Chapter 5

Practical Examples with a Single Observable

5.1 Basic Structural Models

Most models with a single observable use some combination of the local level or trend, possibly with one of the seasonal models. Harvey (1989) calls these *Basic Structural Models* or BSM. We've already seen quite a few examples of the local level and local trend model. We'll now combine with the seasonal model.

The data set is the famous "airline" data used by Box and Jenkins. This is the basis for the "Airline" model, which is a multiplicative seasonal ARIMA $(0,1,1)\times(0,1,1)$ (one regular difference, one regular MA, one seasonal difference, one seasonal MA). That's the best fitting small model for this data set, and is often used as a first guess for a fit for highly seasonal series more generally.

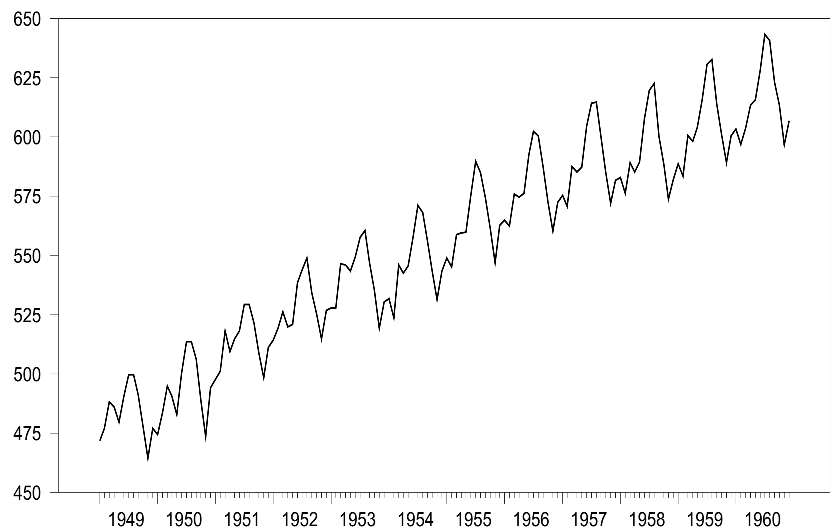

Figure 5.1: (Log) Airline Passenger Data

This has both a decided trend and a seasonal so the two obvious components are the local trend and the additive seasonal. The additive seasonal seems most appropriate since it's the secular calendar rather than weather that governs most of the seasonality.

In the local trend model, we'll allow for both shocks, although it's likely that at least one will have to go. With both shocks included, we'll have one observable being called upon to determine four shocks.

Set up the two component models and combine them to make the full model:

```
@localdlm(type=trend,shocks=both,a=at,c=ct,f=ft)
@seasonaldlm(type=additive,a=as,c=cs,f=fs)
compute a=at~\as,c=ct~~cs,f=ft~\fs
```

You can use **@LocalDLMInit** with the DESEASONALIZE option to get the guess values. This removes a seasonal by regression, allowing guess values to be computed for the other parts. That corresponds to a zero variance for the seasonal.

```
nonlin sigsqeps sigsqxi sigsqzeta sigsqomega
@localdlminit(irreg=sigsqeps,trend=sigsqzeta,deseasonalize) lpass
compute sigsqxi=.01*sigsqeps,sigsqomega=0.0
```

The SW matrix now has to cover the three state shocks. Since those are assumed to be independent, it will be done with %DIAG.

```
dlm(y=lpass,a=a,c=c,f=f,exact,$
  sv=sigsqeps,sw=%diag(||sigsqxi,sigsqzeta,sigsqomega||),$
  method=bfgs) / xstates
```

When estimated unconstrained, the variance on the trend rate (SIGSQZETA) comes in negative. This is re-estimated first with that pegged to zero, then with SIGSQXI at zero. The one with the fixed trend rate fits much better, so that's what we choose. We Kalman smooth with that model to get the components. The local trend is captured by the first state, the seasonal by the third. To get the "deseasonalized" data, we subtract the seasonal from the actual data:

```
set trend    = xstates(t)(1)
set seasonal = xstates(t)(3)
set deseason = lpass-seasonal
```

The graphs of these show that the seasonal seems to change over the years, with a "mini-peak" in March eventually fading away. The one potential problem with these from a practical standpoint is that the "trend" seems to be a bit more uneven than we would like. It's possible that the more flexible Fourier seasonal model would help shift some of that into the seasonal.

5.2 Trend plus Stationary Cycle

In the local trend model, the "cycle" (the gap between the series and the estimated trend) isn't modeled explicitly. Instead, it's just treated as the "uncorrelated" measurement error. In practice, of course, it isn't actually uncorrelated, and we really only get reasonable results out of the model when we do something to constrain the ability of the trend to pick up short-term movements.

An alternative approach is to allow the cycle to be modeled more explicitly. The typical choice for this is a stationary AR(2) process, since it has enough

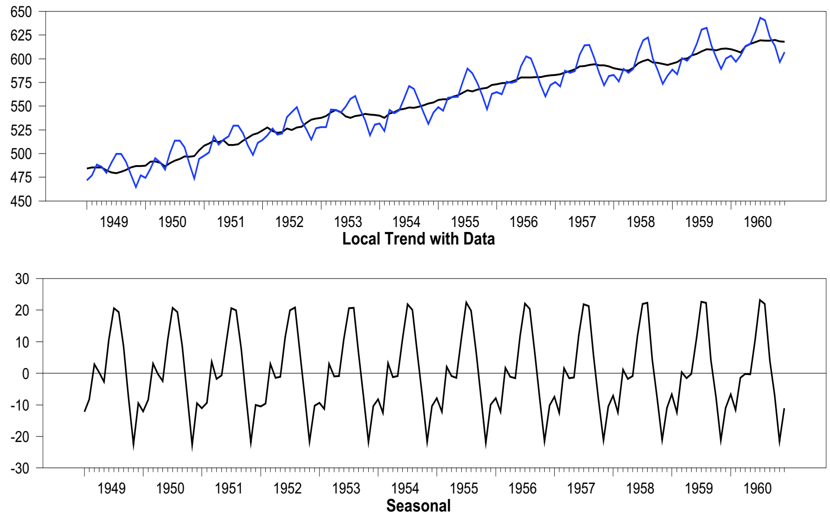

Figure 5.2: Airline Model Components

flexibility to model actual (damped) sine wave cycles. The observable series is then the sum of a (non-stationary) trend model and the (stationary) cycle.

An example of this is in Perron and Wada (2009). They fit several models to US GDP. Their basic model is:

$$\begin{aligned} y_t &= \tau_t + c_t \\ \tau_t &= \tau_{t-1} + \mu + \eta_t \\ c_t &= \varphi_1 c_{t-1} + \varphi_2 c_{t-2} + \varepsilon_t \end{aligned}$$

The trend is a local trend with fixed rate. This translates into the following state-space representation:

$$\begin{bmatrix} \tau_t \\ c_t \\ c_{t-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \varphi_1 & \varphi_2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \tau_{t-1} \\ c_{t-1} \\ c_{t-2} \end{bmatrix} + \begin{bmatrix} \mu \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix}$$

$$y_t = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \mathbf{X}_t$$

One possible setup for this is:

```
dec frml[rect] af
dec frml[vect] zf
dec frml[symm] swf
*
nonlin mu sn ph1 ph2 se
*
frml af = ||1.0,0.0,0.0|$
            0.0,ph1,ph2|$
            0.0,1.0,0.0||
frml zf = ||mu,0.0,0.0||
frml swf = %diag(||sn^2,se^2||)
*
* Fixed components
*
compute [vect] c=||1.0,1.0,0.0||
compute [rect] f=%identity(2)~~%zeros(1,2)
```

In Example **5.2**, the following is used for getting guess values for the various parameters. This does an initial decomposition of the data into trend and cycle using an HP filter, estimating an AR(2) on the crude cycle and a linear trend on the trend part to get guess values.

```
filter(type=hp) lgdp / gdp_hp
set gap_hp = lgdp - gdp_hp
linreg gap_hp
# gap_hp{1 2}
compute ph1=%beta(1),ph2=%beta(2),se=sqrt(%seesq)
set trend = t
linreg gdp_hp
# constant trend
compute mu=%beta(2)
compute sn=sqrt(.1*%seesq)
```

The model has one unit root and two stationary roots. The authors handled this by:

1. computing the ergodic solution for the final two states (the cycle)
2. adding a large (1,1) element to allow for the unit root in state 1 and
3. dropping the first observation from the calculation of the likelihood.

This is more sophisticated than the typical handling of such models. More common is to just use a big diagonal matrix for all states, and condition on additional data points. However, **DLM** with PRESAMPLE=ERGODIC takes care of all of that, and also avoids the problem of picking an appropriate "large" value.

```
dlm(presample=ergodic,a=af,z=zf,sw=swf,c=c,f=f,y=lgdp,$
   method=bfgs,type=filter)
```

The output from this includes the following:

```
Usable Observations    206
Rank of Observables    205
Log Likelihood            -286.60535
```

`Usable Observations` shows the number of actual values for y. `Rank of Observables` shows the number of observations actually used for computing the likelihood.

The principal models of interest in the paper have a break in the trend rate at 1973:1. That requires only the minor change of

```
set d1973 = t>1973:1
frml zf = ||mu+d*d1973,0.0,0.0||
```

and the addition of the trend shift coefficient `D` to the parameter set. They also, however, allow for a correlation between the two shocks, which we have not yet done. Their parameterization for this is to estimate the two standard errors and the correlation coefficient, thus:

```
frml swf = ||sn^2|rhone*abs(sn)*abs(se),se^2||
```

In most cases, we wouldn't recommend that way of handling this—there's nothing preventing this from estimating $|\rho| > 1$, which, in fact, happens in this case. It also starts to get rather ugly if you try to extend it beyond two shocks. To get a general positive semi-definite matrix, it's better to parameterize it as a packed lower triangular matrix. (`SYMMETRIC` doesn't work because it doesn't force the matrix to be p.s.d.). The following estimates the more general model; the `%LTOUTERXX` function takes the outer product ($\mathbf{LL}'$) of a packed lower triangular matrix.

```
dec packed swlower(2,2)
frml swf = %ltouterxx(swlower)
nonlin mu d ph1 ph2 swlower
dlm(presample=ergodic,a=af,z=zf,sw=swf,c=c,f=f,y=lgdp,method=bfgs,$
  type=filter) / statesur
```

If the two shocks collapse to a single one, you'll see something like the following:

```
5.  SWLOWER(1,1)        0.104212082  0.223075475       0.46716  0.64038495
6.  SWLOWER(2,1)        0.842690299  0.224098032       3.76036  0.00016967
7.  SWLOWER(2,2)       -0.000002419  1.245301185 -1.94245e-006  0.99999845
```

with the (2,2) element coming in effectively zero.

Example 5.1 Airline Data

This fits several models to the Box-Jenkins airline data.

```
cal(m) 1949
open data airline.prn
data(format=prn,org=columns,compact=sum) 1949:1 1960:12 pass
set lpass = 100.0*log(pass)
*
graph(footer="Box-Jenkins airline passenger data")
# lpass
*
* (a) Deterministic trend and seasonal
*
seasonal seasons
set trend = t
linreg lpass
# constant trend seasons{0 to -10}
set olsresids = %resids
*
* (b) "Airline" ARIMA model
*
boxjenk(diffs=1,sdiffs=1,sma=1,ma=1,maxl) lpass
set arimaresids = %resids
*
* Components for BSM
*
@localdlm(type=trend,shocks=both,a=at,c=ct,f=ft)
@seasonaldlm(type=additive,a=as,c=cs,f=fs)
compute a=at~\as,c=ct~~cs,f=ft~\fs
*
* (c) BSM unconstrained.
*
nonlin sigsqeps sigsqxi sigsqzeta sigsqomega
@localdlminit(irreg=sigsqeps,trend=sigsqzeta,deseasonalize) lpass
compute sigsqxi=.01*sigsqeps,sigsqomega=0.0
*
dlm(y=lpass,a=a,c=c,f=f,exact,$
  sv=sigsqeps,sw=%diag(||sigsqxi,sigsqzeta,sigsqomega||),$
  method=bfgs) / xstates
*
* (d) BSM constrained to fixed trend rate
*
nonlin sigsqeps sigsqxi sigsqzeta=0.00 sigsqomega
dlm(y=lpass,a=a,c=c,f=f,exact,$
  sv=sigsqeps,sw=%diag(||sigsqxi,sigsqzeta,sigsqomega||),$
  method=bfgs,vhat=vhat,svhat=svhat) / xstates
set resids = %scalar(vhat)/sqrt(%scalar(svhat))
@STAMPDiags resids
@CUSUMTests resids
*
* (e) BSM constrained to no level shock
*
```

```
nonlin sigsqeps sigsqxi=0.00 sigsqzeta sigsqomega
dlm(y=lpass,a=a,c=c,f=f,exact,$
  sv=sigsqeps,sw=%diag(||sigsqxi,sigsqzeta,sigsqomega||),$
  method=bfgs,vhat=vhat,svhat=svhat) / xstates
set resids = %scalar(vhat)/sqrt(%scalar(svhat))
@STAMPDiags resids
@CUSUMTests resids
*
* The fixed trend rate model seems clearly superior. We redo this with
* smoothing.
*
nonlin sigsqeps sigsqxi sigsqzeta=0.00 sigsqomega
dlm(y=lpass,a=a,c=c,f=f,exact,$
  sv=sigsqeps,sw=%diag(||sigsqxi,sigsqzeta,sigsqomega||),$
  method=bfgs,type=smooth) / xstates vstates
*
* Extract the trend and the seasonal
*
set trend    = xstates(t)(1)
set seasonal = xstates(t)(3)
set deseason = lpass-seasonal
*
graph(footer="Local Trend with Data") 2
# trend
# lpass
graph(footer="Deseasonalized Data") 2
# deseason
# lpass
graph(footer="Seasonal")
# seasonal
*
set bsmresids = %scalar(vhat)
*
* Compare the (unscaled) residuals from the three models. These should
* be comparable. However, the ARIMA model is defined only from 1950:2 on
* (because of lags), and the residuals for the state space model are not
* reliable until the same point (while resolving the unit roots) so we
* only graph those.
*
graph(footer="Comparison of Residuals",key=below,$
  klabels=||"State Space","ARIMA","Regression"||) 3
# bsmresids   1950:2 *
# arimaresids 1950:2 *
# olsresids   1950:2 *
```

Example 5.2 Trend plus Stationary Cycle Model

This fits several unobserved components models with trend plus stationary cycle from Perron and Wada (2009).

```
open data lgdp.txt
calendar(q) 1947
data(format=free,org=columns) 1947:01 1998:02 lgdp
*
* Everything is run on 100*log data
*
set lgdp  = 100.0*lgdp
set d1973 = t>1973:1
*
@NBERCycles(down=recession)
*
* UC-0 decomposition with AR(2) cycle, no break in trend rate.
* Uncorrelated component variances.
*
* Components which depend upon free parameters
*
dec frml[rect] af
dec frml[vect] zf
dec frml[symm] swf
*
nonlin mu sn ph1 ph2 se
*
frml af = ||1.0,0.0,0.0|$
            0.0,ph1,ph2|$
            0.0,1.0,0.0||
frml zf = ||mu,0.0,0.0||
frml swf = %diag(||sn^2,se^2||)
*
* Fixed components
*
compute [vect] c=||1.0,1.0,0.0||
compute [rect] f=%identity(2)~~%zeros(1,2)
*
* Get initial guess values. This is probably overkill. It extracts an HP
* trend, and runs an AR(2) on the "gap" to get an estimate of the two
* cycle coefficients and the standard error. It then estimates the trend
* rate by a linear regression of the HP trend on 1 and t. The standard
* error in that is likely to be fairly high relative to the standard
* error in the shock to the trend process (since the latter is a shock
* to a unit root process and so gets accumulated), so it gets scaled down.
*
filter(type=hp) lgdp / gdp_hp
set gap_hp = lgdp - gdp_hp
linreg gap_hp
# gap_hp{1 2}
compute ph1=%beta(1),ph2=%beta(2),se=sqrt(%seesq)
set trend = t
linreg gdp_hp
```

```
# constant trend
compute mu=%beta(2)
compute sn=sqrt(.1*%seesq)
*
dlm(presample=ergodic,a=af,z=zf,sw=swf,c=c,f=f,y=lgdp,$
   method=bfgs,type=filter)
*
* Same with broken trend rate
*
nonlin mu d sn ph1 ph2 se
*
frml zf = ||mu+d*d1973,0.0,0.0||
*
dlm(presample=ergodic,a=af,z=zf,sw=swf,c=c,f=f,y=lgdp,$
   method=bfgs,type=filter) / states0
*
set cycle0 = states0(t)(2)
set trend0 = states0(t)(1)
*
graph(shaded=recession,$
 footer="Figure 1 UC-0 Cycle, US Real GDP. Percent Deviation from Trend")
# cycle0
*
* Unrestricted UC model, allowing for correlation between the shocks.
* The maximum occurs at rho=1 (possible since we have only one
* observable, so having just a single effective shock is OK).
*
compute rhone=0.0,sn=.1
nonlin mu d sn ph1 ph2 se rhone=1.0
*
frml swf = ||sn^2|rhone*abs(sn)*abs(se),se^2||
dlm(presample=ergodic,a=af,z=zf,sw=swf,c=c,f=f,y=lgdp,method=bfgs,$
  type=filter) / statesur
*
* Alternative parameterization for covariance matrix of shocks.
*
dec packed swlower(2,2)
*
frml swf = %ltouterxx(swlower)
compute swlower(1,1)=.7,swlower(2,1)=0.0,swlower(2,2)=.7
nonlin mu d ph1 ph2 swlower
dlm(presample=ergodic,a=af,z=zf,sw=swf,c=c,f=f,y=lgdp,method=bfgs,$
  type=filter) / statesur
```