

Workshop on Macroeconomic Forecasting, Analysis and Policy with Data Revision

Call for Papers: CIRANO and the Bank of Canada invite papers for their joint Workshop on Macroeconomic Forecasting, Analysis and Policy with Data Revision, to be held Friday, October 14th 2005 in Montreal, Canada in collaboration with the University of Toronto Institute for Public Policy.

Theme: The organizers invite the submission of papers related to the broad theme of real-time data analysis in macroeconomics and finance. This includes topics such as

- the creation of new real-time macroeconomic databases
- modeling or forecasting data revision
- constructing and evaluating forecasts with data subject to revision
- reliability studies examining the revision of forecasts or current analysis
- policy formulation or evaluation in the face of uncertainty about current conditions

Both theoretical and applied studies are welcomed. Papers that should be of interest to either statistical agencies or macroeconomic policymakers are particularly encouraged.

Program Committee

Jean Boivin (Columbia Business School)
Dean Croushore (University of Richmond)
Graham Elliott (U.C.S.D.)
Stéphane Gregoir (INSEE and CREST)
Athanasios Orphanides (Federal Reserve Board)
Florian Pelgrin (Bank of Canada)
Simon van Norden (HEC Montréal, CIRANO, CIREQ)

Bryan Campbell (Concordia University, CIRANO, CIREQ), Committee secretary

Submissions : Submissions should be sent to Evelyne Giard (evelyne.giard@cirano.qc.ca) before July 15, 2005 and should include a detailed abstract (one page maximum) and preferably a draft paper (if available.) The workshop programme will be determined by mid-August 2005.

Support : The organizers will provide accommodation for two nights in Montreal along with lunch and dinner during the workshop. Under special circumstances some funding for travel is available. Further inquiries should be directed to the workshop secretary (bcamp@jmsb.concordia.ca).

The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time

Athanasios Orphanides¹
Board of Governors of the Federal Reserve System

Simon van Norden
HEC (Montreal), CIRANO and CIREQ

Forthcoming, JMCB 2005

Abstract: Do Output Gaps improve forecasts of inflation? Maybe not...

1. Usual disclaimers apply.

Motivation

Stylized Fact: Robust empirical link between output gaps and later inflation.

Problem: Output Gap estimates have large revisions *ex post*.

- Orphanides and van Norden (2002), Orphanides (2003), etc.
- Table 1.
- Data revision is only part of the problem.

Empirical link based on conventional (Final) estimates of Gap.

Usefulness for forecasting depends on real-time estimates.

- Are Real-Time estimates of gaps as useful?

Let's check!

Output Gap

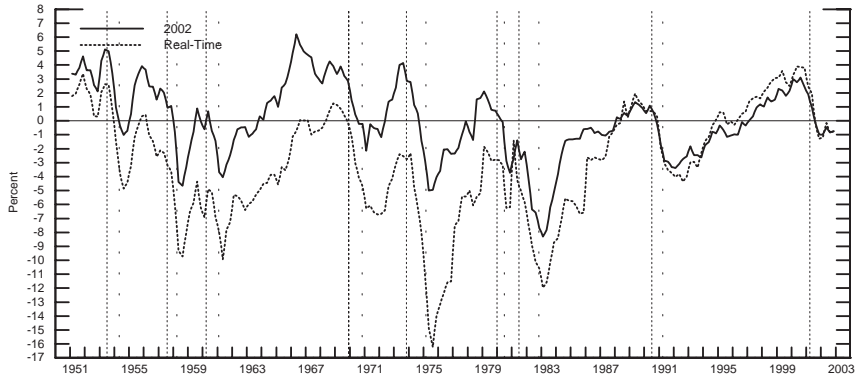
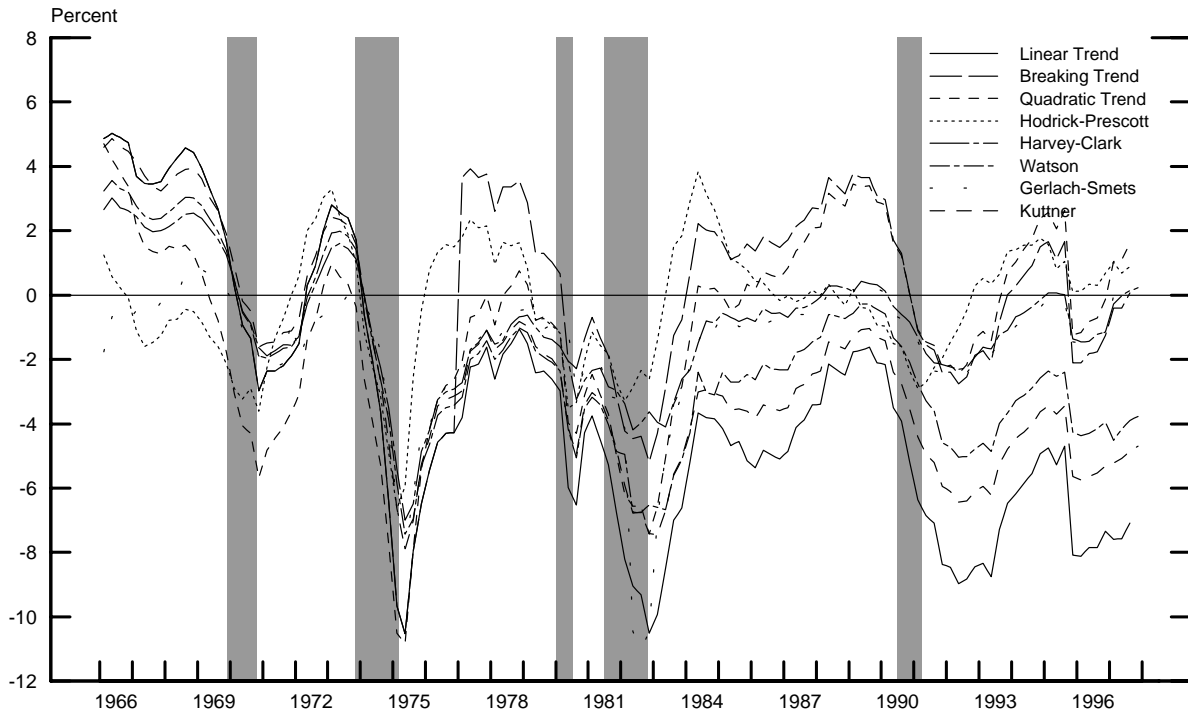


Fig. 3. Real-time and retrospective views of the economy. One-quarter lagged values are plotted. Real-time data as discussed in the text. See also notes to [Fig. 1](#).

Figure 1

Real-Time Estimates of the Business Cycle



Final Estimates of the Business Cycle

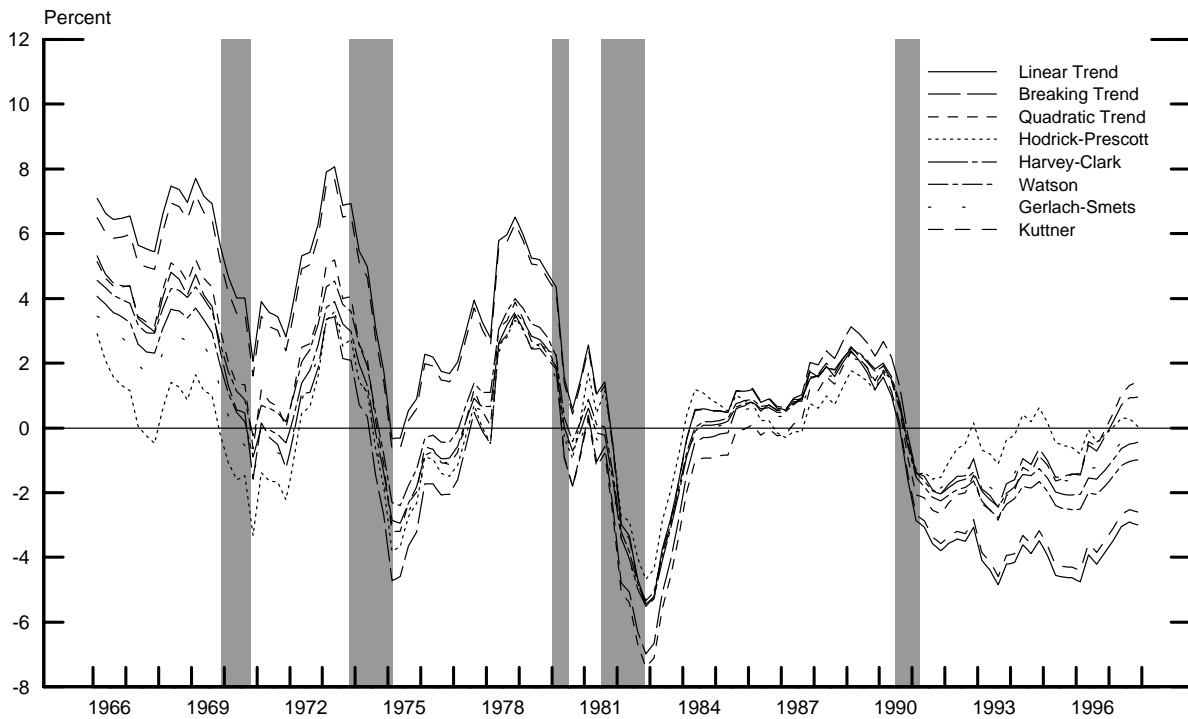
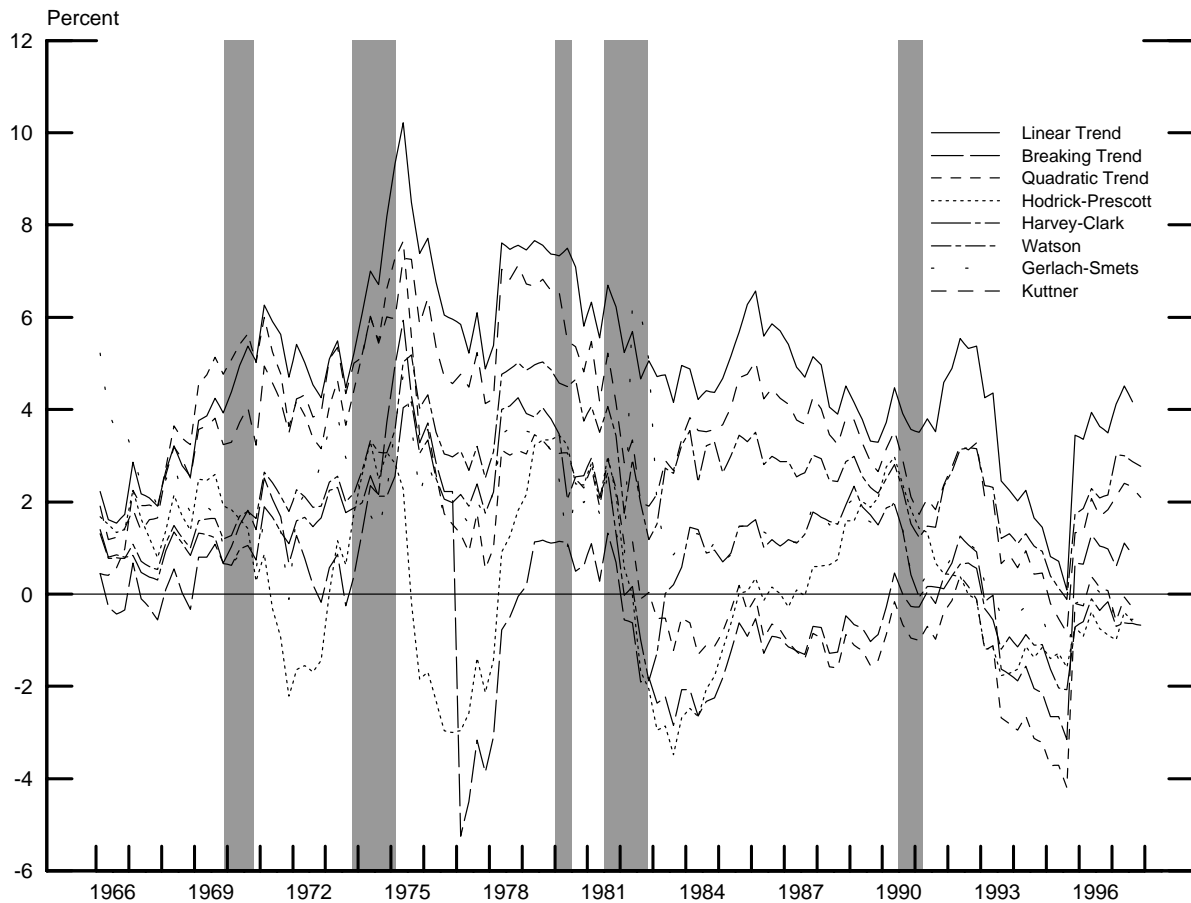


Figure 2

Total Revision in Business Cycle Estimates



U.S. Forecasting Experiment

Final Estimates: Full sample (1947Q1-2003Q2), series published August 2003.

Real-Time: New data vintage & parameter estimates each quarter 65Q1-03Q2

FL-FL: OLS Recursive estimation from 1955Q1-2002Q1:

$$\pi_{t+h}^h = \alpha + \sum_{i=1}^n \beta_i \cdot \pi_{t-i}^1 + \sum_{i=1}^m \gamma_i \cdot y_{t-i} + e_{t+h}$$

- Use final estimates of output gaps y_t .
- Set $m=1$, $n=1$ using full-sample BIC.
- Calculate RMSFE for 1969Q1-2002Q1.

VL-RT: OLS Recursive estimation from 1955Q1-2002Q1.

- Use real-time estimate of y_t (replace entire series each quarter.)
- Re-adjust lag lengths (m,n) each quarter.
- RMSFE for 1969Q1-2002Q1

Benchmark: AR model restricts $\gamma_i = 0$ for all i . (Naive => conservative)

Output Gap Estimation

- | | |
|-----------------------------|---|
| 1. Linear Trend: | $y_t - \alpha - \beta \cdot t$ |
| 2. Quadratic Trend: | $y_t - \alpha - \beta \cdot t - \gamma \cdot t^2$ |
| 3. Breaking Trend: | β changes in '73; change introduced in '77 |
| 4. Hodrick-Prescott: | $\lambda = 1600$ |
| 5. Band-Pass: | 6-32Q, padded with AR forecasts |
| 6. Beveridge Nelson: | univariate ARIMA(1,1,2) |
| 7. SVAR: | trivariate with prices & t-bill rates. |
| 8. Watson: | univariate UC -- local level |
| 9. Harvey-Clark: | univariate UC -- local linear trend |
| 10. Harvey-Jaeger: | univariate UC -- stochastic cycle |
| 11. Kuttner: | bivariate UC -- Watson + Phillips Curve |
| 12. Gerlach-Smets: | bivariate UC -- Harvey-Clark + Phillips Curve |
| 13. TOFU: | replace y_t with real Q/Q output growth |

Table 1

Description of Alternative Output Gap Measures and Summary Reliability Statistics

Method	Data	Method Details	COR	AR	NSR	OPSIGN
Linear Trend	Univariate.		0.88	0.90	1.63	0.58
Quadratic Trend	Univariate.		0.51	0.97	1.06	0.42
Breaking Trend	Univariate.	Trend Break in 1973Q1, starting in 1977Q1.	0.77	0.87	0.81	0.28
Hodrick-Prescott	Univariate.	With $\lambda = 1600$.	0.52	0.93	1.05	0.45
Band Pass	Univariate.	6–32 quarters, series padded with AR forecasts.	0.72	0.77	0.77	0.36
Beveridge-Nelson	Univariate.	Assumes ARIMA(1,1,2).	0.84	0.09	0.63	0.30
Structural VAR	Trivariate.	Imposes long-run restrictions.	0.68	0.85	0.95	0.41
Watson	Univariate.	Local Level and AR(2).	0.88	0.87	1.50	0.55
Harvey-Clark	Univariate.	Local Linear Trend and AR(2).	0.75	0.92	0.91	0.39
Harvey-Jaeger	Univariate.	Local Linear Trend and Cycle.	0.56	0.86	0.92	0.50
Kuttner	Bivariate.	Watson model and inflation equation.	0.87	0.90	1.54	0.61
Gerlach-Smets	Bivariate.	Harvey-Clark model and inflation equation.	0.79	0.82	1.05	0.40

Table 2

Relative Improvement in MSFE

Method	AR	AR p-value	TF	TF p-value
<i>Fixed Lags, Final Gaps</i>				
Benchmark MSFE	0.494		0.436	
Linear Trend	0.302	0.009	0.148	0.164
Quadratic Trend	0.168	0.010	0.030	0.779
Breaking Trend	0.106	0.034	-0.024	0.778
Hodrick-Prescott	0.149	0.000	0.013	0.900
Band-Pass	0.134	0.000	0.000	0.997
Beveridge-Nelson	0.139	0.000	0.004	0.309
SVAR	0.047	0.121	-0.077	0.474
Watson	0.319	0.001	0.163	0.060
Harvey-Clark	0.270	0.002	0.120	0.162
Harvey-Jaeger	0.109	0.001	-0.022	0.811
Kuttner	0.336	0.008	0.178	0.079
Gerlach-Smets	0.362	0.001	0.201	0.052
<i>Fixed Lags, Quasi-Final Gaps</i>				
Watson	0.132	0.043	-0.002	0.979
Harvey-Clark	0.070	0.068	-0.056	0.374
Harvey-Jaeger	-0.032	0.811	-0.146	0.382
Kuttner	0.248	0.030	0.100	0.250
Gerlach-Smets	0.091	0.070	-0.038	0.414
<i>Variable Lags, Real-time Gaps</i>				
Benchmark MSFE	0.559		0.416	
Linear Trend	0.045		-0.219	
Quadratic Trend	0.021		-0.237	
Breaking Trend	0.043		-0.221	
Hodrick-Prescott	0.132		-0.154	
Band-Pass	0.283		-0.042	
Beveridge-Nelson	0.211		-0.095	
SVAR	-0.093		-0.323	
Watson	0.121		-0.163	
Harvey-Clark	0.147		-0.143	
Harvey-Jaeger	0.080		-0.193	
Kuttner	0.107		-0.173	
Gerlach-Smets	0.099		-0.179	

Main Results - Table 2

Results vary critically depending on benchmark and forecast construction.

FL-FL

1. All improve on AR model's RMSFE
All but 1 by $>10\%$, $4/13 > 30\%$
All but 1 p-value $< 5\%$.
2. All but 3 improve on TF benchmark RMSFE
4/12 have p-value $< 5\%$.

VL-RT

1. All but 1 improve on AR model's RMSFE
6/12 by $> 10\%$.
Bootstrap simulations suggest perhaps 4/12 have p-values $< 5\%$.
2. None improve on TF benchmark RMSFE
9/12 $< -15\%$
Bootstrap simulations suggest perhaps 3/12 have p-values $< 5\%$.

Table 3

The Effect of Lag Selection and Data Vintage

Method	MSFE			Change in MSFE (percent)		
	FL-FL	VL-FL	VL-RT	FL to VL	FL to RT	Total
AR benchmark	0.494	0.559	0.559	-13.0	0.0	-13.0
TF benchmark	0.436	0.496	0.416	-13.7	16.0	4.6
Linear Trend	0.380	0.438	0.533	-15.4	-21.7	-40.4
Quadratic Trend	0.423	0.500	0.545	-18.1	-9.0	-28.8
Breaking Trend	0.447	0.494	0.534	-10.6	-8.0	-19.5
Hodrick-Prescott	0.430	0.556	0.492	-29.2	11.5	-14.4
Band-Pass	0.436	0.502	0.434	-15.2	13.5	0.4
Beveridge-Nelson	0.434	0.482	0.460	-11.0	4.5	-6.0
SVAR	0.472	0.502	0.614	-6.4	-22.3	-30.1
Watson	0.375	0.433	0.497	-15.4	-14.9	-32.6
Harvey-Clark	0.389	0.448	0.486	-15.1	-8.4	-24.7
Harvey-Jaeger	0.446	0.577	0.516	-29.5	10.7	-15.7
Kuttner	0.370	0.402	0.503	-8.5	-25.3	-36.0
Gerlach-Smets	0.363	0.426	0.507	-17.3	-19.0	-39.6
Mean				-15.6	-5.2	-21.1
Std Dev				6.7	14.5	14.4

Notes:

MSFE denotes the mean squared forecast error (shown multiplied by 1000.)

FL-FL refers to forecasts using fixed lag lengths and final output gap estimates.

VL-FL refers to forecasts using variable lag lengths and final output gap estimates.

VL-RT refers to forecasts using variable lag lengths and real-time output gap estimates.

FL to VL refers to the change from FL-FL to VL-FL.

FL to RT refers to the change from VL-FL to VL-RT.

Total refers to the change from FL-FL to VL-RT.

Why Do the Results Change Dramatically?

See Table 3.

1. Lag length uncertainty makes *all* models do worse.
2. Data revision makes *most* (but not all) models do worse.
3. TF benchmark *improves* under these conditions.

How should we interpret the TF benchmark?

Univariate gap estimators can be expressed as weighted moving average of Δy_t .

- Bivariate gap estimators also weigh π_t .
- Different models impose different implicit restrictions on weights.

TF benchmark relaxes all restrictions on weights of Δy_t .

- Do these implicit restrictions help forecast inflation?
Perhaps not in the simple models we examined.

Robustness

We checked....

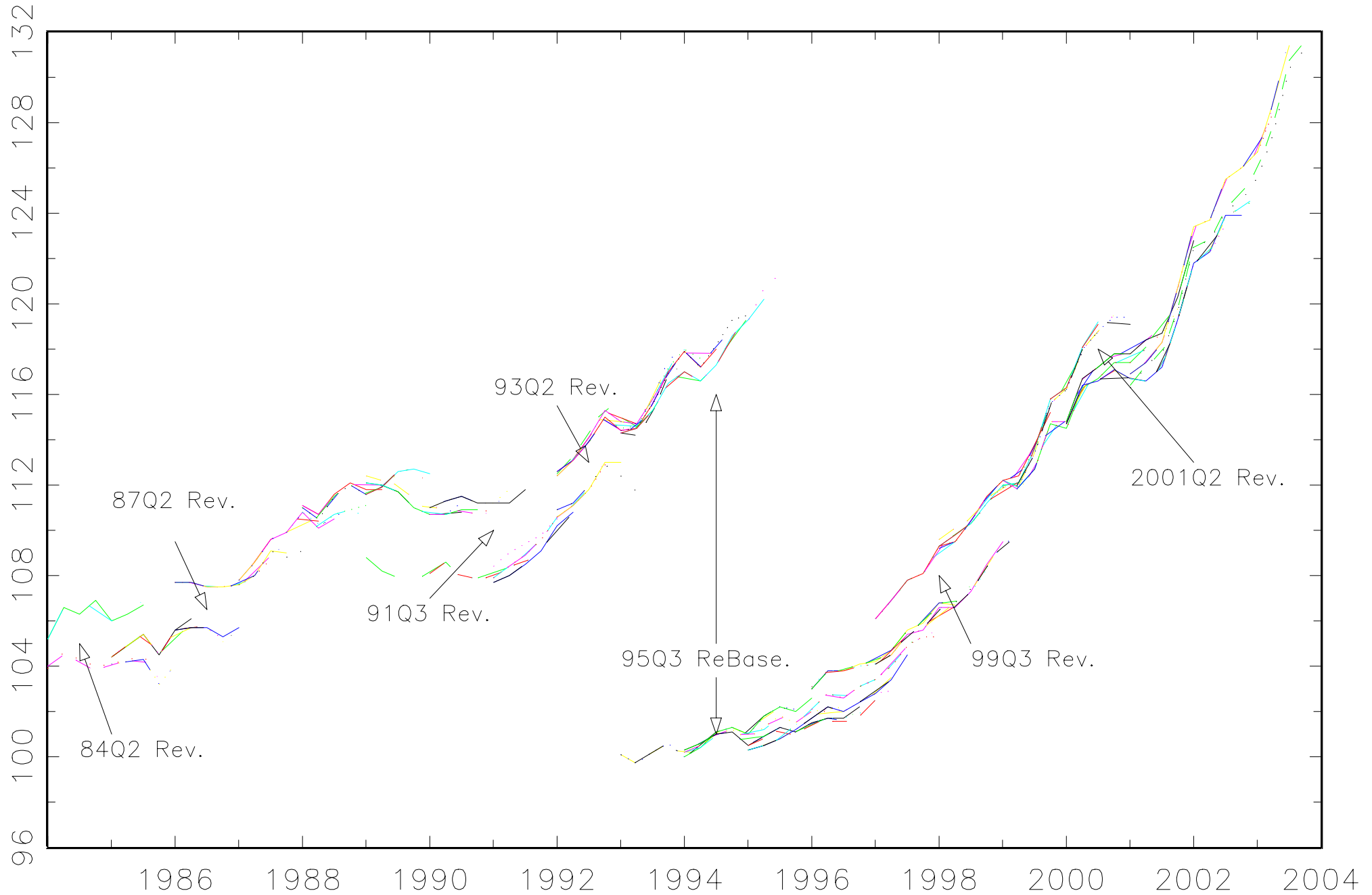
- 1) Forecast Horizon (2-9Q)
- 2) Sample period for estimating forecast equation
- 3) Sample period for forecast evaluation
- 4) Benchmark Model (TOFU, YN, RW)
- 5) Lag Selection Method (robust std. errs., AIC, BIC)
- 6) Forecast accuracy vs encompassing
- 7) Structure of Data vintages (Koenig, Dolmas & Piger critique)
- 8) Country (Canada -- Cayen and van Norden)
- 9) Finite-sample properties of forecast tests

Check for yourself: <http://www.hec.ca/pages/simon.van-norden/wps/summary.xls>

So What?

1. What about the many papers that show output gaps help forecast π ?
*Almost all use (1) Final gaps or (2) fixed lag structure or (3) RW benchmark
The latter is easily beat by an AR benchmark (this is 4Q inflation!)*
2. Failure to forecast inflation means the Fed is using information efficiently.
We would not have expected to find such results at all forecast horizons.
3. Need to use production-function approach to output gaps.
*Replaces potential output uncertainty with uncertainty about NAIRU, trend
TFP & participation rates, full-capacity level of CAPU,....
Data revision in productivity looks worse than in output;
is capital stock worse still?*
4. Need to use markups instead of output gap.
*Koenig (2003): data revision in mark-ups is more severe ($\rho = 0.27$)
Finds no evidence that real-time mark-ups help explain inflation.*
5. Is the U.S. typical? *Canada similar. UK?*
6. Need to allow for breaks in Phillips Curve.
Reduces test power relative to full-sample estimation.

Output per Person-Hour US Non-Farm Business Sector



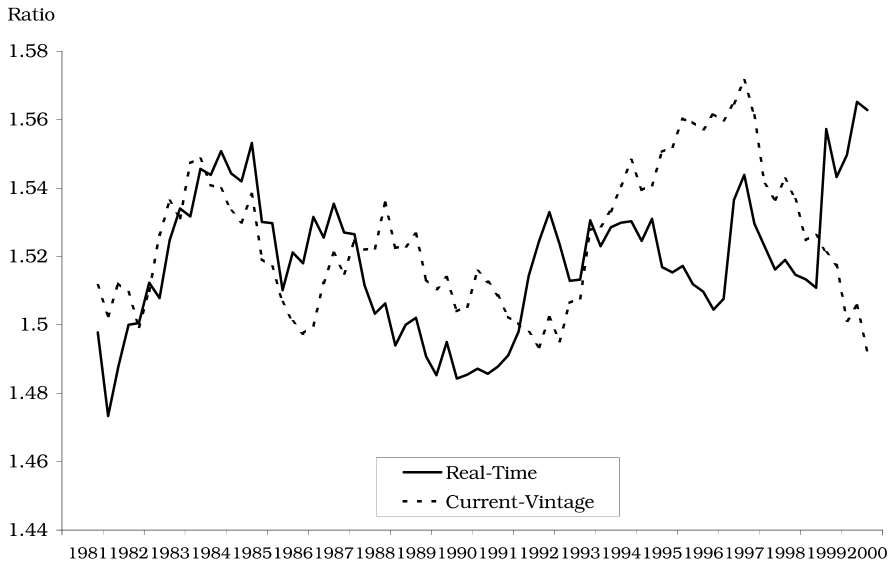


Fig. 1. Real-time and current-vintage markups.

Summary and Extensions

Conventional estimates of gap do not appear to improve inflation forecasts very much, if at all.

Euro Area?

Better gaps?

- Diffusion indices? Dynamic factor models?

Conditional forecasting.

- Other factors in Phillips Curve (expectations, import prices, ...)
- Control for policy stance (projection vs forecast)

Growth Rate Rules vs Taylor Rules

Econometric Refinements

- Model Confidence Sets? Inoue & Rossi?

Public Accountability & Due Diligence: How can a policymaker responsibly decide on policy without simulated real-time analysis of its performance?