Regression Models

- Bivariate data (y,x)
- Multivariate $(y,x_1,\ldots,x_k)$
- Suppose the conditional mean of *y* is a function of x

$$E(y_t \mid x_t) = \alpha + \beta x_t$$

- Then the regression function is the optimal forecast of *y* given *x*

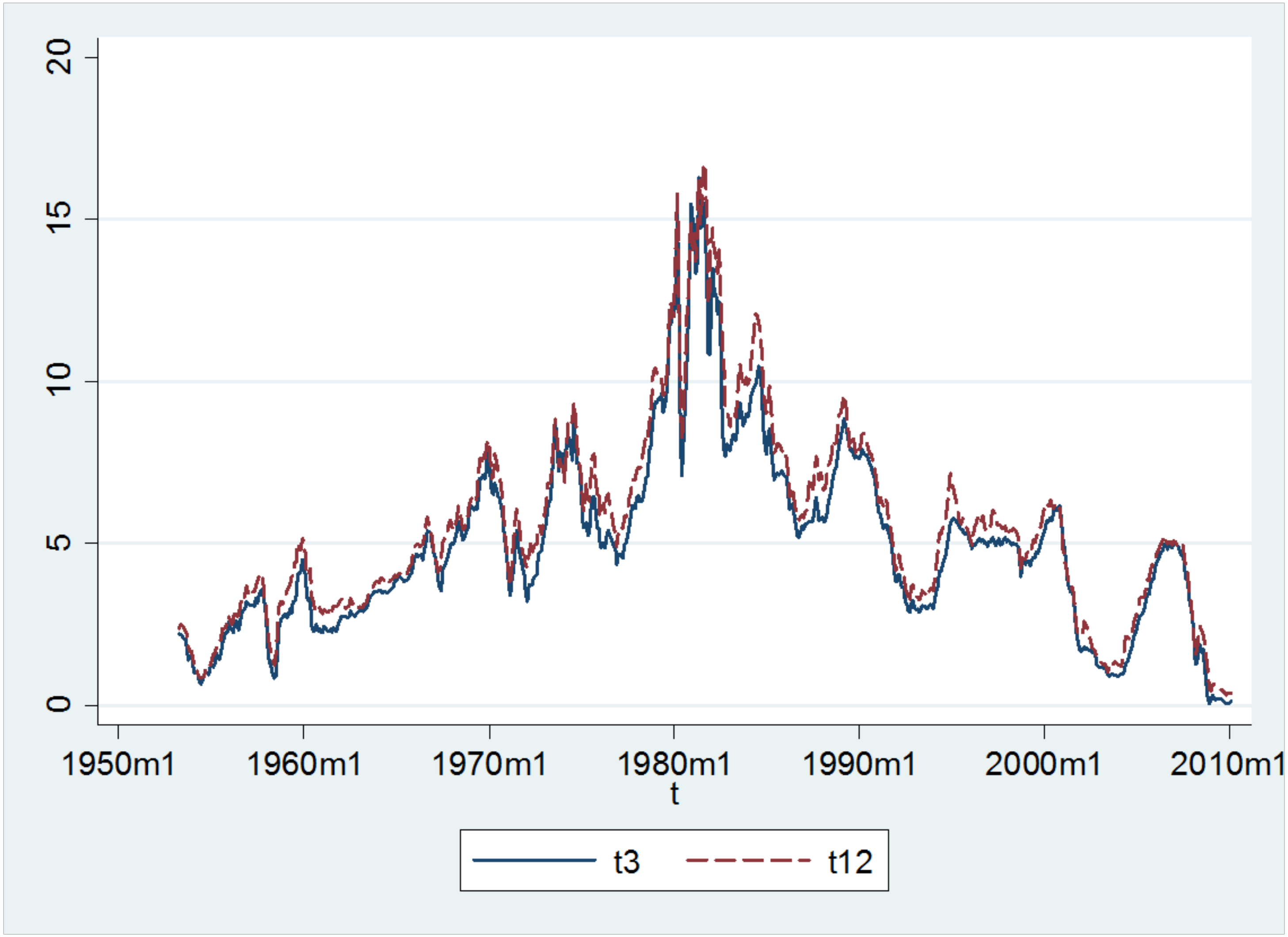
Regression

- Model

$$y_t = \alpha + \beta x_t + e_t$$

- Estimation: Least-Squares
- Example: Interest Rates
  - Monthly
  - Rates on 3-month and 12-month T-Bills

3-month and 12-month T-Bill



Least-Squares, 3-month on 12-month

```
. reg t3 t12,r

Linear regression                                      Number of obs =     683
                                                       F(  1,   681) =15530.05
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9839
                                                       Root MSE      =  .36402
```

t3	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
t12	.9387454	.0075329	124.62	0.000	.923955	.9535359
_cons	-.2139479	.0347796	-6.15	0.000	-.2822362	-.1456597

Forecast

$$y_{n+h} = \alpha + \beta x_{n+h} + e_{n+h}$$

$$\hat{y}_{n+h|n} = \alpha + \beta x_{n+h}$$

- A forecast of y_{n+h} requires x_{n+h}
- This is not typically feasible

Regressor forecast

- Suppose we have a forecast for x
- Then

$$\hat{y}_{n+h|n} = \alpha + \beta \hat{x}_{n+h}$$

- For example, if

$$E\left(x_t \mid \Omega_{t-h}\right) = \gamma + \phi x_{t-h}$$

then

$$\hat{x}_{n+h|n} = \gamma + \phi x_n$$

12-month t-bill on Lagged Value

- Regress x_t on x_{t-12} (12-month ahead forecast)

```
. reg t12 L12.t12,r

Linear regression                               Number of obs =     671
                                                F(  1,   669) =  856.18
                                                Prob > F      =  0.0000
                                                R-squared     =  0.6964
                                                Root MSE      =  1.6663

------------------------------------------------------------------------------
             |               Robust
         t12 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         t12 |
        L12. |   .8450941   .0288817    29.26   0.000     .7883844    .9018038
             |
       _cons |   .8391917   .1427686     5.88   0.000     .5588633     1.11952
------------------------------------------------------------------------------

. predict x
(option xb assumed; fitted values)
(423 missing values generated)
```

3-month on 12-month

- Prediction using regression and fitted value

```
. reg t3 t12,r

Linear regression                                      Number of obs =     683
                                                       F(  1,   681) =15530.05
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9839
                                                       Root MSE      =  .36402

------------------------------------------------------------------------------
             |               Robust
          t3 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         t12 |   .9387454   .0075329   124.62   0.000      .923955    .9535359
       _cons |  -.2139479   .0347796    -6.15   0.000    -.2822362   -.1456597
------------------------------------------------------------------------------

. gen f=_b[_cons]+_b[t12]*x
(423 missing values generated)
```

Interest Rate Forecast

- Estimates

$$\hat{y}_t = -0.21 + 0.94 x_t$$

$$\hat{x}_t = 0.84 + 0.85 x_{t-12}$$

- Current: $x_{2010M2} = 0.35\%$

$$\hat{x}_{2011M2} = 0.84 + 0.85 \cdot 0.35 = 1.14$$

$$\hat{y}_t = -0.21 + 0.94 \cdot 1.14 = 0.86$$

Example

- The AR(1) forecasts the 12-month T-bill next February to rise to 1.14%
- The regression model forecasts the 3-month T-bill next February to be 0.85%
  - Currently 0.11%

Direct Method (preferred)

- Combine

$$E\left(y_t \mid x_t\right) = \alpha + \varphi x_t$$

$$E\left(x_t \mid \Omega_{t-h}\right) = \gamma + \phi x_{t-h}$$

- We obtain

$$\begin{aligned} E\left(y_t \mid \Omega_{t-h}\right) &= \alpha + \varphi E\left(x_t \mid \Omega_{t-h}\right) \\ &= \alpha + \varphi\left(\gamma + \phi x_{t-h}\right) \\ &= \mu + \beta x_{t-h} \end{aligned}$$

Forecast Regression

- h-step-ahead

$$y_t = \mu + \beta x_{t-h} + e_t$$

- Forecast

$$y_{n+h|n} = \mu + \beta x_n$$

3-month on Lagged 12-month

```
. reg t3 L12.t12,r

Linear regression                               Number of obs =     671
                                                F(  1,   669) =  652.61
                                                Prob > F      =  0.0000
                                                R-squared     =  0.6724
                                                Root MSE      =  1.6393
```

t3	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
t12 L12.	.7865454	.030789	25.55	0.000	.7260906	.8470001
_cons	.6103172	.1492087	4.09	0.000	.3173434	.9032909

```
. predict p
(option xb assumed; fitted values)
(423 missing values generated)
```

$$y_{n+h|n} = .61 + .79x_n$$

$$\hat{y}_{n+h|n} = .61 + .79 \cdot 0.35 = 0.89$$

AR(q) Regressors

- Suppose x is an AR(q)

$$y_t = \alpha + \beta x_t + e_t$$

$$x_t = \gamma + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_q x_{t-q} + u_t$$

- Then a one-step forecasting equation for y is

$$y_t = \mu + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q} + e_t$$

- And an h-step is

$$y_t = \mu + \beta_1 x_{t-h} + \beta_2 x_{t-h-1} + \cdots + \beta_q x_{t-h+1-q} + e_t$$

T-Bill example: AR(12) for 12-month

```
. reg t12 L(12/23).t12,r

Linear regression                                      Number of obs =     660
                                                       F( 12,   647) =   86.63
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6997
                                                       Root MSE      =  1.6539
```

t12	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
t12						
L12.	1.344075	.2830537	4.75	0.000	.7882605	1.89989
L13.	-.705637	.4080853	-1.73	0.084	-1.506969	.0956945
L14.	.4306155	.5036102	0.86	0.393	-.5582923	1.419523
L15.	-.2176613	.531493	-0.41	0.682	-1.261321	.8259981
L16.	.2448776	.5273973	0.46	0.643	-.7907394	1.280495
L17.	-.3206738	.4990621	-0.64	0.521	-1.300651	.6593032
L18.	.3327822	.4753309	0.70	0.484	-.6005953	1.26616
L19.	.0635703	.4867974	0.13	0.896	-.8923233	1.019464
L20.	-.3449998	.4765642	-0.72	0.469	-1.280799	.5907994
L21.	-.0145304	.4438058	-0.03	0.974	-.8860041	.8569433
L22.	-.054846	.3902926	-0.14	0.888	-.8212391	.7115471
L23.	.0611679	.2093981	0.29	0.770	-.350014	.4723499
_cons	1.012773	.1487554	6.81	0.000	.7206718	1.304875

Regress 3-month on 12 lags of 12-month

```
. reg t3 L(12/23).t12,r

Linear regression                                      Number of obs =     660
                                                       F( 12,   647) =   71.52
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6819
                                                       Root MSE      =  1.6148
```

t3	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
t12						
L12.	1.264231	.2910707	4.34	0.000	.6926733	1.835788
L13.	-.7090496	.4199322	-1.69	0.092	-1.533644	.115545
L14.	.506987	.5345196	0.95	0.343	-.5426157	1.55659
L15.	-.177356	.5835742	-0.30	0.761	-1.323284	.968572
L16.	.192489	.5757508	0.33	0.738	-.9380768	1.323055
L17.	-.2595061	.5211958	-0.50	0.619	-1.282946	.7639333
L18.	.3006329	.4759674	0.63	0.528	-.6339945	1.23526
L19.	.0498926	.4801796	0.10	0.917	-.893006	.9927912
L20.	-.3123886	.4564153	-0.68	0.494	-1.208623	.5838454
L21.	-.0645272	.4131703	-0.16	0.876	-.8758438	.7467894
L22.	-.072424	.3645739	-0.20	0.843	-.7883149	.6434668
L23.	.0346474	.2031079	0.17	0.865	-.3641829	.4334777
_cons	.8176276	.1495144	5.47	0.000	.5240356	1.11122

Forecast

- Predicted value for 2011M2=1.06
- Predicted value using 3 lags=0.91

Distributed Lags

- This class of models is called *distributed lags*

$$y_t = \mu + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_q x_{t-q} + e_t$$
$$= \mu + B(L) x_{t-1} + e_t$$

- If we interpret the coefficients as the effect of *x* on *y*, we sometimes say
  - β_1 is the immediate impact
  - $\beta_1 + \ldots + \beta_n = B(1)$ is the long-run impact

Regressors and Dynamics

- We have seen AR forecasting models
- And now distributed lag model
- Add both together!

$$A(L)y_t = \mu + B(L)x_{t-1} + e_t$$

- or

$$y_t = \mu + \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p}$$
$$+ \beta_1 x_{t-1} + \cdots + \beta_q x_{t-q} + e_t$$

h-step

- Regress on lags of y and x, h periods back
- Estimate by least squares
- Forecast using estimated coefficients and final values

$$y_t = \mu + \alpha_1 y_{t-h} + \cdots + \alpha_p y_{t-h+1-p} + \beta_1 x_{t-h} + \cdots + \beta_q x_{t-h+1-q} + e_t$$

3-month t-bill forecast

t3	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
t3						
L12.	.2302162	.5728414	0.40	0.688	-.8946763	1.355109
L13.	.314181	.7814376	0.40	0.688	-1.220333	1.848695
L14.	-.2879817	.7375502	-0.39	0.696	-1.736314	1.160351
L15.	-.275658	.7883744	-0.35	0.727	-1.823794	1.272478
L16.	.0095163	.7952381	0.01	0.990	-1.552098	1.571131
L17.	-.0485096	.780273	-0.06	0.950	-1.580737	1.483718
L18.	.543673	.7693474	0.71	0.480	-.9670999	2.054446
L19.	-.5602249	.7119881	-0.79	0.432	-1.958361	.837911
L20.	-.2627815	.6343378	-0.41	0.679	-1.508435	.982872
L21.	-.4041732	.5640965	-0.72	0.474	-1.511893	.703547
L22.	.0995731	.546212	0.18	0.855	-.9730272	1.172173
L23.	-.6719082	.3712314	-1.81	0.071	-1.400898	.0570815
t12						
L12.	1.036348	.4892947	2.12	0.035	.0755169	1.997179
L13.	-.9891212	.7343545	-1.35	0.178	-2.431178	.4529358
L14.	.7258278	.7013458	1.03	0.301	-.6514098	2.103065
L15.	.0943665	.7313066	0.13	0.897	-1.341705	1.530438
L16.	.2231666	.7806436	0.29	0.775	-1.309789	1.756122
L17.	-.2028072	.7383567	-0.27	0.784	-1.652723	1.247109
L18.	-.218164	.7335827	-0.30	0.766	-1.658705	1.222377
L19.	.5546348	.7087097	0.78	0.434	-.8370634	1.946333
L20.	-.0436014	.6599552	-0.07	0.947	-1.33956	1.252357
L21.	.3333265	.6025965	0.55	0.580	-.8499963	1.516649
L22.	-.2622156	.554213	-0.47	0.636	-1.350528	.8260964
L23.	.7324488	.3403644	2.15	0.032	.0640728	1.400825
_cons	.5601103	.1315604	4.26	0.000	.3017642	.8184564

Forecast

- Predicted value for 2011M2=1.26

Model Selection

- The dynamic distributed lag model has p lags of *y* and q lags of *x*, a total of 1+p+q estimated coefficients
- Models (p and q) can be selected by calculating and minimizing the AIC

$$AIC = N \ln\left(\frac{SSR}{T}\right) + 2(p+q+1)$$

- If the sample is, say, 251, the AIC is

.dis ln(e(rss)/e(N))*251+e(rank)*2

Predictive Causality

- The variable x affects a forecast for y if lagged values of x have true non-zero coefficients in the dynamic regression of y on lagged y's and lagged x's
- If one of the β's are non-zero

$$y_t = \mu + \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p}$$

$$+ \beta_1 x_{t-1} + \cdots + \beta_q x_{t-q} + e_t$$

Predictive Causality

- In this case, we say that "x causes y"
  - It does not mean causality in a mechanical sense
  - Only that x "predictively causes" y
  - True causality could actually be the reverse
- In economics, "predictive causality" is frequently called "Granger causality"

Clive Granger

- UCSD econometrician
  - 1934-2009
  - Winner of 2003 Nobel Prize
  - Greatest time-series econometrician of all time
  - Photo from March 2007
- Many accomplishments
  - Granger causality
  - Spurious regression
  - Cointegration

Non-Causality

- Hypothesis:
  - x does not predictively (Granger) cause y
- Test
  - Reject hypothesis of non-causality if joint test of all lags on x are zero
  - F test using robust r option

$$y_t = \mu + \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p}$$
$$+ \beta_1 x_{t-1} + \cdots + \beta_q x_{t-q} + e_t$$

STATA Command

- **.reg t3 L(1/12).t3 L(1/12).t12, r**
- **.testparm L(1/12).t12**

T3 on T12

```
. reg t3 L(1/12).t3 L(1/12).t12,r
```

Linear regression

Number of obs	=	672
F(24, 647)	=	825.67
Prob > F	=	0.0000
R-squared	=	0.9839
Root MSE	=	.36977

t3	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
t3						
L1.	.8974284	.1641167	5.47	0.000	.5751627	1.219694
L2.	.0407911	.2464121	0.17	0.869	-.4430728	.524655
L3.	-.0264118	.2323006	-0.11	0.910	-.4825658	.4297423
L4.	-.0219068	.2055053	-0.11	0.915	-.4254446	.381631
L5.	-.0240415	.2010956	-0.12	0.905	-.4189204	.3708374
L6.	-.2890795	.1773723	-1.63	0.104	-.6373744	.0592155
L7.	.5806952	.1630264	3.56	0.000	.2605703	.90082
L8.	-.2881936	.1863616	-1.55	0.122	-.6541401	.0777529
L9.	.1304009	.1864947	0.70	0.485	-.2358071	.4966089
L10.	-.2427493	.2108571	-1.15	0.250	-.6567961	.1712975
L11.	-.0109552	.2154612	-0.05	0.959	-.4340429	.4121326
L12.	.0932065	.1316611	0.71	0.479	-.1653281	.3517412

Lags on T12

t12						
L1.	.5909867	.1667627	3.54	0.000	.2635253	.9184482
L2.	-.7507476	.246725	-3.04	0.002	-1.235226	-.2662692
L3.	.3040546	.2442302	1.24	0.214	-.175525	.7836342
L4.	-.1351343	.2480177	-0.54	0.586	-.6221511	.3518825
L5.	.2764384	.2004273	1.38	0.168	-.1171281	.6700049
L6.	-.0746317	.1668985	-0.45	0.655	-.4023597	.2530964
L7.	-.4188874	.1641157	-2.55	0.011	-.7411511	-.0966237
L8.	.3884508	.2071626	1.88	0.061	-.0183414	.7952429
L9.	-.0715794	.2095218	-0.34	0.733	-.4830042	.3398453
L10.	.09139	.2050567	0.45	0.656	-.3112669	.494047
L11.	.0928276	.1716054	0.54	0.589	-.2441431	.4297982
L12.	-.1533698	.0976396	-1.57	0.117	-.3450985	.0383589
_cons	.0269204	.0369971	0.73	0.467	-.0457283	.0995692

Causality Test

```
. testparm L(1/12).t12

 ( 1)  L.t12 = 0
 ( 2)  L2.t12 = 0
 ( 3)  L3.t12 = 0
 ( 4)  L4.t12 = 0
 ( 5)  L5.t12 = 0
 ( 6)  L6.t12 = 0
 ( 7)  L7.t12 = 0
 ( 8)  L8.t12 = 0
 ( 9)  L9.t12 = 0
 (10)  L10.t12 = 0
 (11)  L11.t12 = 0
 (12)  L12.t12 = 0

       F( 12,   647) =    2.91
            Prob > F =    0.0006
```

- P-value is near zero
- Reject hypothesis of non-causality
- Infer that 12-month T-Bill helps predict 3-month T-bill
- Long rates help to predict short rates

Reverse: T12 on T3

- Do short rates help to forecast long rates?
- Regress T12 on lagged values, and lags of T3

T12 on T3

```
. reg t12 L(1/12).t12 L(1/12).t3,r

Linear regression                               Number of obs =     672
                                                F( 24,   647) =  971.95
                                                Prob > F      =  0.0000
                                                R-squared     =  0.9860
                                                Root MSE      =  .36431
```

t12	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
t12						
L1.	1.655442	.1420021	11.66	0.000	1.376601	1.934283
L2.	-1.090959	.2477411	-4.40	0.000	-1.577432	-.6044853
L3.	.5101927	.2306454	2.21	0.027	.0572887	.9630967
L4.	-.2382451	.2021018	-1.18	0.239	-.6350998	.1586097
L5.	.3252625	.1776901	1.83	0.068	-.0236564	.6741814
L6.	-.1519077	.1573297	-0.97	0.335	-.4608463	.1570308
L7.	-.2820363	.1636427	-1.72	0.085	-.6033712	.0392986
L8.	.2813756	.1854223	1.52	0.130	-.0827265	.6454777
L9.	-.0380916	.1928445	-0.20	0.843	-.4167682	.340585
L10.	.0560037	.1762077	0.32	0.751	-.2900044	.4020117
L11.	.2563501	.167069	1.53	0.125	-.0717128	.584413
L12.	-.2783838	.1015364	-2.74	0.006	-.4777644	-.0790031

Lags on T3

t3						
L1.	-.1825098	.1530766	-1.19	0.234	-.4830968	.1180772
L2.	.403259	.2554067	1.58	0.115	-.0982671	.904785
L3.	-.2821056	.2236905	-1.26	0.208	-.7213526	.1571413
L4.	.1418548	.1717628	0.83	0.409	-.1954251	.4791346
L5.	-.0539846	.1705447	-0.32	0.752	-.3888726	.2809034
L6.	-.2983116	.1538922	-1.94	0.053	-.6005002	.0038769
L7.	.5620287	.1591393	3.53	0.000	.2495368	.8745207
L8.	-.2021671	.1673543	-1.21	0.227	-.5307902	.126456
L9.	-.0413945	.1739921	-0.24	0.812	-.3830518	.3002628
L10.	-.1121067	.1804495	-0.62	0.535	-.4664441	.2422308
L11.	-.1303389	.1907979	-0.68	0.495	-.5049967	.244319
L12.	.178009	.1232592	1.44	0.149	-.0640275	.4200454
_cons	.060513	.0350361	1.73	0.085	-.0082853	.1293112

Causality Test

```
. testparm L(1/12).t3

 ( 1)  L.t3 = 0
 ( 2)  L2.t3 = 0
 ( 3)  L3.t3 = 0
 ( 4)  L4.t3 = 0
 ( 5)  L5.t3 = 0
 ( 6)  L6.t3 = 0
 ( 7)  L7.t3 = 0
 ( 8)  L8.t3 = 0
 ( 9)  L9.t3 = 0
 (10)  L10.t3 = 0
 (11)  L11.t3 = 0
 (12)  L12.t3 = 0

       F( 12,   647) =    1.71
            Prob > F =    0.0612
```

- P-value is nearly significant
- Not clear if we reject hypothesis of non-causality
- Unclear if 3-month T-Bill helps predict 12-month T-bill
  - If short rates help to predict long rates

Term Structure Theory

- This is not surprising, given the theory of the **term structure of interest rates**
- Helpful to review interest rate theory

Bonds

- A bond with face value $1000 is a promise to pay $1000 at a specific date in the future
  - If that date is 3 months from today, it is a 3-month bonds
  - If that date is 12 months from today, it is a 12-month bond
- Rate: If a 3-month $1000 bond sells for $980, the interest percentage for the 3-month period is 100*20/980=2.04%, or 8.16% annual rate

Term Structure

- Suppose an investor has a 2-period horizon
  - They can purchase a 2-period bond
  - Or a sequence of one-period bonds
- Competitive equilibrium sets the prices of the bonds so they have equal expected returns.
  - The average expected one-period returns equal the two-period return
  - The two-period return is an expectation of future short rates

$$Long_t = \frac{Short_t + E\left(Short_{t+1} \mid \Omega_t\right)}{2}$$

Term Structure Regression

- This implies

$$E\left(Short_{t+1} \mid \Omega_t\right) = 2Long_t - Short_t$$

- Thus a predictive regression for short-term interest rates is a function of lagged long-term interest rates
- Long-term interest rates help forecast short term rates because long-term rates are themselves market forecast of future short rates
  - High long-term rates mean that investors expect short rates to rise in the future

Causality

- The theory of the term structure predicts that long-term rates will help predict short-term rates
- It does not predict the reverse
- This is consistent with our hypothesis tests
  - 12-month T-Bill predicted 3-month T-Bill
  - Unclear if 3-month predicts 12-month.

Selection of Causal Variables

- Even if we don’t reject non-causality of y by x, we still might want to include x in forecast regression
  - Testing is not a good selection method
  - AIC is a better for selection

Example

- Prediction of 3-month rate
  - AR(12) only: AIC=-1249
  - AR(12) plus T12(12 lags): AIC=-1313
  - Full model has smaller AIC, so is preferred for forecasting
    - This is consistent with causality test
- Prediction of 12-month rate
  - AR(12) only: AIC=-1309
  - AR(12) plus T3(12 lags): AIC=-1333
  - Full model has smaller AIC, so is preferred for forecasting
    - Even though we cannot reject non-causality, AIC recommends using the short rate to forecast the long rate