

Data Revisions are not Well-Behaved

S. Boragan Aruoba

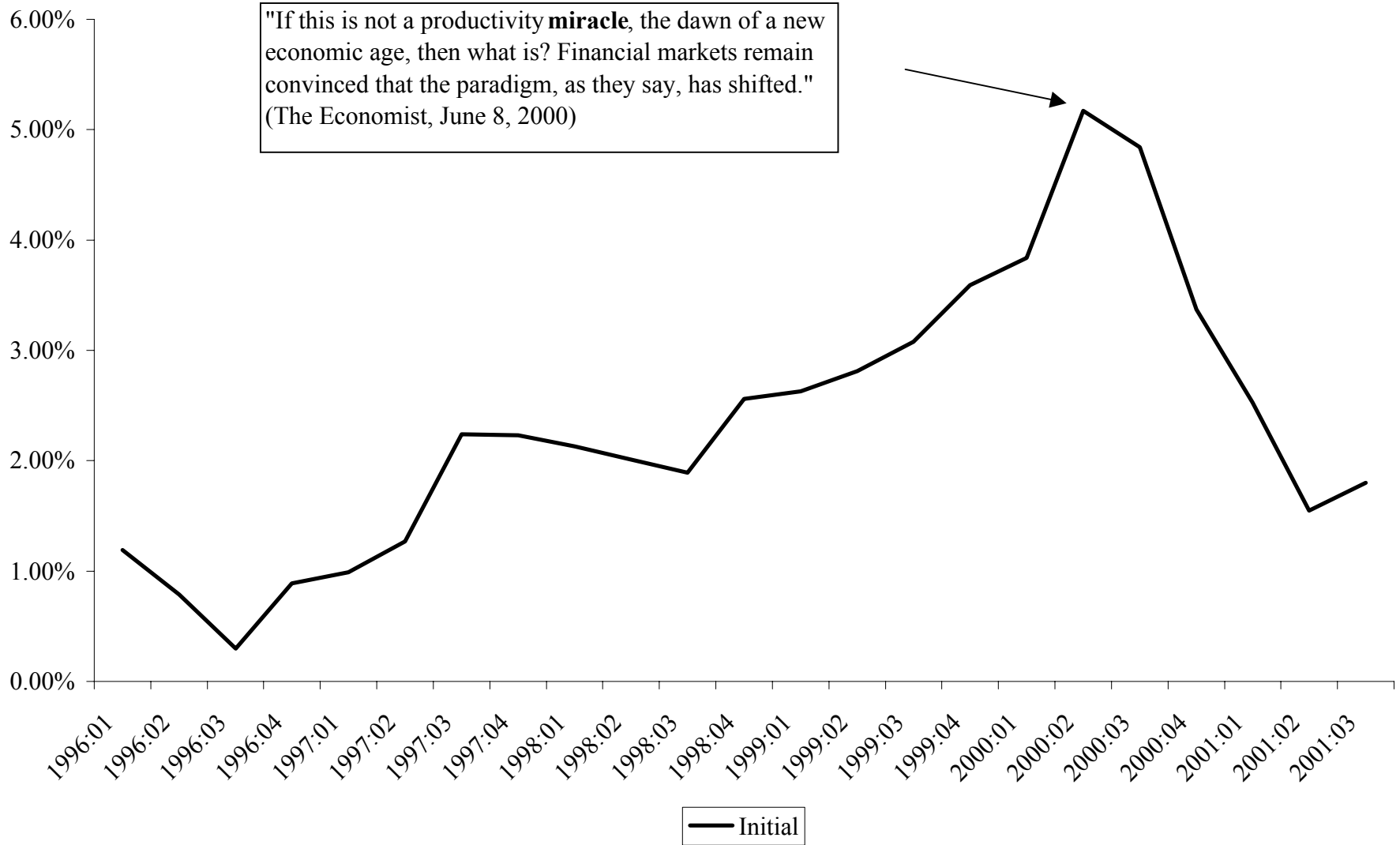
University of Maryland

June 14, 2005

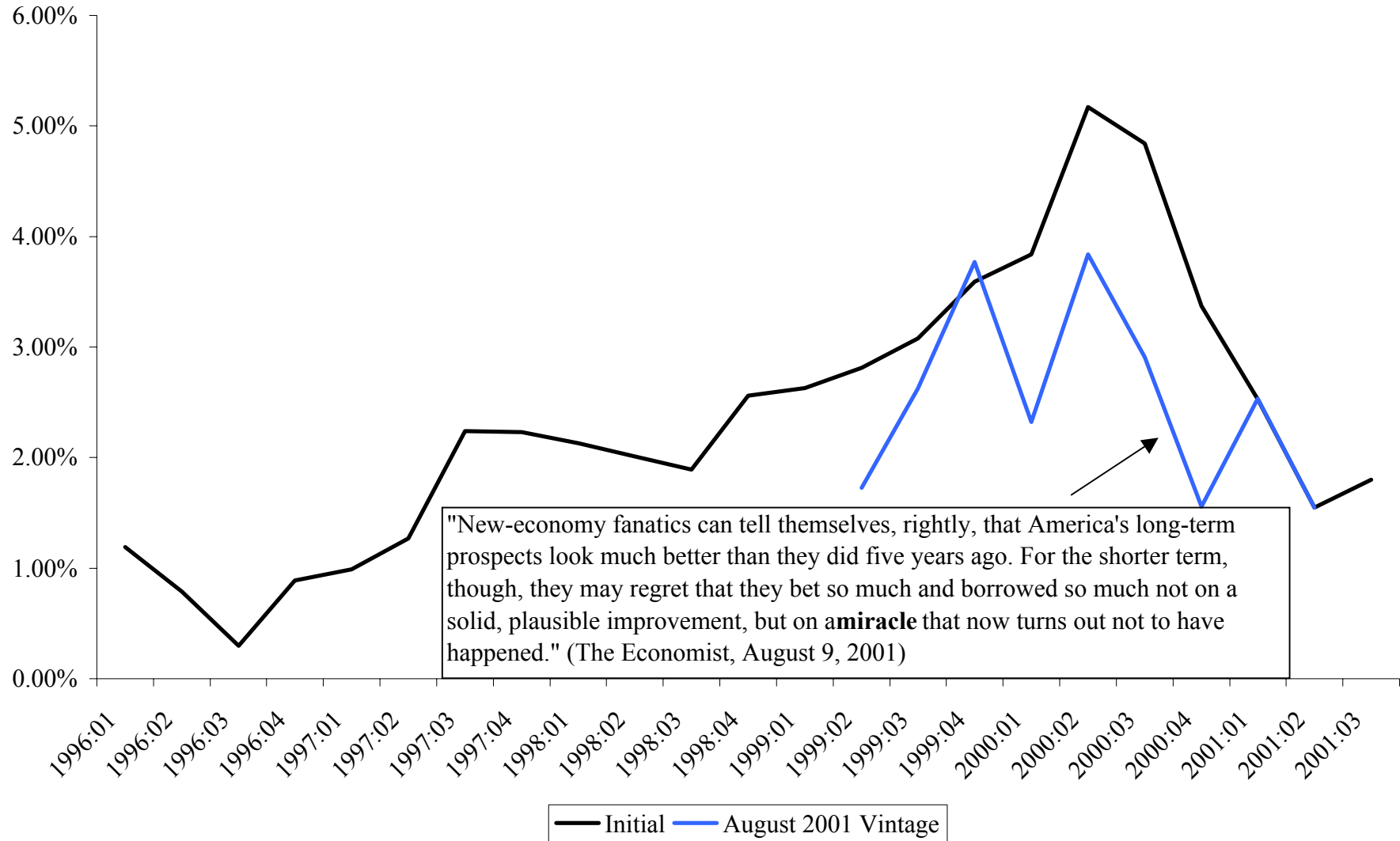
EABCN RTBD Workshop

Brussels, Belgium

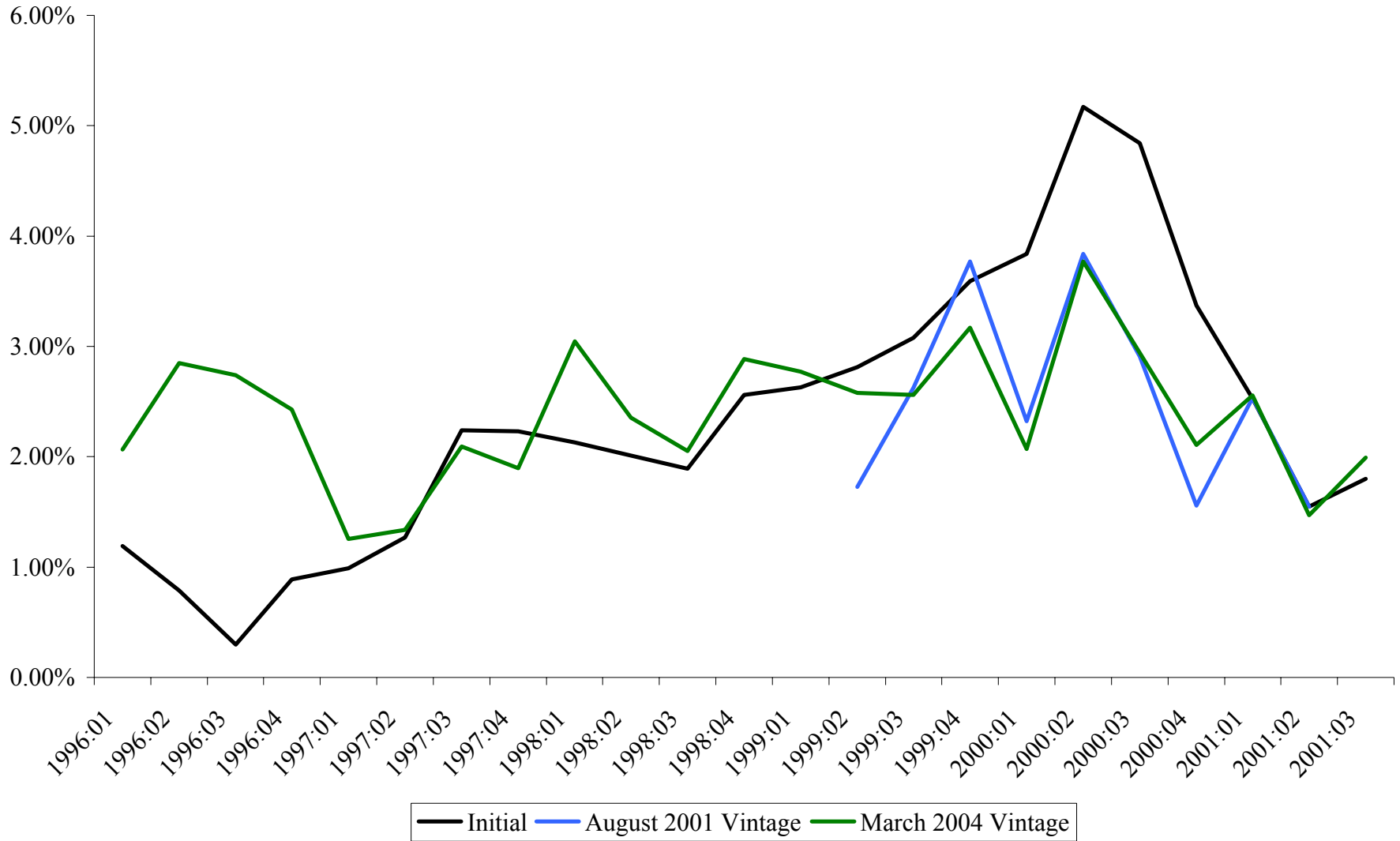
Labor Productivity Growth "Miracle" in the Second Half of 1990s



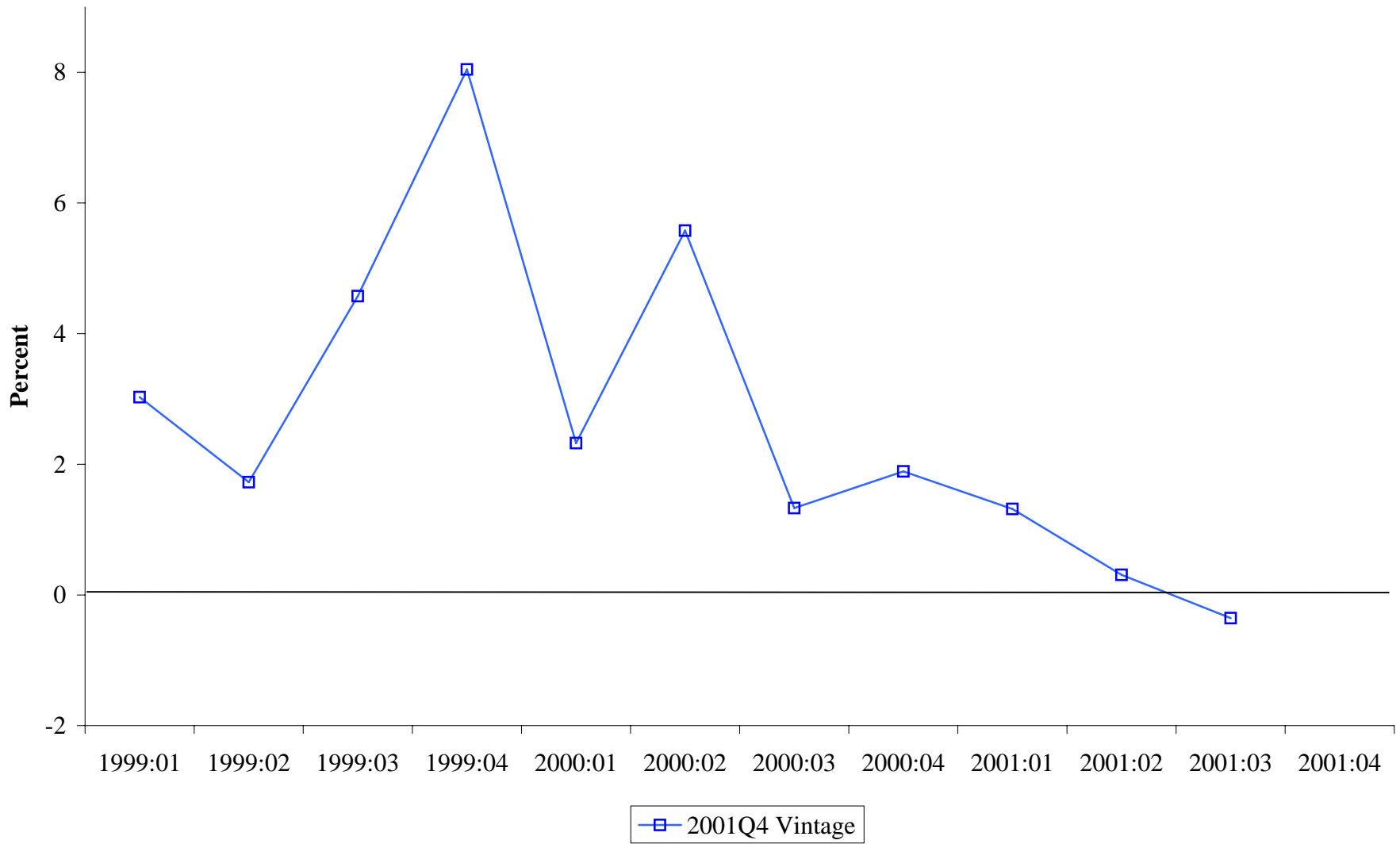
Labor Productivity Growth "Miracle" in the Second Half of 1990s



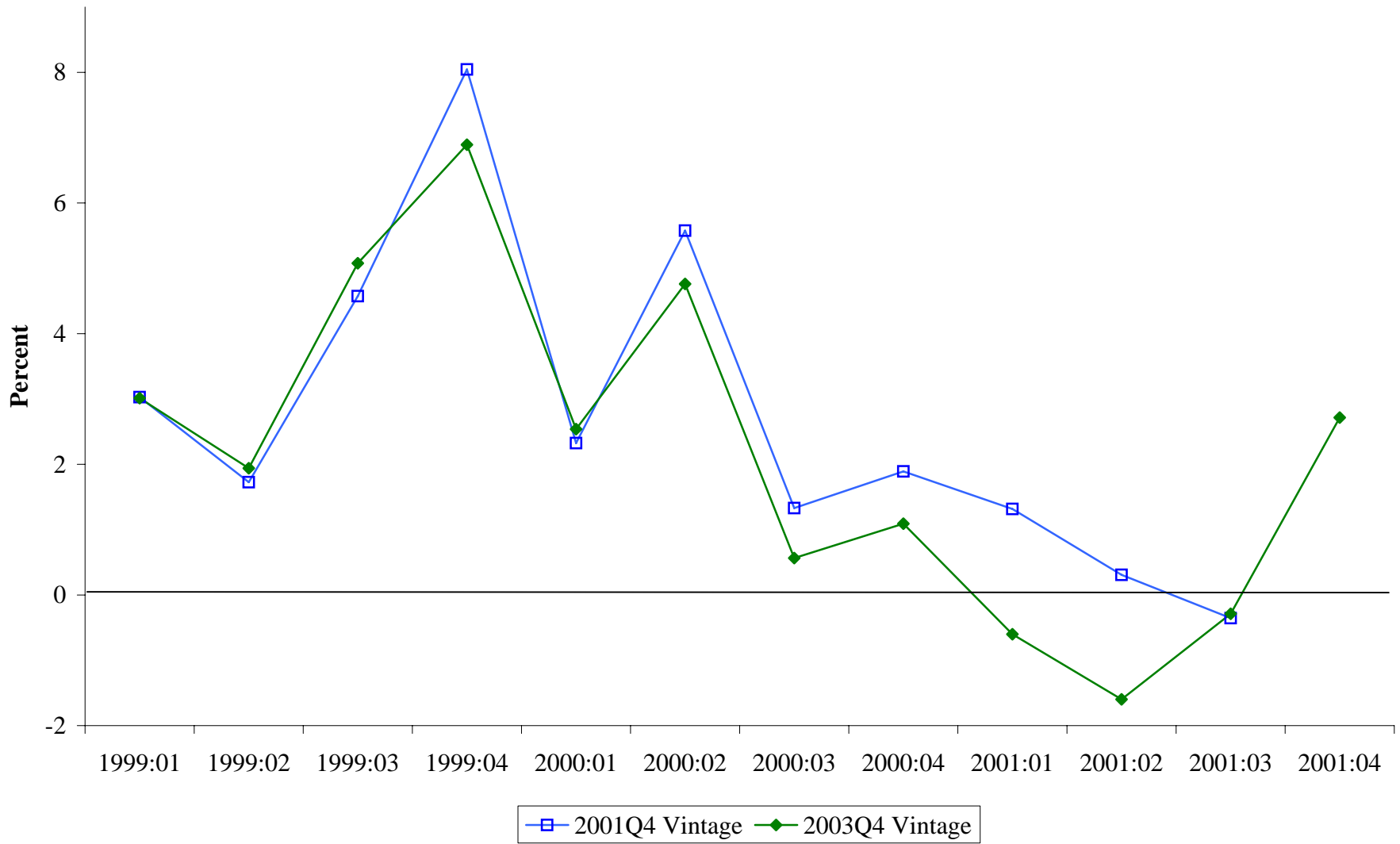
Labor Productivity Growth "Miracle" in the Second Half of 1990s



Quarterly Real GDP Growth (Different Vintages / Annualized)



Quarterly Real GDP Growth (Different Vintages / Annualized)



Question

$$\underbrace{y_t^f}_{\text{True Value}} = \underbrace{y_t^{t+1}}_{\text{Initial Announcement}} + \underbrace{r_t^f}_{\text{Final Revision}}$$

Question

$$\underbrace{y_t^f}_{\text{True Value}} = \underbrace{y_t^{t+1}}_{\text{Initial Announcement}} + \underbrace{r_t^f}_{\text{Final Revision}}$$

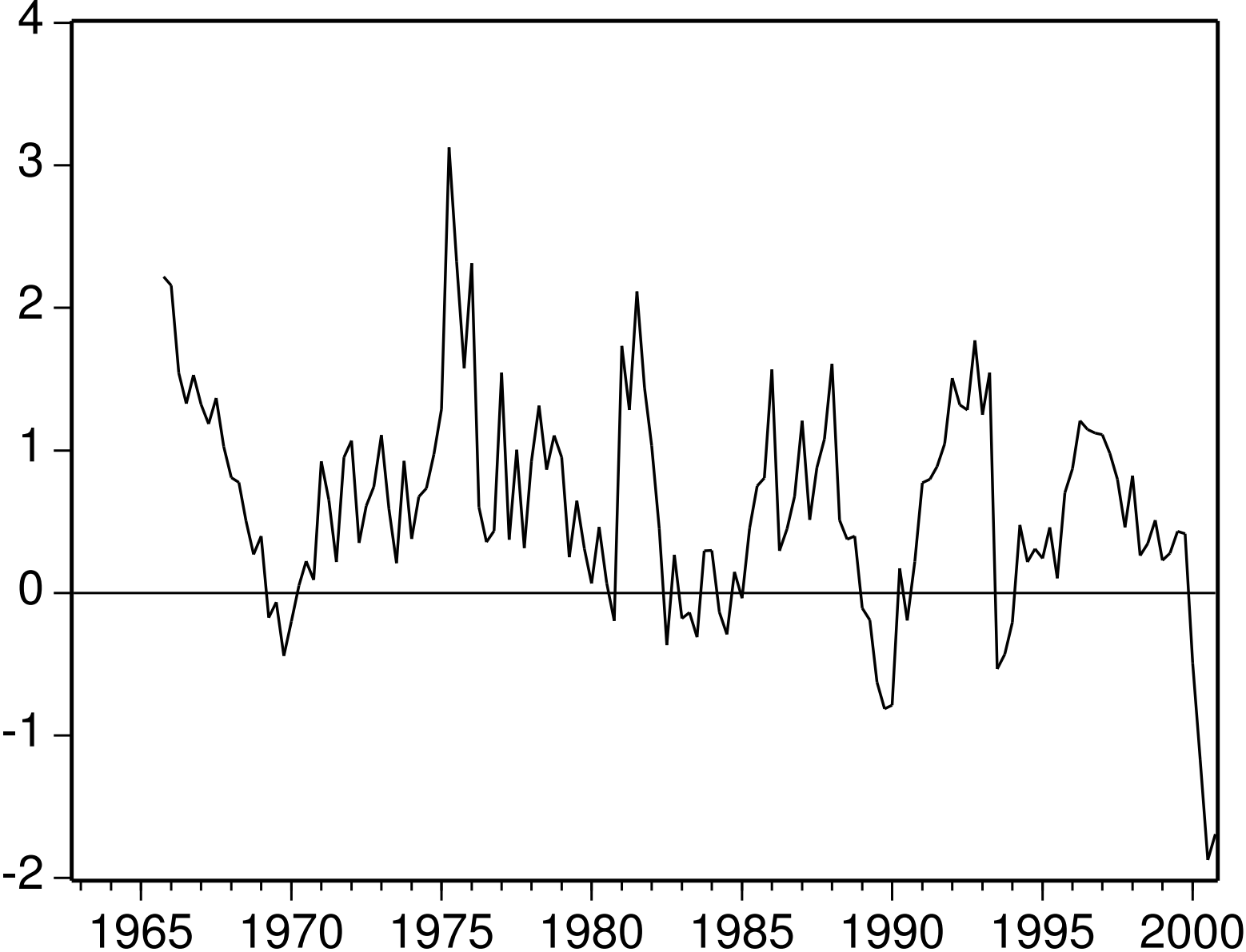
“Well-behaved” data revisions must satisfy three properties:

$$(P1) : E \left(r_t^f \right) = 0$$

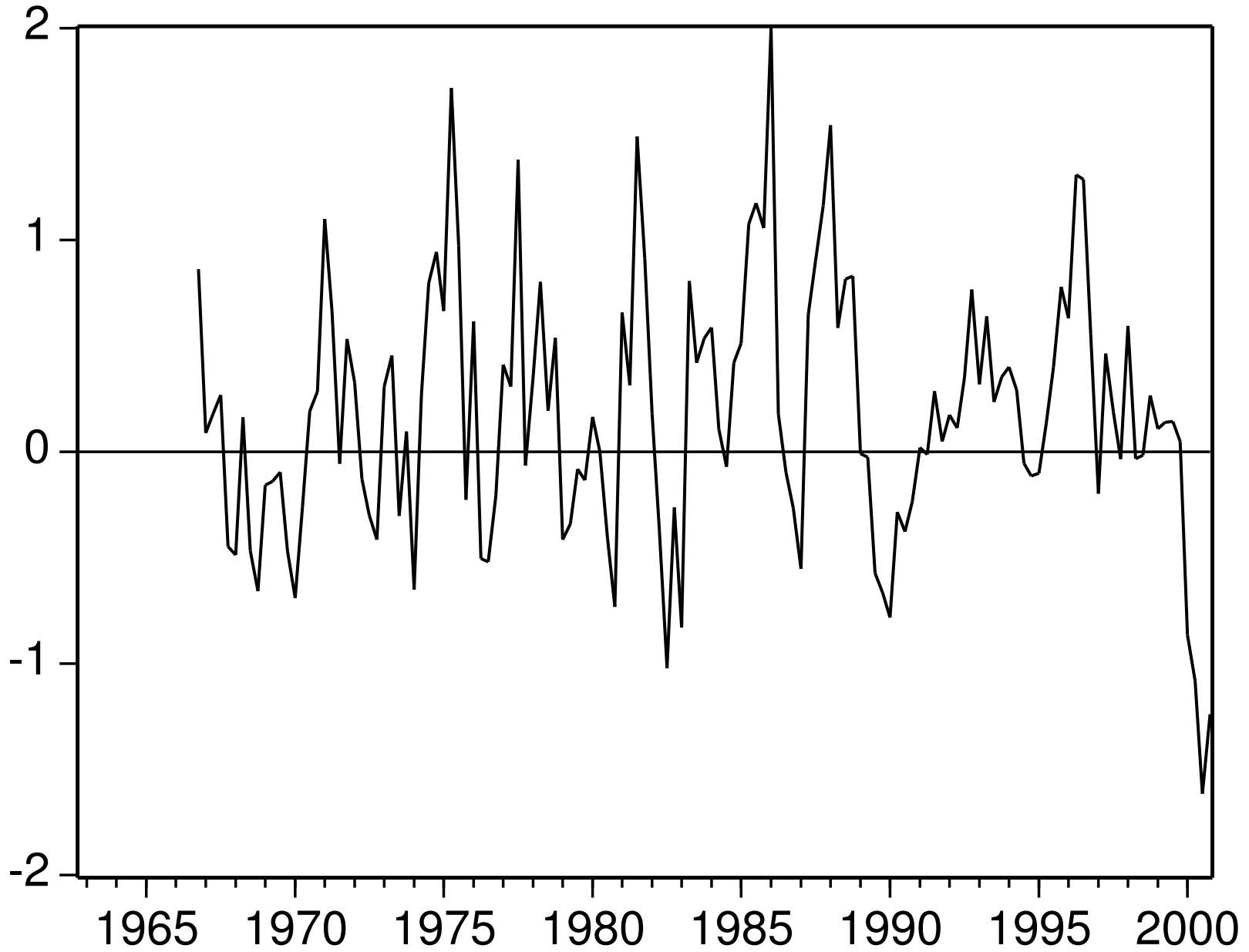
$$(P2) : \text{var} \left(r_t^f \right) \text{ is small}$$

$$(P3) : E \left(r_t^f | I_{t+1} \right) = 0$$

Final Revision to Nominal Output (As Percentage of Initial Announcement)



Final Revisions to Annual Growth of Real Output



Summary of Findings

- Positive mean revision
- Large revisions
- Predictable revisions
- Robust (if not stronger) in subsamples
- Evidence that private forecasters ignore predictability

Data Sources

- Real-time Data Set (Federal Reserve Bank of Philadelphia)
- Own Data Set for Labor Productivity

Data Sources

- Real-time Data Set (Federal Reserve Bank of Philadelphia)
- Own Data Set for Labor Productivity

Variable Types

- Quarterly / Monthly
- “Relative” / “Level”

Data Sources

- Real-time Data Set (Federal Reserve Bank of Philadelphia)
- Own Data Set for Labor Productivity

Variable Types

- Quarterly / Monthly
- “Relative” / “Level”

Defining Revisions

- Cumulative Revision after h periods : $r_t^{(1)h} = y_t^{(t+1)+h} - y_t^{t+1}$
- Relative Cumulative Revision after h periods : $r_t^{(2)h} = \frac{r_t^{(1)h}}{y_t^{t+1}}$

Data Sources

- Real-time Data Set (Federal Reserve Bank of Philadelphia)
- Own Data Set for Labor Productivity

Variable Types

- Quarterly / Monthly
- “Relative” / “Level”

Defining Revisions

- Cumulative Revision after h periods : $r_t^{(1)h} = y_t^{(t+1)+h} - y_t^{t+1}$
- Relative Cumulative Revision after h periods : $r_t^{(2)h} = \frac{r_t^{(1)h}}{y_t^{t+1}}$

Removing “Uninformative” Revisions

Data Sources

- Real-time Data Set (Federal Reserve Bank of Philadelphia)
- Own Data Set for Labor Productivity

Variable Types

- Quarterly / Monthly
- “Relative” / “Level”

Defining Revisions

- Cumulative Revision after h periods : $r_t^{(1)h} = y_t^{(t+1)+h} - y_t^{t+1}$
- Relative Cumulative Revision after h periods : $r_t^{(2)h} = \frac{r_t^{(1)h}}{y_t^{t+1}}$

Removing “Uninformative” Revisions

Defining the Final Revision

Table I - Summary Statistics of Final Revisions - Full Sample

	N	Mean	Minimum	Maximum	Std. Dev.	Noise / Signal	Corr. with Initial	A/C (1)	Q-stat (20)
Revisions as Percentage of Initial Announcements									
Nominal Output	141	0.60%	-1.87%	3.12%	0.75%	-	-	0.68	0.00
Real Output	141	0.43%	-1.74%	2.82%	0.79%	-	-	0.72	0.00
Non-Farm Payroll Employment	422	0.26%	-1.08%	2.22%	0.70%	-	-	0.95	0.00
Industrial Production Index (Total Industry)	445	0.94%	-3.52%	7.17%	1.86%	-	-	0.94	0.00
Industrial Production Index (Manufacturing)	247	0.74%	-4.13%	2.33%	2.33%	-	-	0.95	0.00
Level of Revisions									
Annual Growth of Real Output	137	0.18%	-1.62%	2.00%	0.61%	0.24	-0.14	0.55	0.00
Annual Growth of Nominal Output	137	0.29%	-1.74%	2.57%	0.67%	0.25	-0.01	0.60	0.00
Annual Inflation (Output Deflator)	137	0.10%	-0.65%	1.12%	0.33%	0.13	0.32	0.67	0.00
Annual Growth of Labor Productivity	123	0.35%	-3.12%	3.34%	1.34%	0.82	-0.51	0.67	0.00
Annual Growth of Non-Farm Payroll Employment	414	0.15%	-0.83%	1.22%	0.39%	0.14	0.23	0.92	0.00
Annual Growth of Industrial Production (Total Industry)	433	0.48%	-2.66%	5.40%	0.80%	0.20	0.06	0.82	0.00
Annual Growth of Industrial Production (Manufacturing)	235	0.56%	-2.48%	2.93%	1.11%	0.23	-0.10	0.83	0.00
Quarterly Growth of Real Output	137	0.25%	-2.85%	5.12%	1.51%	0.41	-0.02	-0.15	0.15
Quarterly Growth of Nominal Output	137	0.43%	-3.60%	6.66%	1.59%	0.44	-0.07	-0.06	0.19
Quarterly Inflation (Output Deflator)	137	0.17%	-1.90%	3.15%	0.73%	0.27	0.13	0.15	0.84
Quarterly Growth of Labor Productivity	123	0.30%	-8.94%	7.02%	2.97%	0.94	-0.42	-0.17	0.50
Monthly Growth of Non-Farm Payroll Employment	414	0.37%	-4.85%	5.19%	1.42%	0.64	0.00	0.10	0.00
Monthly Growth of Industrial Production (Total Industry)	433	1.11%	-20.28%	24.12%	4.70%	0.48	-0.08	0.11	0.00
Monthly Growth of Industrial Production (Manufacturing)	235	1.21%	-12.81%	14.61%	4.58%	0.52	-0.24	0.01	0.41
Civilian Unemployment Rate	131	0.01%	-0.20%	0.17%	0.08%	0.05	0.04	-0.06	0.00
Capacity Utilization (Total Industry)	202	0.14%	-1.50%	2.30%	0.84%	0.46	-0.38	0.85	0.00
Capacity Utilization (Manufacturing)	249	0.14%	-2.10%	2.40%	0.94%	0.28	-0.41	0.87	0.00

Notes: All monthly and quarterly growth variables are annualized. Boldface denote significance at 10% level. A/C(1) column reports the first order autocorrelation coefficient. Q-stat(20) reports the p -value associated with the Q -statistic at 20 lags.

- Mankiw and Shapiro (1986) and Mankiw et al. (1984)
- Revisions to money supply are noise.
- Revisions to GNP (level/growth) are news.

- Mankiw and Shapiro (1986) and Mankiw et al. (1984)
- Revisions to money supply are noise.
- Revisions to GNP (level/growth) are news.
- Replicate their results using our data set in their sample.
- Reject both hypotheses in full sample.
- Reject both hypotheses in their sample when r_{t-3}^3 is used.

Relative Variables

$$r_t^f = \alpha + \sum_{i=1}^s \beta_i r_{t-i}^{(2)i} + I_Q \sum_{i=1}^4 \lambda_i Q_t^i + \delta t + \varepsilon_t$$

Level Variables

$$r_t^f = \alpha + \gamma y_t^{t+1} + \sum_{i=1}^s \beta_i r_{t-i}^{(1)i} + I_Q \sum_{i=1}^4 \lambda_i Q_t^i + \delta t + \varepsilon_t$$

Choose best model according to AIC and SIC.

TABLE II - Results of the Ex-Post Forecasting Exercise - Full Sample

Dependent Variable (Final Revision of)	Criterion	N	Explanatory Variables	R^2	\bar{R}^2	Wald Test <i>p</i>-value	RMSE1 / RMSE3	RMSE2 / RMSE3
Revisions as Percentage of Initial Announcements								
Nominal Output	SIC	125	Cons, R5, R10, Trend	0.14	0.12	0.00	0.75	0.78
	AIC	125	Cons, R7, R10, Q1, Trend	0.16	0.13	0.00	0.74	
Real Output	SIC	125	Cons, R10, Trend	0.09	0.07	0.00	0.86	0.88
	AIC							
Non-Farm Payroll Employment	SIC	412	Cons, Trend	0.03	0.03	0.00	0.93	0.82
	AIC	412	Cons, R6, Trend	0.03	0.03	0.00	0.93	
Industrial Production Index (Total Industry)	SIC	427	Cons, R14	0.05	0.04	0.00	0.87	0.47
	AIC	427	Cons, R14, Trend	0.06	0.05	0.00	0.86	
Industrial Production Index (Manufacturing)	SIC	230	Cons, R12, Trend	0.36	0.35	0.00	0.76	0.47
	AIC	230	Cons, R4, R12, R14, Trend	0.37	0.36	0.00	0.76	
Level of Revisions								
Annual Growth of Real Output	SIC	121	Cons, R10	0.00	-0.01	0.00	0.95	0.95
	AIC	121	Cons, Init, R4, R10	0.04	0.01	0.00	0.93	
Annual Growth of Nominal Output	SIC	121	Cons, R10, Trend	0.05	0.03	0.00	0.90	0.90
	AIC	121	Cons, R2, R4, R10, Trend	0.08	0.04	0.00	0.88	
Annual Inflation (Output Deflator)	SIC	125	Init, R5	0.10	0.09	0.00	0.91	0.93
	AIC	122	Init, R5, R9, Trend	0.13	0.10	0.00	0.90	
Annual Growth of Labor Productivity	SIC	110	Cons, Init, R6	0.28	0.27	0.00	0.83	0.97
	AIC	110	Cons, Init, R3, R6	0.30	0.28	0.00	0.82	
Annual Growth of Non-Farm Payroll Employment	SIC	400	Init	0.08	0.08	0.00	0.90	0.78
	AIC	400	Init, R12, R14	0.09	0.08	0.00	0.90	
Annual Growth of Industrial Production (Total Industry)	SIC	415	Trend	0.02	0.02	0.00	0.89	0.70
	AIC	415	Cons, Init, R14, Trend	0.04	0.03	0.00	0.88	
Annual Growth of Industrial Production (Manufacturing)	SIC	218	Cons, R8, R10, R14, Trend	0.40	0.38	0.00	0.69	0.73
	AIC	218	Cons, Init, R8, R10, R12, R14, Trend	0.41	0.40	0.00	0.68	
Quarterly Growth of Real Output	SIC	122	Cons, R1, R9	0.09	0.07	0.00	0.93	0.96
	AIC	121	Cons, R1, R3, R9, R10, Q3	0.15	0.12	0.00	0.90	
Quarterly Growth of Nominal Output	SIC	125	Cons, R2, Q3	0.06	0.04	0.00	0.93	0.93
	AIC	122	Cons, R1, R2, R9, Q3, Trend	0.11	0.07	0.00	0.91	
Quarterly Inflation (Output Deflator)	SIC	125	Init, R1	0.05	0.04	0.00	0.95	0.96
	AIC	122	Cons, R1, R9, Trend	0.07	0.05	0.00	0.94	
Quarterly Growth of Labor Productivity	SIC	114	Init, R1, Q2	0.21	0.19	0.00	0.89	0.96
	AIC	114	Cons, Init, R1, Q2	0.23	0.21	0.00	0.88	
Monthly Growth of Non-Farm Payroll Employment	SIC	400	Cons, Init, R6, R12	0.15	0.14	0.00	0.90	0.80
	AIC	400	Cons, Init, R5, R6, R12	0.15	0.14	0.00	0.89	
Monthly Growth of Industrial Production (Total Industry)	SIC	415	Cons	0.00	0.00	0.00	0.97	0.92
	AIC	415	Cons, R1, R2, R4, R7, R9, R12	0.05	0.03	0.00	0.95	
Monthly Growth of Industrial Production (Manufacturing)	SIC	218	Init, R4, Trend	0.08	0.08	0.00	0.93	0.97
	AIC	218	Cons, Init, R1, R2, R3, R4, R7, R14	0.14	0.11	0.00	0.90	
Civilian Unemployment Rate	SIC	121	R6, R7, Trend	0.17	0.16	0.00	0.90	0.97
	AIC	121	Cons, R4, R5, R8, R9, Trend	0.23	0.19	0.00	0.87	
Capacity Utilization (Total Industry)	SIC	188	Cons, Init, R7, Trend	0.29	0.28	0.00	0.82	0.96
Capacity Utilization (Manufacturing)	AIC							
Capacity Utilization (Manufacturing)	SIC	235	Cons, Init, R7, R14, Trend	0.28	0.27	0.00	0.84	0.94
	AIC							

Final Revision to Nominal Output and Various Forecasts

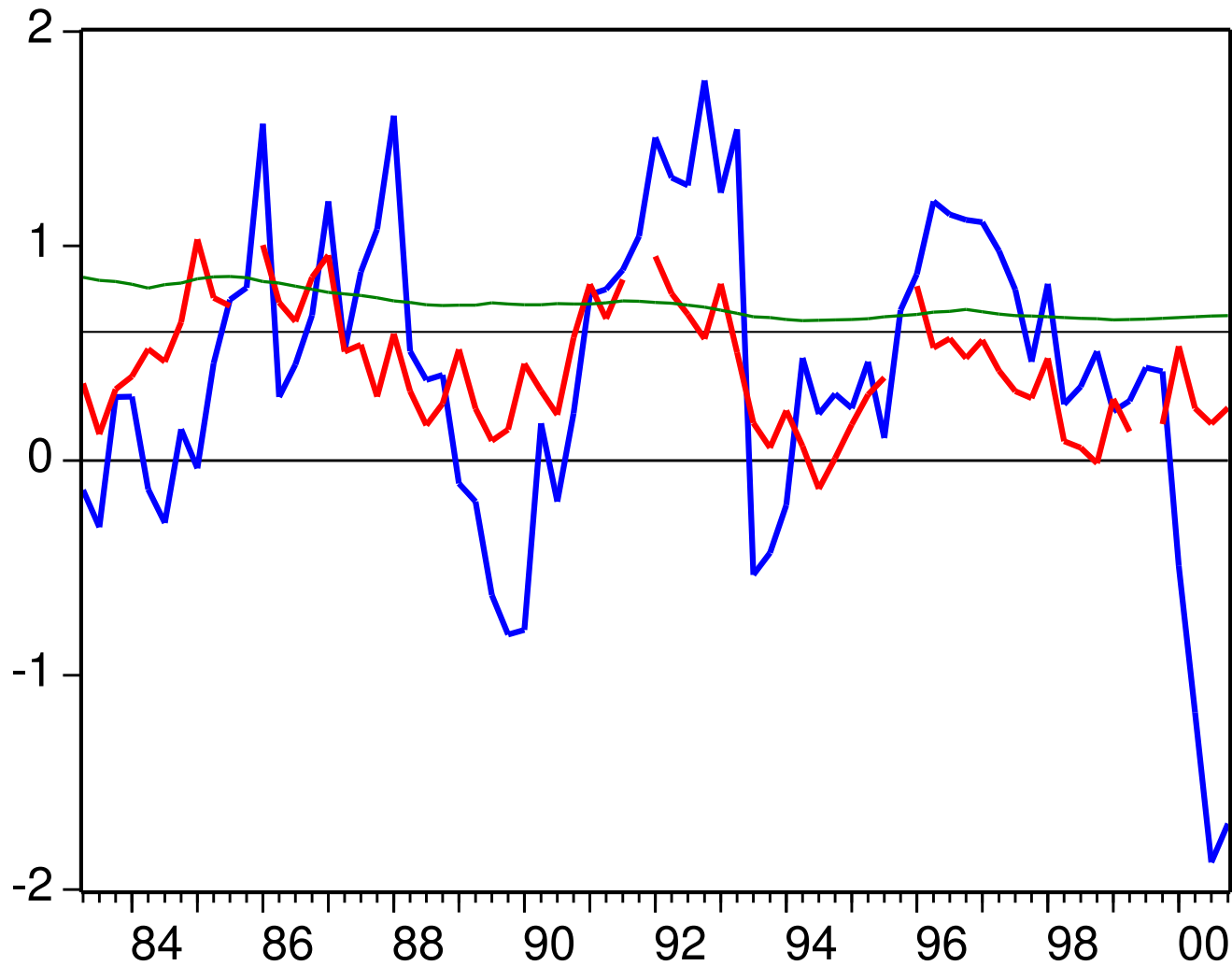
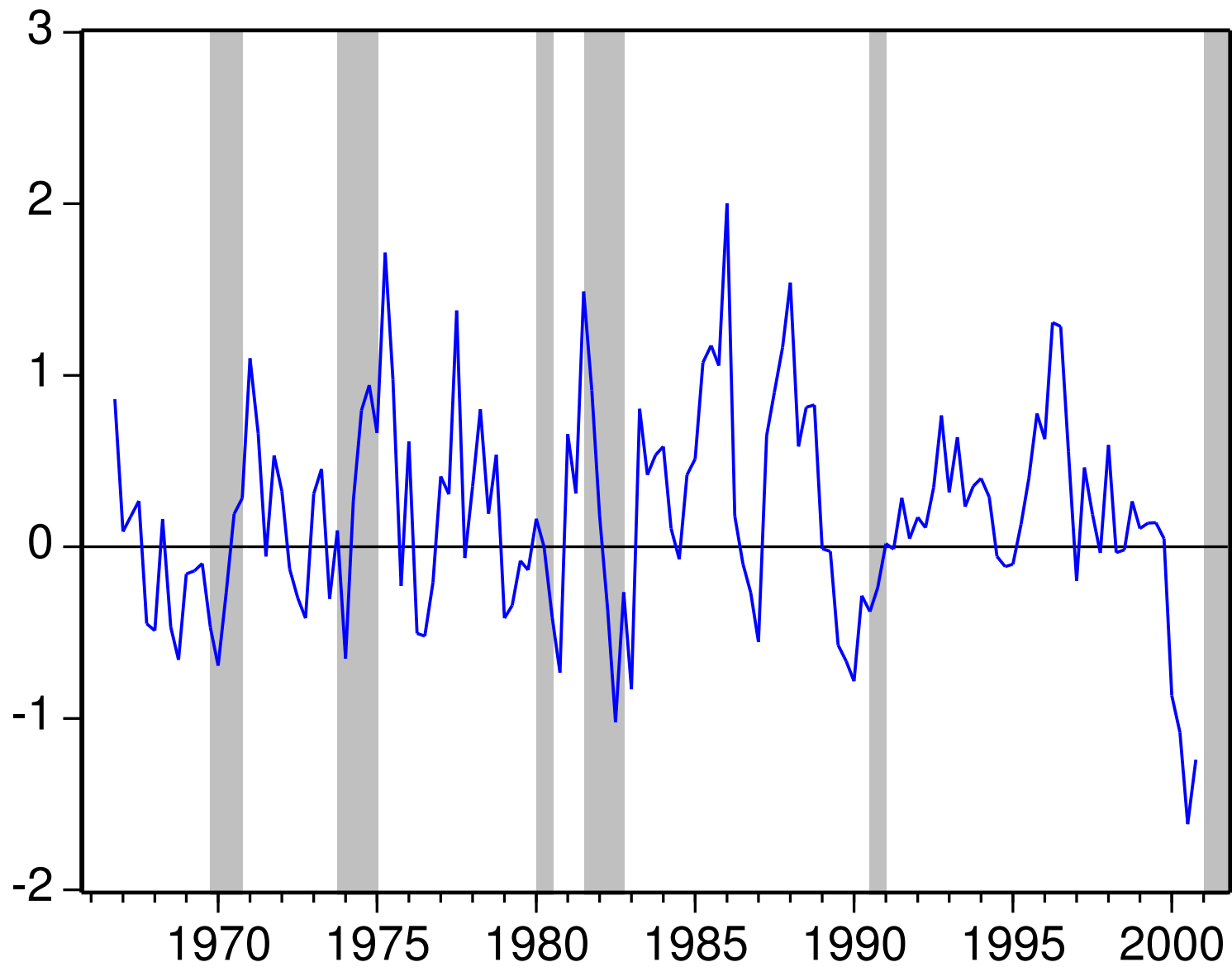


TABLE IV - Results from the Survey of Professional Forecasters

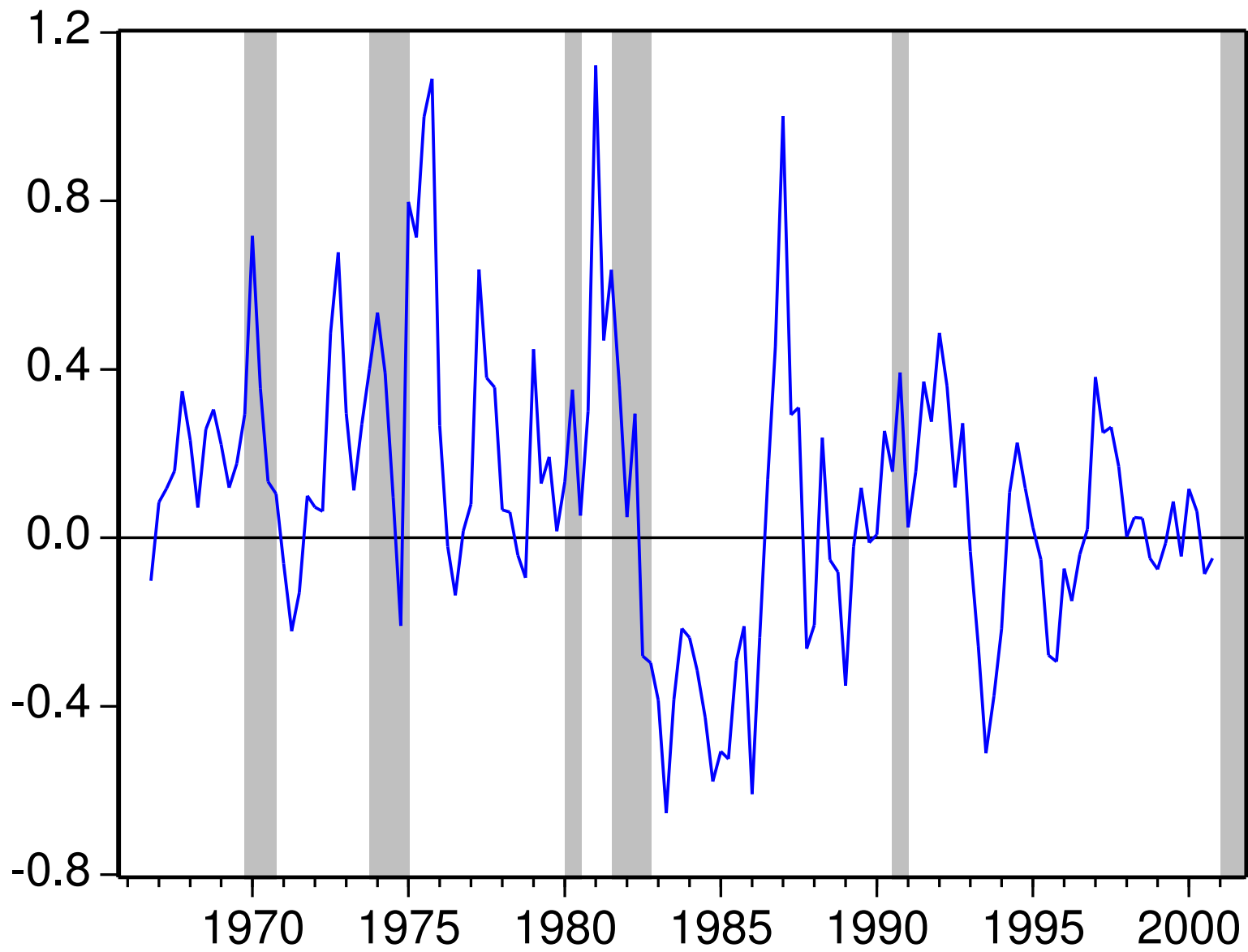
	Nominal Output	Real Output	Real Consumption	Price Index
Pre-1990				
Median Equals Initial	88.5%	86.5%	92.6%	92.0%
Within One Point	77.0%	67.9%	76.2%	97.4%
Correct Sign	20.7%	23.7%	22.1%	38.9%
Mean Revision (SPF)	0.03%	0.00%	0.08%	0.07%
Mean Revision (RTDS)	0.60%	0.46%	0.47%	0.14%
Post-1990				
Median Equals Initial	97.8%	95.4%	95.5%	91.1%
Within One Point	92.0%	94.0%	94.4%	100.0%
Correct Sign	26.1%	35.2%	34.8%	43.3%
Mean Revision (SPF)	0.00%	0.00%	0.00%	0.04%
Mean Revision (RTDS)	0.60%	0.51%	0.40%	0.08%

Note: See the text for the definitions of each row.

Revision to Annual Real Output Growth



Revision to Annual Inflation



- **Subsamples**

- Mean revisions larger in second half
- Noise-to-signal ratios larger in second half
- Increased predictability in second half

- **Subsamples**

- Mean revisions larger in second half
- Noise-to-signal ratios larger in second half
- Increased predictability in second half

- **Components of output**

- Consumption (especially durables) and exports are important for positive mean
- All components have large noise-to-signal ratios
- All components show predictability

Conclusions and Further Research

- Data revisions to macro variables in US violate three basic principles:
 - Zero mean
 - Small variance
 - Unpredictable
- Professional forecasters seem to ignore these results.

Conclusions and Further Research

- Data revisions to macro variables in US violate three basic principles:
 - Zero mean
 - Small variance
 - Unpredictable
- Professional forecasters seem to ignore these results.
- Aruoba (2005) : Benefit of reducing predictability : \$5 billion.

Conclusions and Further Research

- Data revisions to macro variables in US violate three basic principles:
 - Zero mean
 - Small variance
 - Unpredictable
- Professional forecasters seem to ignore these results.
- Aruoba (2005) : Benefit of reducing predictability : \$5 billion.
- Remaining questions:
 - Link between “great moderation” and increased data uncertainty?
 - Policy makers and data revisions.
 - Financial markets and data revisions.