Additional time series models

Christopher F Baum

ECON 8823: Applied Econometrics

Boston College, Spring 2015

State-space models

Many linear time-series models can be written as linear state-space models, including vector autoregressive moving-average (VARMA) models, dynamic-factor (DF) models, and structural time series (STS) models. The solutions to some stochastic dynamic-programming problems can also be written in the form of linear state-space models.

We can estimate the parameters of a linear state-space model by maximum likelihood (ML). The Kalman filter or a diffuse Kalman filter is used to write the likelihood function in prediction-error form, assuming normally distributed errors. The quasi-maximum likelihood (QML) estimator, which drops the normality assumption, is consistent and asymptotically normal when the model is stationary.

State-space models

Many linear time-series models can be written as linear state-space models, including vector autoregressive moving-average (VARMA) models, dynamic-factor (DF) models, and structural time series (STS) models. The solutions to some stochastic dynamic-programming problems can also be written in the form of linear state-space models.

We can estimate the parameters of a linear state-space model by maximum likelihood (ML). The Kalman filter or a diffuse Kalman filter is used to write the likelihood function in prediction-error form, assuming normally distributed errors. The quasi-maximum likelihood (QML) estimator, which drops the normality assumption, is consistent and asymptotically normal when the model is stationary.

The Stata `sspace` command estimates linear state-space models with time-invariant coefficients, which include the models just listed and a number of others. These can be expressed as

$$
\begin{aligned}
z_t &= A z_{t-1} + B x_t + C \epsilon_t \\
y_t &= D z_t + F w_t + G \nu_t
\end{aligned}
$$

where z_t is a m-vector of unobserved state variables, y_t is a n-vector of observed endogenous variables, x_t and w_t are k_x and k_w vectors of exogenous variables, ϵ_t is a q-vector of state-error terms, ν_t is a r-vector of observation-error terms, and A, B, C, D, F, G are parameter matrices.

In this framework, the equations for z_t are known as the state equations, and the equations for y_t are known as the observation equations. The error terms are assumed to be zero mean, normally distributed, serially uncorrelated and independent of one another:

$$\begin{aligned} \epsilon_t &\sim N(0, Q) \\ \nu_t &\sim N(0, R) \end{aligned}$$

The state-space form is used to derive the log-likelihood of the observed endogenous variables conditional on their own past and any exogenous variables. When the model is stationary, a method for recursively predicting the current values of the states and the endogenous variables, known as the Kalman filter, is used to obtain the prediction error form of the log-likelihood function. When the model is nonstationary, a diffuse Kalman filter is used.

In this framework, the equations for z_t are known as the state equations, and the equations for y_t are known as the observation equations. The error terms are assumed to be zero mean, normally distributed, serially uncorrelated and independent of one another:

$$\begin{aligned} \epsilon_t &\sim N(0, Q) \\ \nu_t &\sim N(0, R) \end{aligned}$$

The state-space form is used to derive the log-likelihood of the observed endogenous variables conditional on their own past and any exogenous variables. When the model is stationary, a method for recursively predicting the current values of the states and the endogenous variables, known as the Kalman filter, is used to obtain the prediction error form of the log-likelihood function. When the model is nonstationary, a diffuse Kalman filter is used.

Example: a stationary state-space model

Following Hamilton's text (1994, pp. 372–374), we can write a standard *AR*(1) model

$$y_t - \mu = \alpha(y_{t-1} - \mu) + \epsilon_t$$

as a state-space model with state and observation equations

$$\begin{aligned} u_t &= \alpha u_{t-1} + \epsilon_t \\ y_t &= \mu + u_t \end{aligned}$$

where the unobserved state is $u_t = y_t - \mu$.

To implement this model on a univariate time series (in this case, the growth rate of the US manufacturing sector's capacity utilization rate), we specify the state and observation equations' components, imposing the constraint that u_t enters the observation equation with a unit coefficient.

```
. webuse manufac
(St. Louis Fed (FRED) manufacturing data)

. constraint 1 [D.lncaputil]u = 1

. sspace (u L.u, state noconstant) (D.lncaputil u, noerror), const(1) nolog vsq
> uish

State-space model

Sample: 1972m2 - 2008m12                        Number of obs     =        443
                                                Wald chi2(1)      =      61.73
Log likelihood =    1516.44                     Prob > chi2       =     0.0000
 ( 1)  [D.lncaputil]u = 1

------------------------------------------------------------------------------
             |                 OIM
  lncaputil  |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
u            |
           u |
         L1. |   .3523983   .0448539     7.86   0.000     .2644862    .4403104
-------------+----------------------------------------------------------------
D.lncaputil  |
           u |          1  (constrained)
       _cons |  -.0003558   .0005781    -0.62   0.538     -.001489    .0007773
-------------+----------------------------------------------------------------
Variance     |
  u          |   .0000622   4.18e-06    14.88   0.000      .000054    .0000704
------------------------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

The estimated autoregressive coefficient of 0.353 indicates that there is persistence in the growth rate of the CU rate series. The estimated mean of the differenced series is not distinguishable from zero, indicating the absence of a deterministic linear trend in the CU series.

As Hamilton shows, any univariate $AR(p)$ process can be placed in state-space form, with the number of state equations equal to p, the order of the autoregression, and a single observation equation for the contemporaneous level of the process. Likewise, any univariate $MA(q)$ process can be written as a set of $(q + 1)$ state equations in the errors and a single observation equation.

The estimated autoregressive coefficient of 0.353 indicates that there is persistence in the growth rate of the CU rate series. The estimated mean of the differenced series is not distinguishable from zero, indicating the absence of a deterministic linear trend in the CU series.

As Hamilton shows, any univariate $AR(p)$ process can be placed in state-space form, with the number of state equations equal to p, the order of the autoregression, and a single observation equation for the contemporaneous level of the process. Likewise, any univariate $MA(q)$ process can be written as a set of $(q+1)$ state equations in the errors and a single observation equation.

As a logical extension, any linear $ARMA(p, q)$ process can be written in state-space form by defining $r = \max(p, q + 1)$, which then gives rise to r state equations and a single observation equation. For example, consider a zero-mean $ARMA(1, 1)$ model:

$$y_t = \alpha y_{t-1} + \theta \epsilon_{t-1} + \epsilon_t$$

with state equations

$$\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \begin{pmatrix} y_t \\ \theta \epsilon_t \end{pmatrix} = \begin{pmatrix} \alpha & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ \theta \epsilon_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ \theta \end{pmatrix} \epsilon_t$$

and observation equation

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \theta \epsilon_t \end{pmatrix}$$

with u_{1t} and u_{2t} as the unobserved states.

We may estimate this model by writing down the state and observation equations, providing constraints for those coefficients which should be unity. As the previous example has shown that there is no deterministic trend in the level series, we set the mean of the differenced series to zero by excluding the constant from the observation equation.

```
. constraint 2 [u1]L.u2 = 1
. constraint 3 [u1]e.u1 = 1
. constraint 4 [D.lncaputil]u1 = 1
. sspace (u1 L.u1 L.u2 e.u1, state noconstant) (u2 e.u1, state noconstant) ///
>        (D.lncaputil u1, noconstant), constraints(2/4) covstate(diagonal) nolog vsq
State-space model
Sample: 1972m2 - 2008m12                          Number of obs    =       443
                                                  Wald chi2(2)     =    333.84
Log likelihood =   1531.255                       Prob > chi2      =    0.0000
 ( 1)  [u1]L.u2 = 1
 ( 2)  [u1]e.u1 = 1
 ( 3)  [D.lncaputil]u1 = 1
------------------------------------------------------------------------------
             |                 OIM
  lncaputil  |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
u1           |
          u1 |
         L1. |   .8056815   .0522661    15.41   0.000     .7032418    .9081212
          u2 |
         L1. |          1  (constrained)
        e.u1 |          1  (constrained)
-------------+----------------------------------------------------------------
u2           |
        e.u1 |  -.5188453   .0701985    -7.39   0.000    -.6564317   -.3812588
-------------+----------------------------------------------------------------
...
```

```
 ...
D.lncaputil |
         u1 |             1  (constrained)
------------+----------------------------------------------------------------
Variance    |
  u1        |      .0000582   3.91e-06    14.88   0.000      .0000505     .0000659
------------+----------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

In this 'error-form' representation, the coefficient on `L1.u1` in the `u1` equation is our estimate of α, and the coefficient on `e.u1` in the `u2` equation is our estimate of θ.

```
 ...
D.lncaputil  |
          u1 |            1   (constrained)
-------------+----------------------------------------------------------------
Variance     |
  u1         |     .0000582   3.91e-06    14.88   0.000     .0000505    .0000659
------------------------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

In this ‘error-form’ representation, the coefficient on `L1.u1` in the `u1` equation is our estimate of α, and the coefficient on `e.u1` in the `u2` equation is our estimate of θ.

Example: a bivariate state-space model

In the `manufac` dataset, the `lnhours` variable represents the log of manufacturing hours per week, which we treat as stationary in first differences. If we hypothesize that the process driving the growth rate in capacity utilization affects the growth rate of hours worked, but not vice versa, then we want to express the comovements of these variables in a triangular linear system: essentially a $VAR(1)$ subject to constraints.

$$\begin{pmatrix} \Delta lncaputil_t \\ \Delta lnhours_t \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 \\ \alpha_2 & \alpha_3 \end{pmatrix} \begin{pmatrix} \Delta lncaputil_{t-1} \\ \Delta lnhours_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

We can write this in state-space form with state equations

$$\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 \\ \alpha_2 & \alpha_3 \end{pmatrix} \begin{pmatrix} u_{1,t-1} \\ u_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

with Var(ϵ) = Σ and observation equations

$$\begin{pmatrix} \Delta lncaputil_t \\ \Delta lnhours_t \end{pmatrix} = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

$$\begin{pmatrix} \Delta lncaputil_t \\ \Delta lnhours_t \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 \\ \alpha_2 & \alpha_3 \end{pmatrix} \begin{pmatrix} \Delta lncaputil_{t-1} \\ \Delta lnhours_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

We can write this in state-space form with state equations

$$\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 \\ \alpha_2 & \alpha_3 \end{pmatrix} \begin{pmatrix} u_{1,t-1} \\ u_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

with Var(ϵ) = Σ and observation equations

$$\begin{pmatrix} \Delta lncaputil_t \\ \Delta lnhours_t \end{pmatrix} = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

To estimate the model, we specify each of the state equations and observation equations, keeping in mind that the latter are trivial identities. The `covstate(unstructured)` option specifies that the covariance structure for the state errors (ϵ in this example) should be symmetric and positive definite, with parameters for all variances and covariances to be estimated.

```
 . constraint 5 [D.lncaputil]u1 = 1
. constraint 6 [D.lnhours]u2 = 1
. sspace (u1 L.u1, state noconstant) ///
>        (u2 L.u1 L.u2, state noconstant) ///
>        (D.lncaputil u1, noconstant noerror) ///
>        (D.lnhours u2, noconstant noerror), ///
>        constraints(5/6) covstate(unstructured) nolog vsquish
State-space model
Sample: 1972m2 - 2008m12                          Number of obs   =        443
                                                  Wald chi2(3)    =     166.87
Log likelihood =  3211.7532                       Prob > chi2     =     0.0000
 ( 1)  [D.lncaputil]u1 = 1
 ( 2)  [D.lnhours]u2 = 1
------------------------------------------------------------------------------
             |                 OIM
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
u1           |
          u1 |
         L1. |    .353257   .0448456     7.88   0.000     .2653612    .4411528
-------------+----------------------------------------------------------------
u2           |
          u1 |
         L1. |   .1286218   .0394742     3.26   0.001     .0512537    .2059899
          u2 |
         L1. |  -.3707083   .0434255    -8.54   0.000    -.4558208   -.2855959
------------------------------------------------------------------------------
...
```

```
 ...
D.lncaputil  |
          u1 |          1  (constrained)
-------------+------------------------------------------------------------------
D.lnhours    |
          u2 |          1  (constrained)
-------------+------------------------------------------------------------------
Variance     |
  u1         |   .0000623   4.19e-06    14.88   0.000      .0000541    .0000705
-------------+------------------------------------------------------------------
Covariance   |
  u1         |
          u2 |    .000026   2.67e-06     9.75   0.000      .0000208    .0000312
-------------+------------------------------------------------------------------
Variance     |
  u2         |   .0000386   2.61e-06    14.76   0.000      .0000335    .0000437
--------------------------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

The estimated parameter in the `u1` equation is α_1. The estimated parameters in the `u2` equation are α_2, α_3 respectively. The estimated autoregressive coefficient α_1 is similar to that produced in the univariate model for `D.lncaputil` in the earlier example. Both the effect of `D.lncaputil` on `D.lnhours` and the autoregressive coefficient for `D.lnhours` are statistically significant.

We could also impose constraints on the covariance matrix of state errors, such as restricting the covariance of the errors to zero with an additional `constraint` command.

The estimated parameter in the `u1` equation is α_1. The estimated parameters in the `u2` equation are α_2, α_3 respectively. The estimated autoregressive coefficient α_1 is similar to that produced in the univariate model for `D.lncaputil` in the earlier example. Both the effect of `D.lncaputil` on `D.lnhours` and the autoregressive coefficient for `D.lnhours` are statistically significant.

We could also impose constraints on the covariance matrix of state errors, such as restricting the covariance of the errors to zero with an additional `constraint` command.

We may add additional structure to this bivariate example by allowing the error process to be non-*i.i.d.*. While still maintaining the triangular structure of the system, we add a $MA(1)$ component to the CU equation, but continue to model `D.lnhours` as an autoregressive process:

$$\begin{pmatrix} \Delta lncaputil_t \\ \Delta lnhours_t \end{pmatrix} = \begin{pmatrix} \alpha_1 & 0 \\ \alpha_2 & \alpha_3 \end{pmatrix} \begin{pmatrix} \Delta lncaputil_{t-1} \\ \Delta lnhours_{t-1} \end{pmatrix} + \begin{pmatrix} \theta_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \epsilon_{1,t-1} \\ \epsilon_{2,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

A vector autoregressive moving-average, or $VARMA(1,1)$, process.

This can be written in state-space form with state equations

$$\begin{pmatrix} s_{1t} \\ s_{2t} \\ s_{3t} \end{pmatrix} = \begin{pmatrix} \alpha_1 & 1 & 0 \\ 0 & 0 & 0 \\ \alpha_2 & 0 & \alpha_3 \end{pmatrix} \begin{pmatrix} s_{1,t-1} \\ s_{2,t-1} \\ s_{3,t-1} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ \theta_1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

with states

$$\begin{pmatrix} s_{1t} \\ s_{2t} \\ s_{3t} \end{pmatrix} = \begin{pmatrix} \Delta lncaputil_t \\ \theta_1 \epsilon_{1t} \\ \Delta lnhours_t \end{pmatrix}$$

We assume the VCE of the state errors is diagonal, so that only the two variances are to be estimated.

This can be written in state-space form with state equations

$$\begin{pmatrix} s_{1t} \\ s_{2t} \\ s_{3t} \end{pmatrix} = \begin{pmatrix} \alpha_1 & 1 & 0 \\ 0 & 0 & 0 \\ \alpha_2 & 0 & \alpha_3 \end{pmatrix} \begin{pmatrix} s_{1,t-1} \\ s_{2,t-1} \\ s_{3,t-1} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ \theta_1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

with states

$$\begin{pmatrix} s_{1t} \\ s_{2t} \\ s_{3t} \end{pmatrix} = \begin{pmatrix} \Delta lncaputil_t \\ \theta_1 \epsilon_{1t} \\ \Delta lnhours_t \end{pmatrix}$$

We assume the VCE of the state errors is diagonal, so that only the two variances are to be estimated.

To estimate this *VARMA*$(1, 1)$ process, we spell out each of the state equations and observation equations, with the latter as trivial identities. Note that in this expanded form of the model, we have three state equations, but still have only two observation equations. The `covstate(diagonal)` option allows us to specify that only the variances in the state errors' VCE are to be estimated.

```
. constraint 7 [u1]L.u2 = 1
. constraint 8 [u1]e.u1 = 1
. constraint 9 [u3]e.u3 = 1
. constraint 10 [D.lncaputil]u1 = 1
. constraint 11 [D.lnhours]u3 = 1
. sspace (u1 L.u1 L.u2 e.u1, state noconstant) ///
>        (u2 e.u1, state noconstant) ///
>        (u3 L.u1 L.u3 e.u3, state noconstant) ///
>        (D.lncaputil u1, noconstant) (D.lnhours u3, noconstant), ///
> constraints(7/11) technique(nr) covstate(diagonal) nolog vsquish
State-space model
Sample: 1972m2 - 2008m12                          Number of obs  =       443
                                                  Wald chi2(4)   =    427.55
Log likelihood =  3156.0564                       Prob > chi2    =    0.0000
 ( 1)  [u1]L.u2 = 1
 ( 2)  [u1]e.u1 = 1
 ( 3)  [u3]e.u3 = 1
 ( 4)  [D.lncaputil]u1 = 1
 ( 5)  [D.lnhours]u3 = 1
------------------------------------------------------------------------------
             |                 OIM
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
u1           |
          u1 |
         L1. |   .8058031   .0522493    15.42   0.000     .7033964    .9082098
...
```

```
 ...
           u2 |
          L1. |          1  (constrained)
         e.u1 |          1  (constrained)
--------------+----------------------------------------------------------------
u2            |
         e.u1 |   -.518907   .0701848    -7.39   0.000    -.6564667   -.3813474
--------------+----------------------------------------------------------------
u3            |
           u1 |
          L1. |   .1734868   .0405156     4.28   0.000     .0940776     .252896
           u3 |
          L1. |  -.4809376   .0498574    -9.65   0.000    -.5786563   -.3832188
         e.u3 |          1  (constrained)
--------------+----------------------------------------------------------------
D.lncaputil   |
           u1 |          1  (constrained)
--------------+----------------------------------------------------------------
D.lnhours     |
           u3 |          1  (constrained)
--------------+----------------------------------------------------------------
Variance      |
  u1          |   .0000582   3.91e-06    14.88   0.000     .0000505    .0000659
  u3          |   .0000382   2.56e-06    14.88   0.000     .0000331    .0000432
-------------------------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

Not surprisingly, the `D.lnhours` equation indicates that the lagged value of `D.lncaputil` has a positive effect (0.173) on hours worked.

Example: a latent factor state-space model

Following Stock and Watson (*NBER Macro Annual*, 1989), we estimate the parameters of a latent factor model, using four observed series: an industrial production index, aggregate weekly hours, aggregate unemployment and real disposable income. We consider that these variables are jointly driven by a latent factor, f_t, that follows an $AR(2)$ process. In state-space form, the model becomes

$$\begin{pmatrix} f_t \\ f_{t-1} \end{pmatrix} = \begin{pmatrix} \theta_1 & \theta_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f_{t-1} \\ f_{t-2} \end{pmatrix} + \begin{pmatrix} \nu_t \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} \Delta IP_t \\ \Delta Income_t \\ \Delta hours_t \\ \Delta unemp_t \end{pmatrix} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma 4 \end{pmatrix} f_t + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \\ \epsilon_{4t} \end{pmatrix}$$

Assuming a diagonal covariance matrix, we specify the state equations for f_t and its lag, with each observation equation depending linearly on the latent factor f_t.

```
. webuse dfex, clear
(St. Louis Fed (FRED) macro data)
. constraint 12 [lf]L.f = 1
. sspace (f L.f L.lf, state noconstant) (lf L.f, state noconstant noerror) ///
>    (D.ipman f, noconstant) (D.income f, noconstant) (D.hours f, noconstant) //
> /
>    (D.unemp f, noconstant), covstate(identity) constraints(12) nolog vsquish
State-space model
Sample: 1972m2 - 2008m11                          Number of obs   =        442
                                                  Wald chi2(6)    =     751.95
Log likelihood = -662.09507                       Prob > chi2     =     0.0000
 ( 1)  [lf]L.f = 1
```

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
f						
f						
L1.	.2651932	.0568663	4.66	0.000	.1537372	.3766491
lf						
L1.	.4820398	.0624635	7.72	0.000	.3596136	.604466
lf						
f						
L1.	1	(constrained)				

...

```
  ...
D.ipman      |
           f |   .3502249   .0287389    12.19   0.000     .2938976    .4065522
-------------+----------------------------------------------------------------
D.income     |
           f |   .0746338   .0217319     3.43   0.001     .0320401    .1172276
-------------+----------------------------------------------------------------
D.hours      |
           f |   .2177469   .0186769    11.66   0.000     .1811407     .254353
-------------+----------------------------------------------------------------
D.unemp      |
           f |  -.0676016   .0071022    -9.52   0.000    -.0815217   -.0536816
-------------+----------------------------------------------------------------
Variance     |
  D.ipman    |   .1383158   .0167086     8.28   0.000     .1055675    .1710641
  D.income   |   .2773808   .0188302    14.73   0.000     .2404743    .3142873
  D.hours    |   .0911446   .0080847    11.27   0.000     .0752988    .1069903
  D.unemp    |   .0237232   .0017932    13.23   0.000     .0202086    .0272378
------------------------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

The sizable autoregressive coefficients (0.265, 0.482) on the latent factor indicate that it is quite persistent. The IP, income and hours variables all load positively on the factor, while the unemployment rate variable has a significant negative coefficient. The unobserved factor has predictive power for each of the observed variables.

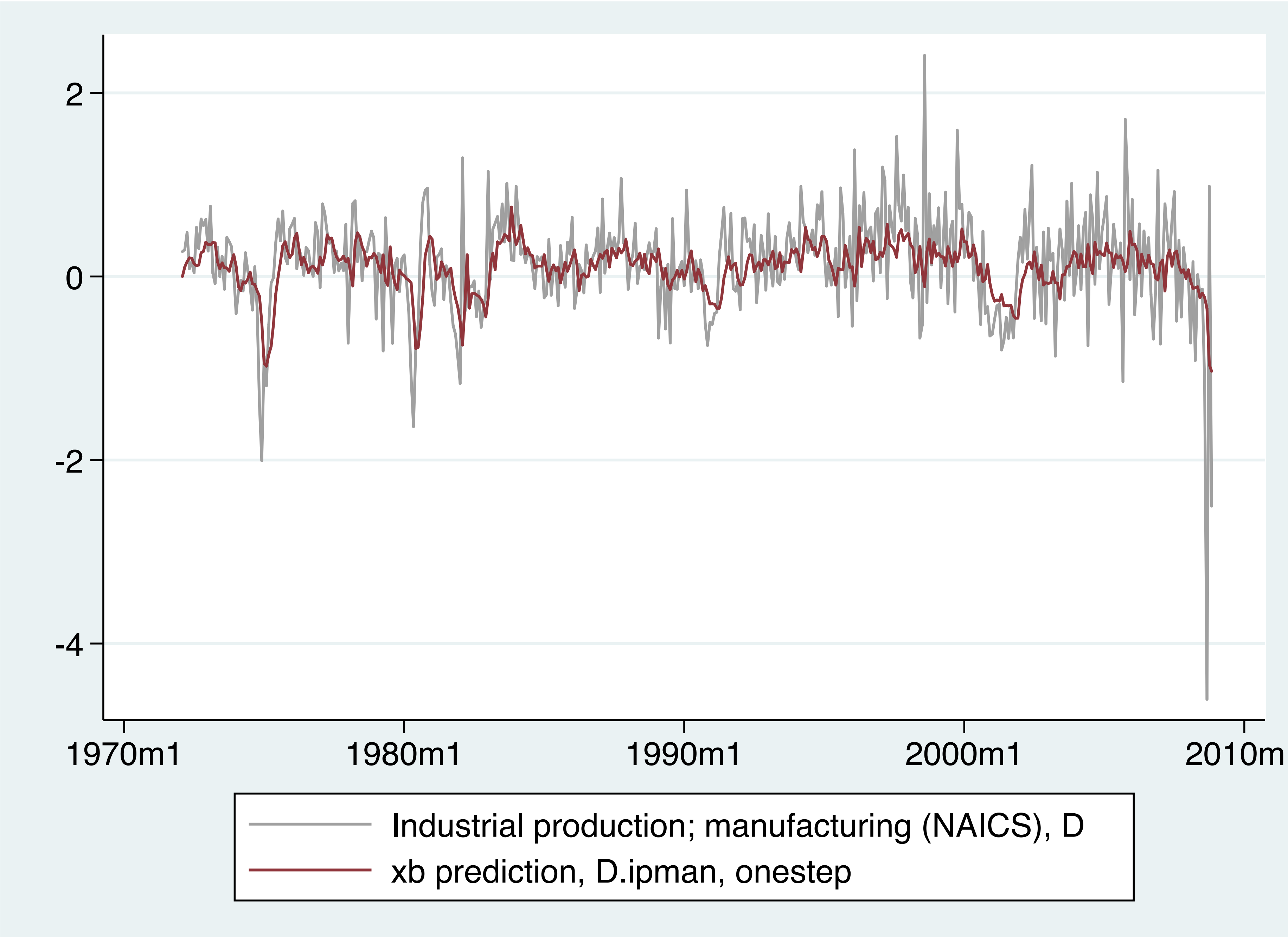
After estimating the model, we can obtain the one-step predictions for each of the four observed variables, and plot them against their actual values.

```
 . predict dep*
(option xb assumed; fitted values)
. tsline D.ipman dep1, lcolor(gs10) xtitle("") legend(rows(2)) ylab(,angle(0))
. gr export 82311-6.pdf, replace
(file /Users/cfbaum/Dropbox/baum/EC823 S2013/82311-6.pdf written in PDF format)
```

The sizable autoregressive coefficients (0.265, 0.482) on the latent factor indicate that it is quite persistent. The IP, income and hours variables all load positively on the factor, while the unemployment rate variable has a significant negative coefficient. The unobserved factor has predictive power for each of the observed variables.

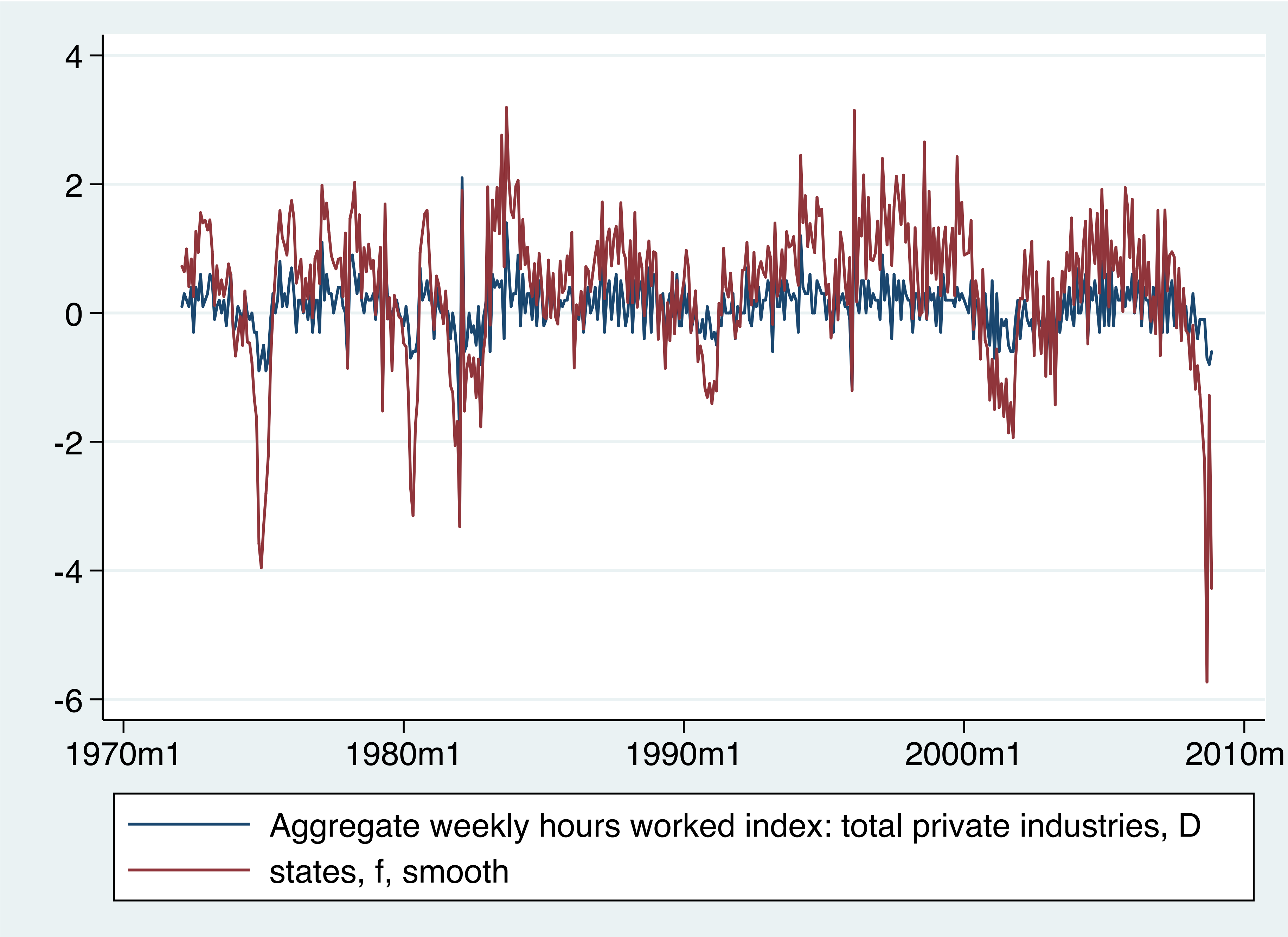
After estimating the model, we can obtain the one-step predictions for each of the four observed variables, and plot them against their actual values.

```
 . predict dep*
(option xb assumed; fitted values)
. tsline D.ipman dep1, lcolor(gs10) xtitle("") legend(rows(2)) ylab(,angle(0))
. gr export 82311-6.pdf, replace
(file /Users/cfbaum/Dropbox/baum/EC823 S2013/82311-6.pdf written in PDF format)
```


2
0
-2
-4
1970m1
1980m1
1990m1
2000m1
2010m1
Industrial production; manufacturing (NAICS), D
xb prediction, D.ipman, onestep

We may also estimate the unobserved (latent) factor, specifying `method(smooth)` in the `predict` command to produce this series. We graph the $\widehat{f_t}$ series along with the change in hours worked, one of the observed series used in the model. Dynamic (out-of-sample) forecasts can also be made from an estimated state-space model.

```
 . predict fac if e(sample), states smethod(smooth) equation(f)
. tsline D.hours fac, xtitle("") legend(rows(2)) ylab(,angle(0))
. gr export 82311-7.pdf, replace
(file /Users/cfbaum/Dropbox/baum/EC823 S2013/82311-7.pdf written in PDF format)
```


4
2
0
-2
-4
-6
1970m1
1980m1
1990m1
2000m1
2010m1
Aggregate weekly hours worked index: total private industries, D
states, f, smooth

Nonstationary state-space models

State-space models can also be applied to nonstationary time series, as proposed by Andrew Harvey. These models parameterize the trend and seasonal components of a set of time series. For instance, the local-level model:

$$
\begin{aligned}
y_t &= \mu_t + \epsilon_t \\
\mu_t &= \mu_{t-1} + \nu_t
\end{aligned}
$$

Here the level of the series is modeled as a random walk plus idiosyncratic noise. It is thus nonstationary. If the variance of ϵ is zero and the variance of ν is positive, the model reduces to a pure random walk. In the opposite case, we have a simple regression with a constant mean.

Nonstationary state-space models

State-space models can also be applied to nonstationary time series, as proposed by Andrew Harvey. These models parameterize the trend and seasonal components of a set of time series. For instance, the local-level model:

$$\begin{aligned} y_t &= \mu_t + \epsilon_t \\ \mu_t &= \mu_{t-1} + \nu_t \end{aligned}$$

Here the level of the series is modeled as a random walk plus idiosyncratic noise. It is thus nonstationary. If the variance of ϵ is zero and the variance of ν is positive, the model reduces to a pure random walk. In the opposite case, we have a simple regression with a constant mean.

We fit this model to weekly closing values of S&P 500 Index.

```
. webuse sp500w, clear
. constraint 13 [z]L.z = 1
. constraint 14 [close]z = 1
. sspace (z L.z, state nocons) (close z, nocons), const(13 14) nolog vsquish
State-space model
Sample: 1 - 3093                                  Number of obs    =       3093
Log likelihood =  -12576.99
 ( 1)   [z]L.z = 1
 ( 2)   [close]z = 1
------------------------------------------------------------------------------
             |                 OIM
       close |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
z            |
           z |
         L1. |          1  (constrained)
-------------+----------------------------------------------------------------
close        |
           z |          1  (constrained)
-------------+----------------------------------------------------------------
Variance     |
  z          |   170.3456   7.584909    22.46   0.000     155.4794    185.2117
  close      |   15.24858   3.392457     4.49   0.000     8.599486    21.89767
------------------------------------------------------------------------------
Note: Model is not stationary.
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

As both components have nonzero variances, the model is nonstationary.

An extension of this model is the local linear-trend model, in which both the level and slope of a linear time trend are assumed to follow a random walk.

$$\begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \beta_{t-1} \end{pmatrix} + \begin{pmatrix} \nu_{1t} \\ \nu_{2t} \end{pmatrix}$$

where $y_t = \mu_t + \epsilon_t$ is the observation equation.

We may fit this model to the industrial production series:

As both components have nonzero variances, the model is nonstationary.

An extension of this model is the local linear-trend model, in which both the level and slope of a linear time trend are assumed to follow a random walk.

$$\begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \beta_{t-1} \end{pmatrix} + \begin{pmatrix} \nu_{1t} \\ \nu_{2t} \end{pmatrix}$$

where $y_t = \mu_t + \epsilon_t$ is the observation equation.

We may fit this model to the industrial production series:

As both components have nonzero variances, the model is nonstationary.

An extension of this model is the local linear-trend model, in which both the level and slope of a linear time trend are assumed to follow a random walk.

$$\begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \beta_{t-1} \end{pmatrix} + \begin{pmatrix} \nu_{1t} \\ \nu_{2t} \end{pmatrix}$$

where $y_t = \mu_t + \epsilon_t$ is the observation equation.

We may fit this model to the industrial production series:

```
 . webuse dfex, clear
(St. Louis Fed (FRED) macro data)
. constraint 15 [f1]L.f1 = 1
. constraint 16 [f1]L.f2 = 1
. constraint 17 [f2]L.f2 = 1
. constraint 18 [ipman]f1 = 1
. sspace (f1 L.f1 L.f2, state noconstant) (f2 L.f2, state noconstant) ///
>  (ipman f1, noconstant), constraints(15/18) nolog vsquish
State-space model
Sample: 1972m1 - 2008m11                          Number of obs    =        443
Log likelihood =  -359.1266
 ( 1)  [f1]L.f1 = 1
 ( 2)  [f1]L.f2 = 1
 ( 3)  [f2]L.f2 = 1
 ( 4)  [ipman]f1 = 1
------------------------------------------------------------------------------
             |                 OIM
       ipman |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
f1           |
          f1 |
         L1. |          1  (constrained)
          f2 |
         L1. |          1  (constrained)
------------------------------------------------------------------------------
             |
...
```

```
 ...
f2         |
        f2 |
       L1. |          1  (constrained)
-----------+--------------------------------------------------------------------
ipman      |
        f1 |          1  (constrained)
-----------+--------------------------------------------------------------------
Variance   |
  f1       |   .1473071   .0407156     3.62   0.000       .067506    .2271082
  f2       |   .0178752   .0065743     2.72   0.003      .0049898    .0307606
  ipman    |   .0354429   .0148186     2.39   0.008      .0063989    .0644868
--------------------------------------------------------------------------------
Note: Model is not stationary.
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

The estimation results suggest that both of the variance parameters are nonzero, providing support for the local linear-trend model.

```
 ...
f2           |
          f2 |
         L1. |          1  (constrained)
-------------+----------------------------------------------------------------
ipman        |
          f1 |          1  (constrained)
-------------+----------------------------------------------------------------
Variance     |
  f1         |   .1473071   .0407156     3.62   0.000      .067506    .2271082
  f2         |   .0178752   .0065743     2.72   0.003     .0049898    .0307606
  ipman      |   .0354429   .0148186     2.39   0.008     .0063989    .0644868
------------------------------------------------------------------------------
Note: Model is not stationary.
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

The estimation results suggest that both of the variance parameters are nonzero, providing support for the local linear-trend model.

Unobserved components models

A specification that is closely related to the nonstationary state-space model is the unobserved component model (UCM). These models decompose a time series into trend, seasonal, cyclical, and idiosyncratic components, allowing for exogenous factors as well:

$$y_t = \tau_t + \gamma_t + \psi_t + \beta x_t + \epsilon_t$$

where τ_t, γ_t, and ψ_t are the trend, seasonal and cyclical components, respectively. β is a vector of fixed parameters. These models can be expressed in the state-space framework and estimated via maximum likelihood.

To parameterize the UCM, a specification must be made for the trend and idiosyncratic components. Additional factors: a cyclical component, seasonal component, or exogenous variables, may also be added.

Harvey (1989) defines 11 flexible models that jointly specify τ_t and ϵ_t. These models are constructed from a common set of building blocks:

To parameterize the UCM, a specification must be made for the trend and idiosyncratic components. Additional factors: a cyclical component, seasonal component, or exogenous variables, may also be added.

Harvey (1989) defines 11 flexible models that jointly specify τ_t and ϵ_t. These models are constructed from a common set of building blocks:

1. No trend or idiosyncratic component (for other components)
2. No trend: $y_t = \epsilon_t$ (for other components)
3. Deterministic constant: $y_t = \mu + \epsilon_t$
4. Local level: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \eta_t$
5. Random walk: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \eta_t$
6. Deterministic trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta$
7. Local level / det. trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta + \eta_t$
8. Random walk with drift: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \beta + \eta_t$
9. Local linear trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \beta_t = \beta_{t-1} + \xi_t$
10. Smooth trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$
11. Random trend: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$

1. No trend or idiosyncratic component (for other components)
2. No trend: $y_t = \epsilon_t$ (for other components)
3. Deterministic constant: $y_t = \mu + \epsilon_t$
4. Local level: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \eta_t$
5. Random walk: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \eta_t$
6. Deterministic trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta$
7. Local level / det. trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta + \eta_t$
8. Random walk with drift: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \beta + \eta_t$
9. Local linear trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \beta_t = \beta_{t-1} + \xi_t$
10. Smooth trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$
11. Random trend: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$

1. No trend or idiosyncratic component (for other components)
2. No trend: $y_t = \epsilon_t$ (for other components)
3. Deterministic constant: $y_t = \mu + \epsilon_t$
4. Local level: $y_t = \mu_t + \epsilon_t, \; \mu_t = \mu_{t-1} + \eta_t$
5. Random walk: $y_t = \mu_t, \; \mu_t = \mu_{t-1} + \eta_t$
6. Deterministic trend: $y_t = \mu_t + \epsilon_t, \; \mu_t = \mu_{t-1} + \beta$
7. Local level / det. trend: $y_t = \mu_t + \epsilon_t, \; \mu_t = \mu_{t-1} + \beta + \eta_t$
8. Random walk with drift: $y_t = \mu_t, \; \mu_t = \mu_{t-1} + \beta + \eta_t$
9. Local linear trend: $y_t = \mu_t + \epsilon_t, \; \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \beta_t = \beta_{t-1} + \xi_t$
10. Smooth trend: $y_t = \mu_t + \epsilon_t, \; \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$
11. Random trend: $y_t = \mu_t, \; \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$

1. No trend or idiosyncratic component (for other components)
2. No trend: $y_t = \epsilon_t$ (for other components)
3. Deterministic constant: $y_t = \mu + \epsilon_t$
4. Local level: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \eta_t$
5. Random walk: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \eta_t$
6. Deterministic trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta$
7. Local level / det. trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta + \eta_t$
8. Random walk with drift: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \beta + \eta_t$
9. Local linear trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \beta_t = \beta_{t-1} + \xi_t$
10. Smooth trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$
11. Random trend: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$

1. No trend or idiosyncratic component (for other components)
2. No trend: $y_t = \epsilon_t$ (for other components)
3. Deterministic constant: $y_t = \mu + \epsilon_t$
4. Local level: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \eta_t$
5. Random walk: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \eta_t$
6. Deterministic trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta$
7. Local level / det. trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta + \eta_t$
8. Random walk with drift: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \beta + \eta_t$
9. Local linear trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \beta_t = \beta_{t-1} + \xi_t$
10. Smooth trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$
11. Random trend: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$

1. No trend or idiosyncratic component (for other components)
2. No trend: $y_t = \epsilon_t$ (for other components)
3. Deterministic constant: $y_t = \mu + \epsilon_t$
4. Local level: $y_t = \mu_t + \epsilon_t, \; \mu_t = \mu_{t-1} + \eta_t$
5. Random walk: $y_t = \mu_t, \; \mu_t = \mu_{t-1} + \eta_t$
6. Deterministic trend: $y_t = \mu_t + \epsilon_t, \; \mu_t = \mu_{t-1} + \beta$
7. Local level / det. trend: $y_t = \mu_t + \epsilon_t, \; \mu_t = \mu_{t-1} + \beta + \eta_t$
8. Random walk with drift: $y_t = \mu_t, \; \mu_t = \mu_{t-1} + \beta + \eta_t$
9. Local linear trend: $y_t = \mu_t + \epsilon_t, \; \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \beta_t = \beta_{t-1} + \xi_t$
10. Smooth trend: $y_t = \mu_t + \epsilon_t, \; \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$
11. Random trend: $y_t = \mu_t, \; \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$

1. No trend or idiosyncratic component (for other components)
2. No trend: $y_t = \epsilon_t$ (for other components)
3. Deterministic constant: $y_t = \mu + \epsilon_t$
4. Local level: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \eta_t$
5. Random walk: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \eta_t$
6. Deterministic trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta$
7. Local level / det. trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta + \eta_t$
8. Random walk with drift: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \beta + \eta_t$
9. Local linear trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \beta_t = \beta_{t-1} + \xi_t$
10. Smooth trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$
11. Random trend: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$

1. No trend or idiosyncratic component (for other components)
2. No trend: $y_t = \epsilon_t$ (for other components)
3. Deterministic constant: $y_t = \mu + \epsilon_t$
4. Local level: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \eta_t$
5. Random walk: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \eta_t$
6. Deterministic trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta$
7. Local level / det. trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta + \eta_t$
8. Random walk with drift: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \beta + \eta_t$
9. Local linear trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \beta_t = \beta_{t-1} + \xi_t$
10. Smooth trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$
11. Random trend: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$

1. No trend or idiosyncratic component (for other components)
2. No trend: $y_t = \epsilon_t$ (for other components)
3. Deterministic constant: $y_t = \mu + \epsilon_t$
4. Local level: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \eta_t$
5. Random walk: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \eta_t$
6. Deterministic trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta$
7. Local level / det. trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta + \eta_t$
8. Random walk with drift: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \beta + \eta_t$
9. Local linear trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \beta_t = \beta_{t-1} + \xi_t$
10. Smooth trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$
11. Random trend: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$

1. No trend or idiosyncratic component (for other components)
2. No trend: $y_t = \epsilon_t$ (for other components)
3. Deterministic constant: $y_t = \mu + \epsilon_t$
4. Local level: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \eta_t$
5. Random walk: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \eta_t$
6. Deterministic trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta$
7. Local level / det. trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta + \eta_t$
8. Random walk with drift: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \beta + \eta_t$
9. Local linear trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \beta_t = \beta_{t-1} + \xi_t$
10. Smooth trend: $y_t = \mu_t + \epsilon_t,\ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$
11. Random trend: $y_t = \mu_t,\ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$

1. No trend or idiosyncratic component (for other components)
2. No trend: $y_t = \epsilon_t$ (for other components)
3. Deterministic constant: $y_t = \mu + \epsilon_t$
4. Local level: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \eta_t$
5. Random walk: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \eta_t$
6. Deterministic trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta$
7. Local level / det. trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta + \eta_t$
8. Random walk with drift: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \beta + \eta_t$
9. Local linear trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \beta_t = \beta_{t-1} + \xi_t$
10. Smooth trend: $y_t = \mu_t + \epsilon_t, \ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$
11. Random trend: $y_t = \mu_t, \ \mu_t = \mu_{t-1} + \beta_{t-1}, \beta_t = \beta_{t-1} + \xi_t$

Many of these models are designed to handle nonstationary time series. The local-level, random-walk, local-level with deterministic trend and random-walk-with-drift models incorporate first-order stochastic trends. The local-linear-trend, smooth-trend and random-trend models are used for series with second-order stochastic trends, which would have to be differenced twice to render them stationary.

A seasonal component models cyclical behavior that occurs at known seasonal periodicities. Modeled in the time domain, the period of the cycle is specified as the number of time periods required for the cycle to complete: e.g., four for quarterly seasonality, twelve for monthly seasonality. Seasonal components may be either deterministic or stochastic. If stochastic, one models the variance of the seasonal component, analogous to random effects in a panel context.

Many of these models are designed to handle nonstationary time series. The local-level, random-walk, local-level with deterministic trend and random-walk-with-drift models incorporate first-order stochastic trends. The local-linear-trend, smooth-trend and random-trend models are used for series with second-order stochastic trends, which would have to be differenced twice to render them stationary.

A seasonal component models cyclical behavior that occurs at known seasonal periodicities. Modeled in the time domain, the period of the cycle is specified as the number of time periods required for the cycle to complete: e.g., four for quarterly seasonality, twelve for monthly seasonality. Seasonal components may be either deterministic or stochastic. If stochastic, one models the variance of the seasonal component, analogous to random effects in a panel context.

As a starting point, consider the default UCM of a random walk process, fit to monthly data on the US civilian unemployment rate.

```
. webuse unrate, clear

. ucm unrate, nolog vsquish

Unobserved-components model
Components: random walk

Sample: 1948m1 - 2011m1                         Number of obs   =        757
Log likelihood =  84.401307
```

unrate	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Variance						
level	.0467196	.002403	19.44	0.000	.0420098	.0514294

```
Note: Model is not stationary.
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

The estimated variance relates to the underlying η_t process.

As a starting point, consider the default UCM of a random walk process, fit to monthly data on the US civilian unemployment rate.

```
 . webuse unrate, clear
. ucm unrate, nolog vsquish
Unobserved-components model
Components: random walk
Sample: 1948m1 - 2011m1                         Number of obs     =        757
Log likelihood =  84.401307
------------------------------------------------------------------------------
             |                 OIM
      unrate |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Variance     |
       level |   .0467196    .002403    19.44   0.000     .0420098    .0514294
------------------------------------------------------------------------------
Note: Model is not stationary.
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

The estimated variance relates to the underlying η_t process.

Following Harvey (1989), we expand upon the simple random walk model to incorporate a stationary cyclical component that produces serially correlated shocks around the random-walk trend. This stochastic-cycle model has three parameters:

1. the frequency at which the random components are centered
2. a damping factor describing the dispersion of the random components around that frequency
3. the variance of the stochastic-cycle process, which acts as a scale factor.

Following Harvey (1989), we expand upon the simple random walk model to incorporate a stationary cyclical component that produces serially correlated shocks around the random-walk trend. This stochastic-cycle model has three parameters:

1. the frequency at which the random components are centered
2. a damping factor describing the dispersion of the random components around that frequency
3. the variance of the stochastic-cycle process, which acts as a scale factor.

Following Harvey (1989), we expand upon the simple random walk model to incorporate a stationary cyclical component that produces serially correlated shocks around the random-walk trend. This stochastic-cycle model has three parameters:

1. the frequency at which the random components are centered
2. a damping factor describing the dispersion of the random components around that frequency
3. the variance of the stochastic-cycle process, which acts as a scale factor.

```
. ucm unrate, cycle(1) nolog vsquish
Unobserved-components model
Components: random walk, order 1 cycle
Sample: 1948m1 - 2011m1                         Number of obs    =        757
                                                Wald chi2(2)     =   26650.81
Log likelihood =  118.88421                     Prob > chi2      =     0.0000
```

unrate	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
frequency	.0933466	.0103609	9.01	0.000	.0730397	.1136535
damping	.9820003	.0061121	160.66	0.000	.9700207	.9939798
Variance						
level	.0143786	.0051392	2.80	0.003	.004306	.0244511
cycle1	.0270339	.0054343	4.97	0.000	.0163829	.0376848

```
Note: Model is not stationary.
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

The estimated frequency is small, implying that cycles are centered around low-frequency components. The sizable damping factor indicates that cyclical components are close to this frequency. The estimated variance of the cyclical component is significantly different from zero.

The estimated central frequency may be converted to an estimated central period:

```
. estat period
```

cycle1	Coef.	Std. Err.	[95% Conf.	Interval]
period	67.31029	7.471004	52.66739	81.95319
frequency	.0933466	.0103609	.0730397	.1136535
damping	.9820003	.0061121	.9700207	.9939798

```
Note: Cycle time unit is monthly.
```

The period of 67 months implies a cyclical component with periodicity of about 5.6 years, within conventional business-cycle periodicities.

The estimated frequency is small, implying that cycles are centered around low-frequency components. The sizable damping factor indicates that cyclical components are close to this frequency. The estimated variance of the cyclical component is significantly different from zero.

The estimated central frequency may be converted to an estimated central period:

```
. estat period
```

cycle1	Coef.	Std. Err.	[95% Conf.	Interval]
period	67.31029	7.471004	52.66739	81.95319
frequency	.0933466	.0103609	.0730397	.1136535
damping	.9820003	.0061121	.9700207	.9939798

Note: Cycle time unit is monthly.

The period of 67 months implies a cyclical component with periodicity of about 5.6 years, within conventional business-cycle periodicities.

The estimated frequency is small, implying that cycles are centered around low-frequency components. The sizable damping factor indicates that cyclical components are close to this frequency. The estimated variance of the cyclical component is significantly different from zero.

The estimated central frequency may be converted to an estimated central period:

```
. estat period
```

cycle1	Coef.	Std. Err.	[95% Conf.	Interval]
period	67.31029	7.471004	52.66739	81.95319
frequency	.0933466	.0103609	.0730397	.1136535
damping	.9820003	.0061121	.9700207	.9939798

```
Note: Cycle time unit is monthly.
```

The period of 67 months implies a cyclical component with periodicity of about 5.6 years, within conventional business-cycle periodicities.

Interpreting cycles in the frequency domain

To understand the stochastic-cycle model, consider that any stationary process may be decomposed into random components occurring at frequencies in the $[0, \pi]$ interval. The autocovariances γ_j, $j \in (0, 1, \ldots, , \infty)$ of a covariance stationary process specify its variance and dependence structure.

In the frequency domain, the spectral density describes the importance of the random components that occur at frequency ω relative to the components at other frequencies. The spectral density can be written as a weighted average of the autocorrelations of y_t, normalized by $\gamma_0 = Var(y)$. Multiplying the spectral density by γ_0 defines the power spectrum of y_t.

Interpreting cycles in the frequency domain

To understand the stochastic-cycle model, consider that any stationary process may be decomposed into random components occurring at frequencies in the $[0, \pi]$ interval. The autocovariances γ_j, $j \in (0, 1, \ldots, , \infty)$ of a covariance stationary process specify its variance and dependence structure.

In the frequency domain, the spectral density describes the importance of the random components that occur at frequency ω relative to the components at other frequencies. The spectral density can be written as a weighted average of the autocorrelations of y_t, normalized by $\gamma_0 = Var(y)$. Multiplying the spectral density by γ_0 defines the power spectrum of y_t.

In an *i*.*i*.*d*. process, the components of all frequencies are equally represented, and the spectral density is a flat line over $(0, \pi)$. This represents ‘white noise’. High-frequency components will raise the spectral density nearing π, while low-frequency components will raise the spectral density nearing 0.

For instance, $y_t = \phi y_{t-1} + \epsilon_t$ will have a spectral density (SD) dominated by low-frequency components as $\phi \rightarrow 1$, whereas high-frequency components will be most important as $\phi \rightarrow -1$. Given the simple structure of this process, the SD with $\phi > 0$ will be monotonically declining, and the SD with $\phi < 0$ monotonically increasing, over the $(0, \pi)$ interval.

In an *i.i.d.* process, the components of all frequencies are equally represented, and the spectral density is a flat line over $(0, \pi)$. This represents 'white noise'. High-frequency components will raise the spectral density nearing π, while low-frequency components will raise the spectral density nearing 0.

For instance, $y_t = \phi y_{t-1} + \epsilon_t$ will have a spectral density (SD) dominated by low-frequency components as $\phi \rightarrow 1$, whereas high-frequency components will be most important as $\phi \rightarrow -1$. Given the simple structure of this process, the SD with $\phi > 0$ will be monotonically declining, and the SD with $\phi < 0$ monotonically increasing, over the $(0, \pi)$ interval.

Autoregressive moving-average (ARMA) models parameterize the autocorrelation in a time series by allowing today's value to be a weighted average of past values and a weighted average of past *i.i.d.* shocks. This allows us to rewrite the ARMA model as a weighted average of past i.i.d. shocks to trace how a shock feeds through the system, as in the context of the impulse response function of a VAR.

In contrast, the parameters of the stochastic-cycle parameterization of autocorrelation in a time series directly provide information about the underlying spectral density. The parameter ω_0 is the central frequency around which the random components are clustered.

Autoregressive moving-average (ARMA) models parameterize the autocorrelation in a time series by allowing today's value to be a weighted average of past values and a weighted average of past *i.i.d.* shocks. This allows us to rewrite the ARMA model as a weighted average of past i.i.d. shocks to trace how a shock feeds through the system, as in the context of the impulse response function of a VAR.

In contrast, the parameters of the stochastic-cycle parameterization of autocorrelation in a time series directly provide information about the underlying spectral density. The parameter ω_0 is the central frequency around which the random components are clustered.

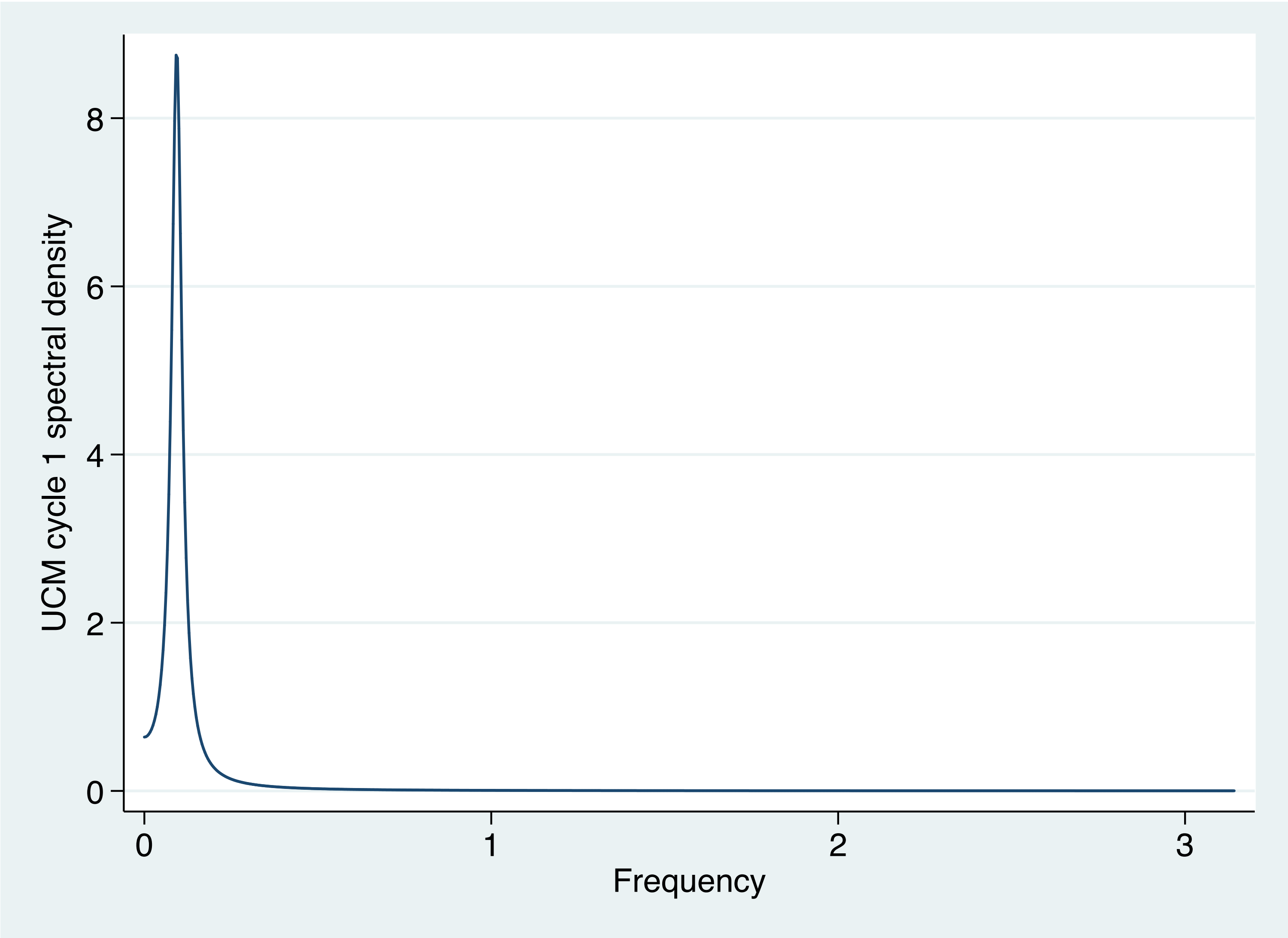
If ω_0 is small, then the model is centered around low-frequency components. If ω_0 is close to π, then the model is centered around high-frequency components.

The parameter ρ is the damping factor that indicates how tightly clustered the random components are around the central frequency ω_0. If ρ is close to zero, there is no clustering of the random components. If ρ is close to one, the random components are tightly clustered around the central frequency ω_0.

If ω_0 is small, then the model is centered around low-frequency components. If ω_0 is close to π, then the model is centered around high-frequency components.

The parameter ρ is the damping factor that indicates how tightly clustered the random components are around the central frequency ω_0. If ρ is close to zero, there is no clustering of the random components. If ρ is close to one, the random components are tightly clustered around the central frequency ω_0.

Returning to our example, where we estimated a period of 5.6 years with a very large damping factor, we may view the spectral density implied by this model.



We may now extend the previous stochastic-cycle model to investigate the possible presence of a high-frequency component in addition to the low-frequency component in the US unemployment rate series. We specify suboptions to `cycle()` to assist in identifying the two components, which can be problematic. The frequency of 0.09 is that estimated in the prior example. A frequency of 2.9, close to π, will be the high-frequency component.

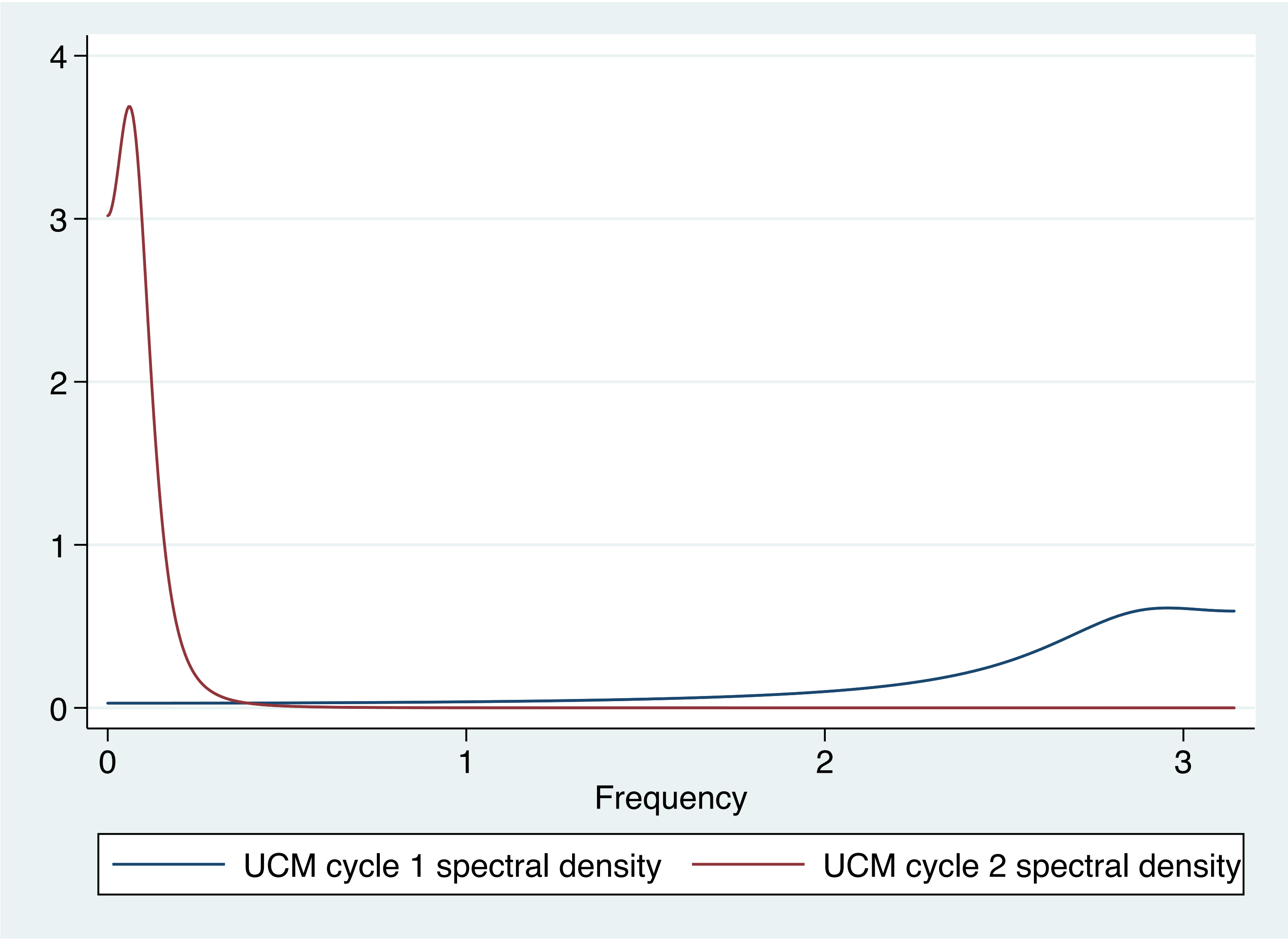
```
. ucm unrate, cycle(1, freq(2.9)) cycle(2, freq(0.09)) nolog vsquish
Unobserved-components model
Components: random walk, 2 cycles of order 1 2
Sample: 1948m1 - 2011m1                         Number of obs    =        757
                                                Wald chi2(4)     =    7681.33
Log likelihood =  146.28326                     Prob > chi2      =     0.0000
```

unrate	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
cycle1						
frequency	2.882382	.0668017	43.15	0.000	2.751453	3.013311
damping	.7004295	.1251571	5.60	0.000	.4551261	.9457329
cycle2						
frequency	.0667929	.0206849	3.23	0.001	.0262513	.1073345
damping	.9074708	.0142273	63.78	0.000	.8795858	.9353559
Variance						
level	.0207704	.0039669	5.24	0.000	.0129953	.0285454
cycle1	.0027886	.0014363	1.94	0.026	0	.0056037
cycle2	.002714	.001028	2.64	0.004	.0006991	.0047289

```
Note: Model is not stationary.
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

The output provides some support for the existence of a second, high-frequency cycle. The high-frequency components are centered around 2.88, whereas the low-frequency components are centered around 0.067. That the estimated damping factor is 0.70 for the high-frequency cycle whereas the estimated damping factor for the low-frequency cycle is 0.91 indicates that the high-frequency components are more diffusely distributed around 2.88 than the low-frequency components are around 0.067.

The distinct spectral densities support the conclusion of two cycles in the data.



The local-level model

We now consider the weekly unemployment claims series (additions to the unemployment rolls). This series appears to be a random walk plus noise, or as often termed the local-level model.

$$\begin{aligned} y_t &= \mu_t + \epsilon_t \\ \mu_t &= \mu_{t-1} + \eta_t \end{aligned}$$

where $\epsilon_t \sim N(0, \sigma^2_\epsilon)$ and $N(0, \sigma^2_\eta)$ are mutually independent.

```
. webuse icsa1, clear
. ucm icsa, model(llevel) nolog vsquish
Unobserved-components model
Components: local level
Sample: 07jan1967 - 19feb2011                    Number of obs   =      2303
Log likelihood = -9893.2469
------------------------------------------------------------------------------
             |                 OIM
        icsa |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Variance     |
  level      |    116.558   8.806587    13.24   0.000     99.29745    133.8186
  icsa       |   124.2715   7.615506    16.32   0.000     109.3454    139.1976
------------------------------------------------------------------------------
Note: Model is not stationary.
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
Note: Time units are in 7 days.
```

The estimation results indicate that both of the stochastic components are statistically significant.

```
. webuse icsa1, clear
. ucm icsa, model(llevel) nolog vsquish
Unobserved-components model
Components: local level
Sample: 07jan1967 - 19feb2011                   Number of obs    =       2303
Log likelihood = -9893.2469
```

icsa	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
Variance						
level	116.558	8.806587	13.24	0.000	99.29745	133.8186
icsa	124.2715	7.615506	16.32	0.000	109.3454	139.1976

```
Note: Model is not stationary.
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
Note: Time units are in 7 days.
```

The estimation results indicate that both of the stochastic components are statistically significant.

We might suspect that there is some serial correlation in the idiosyncratic shock. Alternatively, we could include a cyclical component to model the stationary time-dependence in the series. In the example below, we add a stochastic-cycle model for the stationary cyclical process, but we drop the idiosyncratic term and use a random-walk model instead of the local-level model. We change the model because it is difficult to estimate the variance of the idiosyncratic term along with the parameters of a stationary cyclical component.

```
. ucm icsa, model(rwalk) cycle(1) nolog vsquish
Unobserved-components model
Components: random walk, order 1 cycle
Sample: 07jan1967 - 19feb2011                   Number of obs    =       2303
                                                Wald chi2(2)     =      23.04
Log likelihood = -9881.4441                     Prob > chi2      =     0.0000

------------------------------------------------------------------------------
             |                 OIM
        icsa |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   frequency |   1.469633   .3855657     3.81   0.000     .7139385    2.225328
     damping |   .1644576   .0349537     4.71   0.000     .0959495    .2329656
-------------+----------------------------------------------------------------
Variance     |
  level      |   97.90982   8.320047    11.77   0.000     81.60282    114.2168
  cycle1     |   149.7323   9.980798    15.00   0.000     130.1703    169.2943
------------------------------------------------------------------------------
Note: Model is not stationary.
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
Note: Time units are in 7 days.
```

```
. estat period
```

cycle1	Coef.	Std. Err.	[95% Conf.	Interval]
period	4.275342	1.121657	2.076934	6.47375
frequency	1.469633	.3855657	.7139385	2.225328
damping	.1644576	.0349537	.0959495	.2329656

```
Note: Time units are in 7 days.
. psdensity sdensity3 omega3
. line sdensity3 omega3, ylab(,angle(0))
```

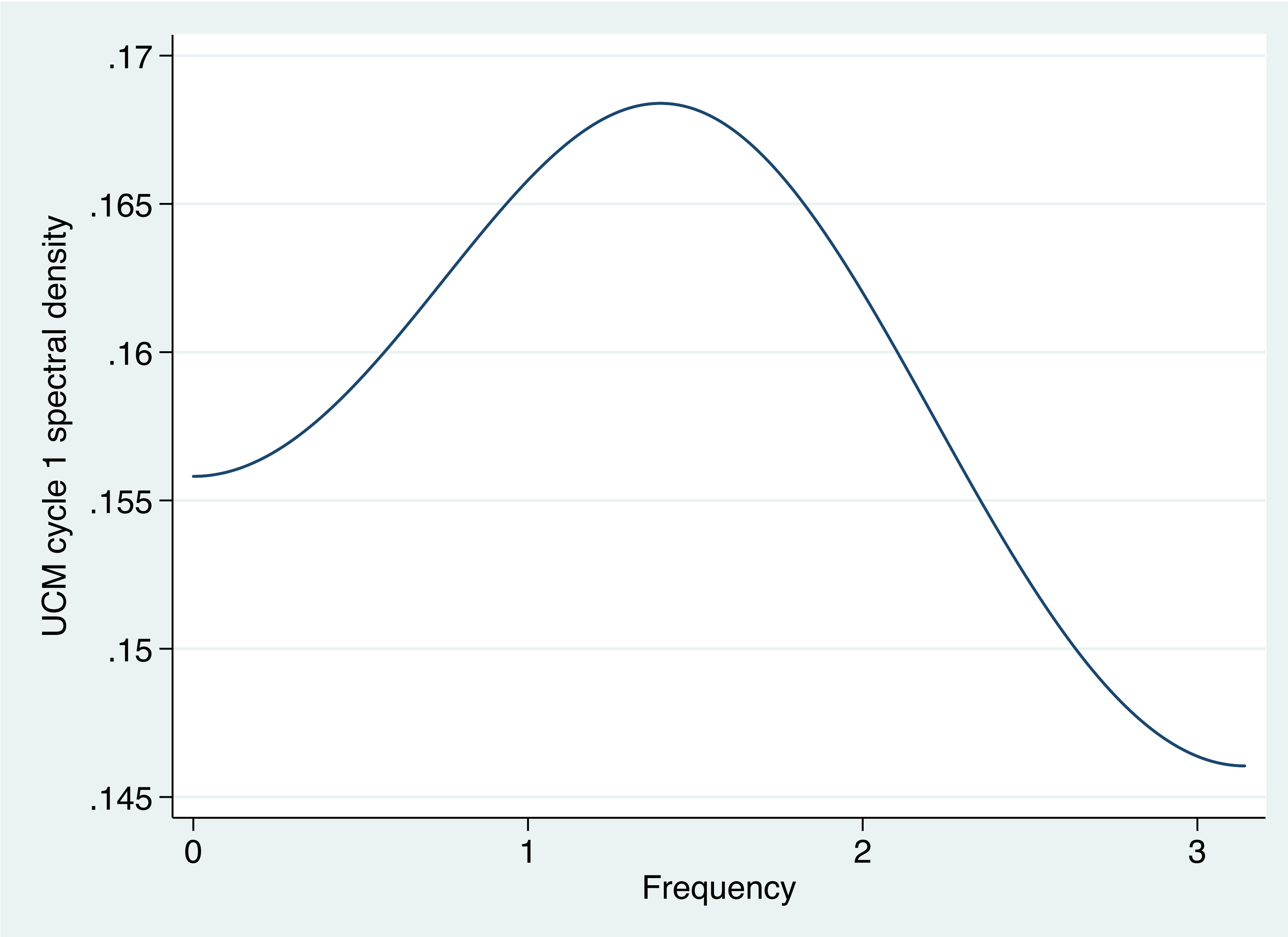
Although the output indicates that the model fits well, the small estimate of the damping parameter indicates that the random components are widely distributed around the central frequency.

```
. estat period
```

cycle1	Coef.	Std. Err.	[95% Conf.	Interval]
period	4.275342	1.121657	2.076934	6.47375
frequency	1.469633	.3855657	.7139385	2.225328
damping	.1644576	.0349537	.0959495	.2329656

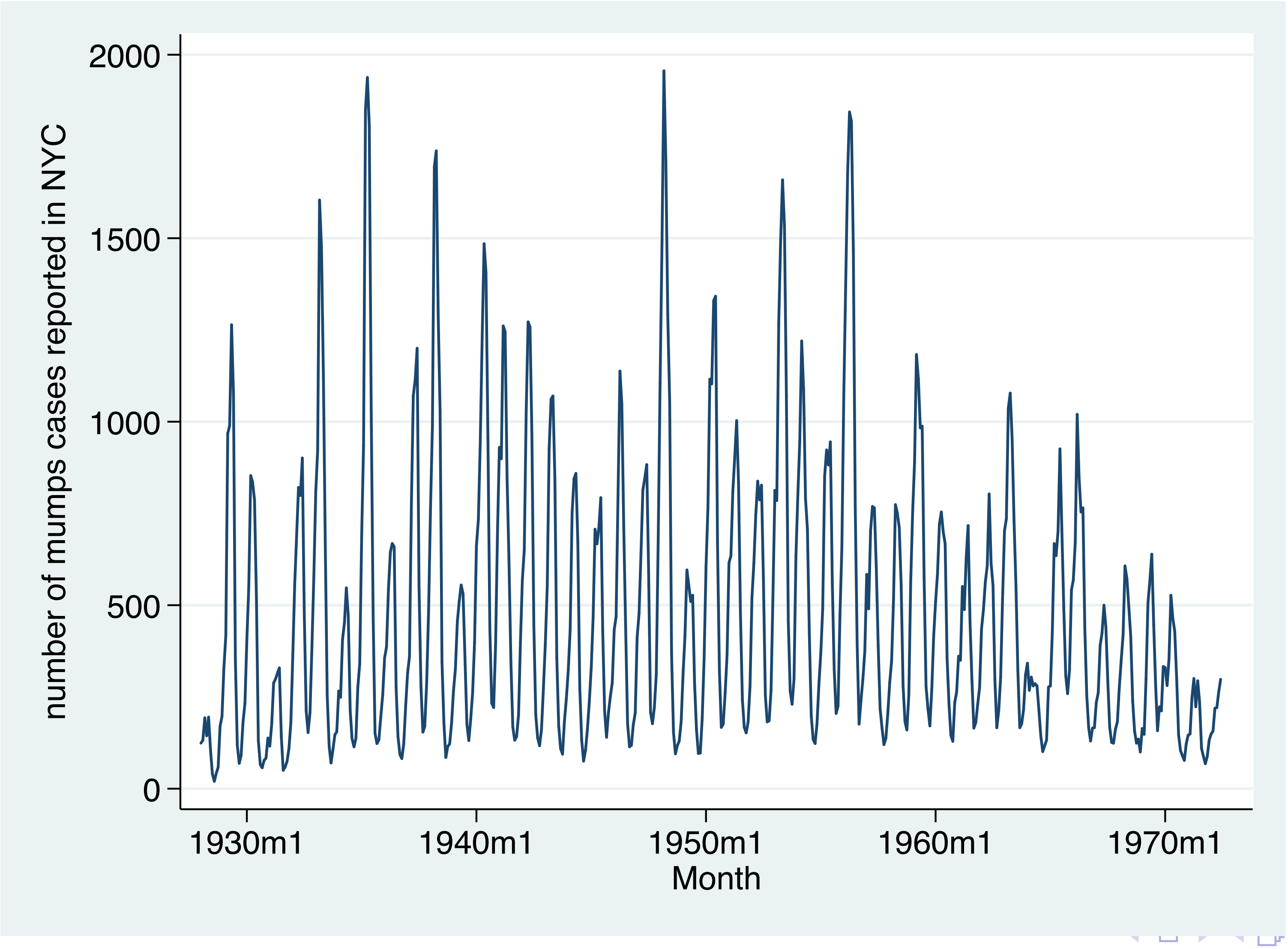
```
Note: Time units are in 7 days.
. psdensity sdensity3 omega3
. line sdensity3 omega3, ylab(,angle(0))
```

Although the output indicates that the model fits well, the small estimate of the damping parameter indicates that the random components are widely distributed around the central frequency.


.17
.165
.16
.155
.15
.145
UCM cycle 1 spectral density
0
1
2
3
Frequency

Modeling seasonality

Consider a series with a seasonal effect, such as this monthly record of new cases of mumps in New York City, 1928–1972.



This could be modeled as a stochastic-seasonal model, allowing for a random walk in the series and a stationary cyclical component.

```
. ucm mumps, seasonal(12) cycle(1) nolog vsquish
Unobserved-components model
Components: random walk, seasonal(12), order 1 cycle
Sample: 1928m1 - 1972m6                          Number of obs    =        534
                                                 Wald chi2(2)     =    2141.69
Log likelihood = -3248.7138                      Prob > chi2      =     0.0000

------------------------------------------------------------------------------
             |                 OIM
       mumps |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   frequency |   .3863607   .0282037    13.70   0.000     .3310824    .4416389
     damping |   .8405622   .0197933    42.47   0.000     .8017681    .8793563
-------------+----------------------------------------------------------------
Variance     |
  level      |   221.2131   140.5179     1.57   0.058            0    496.6231
  seasonal   |   4.151639   4.383442     0.95   0.172            0    12.74303
  cycle1     |   12228.17   813.8394    15.03   0.000     10633.08    13823.27
------------------------------------------------------------------------------
Note: Model is not stationary.
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

These results suggest that the seasonal variation may not be important, and the trend variation (captured by the level variance) is borderline. If the variance of the stochastic seasonal is zero, the seasonal component becomes deterministic, and can be modeled with seasonal dummies. We drop the trend variance, retaining only the cyclical component.

```
. ucm mumps ibn.month, model(none) cycle(1) nolog vsquish
Unobserved-components model
Components: order 1 cycle
Sample: 1928m1 - 1972m6                          Number of obs    =        534
                                                 Wald chi2(14)    =    3404.29
Log likelihood = -3283.0284                      Prob > chi2      =     0.0000
```

mumps	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
cycle1						
frequency	.3272754	.0262922	12.45	0.000	.2757436	.3788071
damping	.844874	.0184994	45.67	0.000	.8086157	.8811322
mumps						
month						
1	480.5095	32.67128	14.71	0.000	416.475	544.544
2	561.9174	32.66999	17.20	0.000	497.8854	625.9494
3	832.8666	32.67696	25.49	0.000	768.8209	896.9122
4	894.0747	32.64568	27.39	0.000	830.0904	958.0591
5	869.6568	32.56282	26.71	0.000	805.8348	933.4787
6	770.1562	32.48587	23.71	0.000	706.4851	833.8274
7	433.839	32.50165	13.35	0.000	370.1369	497.541
8	218.2394	32.56712	6.70	0.000	154.409	282.0698
9	140.686	32.64138	4.31	0.000	76.7101	204.662
10	148.5876	32.69067	4.55	0.000	84.51508	212.6601
11	215.0958	32.70311	6.58	0.000	150.9989	279.1927
12	330.2232	32.68906	10.10	0.000	266.1538	394.2926

...

```
 ...
Variance     |
  cycle1     |    13031.53   798.2719    16.32   0.000     11466.95    14596.11
--------------------------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
. estat period
--------------------------------------------------------------
      cycle1 |      Coef.   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
      period |   19.19847    1.54234      16.17554     22.2214
   frequency |   .3272754   .0262922      .2757436    .3788071
     damping |    .844874   .0184994      .8086157    .8811322
--------------------------------------------------------------
Note: Cycle time unit is monthly.
```

The cyclical variance is an important element. Analysis of its periodicity shows a 19-month cycle, suggesting that new mumps cases peak about every 1.5 years.

```
 ...
Variance     |
  cycle1     |    13031.53   798.2719    16.32   0.000     11466.95    14596.11
------------------------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
. estat period
--------------------------------------------------------------
      cycle1 |      Coef.   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
      period |   19.19847    1.54234      16.17554     22.2214
   frequency |   .3272754   .0262922      .2757436    .3788071
     damping |    .844874   .0184994      .8086157    .8811322
--------------------------------------------------------------
Note: Cycle time unit is monthly.
```

The cyclical variance is an important element. Analysis of its periodicity shows a 19-month cycle, suggesting that new mumps cases peak about every 1.5 years.

Dynamic factor models

Dynamic factor models (DFM) are flexible models for multivariate time series in which unobserved factors have a vector autoregressive structure, exogenous covariates are permitted in both the equations for the latent factors and the equations for observable dependent variables, and the disturbances in the equations for the dependent variables may be autocorrelated.

A DFM contains k endogenous variables, expressed as linear functions of $n_f < k$ unobserved factors and exogenous covariates. Constraints must be imposed for identification of the parameters.

Dynamic factor models

Dynamic factor models (DFM) are flexible models for multivariate time series in which unobserved factors have a vector autoregressive structure, exogenous covariates are permitted in both the equations for the latent factors and the equations for observable dependent variables, and the disturbances in the equations for the dependent variables may be autocorrelated.

A DFM contains k endogenous variables, expressed as linear functions of $n_f < k$ unobserved factors and exogenous covariates. Constraints must be imposed for identification of the parameters.

A DFM can be written as

$$
\begin{aligned}
y_t &= Pf_t + Qx_t + u_t \\
f_t &= Rw_t + A_1 f_{t-1} + \ldots \cdots + A_{t-p} f_{t-p} + \nu_t \\
u_t &= c_1 u_{t-1} + \ldots \cdots + C_{t-q} u_{t-q} + \epsilon_t
\end{aligned}
$$

where y_t, u_t and ϵ_t are $k \times 1$, f_t and ν_t are $n_f \times 1$, x is $n_x \times 1$, and w_t is $n_w \times 1$. In this specification, there are p lags on the factors and q lags on the u error processes.

Several variations of the model may be specified:

	Model	n_f	p	q
Static factors	SF	>0	0	0
Static factors with vector AR errors	SFAR	>0	0	>0
Dynamic factors	DF	>0	>0	0
Dynamic factors with vector AR errors	DFAR	>0	>0	>0
Seemingly unrelated regression	SUR	0	0	0
VAR with vector AR errors	VAR	0	0	>0

The last two are not DFM specifications, but may be estimated to allow for constraints on their error VCE, which cannot be imposed in the standard `sureg` or `var` framework.

Several variations of the model may be specified:

	Model	n_f	p	q
Static factors	SF	>0	0	0
Static factors with vector AR errors	SFAR	>0	0	>0
Dynamic factors	DF	>0	>0	0
Dynamic factors with vector AR errors	DFAR	>0	>0	>0
Seemingly unrelated regression	SUR	0	0	0
VAR with vector AR errors	VAR	0	0	>0

The last two are not DFM specifications, but may be estimated to allow for constraints on their error VCE, which cannot be imposed in the standard `sureg` or `var` framework.

These models are estimated by placing them in state-space form. We have already seen an example of a DFM, in which a single unobserved factor, modeled as an AR(2) process, was related to four observable macro variables. In that example, we used `space` to specify and estimate the model. We could have generated the same results using Stata's `dfactor` command, a bit more parsimoniously:

```
dfactor (D.(ipman income hours unemp) = , nocons) (f =, ar(1/2))
```

We could extend this example to allow for the errors in the observables to be autocorrelated.

These models are estimated by placing them in state-space form. We have already seen an example of a DFM, in which a single unobserved factor, modeled as an AR(2) process, was related to four observable macro variables. In that example, we used `space` to specify and estimate the model. We could have generated the same results using Stata's `dfactor` command, a bit more parsimoniously:

```
dfactor (D.(ipman income hours unemp) = , nocons) (f =, ar(1/2))
```

We could extend this example to allow for the errors in the observables to be autocorrelated.

```
. webuse dfex, clear
(St. Louis Fed (FRED) macro data)

. dfactor (D.(ipman income hours unemp)=, nocons ar(1)) (f=, ar(1/2)), nolog vs
> quish

Dynamic-factor model

Sample: 1972m2 - 2008m11                        Number of obs     =        442
                                                Wald chi2(10)     =     990.91
Log likelihood = -610.28846                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
             |                 OIM
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
f            |
           f |
         L1. |   .4058457   .0906183     4.48   0.000     .2282371    .5834544
         L2. |   .3663499   .0849584     4.31   0.000     .1998344    .5328654
-------------+----------------------------------------------------------------
De.ipman     |
     e.ipman |
         LD. |  -.2772149    .068808    -4.03   0.000    -.4120761   -.1423538
-------------+----------------------------------------------------------------
De.income    |
    e.income |
         LD. |  -.2213824   .0470578    -4.70   0.000    -.3136141   -.1291508
-------------+----------------------------------------------------------------
De.hours     |
     e.hours |
         LD. |  -.3969317   .0504256    -7.87   0.000     -.495764   -.2980994
-------------+----------------------------------------------------------------
...
```

...

De.unemp						
e.unemp						
LD.	-.1736835	.0532071	-3.26	0.001	-.2779675	-.0693995
D.ipman						
f	.3214972	.027982	11.49	0.000	.2666535	.3763408
D.income						
f	.0760412	.0173844	4.37	0.000	.0419684	.110114
D.hours						
f	.1933165	.0172969	11.18	0.000	.1594151	.2272179
D.unemp						
f	-.0711994	.0066553	-10.70	0.000	-.0842435	-.0581553
Variance						
De.ipman	.1387909	.0154558	8.98	0.000	.1084981	.1690837
De.income	.2636239	.0179043	14.72	0.000	.2285322	.2987157
De.hours	.0822919	.0071096	11.57	0.000	.0683574	.0962265
De.unemp	.0218056	.0016658	13.09	0.000	.0185407	.0250704

Note: Tests of variances against zero are one sided, and the two-sided confidence intervals are truncated at zero.

The sizable negative coefficients on each of the `De.` terms imply that incorporating AR(1) errors improves the earlier model. The default for the vector AR structure (the *A* matrices) is a diagonal VCE, with no cross-equation autocorrelations. This can be relaxed by the `arstructure()` option. Allowing for a general matrix, we now estimate a full set of cross-equation autocorrelations.

```
. dfactor (D.(ipman income hours unemp)=, nocons ar(1) arstructure(gen)) ///
> (f=, ar(1/2)), nolog vsquish
Dynamic-factor model
Sample: 1972m2 - 2008m11                          Number of obs    =        442
                                                  Wald chi2(22)    =    1886.33
Log likelihood = -577.02661                       Prob > chi2      =     0.0000
------------------------------------------------------------------------------
             |                 OIM
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
f            |
           f |
         L1. |  -.5931147   .0704447    -8.42   0.000    -.7311838   -.4550455
         L2. |  -.3082691   .0622398    -4.95   0.000    -.4302569   -.1862813
-------------+----------------------------------------------------------------
De.ipman     |
     e.ipman |
         LD. |   .0188223   .0646137     0.29   0.771    -.1078182    .1454628
    e.income |
         LD. |   .2121594   .0483115     4.39   0.000     .1174707    .3068482
     e.hours |
         LD. |    1.02509    .161006     6.37   0.000     .7095238    1.340656
     e.unemp |
         LD. |   -.59724      .16283    -3.67   0.000    -.916381    -.278099
------------------------------------------------------------------------------
...
```

```
 De.income   |
     e.ipman |
         LD. |   .0775566   .0544958     1.42   0.155    -.0292532    .1843664
    e.income |
         LD. |  -.1927469   .0473582    -4.07   0.000    -.2855673   -.0999266
     e.hours |
         LD. |   .2332803   .1295888     1.80   0.072    -.0207091    .4872696
     e.unemp |
         LD. |   .0349881   .1558053     0.22   0.822    -.2703848    .3403609
-------------+----------------------------------------------------------------
De.hours     |
     e.ipman |
         LD. |    .175513    .041344     4.25   0.000     .0944801    .2565458
    e.income |
         LD. |   .0662514   .0301777     2.20   0.028     .0071041    .1253986
     e.hours |
         LD. |   .3987403   .1063789     3.75   0.000     .1902415    .6072391
     e.unemp |
         LD. |  -.4004179   .1054703    -3.80   0.000     -.607136   -.1936998
-------------+----------------------------------------------------------------
De.unemp     |
     e.ipman |
         LD. |  -.0531289   .0194429    -2.73   0.006    -.0912363   -.0150215
    e.income |
         LD. |   -.018593   .0153895    -1.21   0.227    -.0487558    .0115698
     e.hours |
         LD. |  -.2859971   .0510751    -5.60   0.000    -.3861024   -.1858918
     e.unemp |
         LD. |  -.0827445   .0519692    -1.59   0.111    -.1846022    .0191132
...
```

```
...
D.ipman      |
           f |   .1889032   .0228953     8.25   0.000     .1440293    .2337772
-------------+----------------------------------------------------------------
D.income     |
           f |   .0687882   .0264256     2.60   0.009     .0169949    .1205814
-------------+----------------------------------------------------------------
D.hours      |
           f |   .2729581   .0177138    15.41   0.000     .2382396    .3076765
-------------+----------------------------------------------------------------
D.unemp      |
           f |  -.0190063   .0075799    -2.51   0.012    -.0338627   -.0041499
-------------+----------------------------------------------------------------
Variance     |
  De.ipman   |   .1756275   .0144128    12.19   0.000     .1473789    .2038762
  De.income  |   .2642305   .0178817    14.78   0.000      .229183     .299278
  De.hours   |    .022353   .0065214     3.43   0.000     .0095713    .0351346
  De.unemp   |    .023182   .0016716    13.87   0.000     .0199058    .0264582
------------------------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

We may also estimate a static factor model, in which the factors do not have an autoregressive structure. We illustrate with a dataset of monthly unemployment rates across the four US Census regions.

```
. webuse urate, clear
(Monthly unemployment rates in US Census regions)
. dfactor (D.(west south ne midwest) = , noconstant) (z = ), nolog vsquish
Dynamic-factor model
Sample: 1990m2 - 2008m12                    Number of obs    =        227
                                            Wald chi2(4)     =     342.56
Log likelihood =   873.0755                 Prob > chi2      =     0.0000
--------------------------------------------------------------------------
             |                 OIM
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+------------------------------------------------------------
D.west       |
           z |   .0978324   .0065644    14.90   0.000     .0849664    .1106983
-------------+------------------------------------------------------------
D.south      |
           z |   .0859494   .0061762    13.92   0.000     .0738442    .0980546
-------------+------------------------------------------------------------
D.ne         |
           z |   .0918607   .0072814    12.62   0.000     .0775893     .106132
-------------+------------------------------------------------------------
D.midwest    |
           z |   .0861102   .0074652    11.53   0.000     .0714787    .1007417
-------------+------------------------------------------------------------
Variance     |
  De.west    |   .0036887   .0005834     6.32   0.000     .0025453    .0048322
  De.south   |   .0038902   .0005228     7.44   0.000     .0028656    .0049149
  De.ne      |   .0064074   .0007558     8.48   0.000     .0049261    .0078887
  De.midwest |   .0074749   .0008271     9.04   0.000     .0058538     .009096
--------------------------------------------------------------------------
      confidence intervals are truncated at zero.
```

We might want to test whether changes in the latent factor have the same effect on all regional unemployment rates.

```
. test [D.west]z = [D.south]z = [D.ne]z = [D.midwest]z
( 1)   [D.west]z - [D.south]z = 0
( 2)   [D.west]z - [D.ne]z = 0
( 3)   [D.west]z - [D.midwest]z = 0
           chi2(  3) =    3.58
         Prob > chi2 =    0.3109
```

The hypothesis of equality cannot be rejected. We may thus impose those constraints and allow for dynamics in the variables by allowing their errors to follow an AR(1) process.

```
. constraint 2 [D.west]z = [D.south]z
. constraint 3 [D.west]z = [D.ne]z
. constraint 4 [D.west]z = [D.midwest]z
. dfactor (D.(west south ne midwest) = , noconstant ar(1)) (z = ), ///
> constraints(2/4) nolog vsquish
Dynamic-factor model
Sample: 1990m2 - 2008m12                          Number of obs   =        227
                                                  Wald chi2(5)    =     363.34
Log likelihood =  880.97488                       Prob > chi2     =     0.0000
 ( 1)  [D.west]z - [D.south]z = 0
 ( 2)  [D.west]z - [D.ne]z = 0
 ( 3)  [D.west]z - [D.midwest]z = 0
```

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
De.west						
e.west						
LD.	.1297198	.0992663	1.31	0.191	-.0648386	.3242781
De.south						
e.south						
LD.	-.2829014	.0909205	-3.11	0.002	-.4611023	-.1047004
De.ne						
e.ne						
LD.	.2866958	.0847851	3.38	0.001	.12052	.4528715

...

```
  ...
De.midwest   |
   e.midwest |
         LD. |   .0049427   .0782188     0.06   0.950    -.1483634    .1582488
-------------+----------------------------------------------------------------
D.west       |
           z |   .0904724   .0049326    18.34   0.000     .0808047    .1001401
-------------+----------------------------------------------------------------
D.south      |
           z |   .0904724   .0049326    18.34   0.000     .0808047    .1001401
-------------+----------------------------------------------------------------
D.ne         |
           z |   .0904724   .0049326    18.34   0.000     .0808047    .1001401
-------------+----------------------------------------------------------------
D.midwest    |
           z |   .0904724   .0049326    18.34   0.000     .0808047    .1001401
-------------+----------------------------------------------------------------
Variance     |
  De.west    |   .0038959   .0005111     7.62   0.000     .0028941    .0048977
  De.south   |   .0035518   .0005097     6.97   0.000     .0025528    .0045507
  De.ne      |   .0058173   .0006983     8.33   0.000     .0044488    .0071859
  De.midwest |   .0075444   .0008268     9.12   0.000     .0059239     .009165
------------------------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

The AR(1) parameters are not very precisely estimated, with two of the four not significantly different from zero. A dynamic factor specification might be more appropriate. We drop the AR(1) structure on the observed variables' errors and add two lags to the factor equation ($p = 2$).

```
. dfactor (D.(west south ne midwest) = , noconstant) (z =, ar(1/2)), ///
> constraints(2/4) nolog vsquish
Dynamic-factor model
Sample: 1990m2 - 2008m12                        Number of obs   =       227
                                                Wald chi2(3)    =   1077.41
Log likelihood =  959.26145                     Prob > chi2     =    0.0000
 ( 1)   [D.west]z - [D.south]z = 0
 ( 2)   [D.west]z - [D.ne]z = 0
 ( 3)   [D.west]z - [D.midwest]z = 0
```

	Coef.	OIM Std. Err.	z	P>|z|	[95% Conf.	Interval]
z						
z						
L1.	.2280112	.0577456	3.95	0.000	.1148319	.3411904
L2.	.7332268	.0602479	12.17	0.000	.615143	.8513105
D.west						
z	.0513222	.0038618	13.29	0.000	.0437532	.0588913
D.south						
z	.0513222	.0038618	13.29	0.000	.0437532	.0588913
D.ne						
z	.0513222	.0038618	13.29	0.000	.0437532	.0588913
D.midwest						
z	.0513222	.0038618	13.29	0.000	.0437532	.0588913

```
  ...
------------------------------------------------------------------------------
Variance     |
  De.west    |   .0033756     .00043     7.85   0.000     .0025328    .0042183
  De.south   |   .0038912   .0004611     8.44   0.000     .0029874     .004795
  De.ne      |   .0061826   .0006749     9.16   0.000     .0048599    .0075053
  De.midwest |   .0084143   .0008768     9.60   0.000     .0066958    .0101328
------------------------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
    confidence intervals are truncated at zero.
```

This specification is more appealing, with the coefficients on the latent factor summing to nearly unity. We can revisit the issue of using a single coefficient by reestimating without constraints and performing a likelihood ratio test.

```
 ...
-------------+----------------------------------------------------------------
Variance     |
     De.west |   .0033756     .00043     7.85   0.000      .0025328    .0042183
    De.south |   .0038912   .0004611     8.44   0.000      .0029874     .004795
       De.ne |   .0061826   .0006749     9.16   0.000      .0048599    .0075053
  De.midwest |   .0084143   .0008768     9.60   0.000      .0066958    .0101328
------------------------------------------------------------------------------
Note: Tests of variances against zero are one sided, and the two-sided
      confidence intervals are truncated at zero.
```

This specification is more appealing, with the coefficients on the latent factor summing to nearly unity. We can revisit the issue of using a single coefficient by reestimating without constraints and performing a likelihood ratio test.

```
. qui dfactor (D.(west south ne midwest) = , nocons) (z =, ar(1/2)), nolog vsqu
> ish
. lrtest singlecoef .
Likelihood-ratio test                                 LR chi2(3)  =     11.74
(Assumption: singlecoef nested in .)                  Prob > chi2 =    0.0083
```

The test rejects its null, implying that the model allowing for region-specific coefficients is preferred. The `predict` command can be used to compute the estimated factor, which we can graph versus the NBER recession dates.

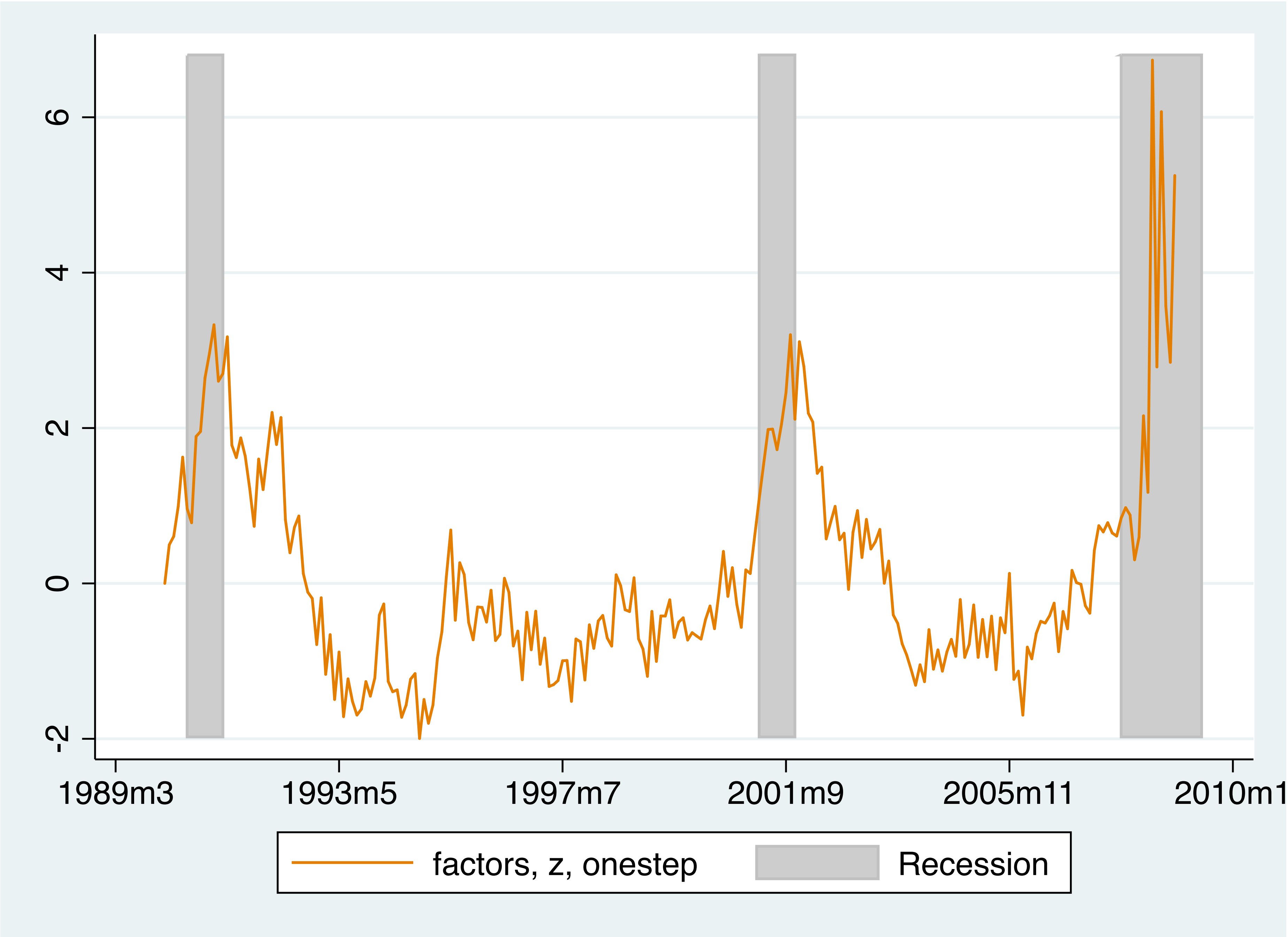
```
. qui dfactor (D.(west south ne midwest) = , nocons) (z =, ar(1/2)), nolog vsqu
> ish

. lrtest singlecoef .

Likelihood-ratio test                                 LR chi2(3)  =     11.74
(Assumption: singlecoef nested in .)                  Prob > chi2 =    0.0083
```

The test rejects its null, implying that the model allowing for region-specific coefficients is preferred. The `predict` command can be used to compute the estimated factor, which we can graph versus the NBER recession dates.

```
 . predict fac29 if e(sample), factor
. nbercycles fac29 if e(sample), file(fac29.do) replace
Code to graph NBER recession dates written to fac29.do
. * append your graph command to this file: e.g.
. * tsline timeseriesvar, xlabel(,format(%tm)) legend(order(4 1 "Recession"))
. twoway function y=6.801925840377808,range(366 374) recast(area) color(gs12) b
> ase(-1.975977147817612) || ///
> function y=6.801925840377808,range(494 502) recast(area) color(gs12) base(-1.
> 975977147817612) || ///
> function y=6.801925840377808,range(575 593) recast(area) color(gs12) base(-1.
> 975977147817612) || ///
> tsline fac29 if e(sample), xlabel(,format(%tm)) legend(order(4 1 "Recession")
> )
.
end of do-file
```


6
4
2
0
-2
1989m3
1993m5
1997m7
2001m9
2005m11
2010m1
factors, z, onestep
Recession