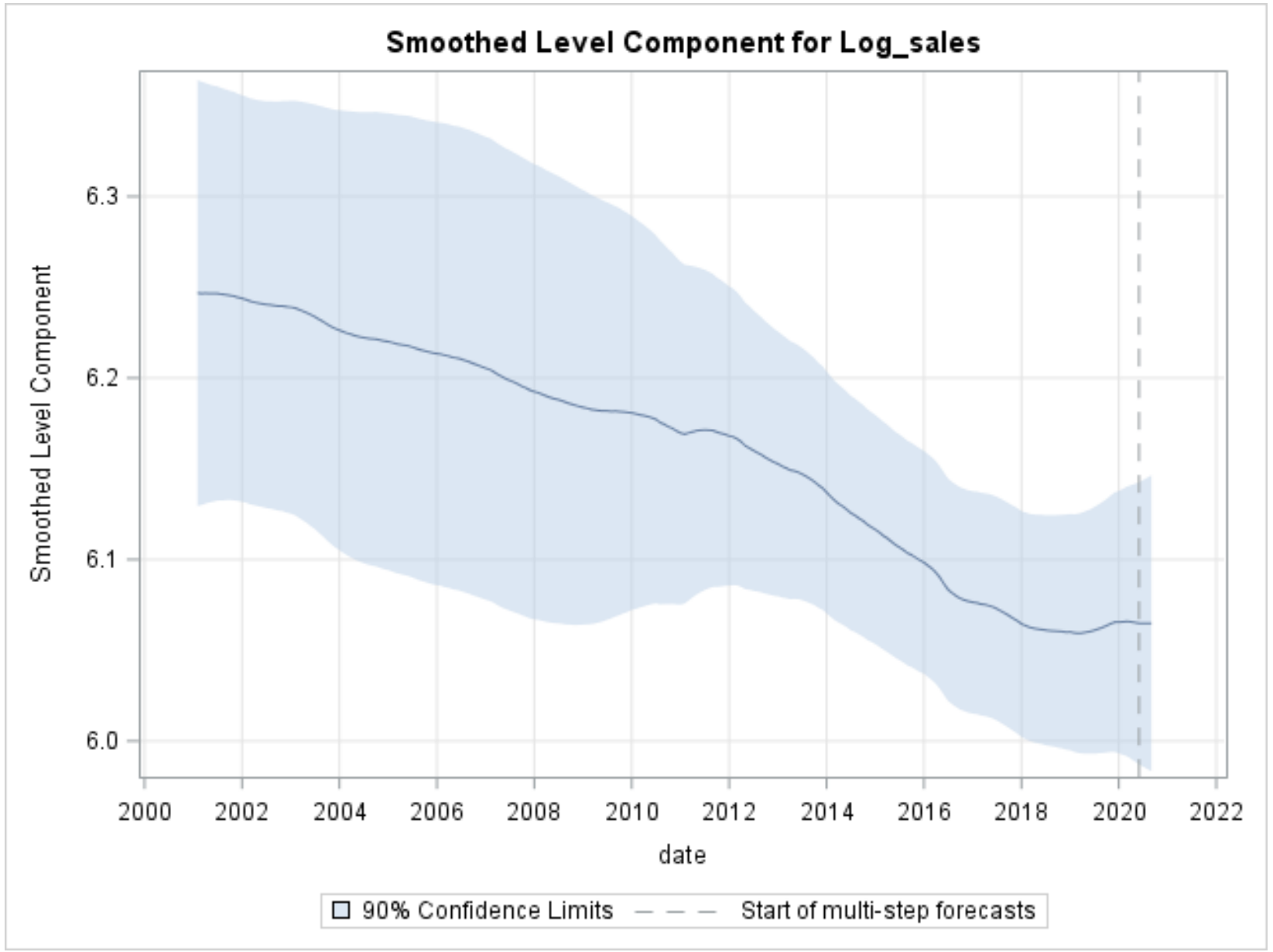
Component	Parameter	Estimate	Approx Std Error	t Value	Approx Pr > |t|
Irregular	**Error Variance**	0.05186	0.0050388	10.29	<.0001
Irregular	**AR_1**	0.14124	0.06902	2.05	0.0407
Irregular	**AR_2**	-0.12847	0.06875	-1.87	0.0617
Level	**Error Variance**	0.00024093	0.0001362	1.77	0.0769


Smoothed Level Component for Log_sales
Smoothed Level Component
6.6
6.4
6.2
6.0
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
2024
date
90% Confidence Limits
Start of multi-step forecasts

```
PROC UCM data=i plots=all;
id date interval=month;
model log_sales=Log_relative_price;
irregular p=2;
level ;
season length=12 print=smooth var=0 noest;
estimate extradiffuse=3;
forecast lead=24 plot=all alpha=0.1;
run;
```

Final Estimates of the Free Parameters

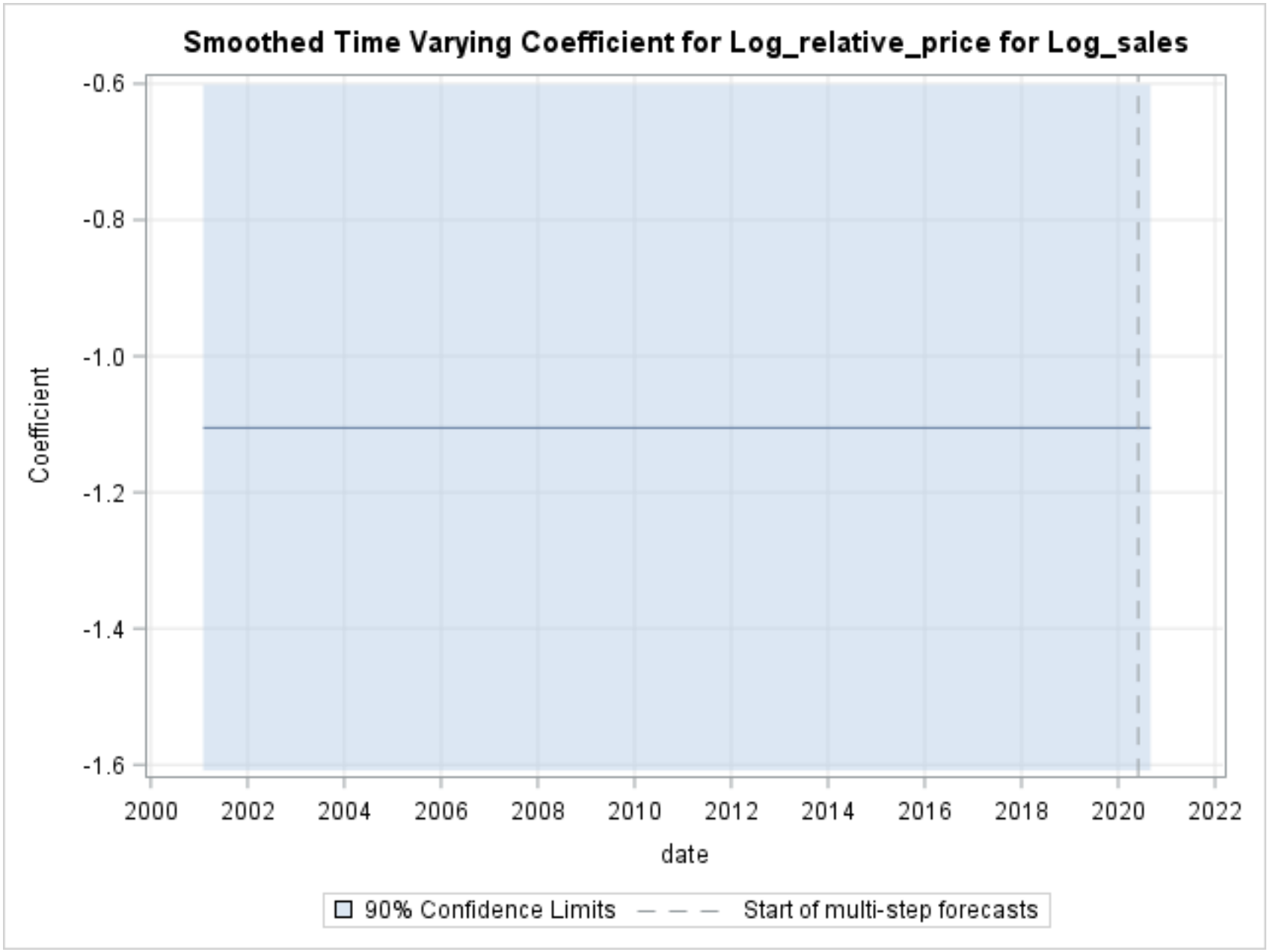
Component	Parameter	Estimate	Approx Std Error	t Value	Approx Pr > |t|
Irregular	**Error Variance**	0.05167	0.0049826	10.37	<.0001
Irregular	**AR_1**	0.13319	0.06807	1.96	0.0504
Irregular	**AR_2**	-0.12977	0.06808	-1.91	0.0566
Level	**Error Variance**	0.00007474	0.00006668	1.12	0.2623
Log_relative_price	**Coefficient**	-1.10507	0.30611	-3.61	0.0003


Smoothed Level Component for Log_sales
Smoothed Level Component
6.3
6.2
6.1
6.0
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
date
90% Confidence Limits
Start of multi-step forecasts

```
PROC UCM data=i plots=all;
id date interval=month;
model log_sales;
irregular p=2;
level ;
season length=12 print=smooth var=0 noest;
randomreg Log_relative_price;
estimate extradiffuse=3;
forecast lead=24 plot=all alpha=0.1;
run;
```

Final Estimates of the Free Parameters

Component	Parameter	Estimate	Approx Std Error	t Value	Approx Pr > |t|
Irregular	**Error Variance**	0.05167	0.0049826	10.37	<.0001
Irregular	**AR_1**	0.13319	0.06807	1.96	0.0504
Irregular	**AR_2**	-0.12977	0.06808	-1.91	0.0566
Level	**Error Variance**	0.00007474	0.00006668	1.12	0.2623
Log_relative_price	**Error Variance**	5.968354E-9	3.60502E-6	0.00	0.9987


Smoothed Time Varying Coefficient for Log_relative_price for Log_sales
-0.6
-0.8
-1.0
-1.2
-1.4
-1.6
Coefficient
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
date
90% Confidence Limits
Start of multi-step forecasts