

Conference on

Forecasting and Monetary Policy

Berlin, 23-24 March 2009

Norman R. Swanson

Rutgers University and the Federal Reserve Bank of Philadelphia

**„Real-Time Datasets Really Do Make a Difference:
Definitional Change, Data Release and Forecasting“**

Real-Time Datasets Really Do Make a Difference: Definitional Change, Data Release, and Forecasting

Valentina Corradi

Andres Fernandez

Norman Rasmus Swanson

January 2009



Outline

- Introduction
- Setup
- Empirical and Theoretical Methodology
- Empirical Results
- Concluding Remarks

Introduction

- The literature on testing for and assessing the rationality of early release economic data is rich and deep.
-
- Howrey (1978)
 - Mankiw, Runkle, and Shapiro (1984)
 - Mankiw and Shapiro (1986)
 - Milbourne and Smith (1989)
 - Keane and Runkle (1989,1990)
 - Kavajecz and Collins (1995)
 - Faust, Rogers, and Wright (2005)
 - Swanson and van Dijk (2006)

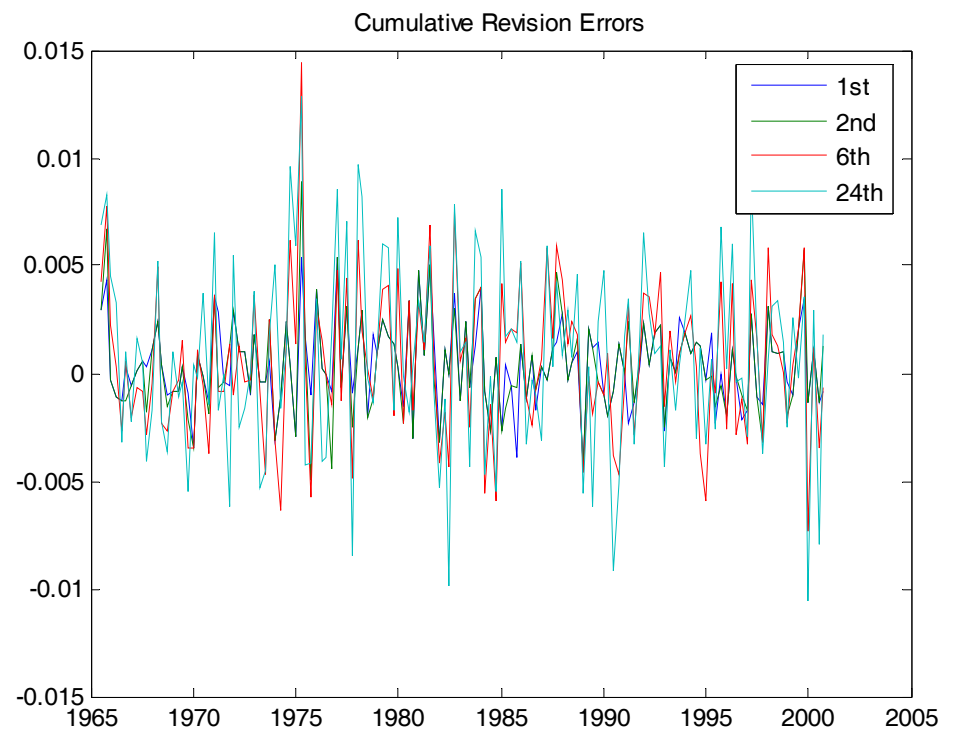
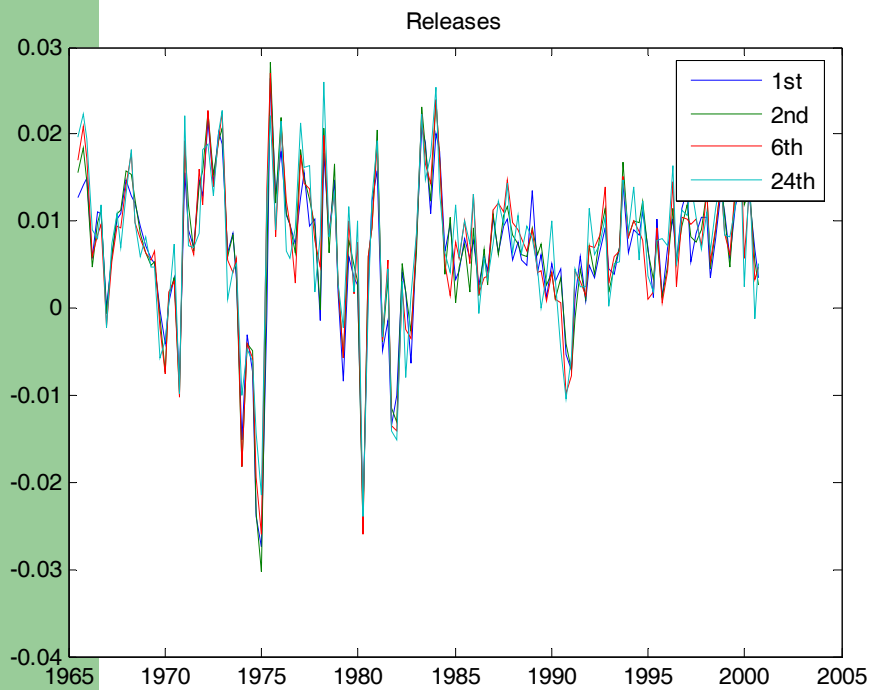
- .

Introduction

- The literature on prediction of real-time data and the usefulness of the revision process is also extensive.
-
- Morgenstern (1963)
 - Stekler (1967)
 - Diebold and Rudebusch (1991)
 - Hamilton and Perez-Quiros (1996)
 - Gallo and Marcellino (1999)
 - Swanson, Ghysels and Callan (1999)
 - Amato and Swanson (2001)
 - Croushore and Stark (2003)
 - Aruoba (2006)
 - Croushore (2006)
 - Faust and Wright (2009)
 - Garratt, Koop, Mise and Vahey (2009)

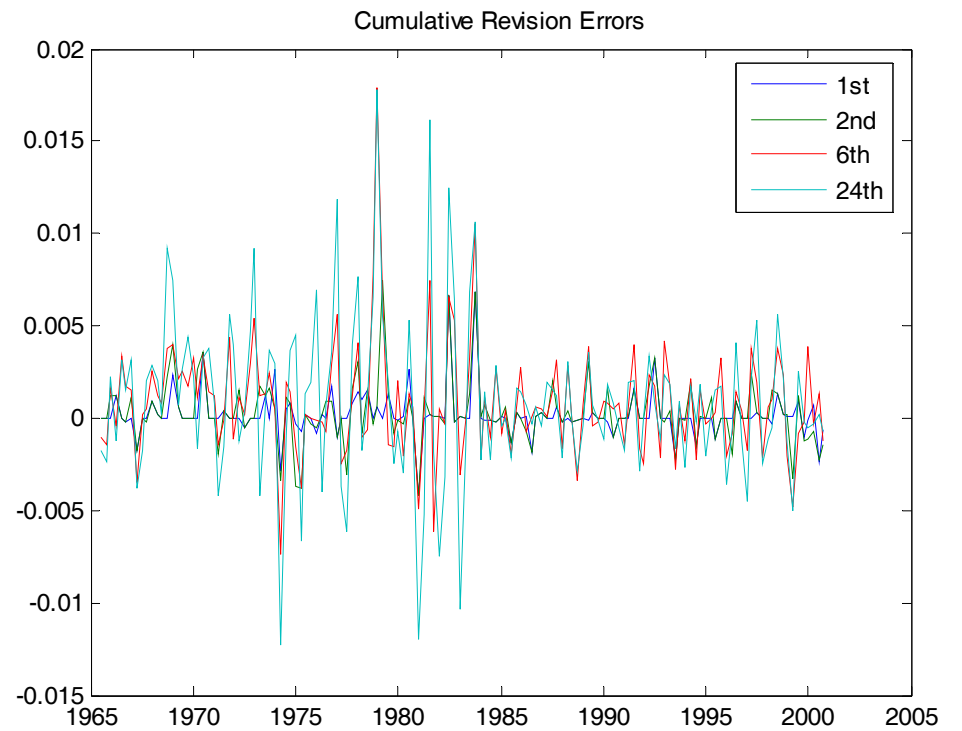
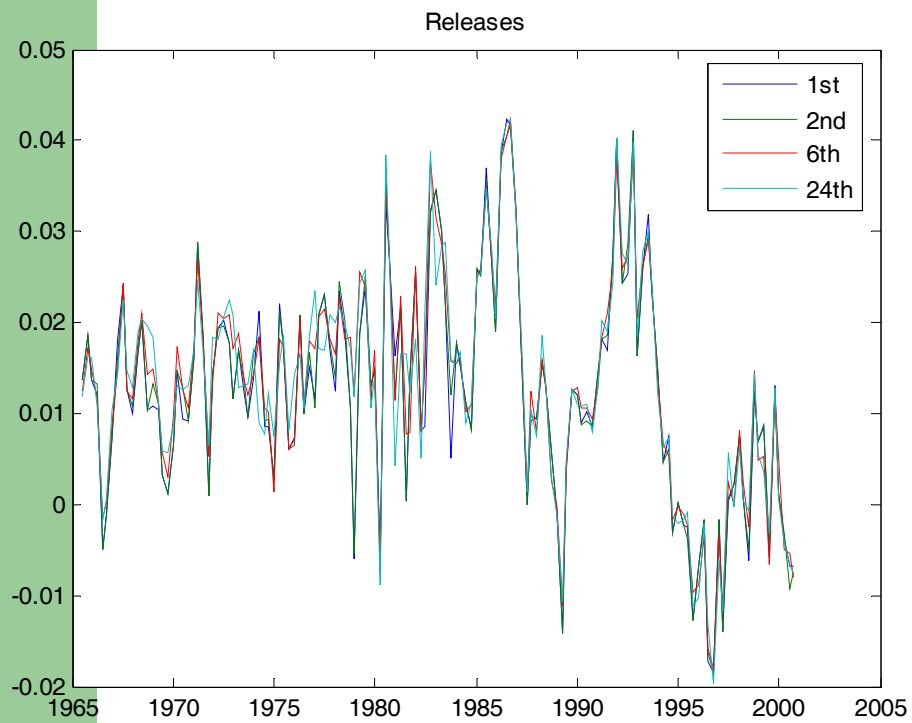
Introduction

Real-time Output



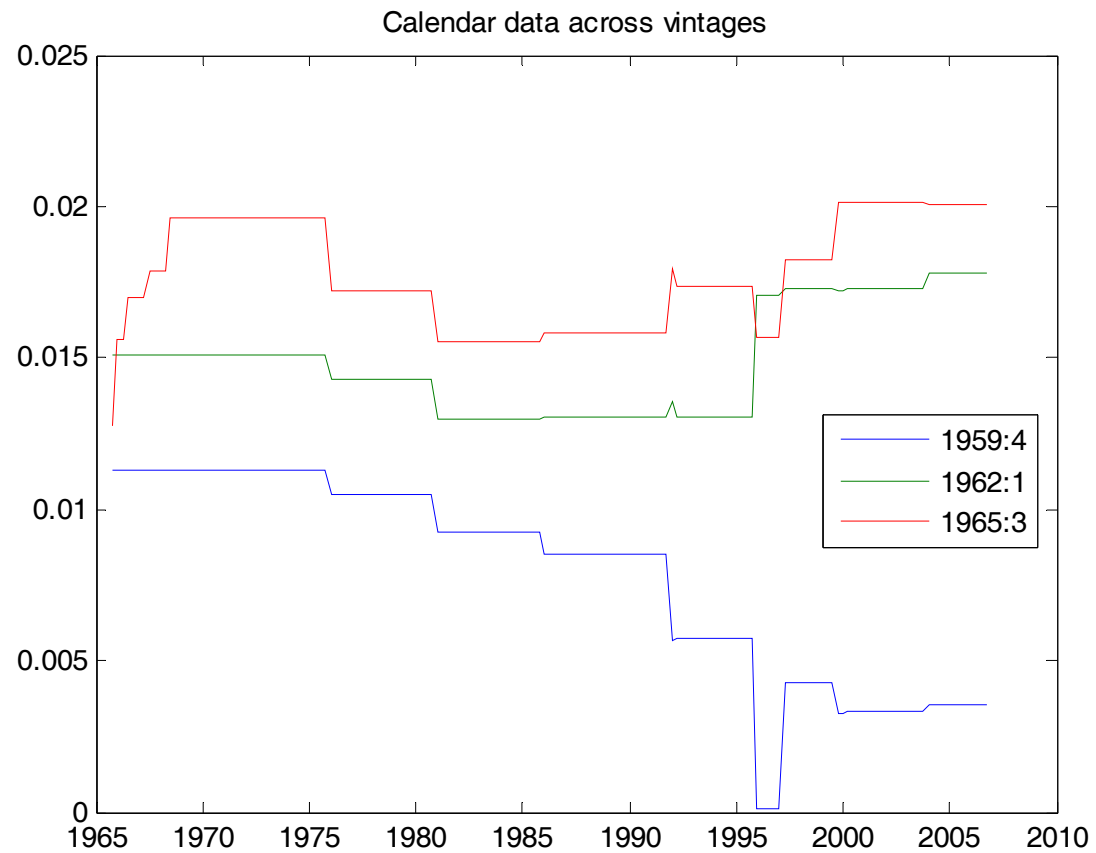
Introduction

Real Time Money



Introduction

Definitional Change



Introduction

- Ex ante prediction experiments –excluding the revision process
 - ➡ What are the trade-offs between mixed release time series (i.e. the latest vintage) in modeling and use of only one release of data?
- Ex ante prediction experiments – including the revision process
 - ➡ Does the revision process help for forecasting?
- Which release should one use for prediction?

Introduction

- “Structural breaks” in the data affect time series of early releases
 - ➡ Do definitional changes matter?
- In-sample, versus quasi prediction based, versus truly ex-ante tests of early release efficiency
 - ➡ Does our new truly prediction based test tell us more than we already knew?
- Does the revision process make a difference for the real-time marginal predictive content of money for income?

Setup

- Let ${}_{t+k}X_t$

denote a variable (reported as an annualized growth rate) for which real-time data are available, where the subscript t denotes the time period to which the datum pertains, and the subscript $t+k$ denotes the time period during which the datum becomes available.

- In addition, let ${}_fX_t$

denote fully revised or "final" data, which is obtained as $k \rightarrow \infty$

- Lastly, let ${}_{t+2}u_t^{t+1} = {}_{t+2}X_t - {}_{t+1}X_t$

denote the errors between the 2nd and the 1st releases at time $t+2$.

Empirical Methodology Prediction

- We consider the issue of prediction using various variable/vintage combinations as defined in the following set of models.

- **Model A (First Available Data):**

$${}_{t+k}X_{t+1} = \alpha_{t+1,t}^A + \sum_{i=1}^{p^A} \beta_{i,t+1,t}^A {}_{t+2-i}X_{t+1-i} + {}_{t+k}\epsilon_{t+1}^A$$

- **Model B (*k*-th Available Data):**

$${}_{t+k}X_{t+1} = \alpha_{t+1,t}^B + \sum_{i=1}^{p^B} \beta_{i,t+1,t}^B {}_{t+2-i}X_{t+3-k-i} + {}_{t+k}\epsilon_{t+1}^B$$

- **Model C (Real-Time Data):**

$${}_{t+k}X_{t+1} = \alpha_{t+1,t}^C + \sum_{i=1}^{p^C} \beta_{i,t+1,t}^C {}_{t+1}X_{t+1-i} + {}_{t+k}\epsilon_{t+1}^C$$

Empirical Methodology Prediction

- Model A (First Available Data):

$$X_{t+k} = \alpha_{t+1,t}^A + \sum_{i=1}^{p^A} \beta_{i,t+1,t}^A X_{t+1-i} + \epsilon_{t+k}^A$$

- Model A Predictions:

$$\hat{X}_{t+k}^f = \hat{\alpha}_{t+1,t}^A + \sum_{i=1}^{p^A} \hat{\beta}_{i,t+1,t}^A X_{t+1-i}$$

- Model A Estimators ($p=1$):

$$\begin{bmatrix} \hat{\alpha}_{t+1,t}^A \\ \hat{\beta}_{1,t+1,t}^A \end{bmatrix} = \begin{bmatrix} t-1 & \sum_{j=2}^t X_{j-1} \\ \sum_{j=2}^t X_{j-1} & \sum_{j=2}^t X_{j-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=2}^t X_j \\ \sum_{j=2}^t X_j X_{j-1} \end{bmatrix}$$

Empirical Methodology Prediction

- Model B (k -th Available Data):

$${}_{t+k}X_{t+1} = \alpha_{t+1,t}^B + \sum_{i=1}^{p^B} \beta_{i,t+1,t}^B {}_{t+2-i}X_{t+3-k-i} + {}_{t+k}\epsilon_{t+1}^B$$

- Model B Predictions:

$${}_{t+k}\hat{X}_{t+1}^f = \hat{\alpha}_{t+1,t}^B + \sum_{i=1}^{p^B} \hat{\beta}_{i,t}^B {}_{t+k-i}X_{t+1-i}$$

- Model B Estimators ($p=1$):

$$\begin{bmatrix} \hat{\alpha}_{t+1,t}^B \\ \hat{\beta}_{1,t+1,t}^B \end{bmatrix} = \begin{bmatrix} t-1 & \sum_{j=2}^t X_{j+1-k} \\ \sum_{j=2}^t X_{j+1-k} & \sum_{j=2}^t X_{j+1-k}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=2}^t X_{j+2-k} \\ \sum_{j=2}^t X_{j+2-k} X_{j+1-k} \end{bmatrix}$$

Empirical Methodology Prediction

- Model C (Real-Time Data):

$$X_{t+1} = \alpha_{t+1,t}^C + \sum_{i=1}^{p^C} \beta_{i,t+1,t}^C X_{t+1-i} + \epsilon_{t+1}^C$$

- Model C Predictions:

$$\hat{X}_{t+1}^f = \hat{\alpha}_{t+1,t}^C + \sum_{i=1}^{p^C} \hat{\beta}_{i,t}^C X_{t+1-i}$$

- Model C Estimators ($p=1$):

$$\begin{bmatrix} \hat{\alpha}_{t+1,t}^C \\ \hat{\beta}_{1,t+1,t}^C \end{bmatrix} = \begin{bmatrix} t-1 & \sum_{j=2}^t X_{j-1} \\ \sum_{j=2}^t X_{j-1} & \sum_{j=2}^t X_{j-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{j=2}^t X_j \\ \sum_{j=2}^t X_j X_{j-1} \end{bmatrix}$$

Empirical Methodology Prediction

(i) $p = 1$; (ii) $p = SIC$; (iii) $p = AIC$; (iv) $p = 0$ (random walk with drift model)

- Different k 's: $k = \{2, 3, 4, 6, 12, 24\}$

- Revision errors as Additional regressors. Four cases:

$${}_{t+1}W'_t = u_{C1} = {}_{t+1}u^t_{t-1}$$

$${}_{t+1}W'_t = u_{C2} = ({}_{t+1}u^t_{t-1}, {}_t u^{t-1}_{t-2})$$

$${}_{t+1}W'_t = u_{C3} = {}_{t+1}u^{t-1}_{t-2}$$

$${}_{t+1}W'_t = u_{C4} = ({}_{t+1}u^t_{t-1}, {}_{t+1}u^t_{t-2})$$

Empirical Methodology Prediction

- **Sample:**
 - Calendar Dates: 1959:4 – 2006:4
 - Vintages: 1965:4 – 2006:4
- **MSFE's associated with 1-step ahead predictions constructed using recursively estimated model:**
 - $T = R + P$
 - $R = 1969:4$
 - Construct sequences of $P-k$ ex-ante predictions.
- **MSFE's examined via the use of DM predictive accuracy tests:**

$$DM = \sqrt{P - k} \frac{\frac{1}{P} \sum_{t=R}^{T-k} \hat{d}_{t,k}}{\frac{1}{P-k} \sum_{j=-\bar{j}}^{\bar{j}} \sum_{t=R+j}^{T-k} K\left(\frac{j}{M}\right) \left(\hat{d}_t - \bar{d}\right) \left(\hat{d}_{t-j} - \bar{d}\right)}$$

Empirical Methodology Efficiency Tests

$$fX_t = \alpha + {}_{t+1}X_t \beta + {}_{t+1}W'_t \gamma + \epsilon_{t+1},$$

- **Swanson and Van Dijk (2006):**

$${}_{t+k}X_t - {}_{t+1}X_t = \alpha + {}_{t+1}X_t \beta + {}_{t+1}W'_t \gamma + \epsilon_{t+k}$$

- **Rationality Test:**

$$H_0 : \alpha = 0, \beta = 1, \text{ and } \gamma = 0,$$

Empirical Methodology Efficiency Tests

- Moment type tests: Corradi, Fernandez and Swanson (2009)

$$M_T = \sup_{\gamma \in \Gamma} |m_{1,T}(\gamma)|$$

$$m_T(\gamma) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T-2} {}_{t+2}u_t^{t+1} w \left(\sum_{j=0}^{t-1} \gamma_j' \Phi_{(t+1-j)} W_{t-j} \right)$$

$$\Gamma = \{ \gamma_j : a_j \leq \gamma_j \leq b_j, j = 1, 2; |a_j|, |b_j| \leq B j^{-\kappa}, \kappa \geq 2 \}$$

$$H_0 : E({}_{t+2}u_t^{t+1} | \mathcal{F}_t^{t+1}) = 0 \text{ a.s.}$$

Empirical Methodology Efficiency Tests

- Let there be no intercept and $p=1$, then:

$$\hat{\beta}_{1,t+1,t}^A = \frac{\sum_{j=2}^t j^{+1} X_j j X_{j-1}}{\sum_{j=2}^t j X_{j-1}^2} \quad \hat{\beta}_{1,t+1,t}^C = \frac{\sum_{j=2}^t {}^{t+1} X_j {}^{t+1} X_{j-1}}{\sum_{j=2}^t {}^{t+1} X_{j-1}^2}$$

- Note that if
 - no issues due to definitional changes;
 - no measurement errors
 - no inefficiency

then

$$\hat{\beta}_{1,t+1,t}^A \equiv \hat{\beta}_{1,t+1,t}^C$$

$$MSFE_A = MSFE_C, \quad \forall k = 2, 3, \dots$$

$$MSFE_X(i) = MSFE_X(j) \quad \forall i, j = 2, 3, \dots, k; \quad \forall X = A, C$$

Empirical Methodology Efficiency Tests

- However, we know it must be the case that there exists measurement error if early releases are efficient, otherwise we would observe revision errors that are identically zero in the data.
- This means that we can construct a test with the null:

$$H_0'' : \textit{measurement error and efficiency}$$

- Note that under this null, even as $P, T \rightarrow \infty$

$$\hat{\beta}_{1,t+1,t}^A \rightarrow \beta^{A,*} \quad \hat{\beta}_{1,t+1,t}^C \rightarrow \beta^{C,*}$$

but clearly

$$\beta^{A,*} \neq \beta^{C,*}$$

Empirical Methodology Efficiency Tests

- It follows that we can write **H_0** and its alternative as follows:

$$H_0' : E(l(\varepsilon_{1,t,k}) - l(\varepsilon_{2,t,k})) = 0, \forall k = 2, 3, \dots$$
$$H_A : \text{the negation of } H_0'.$$

- Assuming quadratic loss, we shall test the following version of **H_0**

$$H_0 : MSFE_A = MSFE_C, \forall k = 2, 3, \dots$$

which can be tested using the DM tests.

Empirical Methodology Efficiency Tests

- Note that if we reject

$$H_0 : MSFE_A = MSFE_C, \quad \forall k = 2, 3, \dots$$

then

(i) No need to test further (we have inefficiency), or

(ii) We test whether we have measurement error in addition to inefficiency, or

(iii) Definitional changes issues are severe.

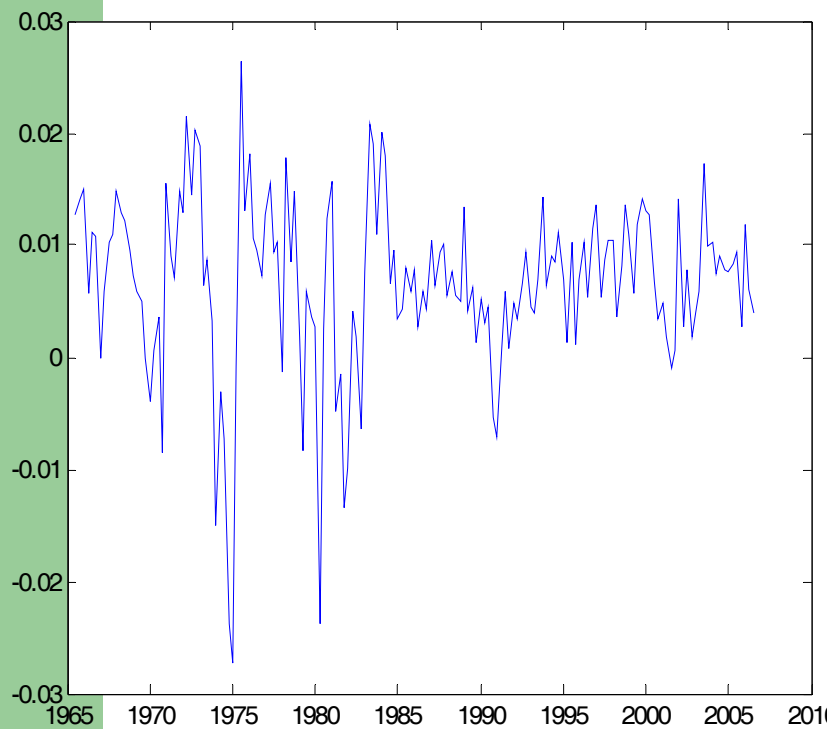
Empirical Results **Data**

- **REAL-TIME DATASET:**
 - real GDP
 - the GDP chain-weighted price index
 - the money supply (measured as M1)
 - the interest rate (measured as the rate on the 3-month Treasury bill)
- **Frequency: Quarterly**
- **Source: *Federal Reserve Bank of Philadelphia's real-time dataset for Macroeconomists, publicly available at***
<http://www.phil.frb.org/econ/forecast/readow.html>

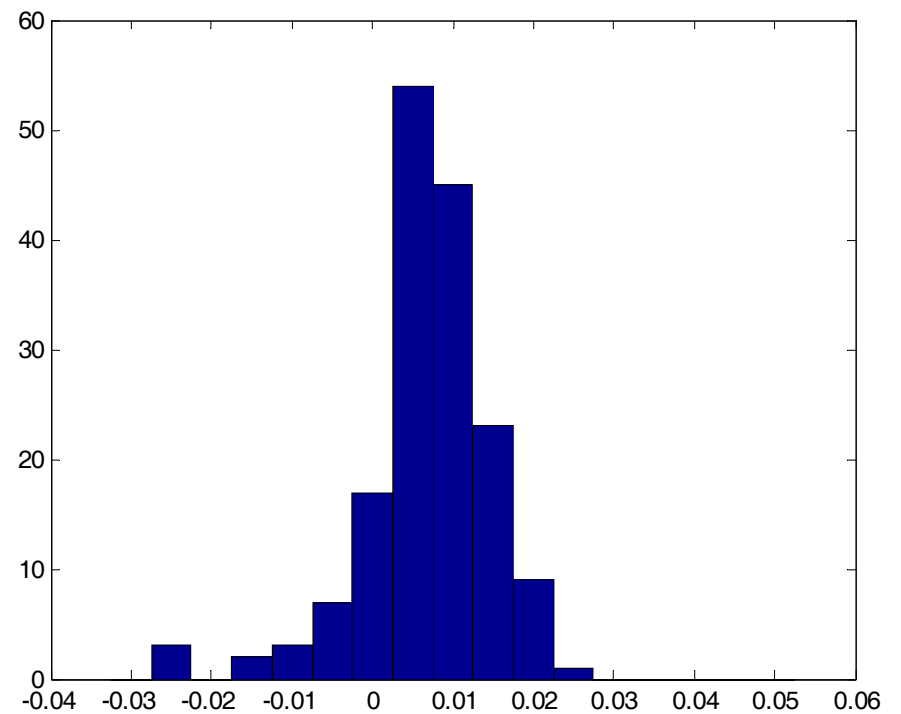
Empirical Results Data

REAL-TIME DATASET: OUTPUT (Growth Rates)

Time Series



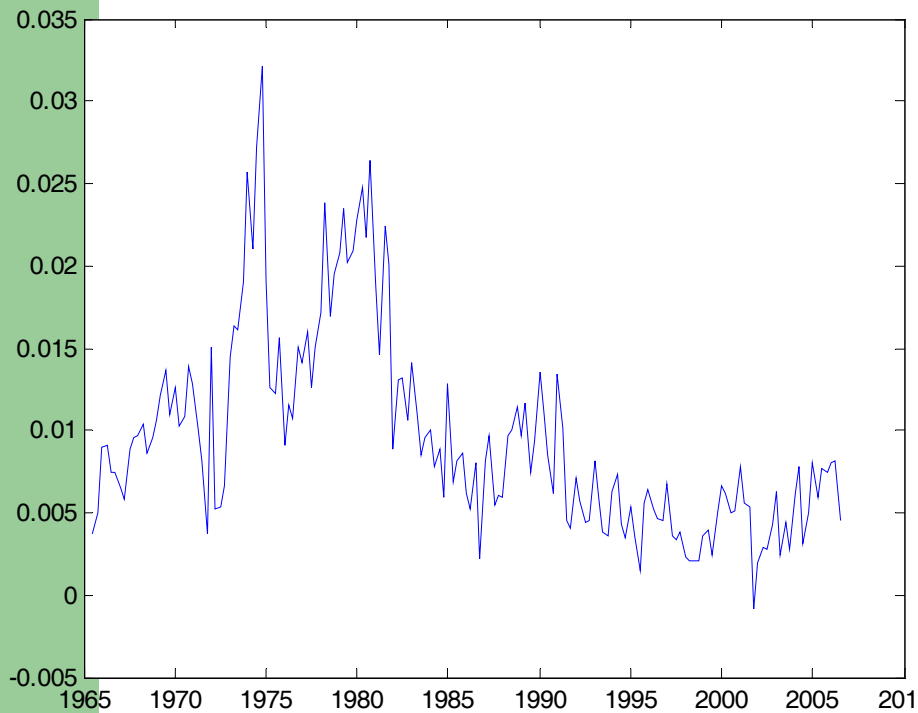
Distribution



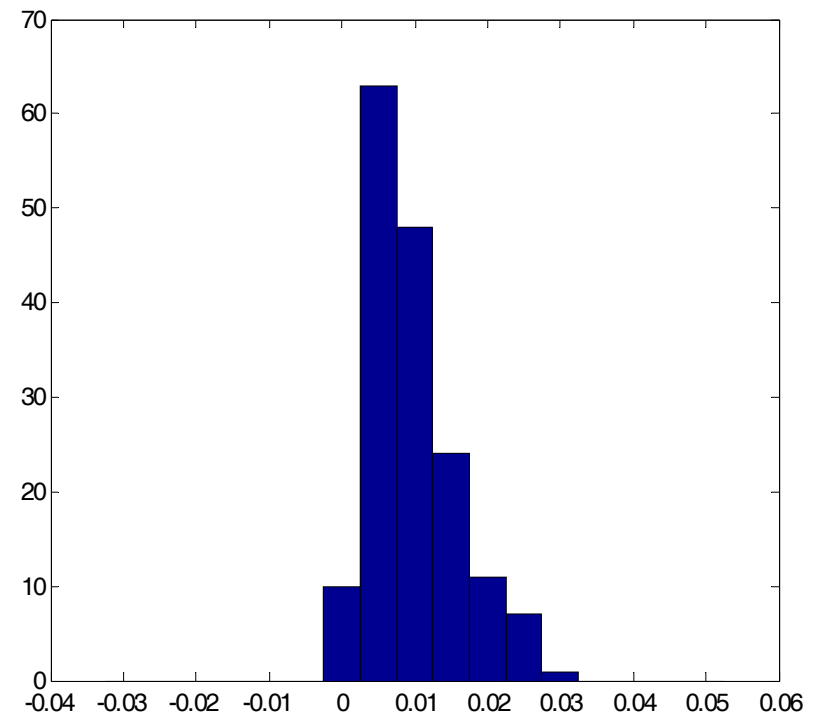
Empirical Results Data

REAL-TIME DATASET: PRICES (Growth Rates)

Time Series



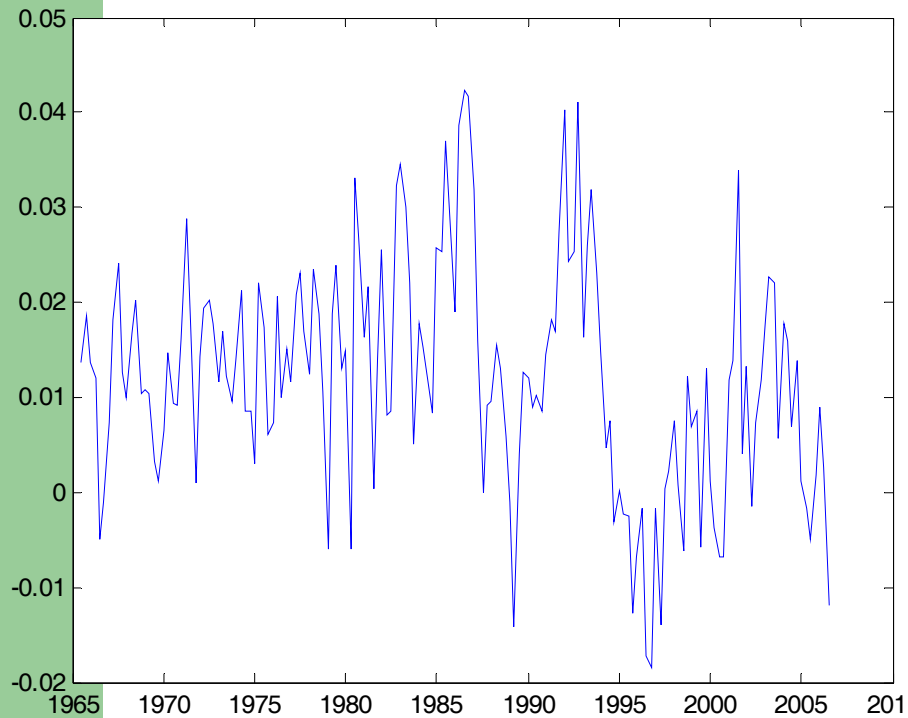
Distribution



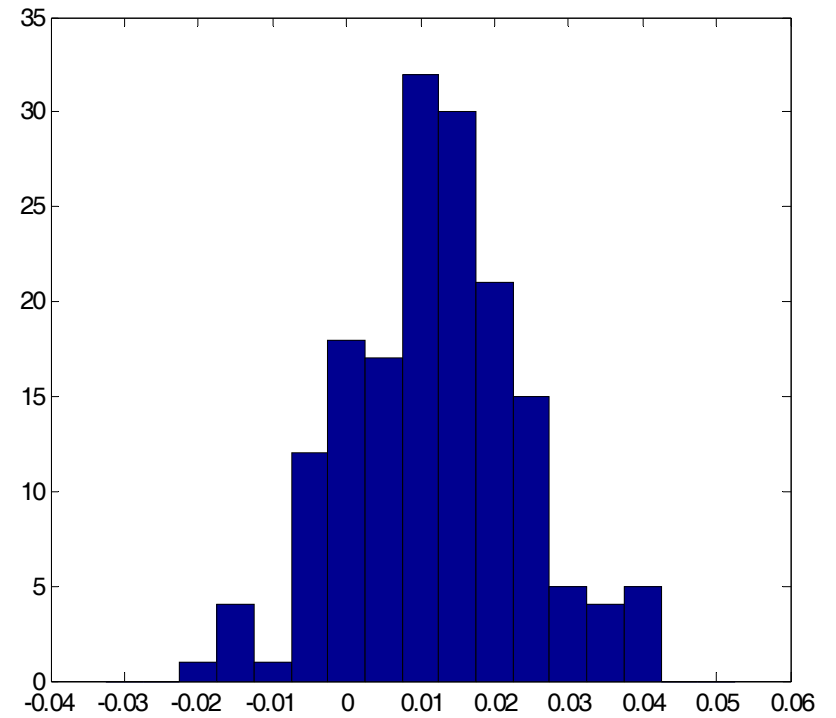
Empirical Results Data

REAL-TIME DATASET: MONEY (Growth Rates)

Time Series

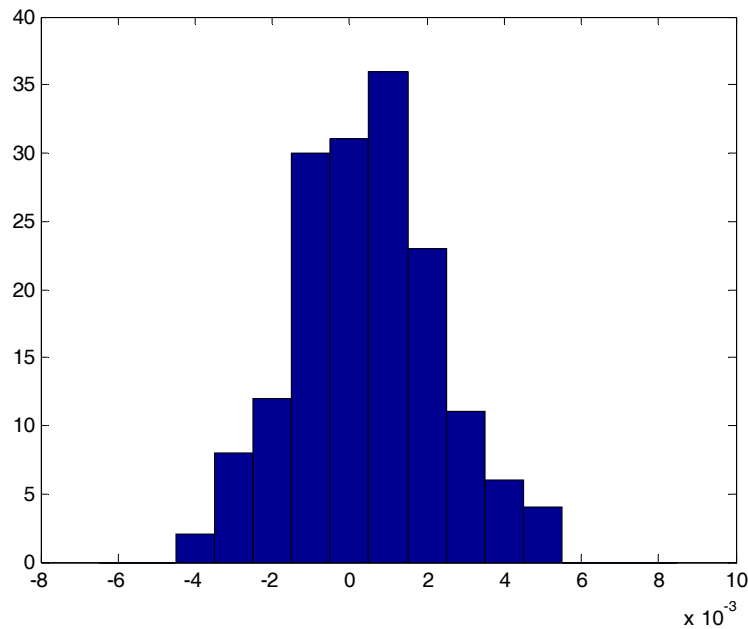


Distribution

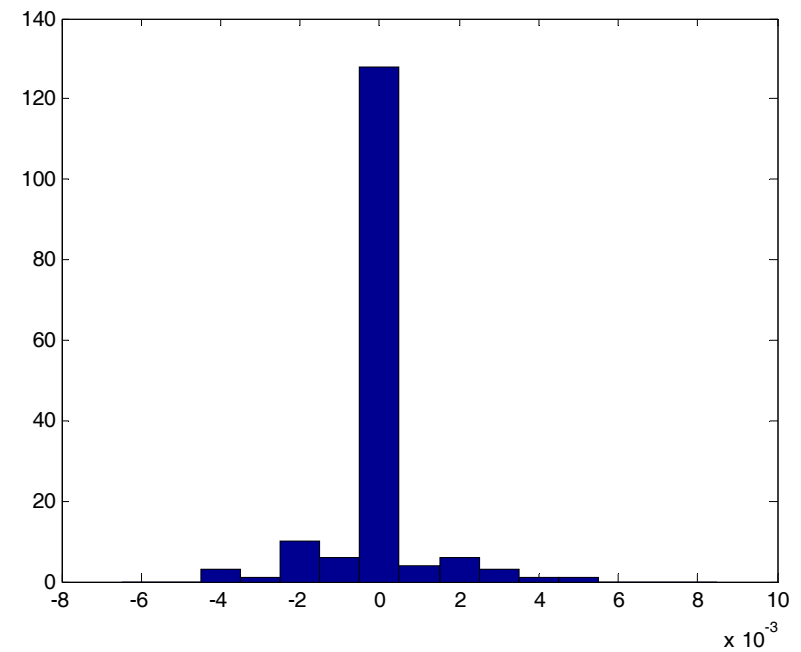


Empirical Results Data

1st Revision Errors Distribution

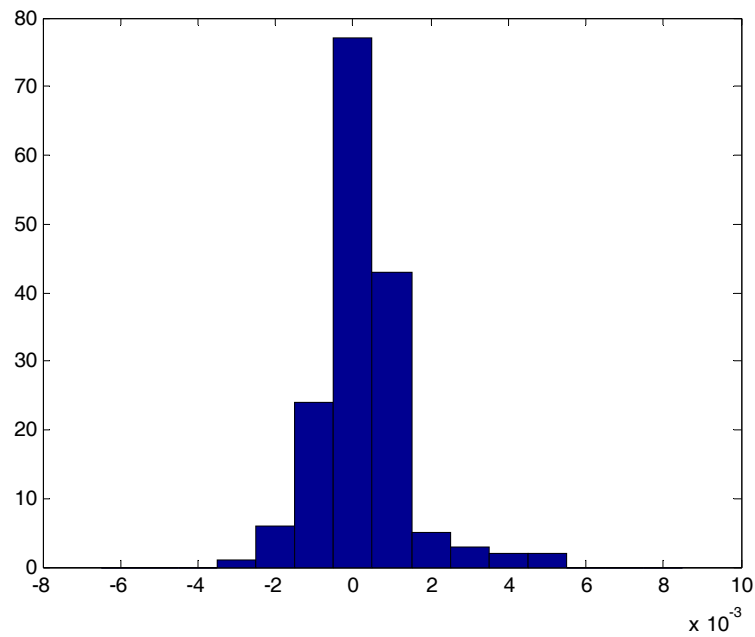


2nd Revision Errors Distribution

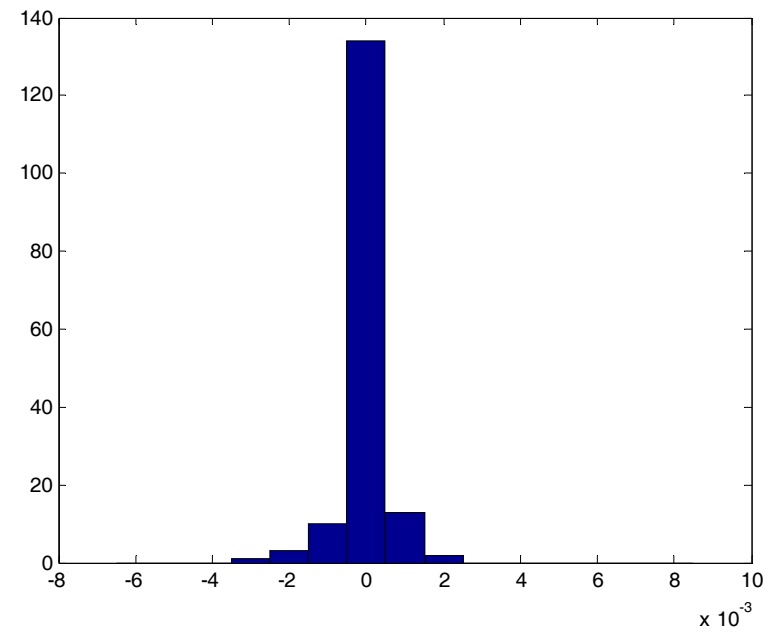


Empirical Results Data

1st Revision Errors Distribution

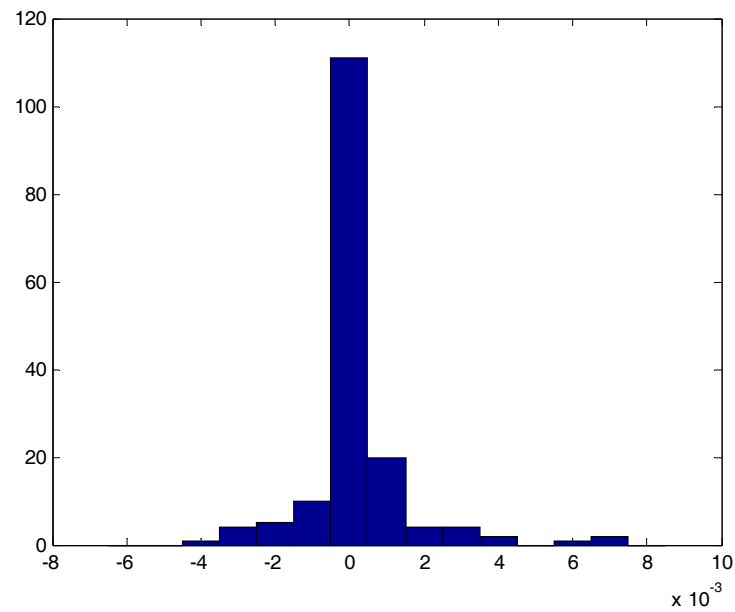


2nd Revision Errors Distribution

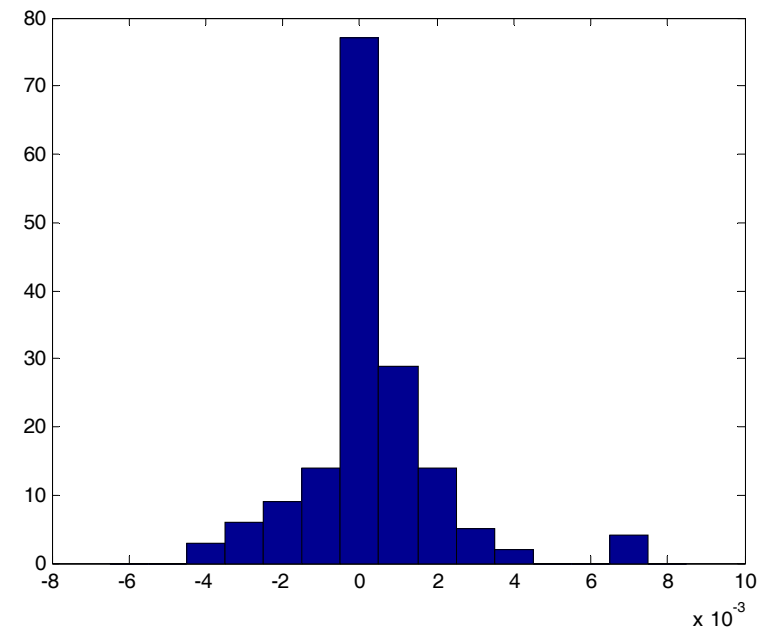


Empirical Results Data

1st Revision Errors Distribution

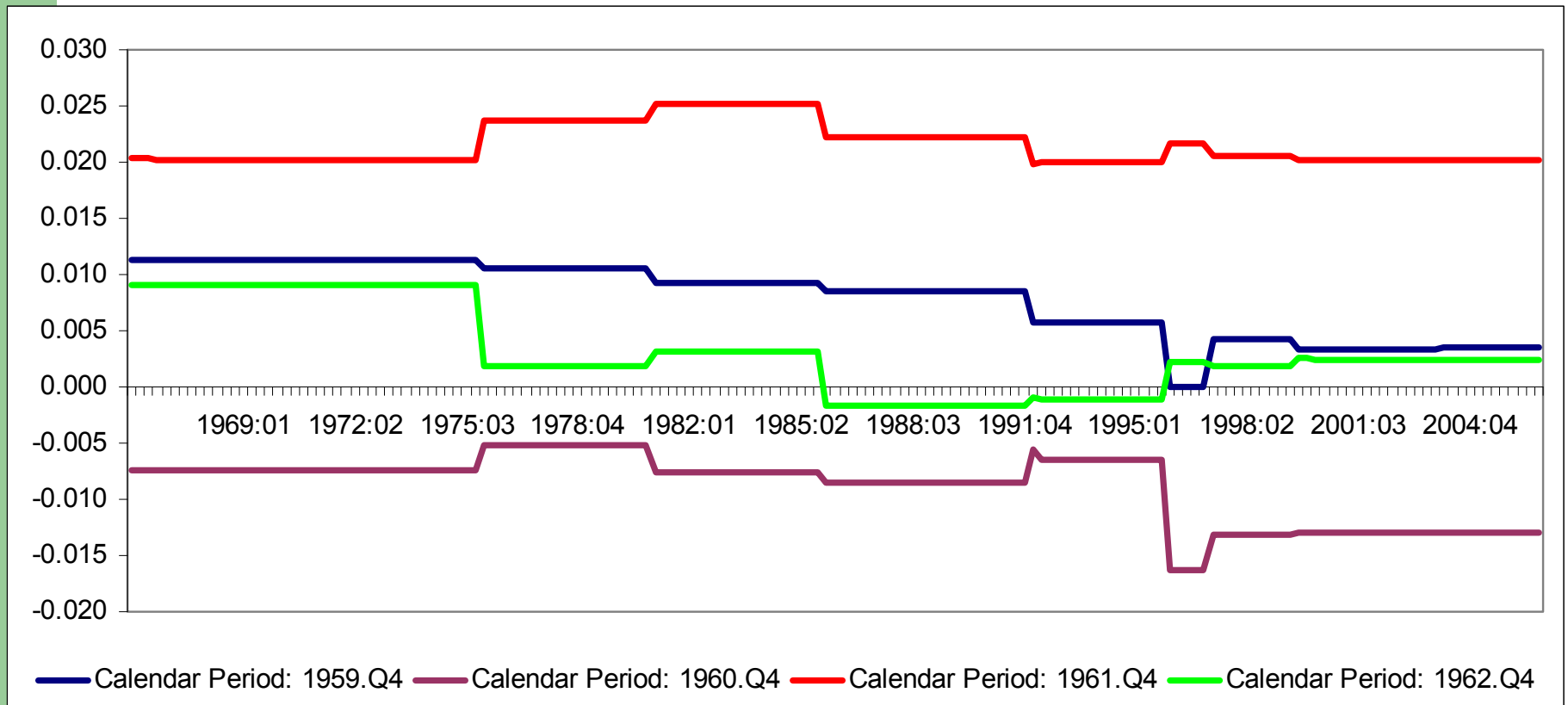


2nd Revision Errors Distribution



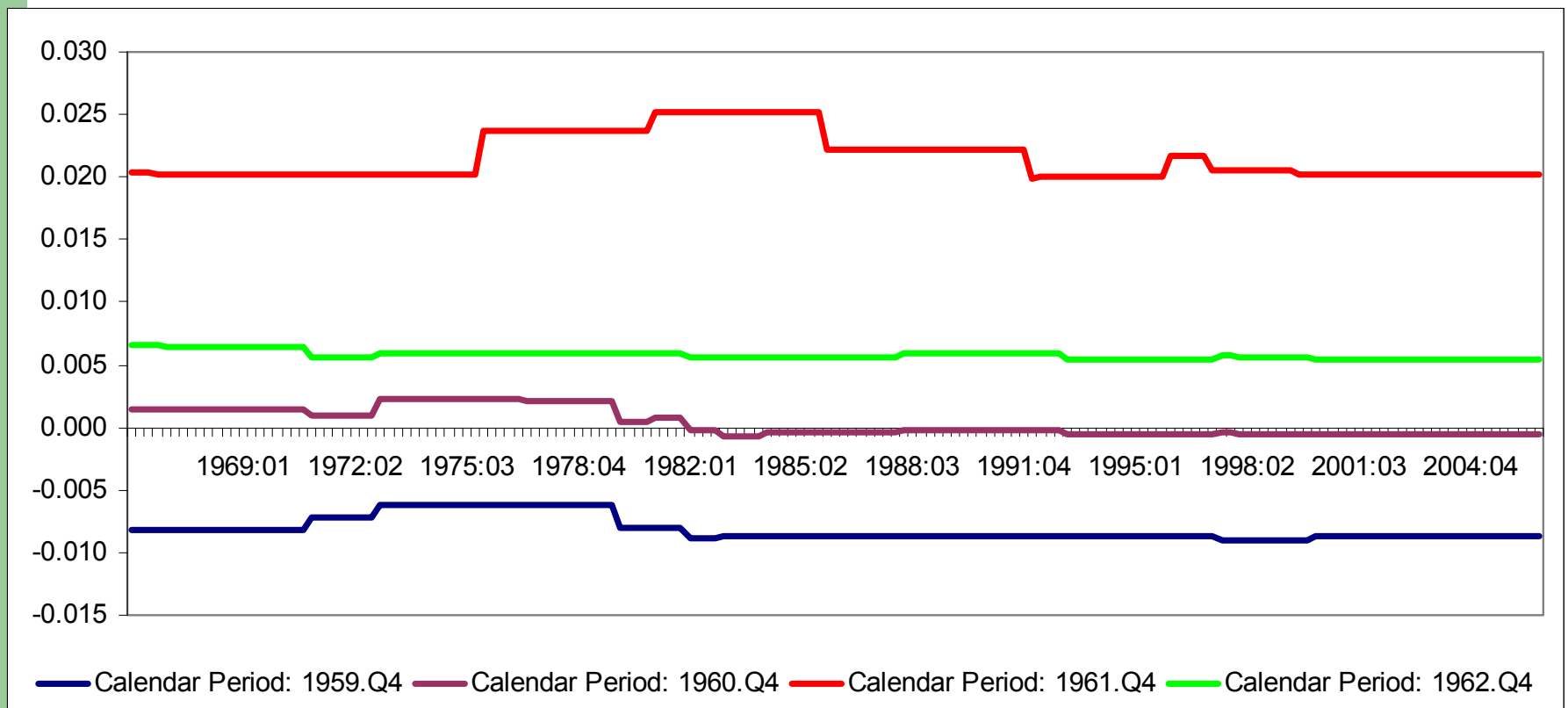
Empirical Results Data

Four Calendar Dates across Vintages: OUTPUT



Empirical Results Data

Four Calendar Dates across Vintages: MONEY





Empirical Results Data

Growth Rate and Revision Error Summary Statistics in Various Subsamples: OUTPUT

<i>Vrbl</i>	<i>Vint</i>	<i>R-Err</i>	<i>smp1</i>	\bar{y}	$\hat{\sigma}_y$	$\hat{\sigma}_{\bar{y}}$	<i>skew</i>	<i>kurt</i>	<i>LB</i>	<i>JB</i>	<i>ADF</i>
1	–	–	65:4-06:4	0.00657	0.00790	0.00061	-1.259	6.734	111.5	134.53	-6.035
2	–	–	65:4-06:4	0.00705	0.00851	0.00066	-1.148	6.678	107.1	123.88	-6.350
–	1	–	65:4-06:4	0.00046	0.00189	0.00015	0.118	2.950	23.58	0.424	-6.050
–	2	–	65:4-06:4	-0.00002	0.00106	0.00008	0.324	8.982	32.01	237.04	-5.471
1	–	–	65:4-75:4	0.00628	0.01114	0.00027	-1.139	4.566	55.42	11.33	-3.424
2	–	–	65:4-75:4	0.00613	0.01169	0.00029	-1.273	4.595	64.44	13.14	-3.082
–	1	–	65:4-75:4	0.00035	0.00191	0.00005	0.498	3.306	8.750	1.572	-3.198
–	2	–	65:4-75:4	-0.00008	0.00094	0.00002	-0.776	7.838	2.908	35.68	-6.000
1	–	–	76:1-80:4	0.00655	0.00976	0.00049	-1.559	5.774	30.36	11.02	-
2	–	–	76:1-80:4	0.00876	0.01139	0.00057	-1.125	5.292	53.24	6.246	-
–	1	–	76:1-80:4	0.00096	0.00199	0.00010	0.478	2.565	9.755	1.002	-
–	2	–	76:1-80:4	-0.00064	0.00134	0.00007	-1.722	4.246	20.36	8.506	-
1	–	–	81:1-85:4	0.00632	0.00998	0.00050	-0.301	2.234	29.88	1.065	-
2	–	–	81:1-85:4	0.00680	0.01044	0.00052	-0.074	2.425	25.95	0.564	-
–	1	–	81:1-85:4	0.00062	0.00259	0.00013	0.150	2.024	7.350	1.169	-
–	2	–	81:1-85:4	-0.00003	0.00072	0.00004	0.470	8.664	8.102	18.65	-

Prediction Experiments: Output

<i>Model</i>	<i>RevErr</i>	<i>k = 2</i>	<i>k = 3</i>	<i>k = 4</i>	<i>k = 6</i>	<i>k = 12</i>	<i>k = 24</i>
<i>Panel A: Mean Square Forecast Errors</i>							
<i>Begin Date of Forecast Period = 1970:1</i>							
A	None	0.642	0.783	0.825	0.841	0.839	0.836
B	None	0.642	0.983	1.028	1.232	1.090	1.134
C	None	0.661	0.792	0.825	0.828	0.817	0.815
RWD - A	None	0.768	0.879	0.890	0.861	0.833	0.843
RWD - B	None	0.768	0.896	0.916	0.905	0.846	0.925
RWD - C	None	0.766	0.874	0.884	0.856	0.829	0.838
<i>Begin Date of Forecast Period = 1983:1</i>							
A	None	0.212	0.259	0.270	0.303	0.322	0.354
B	None	0.212	0.303	0.322	0.490	0.466	0.573
C	None	0.206	0.259	0.267	0.299	0.316	0.346
RWD - A	None	0.275	0.337	0.345	0.374	0.382	0.419
RWD - B	None	0.275	0.326	0.337	0.359	0.345	0.363
RWD - C	None	0.248	0.309	0.314	0.343	0.356	0.387
<i>Begin Date of Forecast Period = 1990:1</i>							
A	None	0.175	0.204	0.216	0.278	0.286	0.332
B	None	0.175	0.247	0.236	0.343	0.432	0.491
C	None	0.176	0.214	0.216	0.275	0.289	0.329
RWD - A	None	0.224	0.275	0.270	0.322	0.327	0.371
RWD - B	None	0.224	0.271	0.266	0.324	0.344	0.379
RWD - C	None	0.208	0.258	0.250	0.304	0.319	0.356

Prediction Experiments: Output

Panel B: Diebold-Mariano Test Statistics Corresponding to Entries in Panel A

Begin Date of Forecast Period = 1970:1

A	None	—	—	—	—	—	—
B	None	—	-1.415	-1.423	-2.216	-1.148	-1.586
C	None	-0.349	-0.159	-0.008	0.213	0.348	0.368
RWD-A	None	-1.077	-0.718	-0.446	-0.125	0.035	-0.047
RWD-B	None	-1.077	-0.811	-0.589	-0.359	-0.041	-0.470
RWD-C	None	-1.070	-0.689	-0.411	-0.097	0.057	-0.012

Begin Date of Forecast Period = 1983:1

A	None	—	—	—	—	—	—
B	None	—	-1.096	-1.056	-2.558	-2.467	-2.542
C	None	0.480	-0.053	0.255	0.344	0.378	0.562
RWD-A	None	-1.525	-1.826	-1.718	-1.688	-1.568	-1.582
RWD-B	None	-1.525	-1.638	-1.576	-1.421	-0.568	-0.200
RWD-C	None	-0.989	-1.366	-1.242	-1.181	-1.001	-0.936

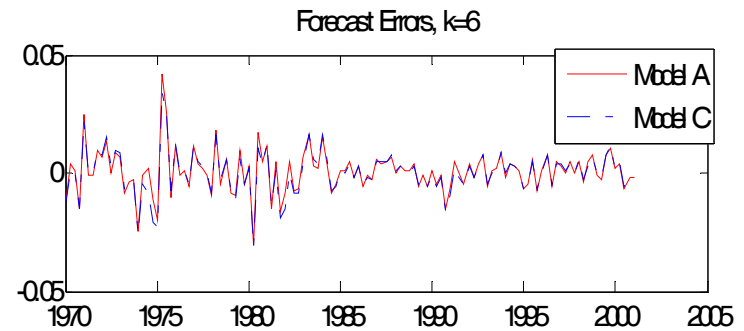
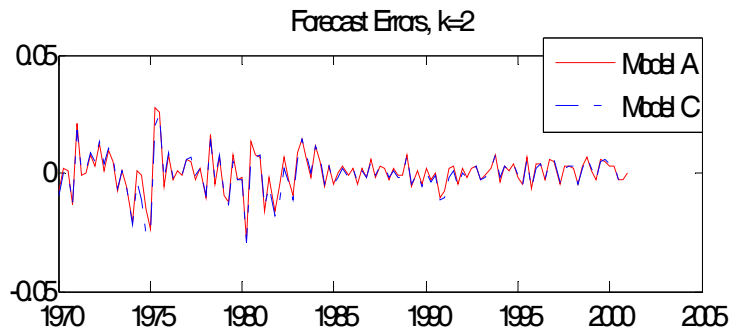
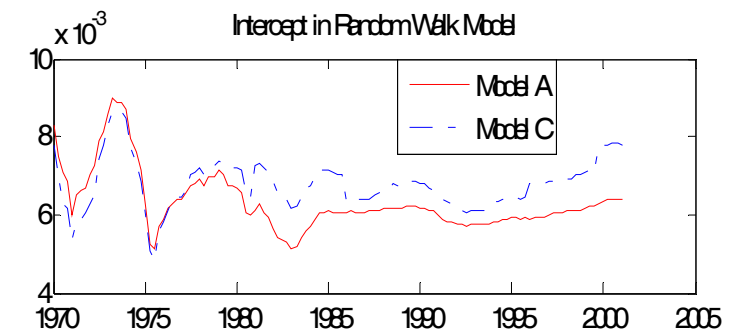
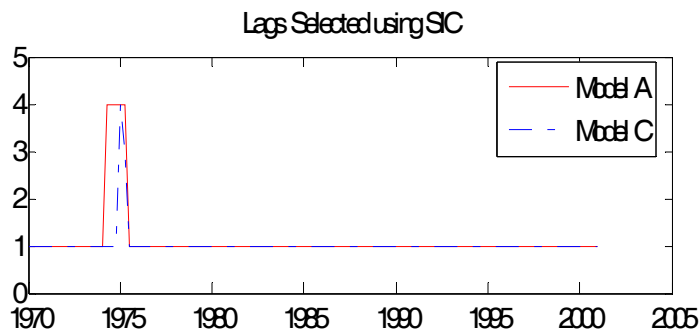
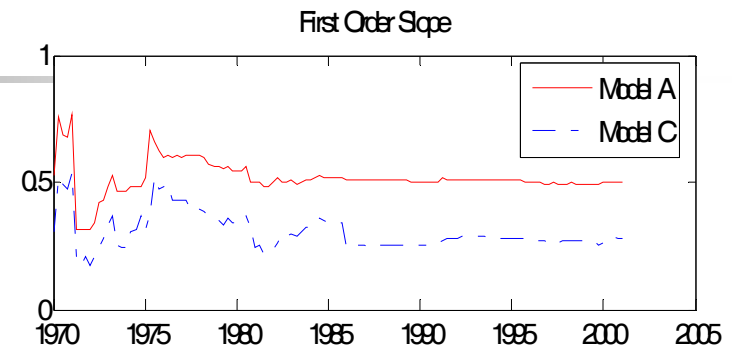
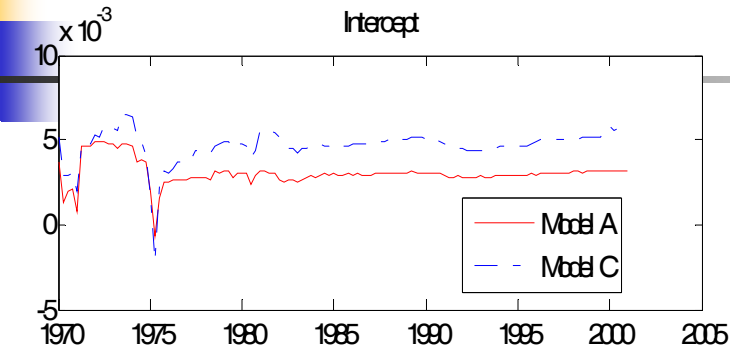
Begin Date of Forecast Period = 1990:1

A	None	—	—	—	—	—	—
B	None	—	-1.419	-0.683	-1.943	-1.743	-2.365
C	None	-0.060	-0.570	-0.039	0.166	-0.149	0.157
RWD-A	None	-1.316	-1.771	-1.499	-1.182	-1.188	-1.157
RWD-B	None	-1.316	-1.557	-1.336	-1.075	-1.027	-0.764
RWD-C	None	-0.852	-1.281	-0.976	-0.711	-0.845	-0.647

Prediction Experiments: Output

Recursive Parameter Estimates, Lags Selected and Forecast Errors in Models A and

C



Prediction Experiments: Output

<i>Model</i>	<i>RevErr</i>	<i>k = 2</i>	<i>k = 3</i>	<i>k = 4</i>	<i>k = 6</i>	<i>k = 12</i>	<i>k = 24</i>
<i>Panel A: Mean Square Forecast Errors</i>							
<i>Begin Date of Forecast Period = 1970:1</i>							
A	<i>u_{C1}</i>	0.636	0.778	0.822	0.848	0.848	0.848
A	<i>u_{C2}</i>	0.729	0.883	0.932	0.962	0.967	0.961
A	<i>u_{C3}</i>	0.681	0.830	0.873	0.889	0.887	0.883
A	<i>u_{C4}</i>	0.645	0.783	0.827	0.862	0.868	0.873
C	<i>u_{C1}</i>	0.656	0.798	0.828	0.845	0.841	0.852
C	<i>u_{C2}</i>	0.781	0.945	0.986	1.023	1.031	1.028
C	<i>u_{C3}</i>	0.618	0.749	0.781	0.827	0.804	0.818
C	<i>u_{C4}</i>	0.659	0.811	0.846	0.886	0.892	0.902
<i>Begin Date of Forecast Period = 1983:1</i>							
A	<i>u_{C1}</i>	0.220	0.268	0.278	0.310	0.332	0.364
A	<i>u_{C2}</i>	0.226	0.274	0.284	0.312	0.335	0.365
A	<i>u_{C3}</i>	0.216	0.263	0.274	0.304	0.322	0.354
A	<i>u_{C4}</i>	0.224	0.272	0.282	0.313	0.336	0.371
C	<i>u_{C1}</i>	0.247	0.301	0.307	0.335	0.366	0.393
C	<i>u_{C2}</i>	0.245	0.300	0.306	0.333	0.365	0.392
C	<i>u_{C3}</i>	0.230	0.285	0.290	0.321	0.337	0.370
C	<i>u_{C4}</i>	0.260	0.316	0.324	0.350	0.389	0.415
<i>Begin Date of Forecast Period = 1990:1</i>							
A	<i>u_{C1}</i>	0.184	0.211	0.222	0.286	0.295	0.342
A	<i>u_{C2}</i>	0.186	0.212	0.224	0.288	0.297	0.343
A	<i>u_{C3}</i>	0.176	0.205	0.217	0.279	0.288	0.332
A	<i>u_{C4}</i>	0.185	0.213	0.224	0.288	0.300	0.351
C	<i>u_{C1}</i>	0.206	0.238	0.241	0.306	0.324	0.366
C	<i>u_{C2}</i>	0.201	0.233	0.237	0.303	0.321	0.365
C	<i>u_{C3}</i>	0.192	0.227	0.231	0.292	0.292	0.340

Prediction Experiments: Prices

<i>Model</i>	<i>RevErr</i>	<i>k = 2</i>	<i>k = 3</i>	<i>k = 4</i>	<i>k = 6</i>	<i>k = 12</i>	<i>k = 24</i>
<i>Panel A: Mean Square Forecast Errors</i>							
<i>Begin Date of Forecast Period = 1970:1</i>							
A	None	0.149	0.166	0.167	0.163	0.139	0.125
B	None	0.149	0.214	0.211	0.339	0.516	0.608
C	None	0.160	0.179	0.182	0.178	0.151	0.137
R W D -A	None	0.430	0.463	0.457	0.442	0.444	0.437
R W D -B	None	0.430	0.486	0.489	0.495	0.557	0.739
R W D -C	None	0.472	0.500	0.496	0.480	0.478	0.469
<i>Begin Date of Forecast Period = 1983:1</i>							
A	None	0.072	0.076	0.075	0.069	0.063	0.063
B	None	0.072	0.077	0.076	0.088	0.257	0.556
C	None	0.075	0.079	0.079	0.073	0.066	0.065
R W D -A	None	0.395	0.376	0.374	0.362	0.349	0.360
R W D -B	None	0.395	0.414	0.416	0.394	0.380	0.303
R W D -C	None	0.512	0.488	0.486	0.474	0.458	0.469
<i>Begin Date of Forecast Period = 1990:1</i>							
A	None	0.057	0.047	0.050	0.049	0.040	0.043
B	None	0.057	0.059	0.054	0.073	0.108	0.191
C	None	0.061	0.051	0.053	0.053	0.042	0.043
R W D -A	None	0.440	0.406	0.409	0.387	0.368	0.378
R W D -B	None	0.440	0.447	0.455	0.426	0.419	0.388
R W D -C	None	0.561	0.522	0.526	0.500	0.478	0.489

Prediction Experiments: Prices

Panel B: Diebold-Mariano Test Statistics Corresponding to Entries in Panel A

Begin Date of Forecast Period = 1970:1

A	None	—	—	—	—	—	—
B	None	—	-1.419	-1.222	-3.003	-4.030	-5.085
C	None	-0.704	-0.758	-0.788	-0.826	-0.807	-0.741
RWD-A	None	-5.584	-5.526	-5.408	-5.565	-5.575	-5.926
RWD-B	None	-5.584	-5.802	-5.647	-5.648	-5.614	-5.647
RWD-C	None	-6.462	-6.544	-6.417	-6.587	-6.677	-6.936

Begin Date of Forecast Period = 1983:1

A	None	—	—	—	—	—	—
B	None	—	-0.061	-0.074	-1.589	-2.627	-3.793
C	None	-0.747	-0.782	-0.715	-0.971	-0.735	-0.452
RWD-A	None	-8.740	-8.298	-8.206	-8.827	-8.798	-8.741
RWD-B	None	-8.740	-8.787	-8.724	-9.020	-8.444	-6.042
RWD-C	None	-9.928	-9.616	-9.594	-10.408	-10.37	-10.06

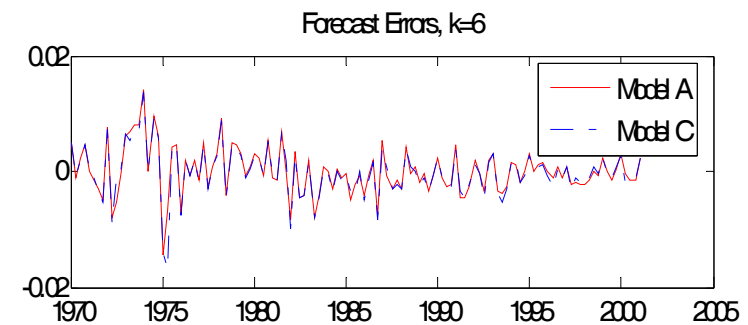
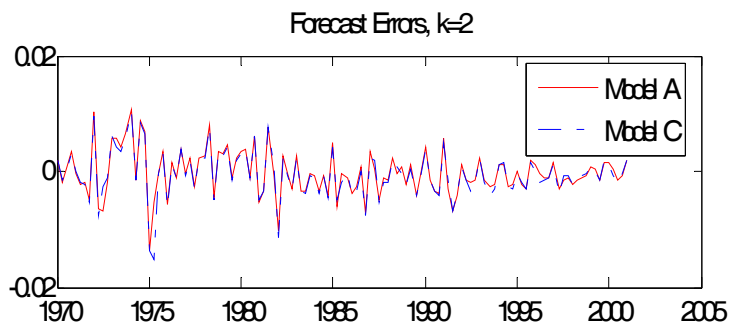
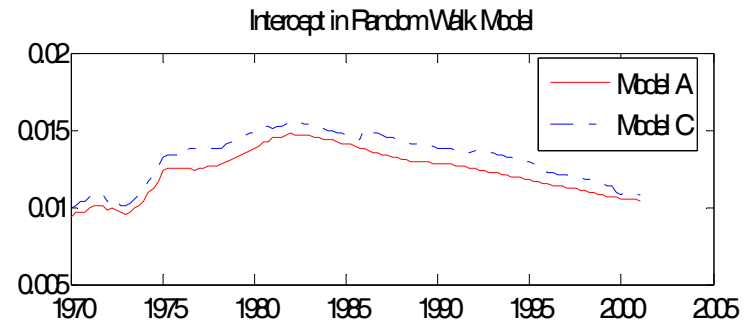
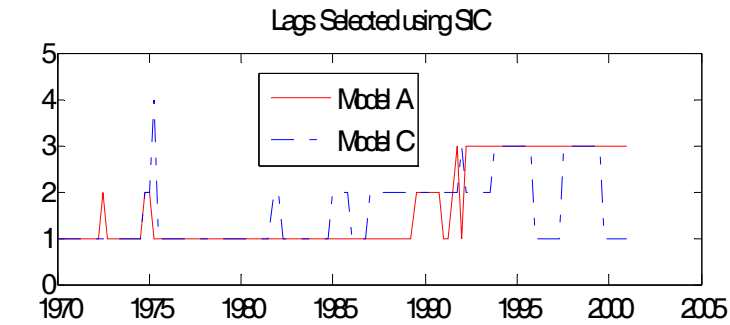
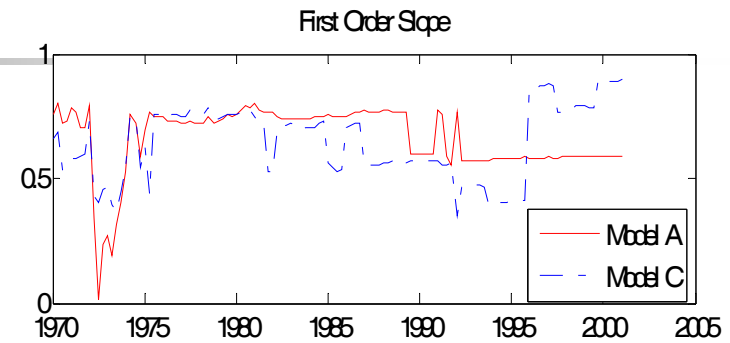
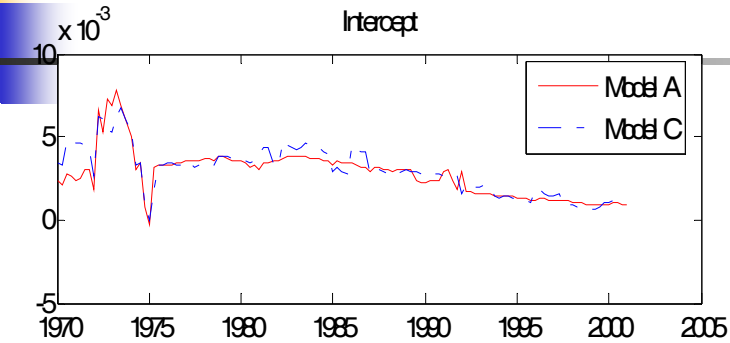
Begin Date of Forecast Period = 1990:1

A	None	—	—	—	—	—	—
B	None	—	-1.669	-0.280	-1.763	-3.278	-3.967
C	None	-0.703	-0.983	-0.429	-0.628	-0.341	-0.171
RWD-A	None	-7.913	-7.974	-7.956	-7.497	-7.640	-7.676
RWD-B	None	-7.913	-8.341	-8.378	-7.720	-7.768	-7.149
RWD-C	None	-8.813	-9.031	-9.185	-8.671	-8.850	-8.730

Prediction Experiments: Prices

Recursive Parameter Estimates, Lags Selected and Forecast Errors in Models A and

C



Prediction Experiments: Prices

<i>Model</i>	<i>RevErr</i>	<i>k = 2</i>	<i>k = 3</i>	<i>k = 4</i>	<i>k = 6</i>	<i>k = 12</i>	<i>k = 24</i>
<i>Panel A: Mean Square Forecast Errors</i>							
<i>Begin Date of Forecast Period = 1970:1</i>							
A	u_{C1}	0.145	0.163	0.164	0.161	0.136	0.124
A	u_{C2}	0.143	0.161	0.162	0.160	0.136	0.121
A	u_{C3}	0.153	0.170	0.170	0.167	0.141	0.124
A	u_{C4}	0.150	0.166	0.167	0.166	0.139	0.122
C	u_{C1}	0.155	0.174	0.177	0.173	0.146	0.132
C	u_{C2}	0.164	0.183	0.187	0.185	0.155	0.137
C	u_{C3}	0.159	0.178	0.181	0.178	0.150	0.136
C	u_{C4}	0.156	0.174	0.177	0.175	0.148	0.133
<i>Begin Date of Forecast Period = 1983:1</i>							
A	u_{C1}	0.072	0.076	0.075	0.069	0.063	0.063
A	u_{C2}	0.070	0.073	0.073	0.069	0.064	0.061
A	u_{C3}	0.072	0.076	0.076	0.070	0.064	0.064
A	u_{C4}	0.073	0.076	0.076	0.071	0.062	0.060
C	u_{C1}	0.076	0.081	0.081	0.076	0.069	0.068
C	u_{C2}	0.077	0.081	0.082	0.080	0.074	0.069
C	u_{C3}	0.080	0.084	0.083	0.079	0.071	0.069
C	u_{C4}	0.078	0.082	0.081	0.077	0.070	0.067
<i>Begin Date of Forecast Period = 1990:1</i>							
A	u_{C1}	0.057	0.046	0.050	0.049	0.041	0.042
A	u_{C2}	0.053	0.043	0.048	0.048	0.042	0.043
A	u_{C3}	0.058	0.047	0.051	0.050	0.041	0.043
A	u_{C4}	0.057	0.045	0.050	0.049	0.041	0.044
C	u_{C1}	0.063	0.053	0.055	0.055	0.045	0.045
C	u_{C2}	0.060	0.050	0.056	0.056	0.048	0.048
C	u_{C3}	0.066	0.055	0.057	0.057	0.046	0.046
C	u_{C4}	0.065	0.054	0.056	0.057	0.046	0.047

Prediction Experiments: Money

<i>Model</i>	<i>RevErr</i>	<i>k = 2</i>	<i>k = 3</i>	<i>k = 4</i>	<i>k = 6</i>	<i>k = 12</i>	<i>k = 24</i>
--------------	---------------	--------------	--------------	--------------	--------------	---------------	---------------

Panel A: Mean Square Forecast Errors

Begin Date of Forecast Period = 1970:1

A	None	1.030	1.024	1.008	0.937	0.984	0.998
B	None	1.030	1.253	1.284	1.253	2.000	1.692
C	None	1.025	1.016	0.999	0.918	0.948	0.971
RWD-A	None	1.633	1.630	1.612	1.556	1.565	1.567
RWD-B	None	1.633	1.650	1.646	1.611	1.666	1.752
RWD-C	None	1.650	1.643	1.623	1.560	1.562	1.561

Begin Date of Forecast Period = 1983:1

A	None	1.168	1.140	1.099	1.007	1.022	1.040
B	None	1.168	1.366	1.475	1.682	2.824	2.308
C	None	1.117	1.088	1.042	0.949	0.955	0.979
RWD-A	None	2.257	2.235	2.199	2.140	2.142	2.148
RWD-B	None	2.257	2.277	2.281	2.285	2.344	2.314
RWD-C	None	2.323	2.300	2.260	2.201	2.198	2.208

Begin Date of Forecast Period = 1990:1

A	None	1.058	1.058	1.035	0.916	0.907	0.955
B	None	1.058	1.155	1.219	1.591	3.001	2.480
C	None	1.023	1.023	0.995	0.877	0.861	0.915
RWD-A	None	2.383	2.395	2.387	2.319	2.279	2.312
RWD-B	None	2.383	2.452	2.499	2.541	2.589	2.487
RWD-C	None	2.539	2.551	2.542	2.475	2.431	2.464

Prediction Experiments: Money

Panel B: Diebold-Mariano Test Statistics Corresponding to Entries in Panel A

Begin Date of Forecast Period = 1970:1

A	None	—	—	—	—	—	—
B	None	—	-2.499	-2.621	-2.391	-3.486	-3.087
C	None	0.155	0.252	0.301	0.666	1.251	0.914
RWD-A	None	-3.635	-3.643	-3.624	-3.671	-3.379	-3.268
RWD-B	None	-3.635	-3.630	-3.583	-3.520	-3.395	-4.090
RWD-C	None	-3.419	-3.390	-3.370	-3.364	-3.078	-2.958

Begin Date of Forecast Period = 1983:1

A	None	—	—	—	—	—	—
B	None	—	-2.099	-2.444	-3.821	-4.059	-3.758
C	None	1.423	1.474	1.622	1.805	1.917	1.802
RWD-A	None	-4.404	-4.439	-4.466	-4.577	-4.482	-4.380
RWD-B	None	-4.404	-4.455	-4.504	-4.562	-4.507	-4.695
RWD-C	None	-4.266	-4.279	-4.294	-4.392	-4.317	-4.232

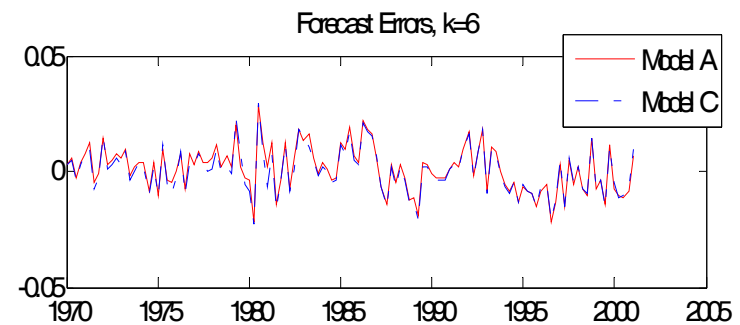
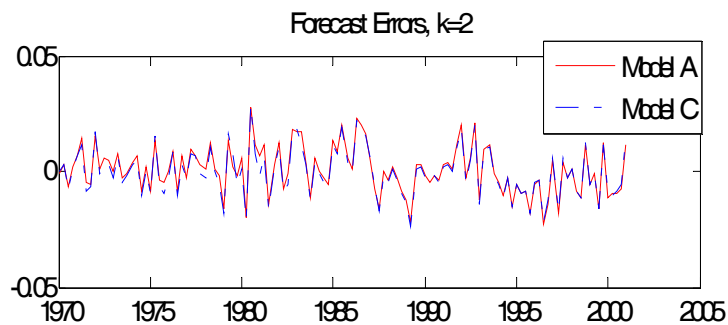
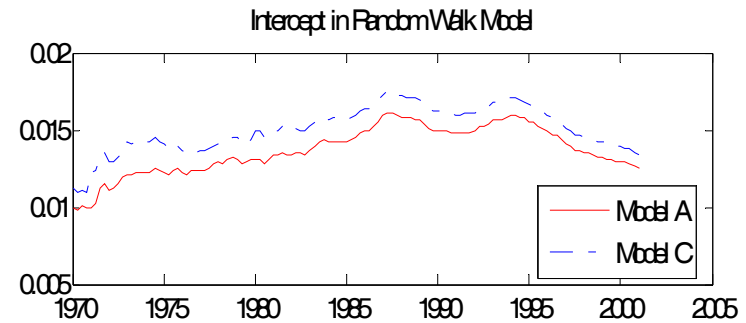
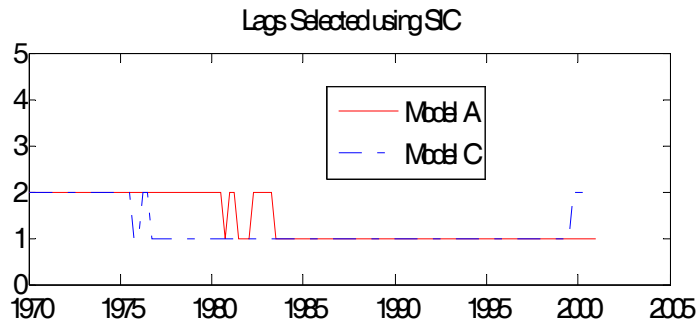
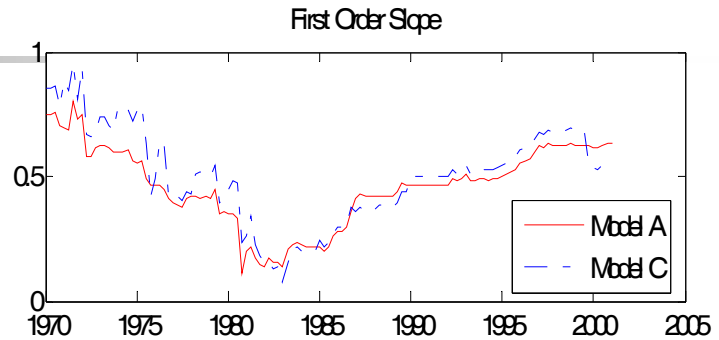
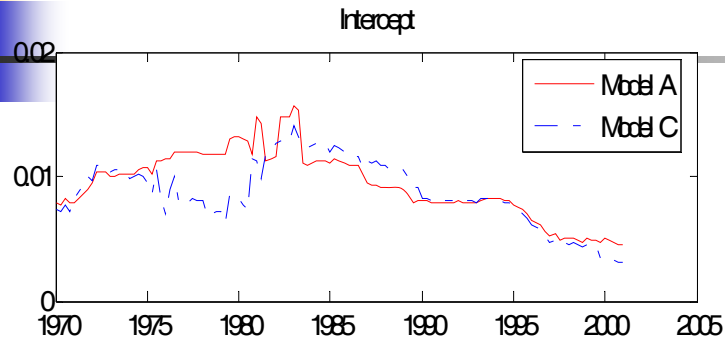
Begin Date of Forecast Period = 1990:1

A	None	—	—	—	—	—	—
B	None	—	-0.710	-1.014	-2.736	-3.411	-3.222
C	None	1.264	1.264	1.366	1.777	1.788	1.553
RWD-A	None	-3.770	-3.803	-3.840	-3.969	-3.850	-3.717
RWD-B	None	-3.770	-3.825	-3.892	-4.042	-4.001	-4.080
RWD-C	None	-3.796	-3.825	-3.858	-3.974	-3.874	-3.753

Prediction Experiments: Money

Recursive Parameter Estimates, Lags Selected and Forecast Errors in Models A and

C



Prediction Experiments: Money

<i>Model</i>	<i>RevErr</i>	<i>k = 2</i>	<i>k = 3</i>	<i>k = 4</i>	<i>k = 6</i>	<i>k = 12</i>	<i>k = 24</i>
<i>Panel A: Mean Square Forecast Errors</i>							
<i>Begin Date of Forecast Period = 1970:1</i>							
A	<i>u_{C1}</i>	1.035	1.030	1.013	0.937	0.982	0.992
A	<i>u_{C2}</i>	1.030	1.026	1.012	0.934	0.979	0.992
A	<i>u_{C3}</i>	1.030	1.024	1.008	0.935	0.983	1.001
A	<i>u_{C4}</i>	1.043	1.038	1.020	0.944	0.988	1.001
C	<i>u_{C1}</i>	1.037	1.029	1.014	0.935	0.961	0.979
C	<i>u_{C2}</i>	1.044	1.036	1.023	0.942	0.969	0.991
C	<i>u_{C3}</i>	1.040	1.030	1.012	0.933	0.963	0.983
C	<i>u_{C4}</i>	1.075	1.067	1.050	0.968	0.992	1.001
<i>Begin Date of Forecast Period = 1983:1</i>							
A	<i>u_{C1}</i>	1.161	1.135	1.093	0.992	1.010	1.033
A	<i>u_{C2}</i>	1.147	1.121	1.080	0.984	1.002	1.031
A	<i>u_{C3}</i>	1.157	1.130	1.089	0.998	1.011	1.036
A	<i>u_{C4}</i>	1.163	1.137	1.095	0.994	1.012	1.035
C	<i>u_{C1}</i>	1.117	1.089	1.043	0.948	0.958	0.984
C	<i>u_{C2}</i>	1.112	1.084	1.039	0.947	0.956	0.987
C	<i>u_{C3}</i>	1.119	1.089	1.042	0.950	0.960	0.985
C	<i>u_{C4}</i>	1.121	1.093	1.046	0.951	0.960	0.987
<i>Begin Date of Forecast Period = 1990:1</i>							
A	<i>u_{C1}</i>	1.064	1.065	1.042	0.905	0.897	0.950
A	<i>u_{C2}</i>	1.051	1.051	1.028	0.892	0.884	0.937
A	<i>u_{C3}</i>	1.050	1.049	1.025	0.907	0.896	0.948
A	<i>u_{C4}</i>	1.068	1.069	1.046	0.907	0.900	0.953
C	<i>u_{C1}</i>	1.028	1.028	1.002	0.875	0.861	0.916
C	<i>u_{C2}</i>	1.021	1.020	0.996	0.870	0.857	0.911
C	<i>u_{C3}</i>	1.028	1.025	0.997	0.877	0.864	0.918
C	<i>u_{C4}</i>	1.031	1.031	1.005	0.877	0.862	0.918

Empirical Results

Summary of Prediction Experiment Findings

- For prices, Model A wins over C, for any k . The "cost" of using mixed releases of data, when the objective is to predict a particular release, outweighs the "costs" associated with "mild" inefficiency and/or definitional change problems.
- For money, Model C wins over Model A, for any k . This suggests that an important role for our new efficiency test is to distinguish between efficiency problems and definitional change problems. In particular, if our efficiency test rejects, we know that one reason for rejection may actually be a failure of the maintained assumption of no definitional change. Thus, if the extant efficiency tests from the literature discussed in the previous section all accept the efficiency null, while our test rejects, then we have evidence that definitional change may be driving our test rejections.
- For output, the evidence can be mixed. Model A "wins" when predicting early releases, while Model C "wins" when predicting later releases.

Empirical Results

Summary of Prediction Experiment Findings

- As expected, Model B performs poorly, as is particularly bad for larger values of k
- MSFEs associated with the "best" models for money (i.e. Model C) largely decrease as k increases. This is consistent with the view that using real-time data that has been revised as much as possible, when forming model coefficient estimates, leads to estimates that are more accurate, when the objective is the prediction of later release or even "final" data. This in turn suggests that later releases of data should be predicted more accurately using Model C (as is indeed the case for money, where Model C is actually the MSFE-best model), but not necessarily when using other models. Indeed, notice that for prices, where Model A wins, the MSFEs actually increase as one increases k from 2 to 3 to 4, before beginning to decrease.

Empirical Results

Summary of Efficiency Test Findings

- Recall that the approach for early release rationality involves testing for inefficiency via examining

$$H_0 : MSFE_A = MSFE_C , \quad \forall k = 2, 3, \dots$$

- From the DM-statistics, only statistics comparing Models A and C that are large in absolute value are those associated with money ... the only series for which our test rejects the null is money.
- This results should be assessed in light of the rationality tests results reported for extant tests from the literature...

Empirical Results

Summary of Efficiency Test Findings

Output	Linear Regression Test		
F-Test Statistic	5% CV	10% CV	
2.893	3.053	2.336	
CCS Test & Chi Square Critical Values			
Statistic	5% CV	10% CV	
0.000078	0.000044	0.000030	
CFS Test Statistic & Bootstrap Critical Values			
Statistic	5%	10%	
0.023944	0.011842	0.009845	

Prices	Linear Regression Test		
F-Test Statistic	5% CV	10% CV	
1.851	3.053	2.336	
CCS Test & Chi Square Critical Values			
Statistic	5% CV	10% CV	
0.000046	0.000030	0.000026	
CFS Test Statistic & Bootstrap Critical Values			
Statistic	5% CV	10% CV	
0.056355	0.043220	0.034450	

Money	Linear Regression Test		
F-Test Statistic	5% CV	10% CV	
-0.52381	3.053	2.336	
CCS Test & Chi Square Critical Values			
Statistic	5% CV	10% CV	
0.000021	0.000042	0.000032	
CFS Test Statistic & Bootstrap Critical Values			
Statistic	5% CV	10% CV	
0.002549	0.015691	0.011907	

Empirical Results

Summary of Efficiency Test Findings

- ...since these tests strongly fail to reject the null of rationality for money, our test is likely rejecting because of definitional change problems associated with using only first release data in Model A.
- DM tests for prices and output fail to reject null of efficiency, but extant tests find inefficiency. Consider prices - there may still be "mild inefficiency" in the data if costs associated with using Model C (i.e. the costs of using mixed releases of data) are sufficiently great, and if there is little problem with definitional change. The reason for this is that in such cases, costs associated with using mixed releases (Model C) can be approximately offset by costs of using relatively inefficient early releases of data when predicting later releases (Model A). In such cases, neither model can significantly "beat" the other. However, this does not mean that one model may not yield smaller point MSFEs, as this indeed happens in the case of prices, for all values of k , with Model A. Table A1.2 confirms that prices are probably "mildly inefficient", as the null of rationality fails to reject based on the linear rationality test implemented using an F-statistic, but is rejected based upon implementation of the CCS and the CFS tests. Moreover, we have additional evidence that prices are indeed "mildly inefficient". Namely, including revision errors (Table A1.2) suggest that MSFEs are almost always slightly better (i.e. u_{C1} and u_{C2}) are included.

Empirical Results

Summary of Efficiency Test Findings

- Taken together, then, we have evidence that definitional changes are an issue in our money series, while efficiency problems are prevalent for our price and output series. Moreover, the way in which these different problems combines in the data determines whether Model A or Model C dominates, when forming predictions for any given release of a variable.
- We thus have strong evidence as to the importance and informational content of real-time datasets.

Empirical Results

Real-Time Marginal Predictive Content of Money for Output

Model A

$${}_{t+k}Y_{t+1} = \alpha + \sum_{i=1}^p \beta_{i,t+2-i}^{A,Y} Y_{t+1-i} + \theta [{}_{t+1}Y_{t-1} - {}_tY_{t-1}] + \sum_{i=1}^p \beta_{i,t+2-i}^{A,P} P_{t+1-i} + \sum_{i=1}^p \beta_{i,t+2-i}^{A,M} M_{t+1-i} + \sum_{i=1}^p \beta_{i,t+2-i}^{A,R} R_{t+1-i} + \theta'_{t+1} W_t + \varepsilon_{t+k}$$

Model C

$${}_{t+k}Y_{t+1} = \alpha + \sum_{i=1}^p \beta_{i,t+1}^{C,Y} Y_{t+1-i} + \theta [{}_{t+1}Y_{t-1} - {}_tY_{t-1}] + \sum_{i=1}^p \beta_{i,t+1}^{C,P} P_{t+1-i} + \sum_{i=1}^p \beta_{i,t+1}^{C,M} M_{t+1-i} + \sum_{i=1}^p \beta_{i,t+1}^{C,R} R_{t+1-i} + \theta'_{t+1} W_t + \varepsilon_{t+k},$$

Empirical Results

Real-Time Marginal Predictive Content of Money for Output

<i>Model</i>	<i>RevErr</i>	<i>k = 2</i>	<i>k = 3</i>	<i>k = 4</i>	<i>k = 6</i>	<i>k = 12</i>	<i>k = 24</i>
<i>Panel A: Mean Square Forecast Errors</i>							
<i>Begin Date of Forecast Period =1971:2</i>							
VAR - NoM Mod A	None	0.866	1.004	1.031	1.049	1.088	1.083
VAR - M Mod A	None	1.069*	1.196*	1.190*	1.200*	1.226*	1.235*
VAR - NoM Mod C	None	0.889	1.027	1.054	1.049	1.074	1.042
VAR - M Mod C	None	1.499	1.629	1.639	1.628	1.662	1.602
VAR - NoM Mod A	u_{C1}	0.719	0.856	0.876	0.891	0.929	0.916
VAR - M Mod A	u_{C1}	1.172	1.296	1.285	1.296	1.330	1.355
VAR - NoM Mod A	u_{C2}	1.036	1.168	1.219	1.235	1.281	1.271
VAR - M Mod A	u_{C2}	1.562	1.691	1.706	1.708	1.731	1.716
VAR - NoM Mod C	u_{C1}	0.840	0.982	1.006	1.009	1.037	1.014
VAR - M Mod C	u_{C1}	2.213	2.322	2.362	2.356	2.406	2.303
VAR - NoM Mod C	u_{C2}	1.107	1.267	1.317	1.327	1.372	1.341
VAR - M Mod C	u_{C2}	1.334	1.499	1.489	1.463	1.482	1.466

Empirical Results

Real-Time Marginal Predictive Content of Money for Output

Begin Date of Forecast Period = 1983:1

VAR - NoM Mod A	None	0.270	0.315*	0.322	0.342	0.338	0.366
VAR - M Mod A	None	0.324*	0.378*	0.377*	0.396	0.389	0.409
VAR - NoM Mod C	None	0.274	0.323	0.326	0.334	0.345	0.364*
VAR - M Mod C	None	0.466	0.524	0.509	0.511	0.500	0.519
VAR - NoM Mod A	u_{C1}	0.269*	0.315*	0.322	0.340	0.339	0.366
VAR - M Mod A	u_{C1}	0.324*	0.378*	0.378	0.396	0.391	0.410
VAR - NoM Mod A	u_{C2}	0.293	0.341	0.348	0.361	0.358	0.382
VAR - M Mod A	u_{C2}	0.390	0.451	0.454	0.447	0.438	0.447
VAR - NoM Mod C	u_{C1}	0.269*	0.319	0.320	0.328	0.346	0.364
VAR - M Mod C	u_{C1}	0.454	0.512	0.496	0.505	0.504	0.518
VAR - NoM Mod C	u_{C2}	0.278	0.325	0.328	0.336	0.354	0.371
VAR - M Mod C	u_{C2}	0.516	0.585	0.567	0.562	0.546	0.561

Empirical Results

Real-Time Marginal Predictive Content of Money for Output

Begin Date of Forecast Period = 1990:1

VAR - NoM Mod A	None	0.227*	0.243	0.253	0.313	0.316	0.343
VAR - M Mod A	None	0.241*	0.252	0.260	0.316	0.319	0.344
VAR - NoM Mod C	None	0.242	0.266	0.265	0.313	0.333	0.344
VAR - M Mod C	None	0.319	0.334	0.329	0.366	0.386	0.392
VAR - NoM Mod A	u_{C1}	0.230	0.245	0.255	0.315	0.318	0.347
VAR - M Mod A	u_{C1}	0.241*	0.253	0.261	0.317	0.319	0.346
VAR - NoM Mod A	u_{C2}	0.239	0.254	0.265	0.323	0.326	0.351
VAR - M Mod A	u_{C2}	0.253	0.264	0.273	0.327	0.329	0.351
VAR - NoM Mod C	u_{C1}	0.235	0.258	0.258	0.308	0.329	0.343
VAR - M Mod C	u_{C1}	0.326	0.338	0.333	0.378	0.404	0.413
VAR - NoM Mod C	u_{C2}	0.246	0.270	0.270	0.316	0.336	0.343
VAR - M Mod C	u_{C2}	0.332	0.349	0.344	0.387	0.412	0.417

Empirical Results

Real-Time Marginal Predictive Content of Money for Output

- Upon inspection of Table 6, it is clear that it is always the case that (regardless of sample period, model, and vintage) the models with money yield higher MSFEs than the models without money (entries that are in bold denote MSFE-best models). This result holds for Models A and C, regardless of whether or not revision errors are included.
- For our longer prediction periods beginning in 1971 and 1983, money does not appear “adequate”, as the AR model yields significantly more accurate predictions. However, for the shortest forecast period from 1990, the MSFE-best models with money are adequate for all releases except first release data.

Concluding Remarks

- Real-time datasets do matter, regardless of release to be predicted.
- Release to be predicted has no impact on variety of real-time model to be used for prediction.
- Revision error process has information for at least one of our variables.
- Definitional change is crucial to understanding interaction between data release used in prediction model construction and forecast accuracy.
- Our efficiency test rejects for money, suggesting that definitional change is an issue for this series.
- Prices mildly inefficient, although Model A always MSFE-best, suggesting little problem with definitional change and bigger problems with using mixed releases of data as in Model C.
- Output appears also mildly inefficient, but evidence more mixed.