National Forecasting with Large International Datasets – an Application to New Zealand

Sandra EICKMEIER* **& Tim NG****

Rome, 27 March 2009

*Deutsche Bundesbank, ** Reserve Bank of New Zealand

This paper ...

- exploits a large international dataset to improve forecasts of New Zealand GDP growth

- uses recent data-rich methods
  - Factor methods: principal components, weighted principal components, targeted predictors, partial least squares
  - Shrinkage methods: elastic net, ridge regression

- compares the forecast accuracy using national data with that using international data

- opens the „black box“ by assessing
  - the marginal predictive content of subsets of international data along economic and regional dimensions
  - historical contributions of different groups of variables to forecasts

Related literature

- Forecasting applications with large datasets
  - PC widely applied: Stock and Watson (2002a, 2002b, 2004), Eickmeier and Ziegler (2008), Matheson (2006)
  - Other data-rich methods less applied in economics and mostly to US data: Boivin and Ng (2006), Bai and Ng (2008), De Mol et al. (2008), Groen and Kapetanios (2008), Lin and Tsay (2006), Schumacher (2009)

- New Zealand forecasts with large datasets
  - Matheson (2006), Bloor and Matheson (2008)

- Only few (PC) applications explicitly investigate predictive content of international information.
  - Banerjee et al. (2006), Gosselin and Tkacz (2001), Cheung and Demers (2007), Brisson et al. (2003), Schumacher (2009)

Data

- 1990Q1-2007Q4
- Target variable: NZL GDP growth
- 446 national predictors taken from Matheson (2006)
- International predictors

	North Amer.	Europe	Asia	Australia	World	Sum
Real	50	41	95	20	-	206
Prices	28	28	41	11	-	108
Monetary + financial	23	34	62	11	-	130
Commodity prices	-	-	-	-	38	38
Sum	101	103	198	42	38	482

- All predictors transformed (seasonally adjusted, stationary, outlier adjusted, mean = 0 and variance = 1)

Forecasting setup

- Univariate benchmark

$$y_{t+h,t} = \alpha W_t + e_{t+h,t}$$

with $y_{t+h,t} = y_{t+h} - y_t$ and W_t being a constant and/or lags of the target $y_{t-i,t-1-i} = y_{t-i} - y_{t-1-i}, i = 0, \ldots, p-1$.

- Additional predictors

$$y_{t+h,t} = \alpha W_t + \Gamma' \underset{N \times 1}{X_t} + e_{t+h,t}$$

- If N is large, reduce dimension of X_t through shrinking Γ parameters and/or summarizing X_t in a few common factors.

Forecasting setup

- Predictors: only national, only international, combined
- Forecast evaluation period starts in 2000Q1
- Recursive, direct out-of-sample forecasts for horizons 1, 2, 4, 8
- Compare RMSE between $y_{t+h,t}$ and realization across models
- Performance-based selection of p

Shrinkage methods

- **Ridge estimator** solves penalized regression problem

$$\min_{\Gamma} \boldsymbol{RSS} + \boldsymbol{\nu} \sum_{\boldsymbol{i}=1}^{\boldsymbol{N}} \Gamma_{\boldsymbol{i}}^2$$

- Estimator $\hat{\Gamma}_{\boldsymbol{h}}^{\boldsymbol{RR}} = (\mathrm{X'X} + \boldsymbol{\nu}\mathrm{I}_N)^{-1}\mathrm{X}'\boldsymbol{y}_{\boldsymbol{h}}$ shrinks LS estimator towards (but not to) zero.

- **Elastic net** (Zhou and Hastie 2005) solves

$$\min_{\Gamma} \boldsymbol{RSS} + \boldsymbol{\nu}_1 \sum_{\boldsymbol{i}=1}^{\boldsymbol{N}} \left|\Gamma_{\boldsymbol{i}}\right| + \boldsymbol{\nu}_2 \sum_{\boldsymbol{i}=1}^{\boldsymbol{N}} \Gamma_{\boldsymbol{i}}^2$$

- It shrinks some coefficients to zero. Solution found with LARS.

Factor methods

- **Principal components** (Chamberlain and Rothschild 1983, Stock and Watson 2002)

$$x_{it} = \lambda_i ' \underset{r \times 1}{F_t} + \xi_{it}$$

- Factor estimates are the first r PCs of X_t. PCs maximize explained variance of X_t.

- **Weighted PC** (Boivin and Ng 2006)
  - WPCBN: drop variables with highly cross-correlated idiosyncratic components and low commonality and estimate PCs from reduced set
  - WPCSW: apply PC to weighted X_t where weights are $1/std(\xi_i)$

Factor methods cont.

- **Targeted predictors** (Bai and Ng 2008)
  - drop variables which are not highly correlated with the target using hard thresholding (TPH) or soft-thresholding (TPS) and estimate PCs from reduced set

- **Partial least squares** (Wold 1982)
  - finds linear orthogonal combinations of predictors that (broadly speaking) give max. covariance between $y_{t+h,t}$ and predictors (PLS factors). Weights are given by covariance between „cleaned“ predictors and target.

- Last step of all factor methods: regress target on factors.

Overview of specifications

trade1	Trade-weighted averages	RBNZ practice
trade2		constant weights based on exports + imports
trade3		2-year moving average weights based on exports + imports
PC	Principal components	r = 1, 2, 3, 5
WPCBN	Weighted PC	Rule 1c in Bai and Ng (2008), r = 1, 2, 3, 5
WPCSW		Rule SWa in Bai and Ng (2008), r = 1, 2, 3, 5
TPH	Targeted predictors	Hard thresholding, r = 1, 2, 3, 5
TPS		Soft thresholding, 30, 100, 200 variables, v_2 = 0.25, r = 1, 2, 3, 5
PLS	Partial least squares	k = 1, 2, 3, 5
RR	Ridge regression	v = 0.25, N, 5N, 10N
EN	Elastic net	30, 100, 200 variables, v_2 = 0.25

Discussion

- Data-rich methods vs. trade-weights
  - simplicity
  - information exploited
  - weights (constant vs. time-variant, linkages accounted for, „optimal“ vs. other)

- PC, WPC vs. other methods
  - covariance between predictors and target

- PC, PLS, RR vs. WPC, TP, EN
  - are coefficients exactly zero?

- Shrinkage vs. factor methods
  - factor structure in the data?

Forecast results - RMSE(model)/RMSE(univariate)

	National				International				National + international			
	h = 1	h = 2	h = 4	h = 8	h = 1	h = 2	h = 4	h = 8	h = 1	h = 2	h = 4	h = 8
Univar. benchm. (RMSE/h)	0.0061	0.0045	0.0034	0.0024	0.0061	0.0045	0.0034	0.0024	0.0061	0.0045	0.0034	0.0024
trade1	-	-	-	-	0.984	1.247	1.657	2.500	-	-	-	-
trade2	-	-	-	-	0.951	1.034	0.970	1.247	-	-	-	-
trade3	-	-	-	-	0.934	1.022	0.933	1.242	-	-	-	-
PC (r = 1)	0.951	0.865	0.985	1.113	1.033	1.090	1.060	0.948	0.984	1.247	1.657	2.500
PC (r = 2)	0.951	0.910	1.119	1.387	1.016	1.056	1.149	1.247	0.951	1.034	0.970	1.247
PC (r = 3)	**0.902**	0.899	1.127	1.381	1.082	1.034	1.134	1.361	0.934	1.022	0.933	1.242
PC (r = 5)	0.951	0.876	1.351	1.361	1.098	1.258	1.493	1.289	1.016	1.067	1.187	1.242
WPCBN (r = 1)	0.951	0.899	0.993	1.124	1.033	1.079	1.045	1.031	1.082	1.180	1.194	1.268
WPCBN (r = 2)	0.951	0.854	1.082	1.392	1.016	1.079	1.149	1.227	1.082	1.180	1.351	1.284
WPCBN (r = 3)	0.984	0.888	1.187	1.232	1.000	1.011	1.425	1.412	1.131	1.202	1.582	1.799
WPCBN (r = 5)	0.967	0.955	1.410	1.485	1.033	1.348	1.373	1.407	1.033	1.090	1.216	1.247
WPCSW (r = 1)	1.033	1.056	1.037	1.026	1.016	1.022	1.000	**0.851**	1.131	1.225	1.291	1.232
WPCSW (r = 2)	0.951	0.854	0.948	1.170	0.967	0.910	0.851	1.093	1.131	1.225	1.381	1.201
WPCSW (r = 3)	**0.902**	0.843	**0.896**	1.093	1.049	0.955	0.918	1.139	1.098	1.169	1.306	1.778
WPCSW (r = 5)	0.984	0.921	**0.896**	1.026	1.066	1.225	1.224	1.139	0.934	0.933	1.045	1.113
TPH (r = 1)	0.951	0.966	0.963	1.021	1.082	1.101	1.127	1.103	0.984	1.045	1.194	1.191
TPH (r = 2)	0.951	1.022	0.948	1.325	1.049	1.112	1.194	1.320	0.984	1.135	1.246	1.196
TPH (r = 3)	1.000	1.056	0.933	1.206	1.197	1.157	1.179	1.521	0.934	0.989	1.194	1.387
TPH (r = 5)	1.180	1.112	1.261	1.237	1.180	1.146	1.373	1.665	0.984	1.034	1.269	1.562
TPS (30 variables, r = 1)	1.016	1.213	1.194	1.211	1.098	1.101	1.216	1.206	1.131	1.022	1.022	1.052
TPS (30 variables, r = 2)	1.033	1.303	1.149	1.289	1.246	1.213	1.373	1.201	1.033	1.022	1.015	1.046
TPS (30 variables, r = 3)	1.033	1.225	1.224	1.273	1.311	1.202	1.261	1.263	1.049	1.045	1.015	1.062
TPS (30 variables, r = 5)	1.000	1.247	1.194	1.289	1.328	1.180	1.306	1.253	1.049	1.056	0.993	0.964
TPS (100 variables, r = 1)	0.934	0.831	1.157	1.139	1.082	0.955	0.963	1.356	1.049	1.191	1.037	1.082
TPS (100 variables, r = 2)	1.098	0.854	1.261	1.160	1.164	1.000	1.112	1.273	1.033	1.157	0.985	1.175
TPS (100 variables, r = 3)	1.066	0.843	1.276	1.206	1.197	1.079	1.060	1.340	1.016	1.101	1.231	1.247
TPS (100 variables, r = 5)	1.115	0.899	1.164	1.196	1.082	1.169	1.134	1.448	1.049	0.989	1.254	1.237
TPS (200 variables, r = 1)	0.934	**0.820**	1.067	1.186	1.066	1.056	1.022	1.129	0.934	0.820	0.955	1.144
TPS (200 variables, r = 2)	0.984	0.899	1.119	1.505	1.066	1.112	1.082	1.258	0.934	0.865	0.978	1.165
TPS (200 variables, r = 3)	0.951	0.944	1.209	1.649	1.000	0.933	0.925	1.418	0.951	0.820	0.970	1.196
TPS (200 variables, r = 5)	1.000	0.933	1.284	1.691	1.131	1.180	1.179	1.521	0.934	0.809	1.045	1.129
PLS (k = 1)	0.951	0.843	0.940	1.108	1.000	1.067	0.784	1.098	0.951	0.865	0.925	1.057
PLS (k = 2)	1.066	1.056	1.336	1.351	0.869	0.944	0.761	1.196	1.000	0.854	0.806	1.021
PLS (k = 3)	0.934	**0.787**	1.007	1.412	0.869	1.045	0.836	1.124	**0.852**	0.831	0.731	0.948
PLS (k = 5)	0.951	1.034	1.209	1.361	**0.787**	0.955	0.761	0.923	**0.852**	0.944	0.784	0.804
RR (v_2 = 0.25)	1.049	0.966	1.239	1.371	0.803	1.011	0.903	1.216	0.885	0.865	0.873	0.969
RR (v_2 = N)	0.918	0.820	0.955	1.186	0.836	**0.899**	**0.716**	1.057	**0.852**	0.775	**0.701**	1.000
RR (v_2 = 5N)	0.934	0.876	0.933	1.046	0.902	0.910	0.806	1.010	0.902	0.854	0.813	0.995
RR (v_2 = 10N)	0.934	0.899	0.940	1.010	0.934	0.933	0.866	1.005	0.934	0.888	0.866	0.995
EN (30 variables, v_2 = 0.25)	1.016	0.910	1.239	1.077	0.951	1.000	0.993	1.021	0.934	**0.764**	0.873	**0.716**
EN (100 variables, v_2 = 0.25)	1.082	1.101	1.500	1.227	0.918	0.921	0.963	1.119	0.918	0.876	1.037	0.778
EN (200 variables, v_2 = 0.25)	1.082	1.011	1.373	1.351	0.836	0.955	0.940	1.206	0.885	0.843	0.963	0.938

What international information is important?

- Forecasts with subsets of international data with regional and economic breakdowns

- Compare
  - national and international forecast models
  - trade and data-rich models
  - forecasts based on subsets of data

- Historical forecast contributions of groups of variables (based on nat + int, RR (N), h=1)

- Best models?

What international information is important?

- Forecasts with subsets of international data with regional and economic breakdowns

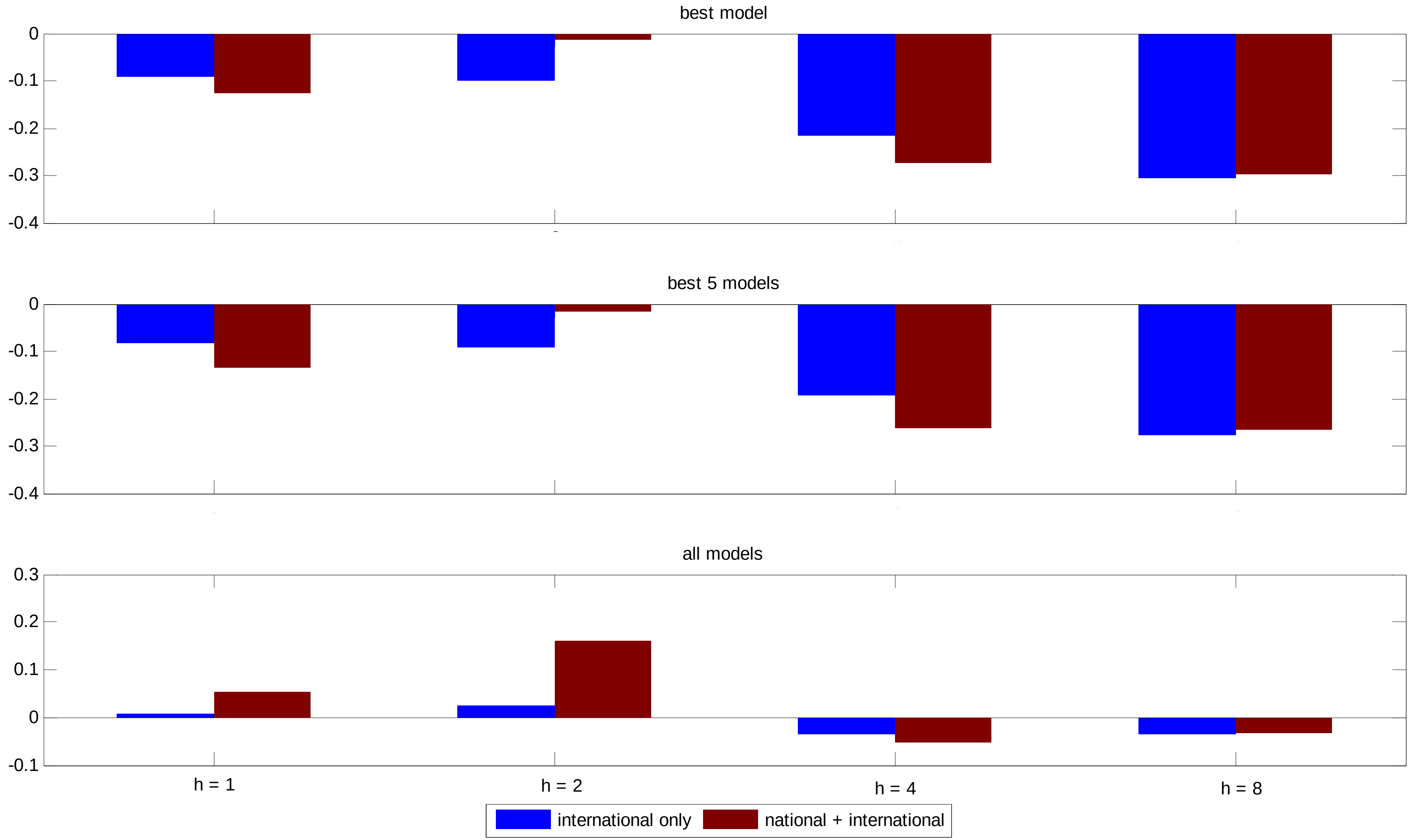
- Compare
  - national and international forecast models
  - trade and data-rich models
  - forecasts based on subsets of data (not in this presentation)

- Historical forecast contributions of groups of variables (based on nat + int, RR (N), h=1)

- Best models? (not in this presentation)

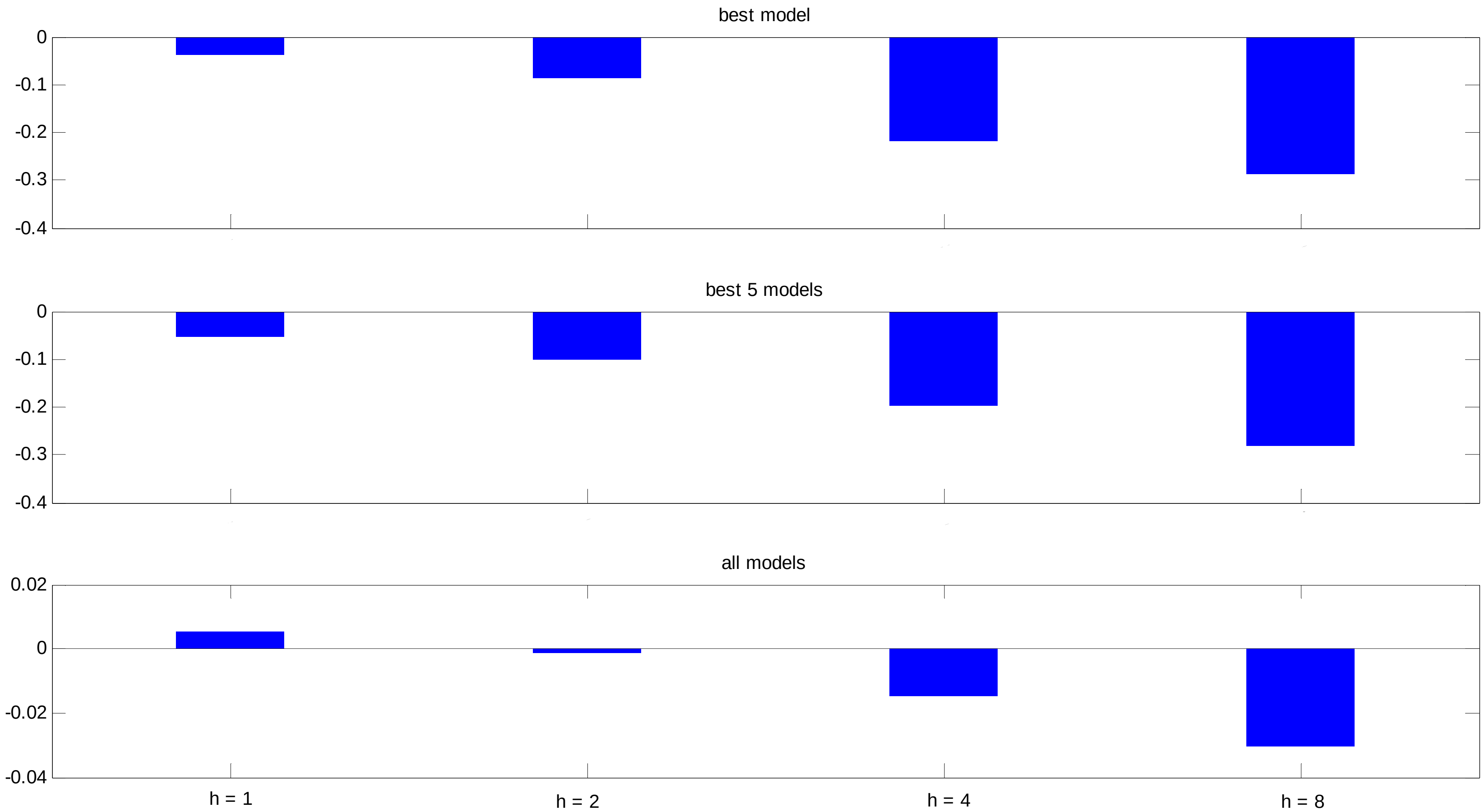
Forecast results – national vs. international

- RMSE(international (+ national))/RMSE(national) - 1



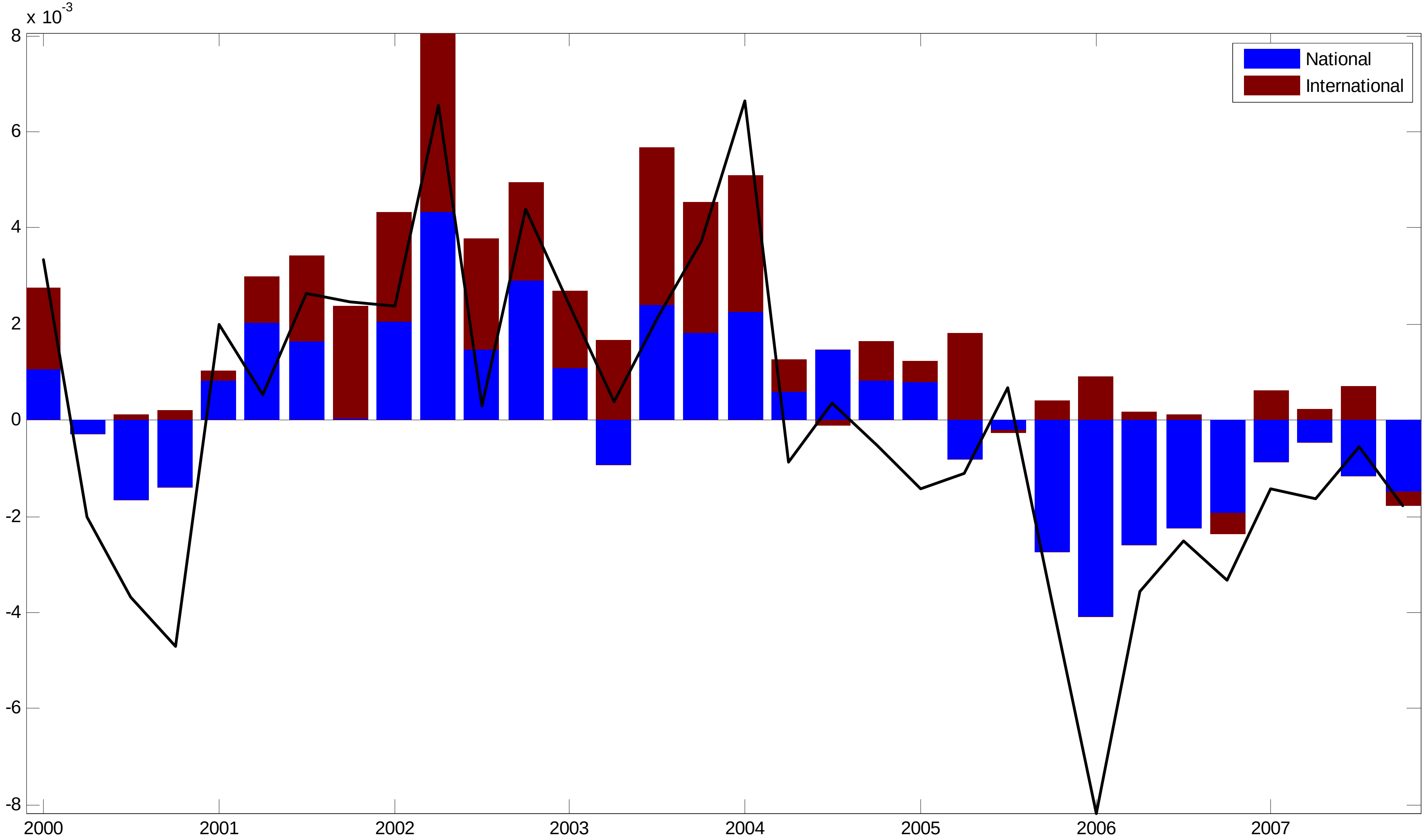
Forecast results – trade aggregates vs. data-rich

- RMSE(data-rich (int. only))/RMSE(trade) - 1



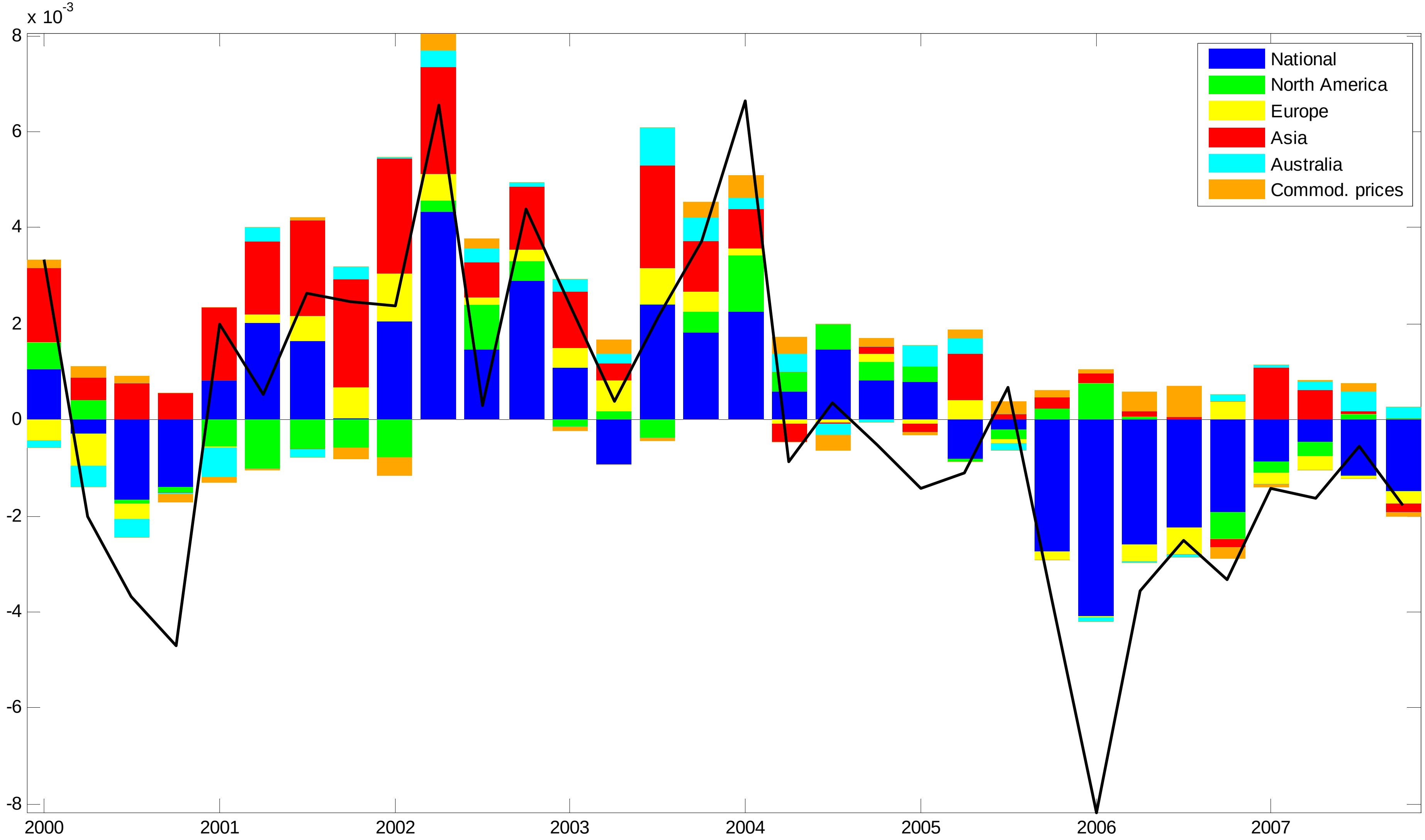
Contribution to NZL's GDP growth forecast

- National vs. international (model: nat + int, RR (N), h = 1)



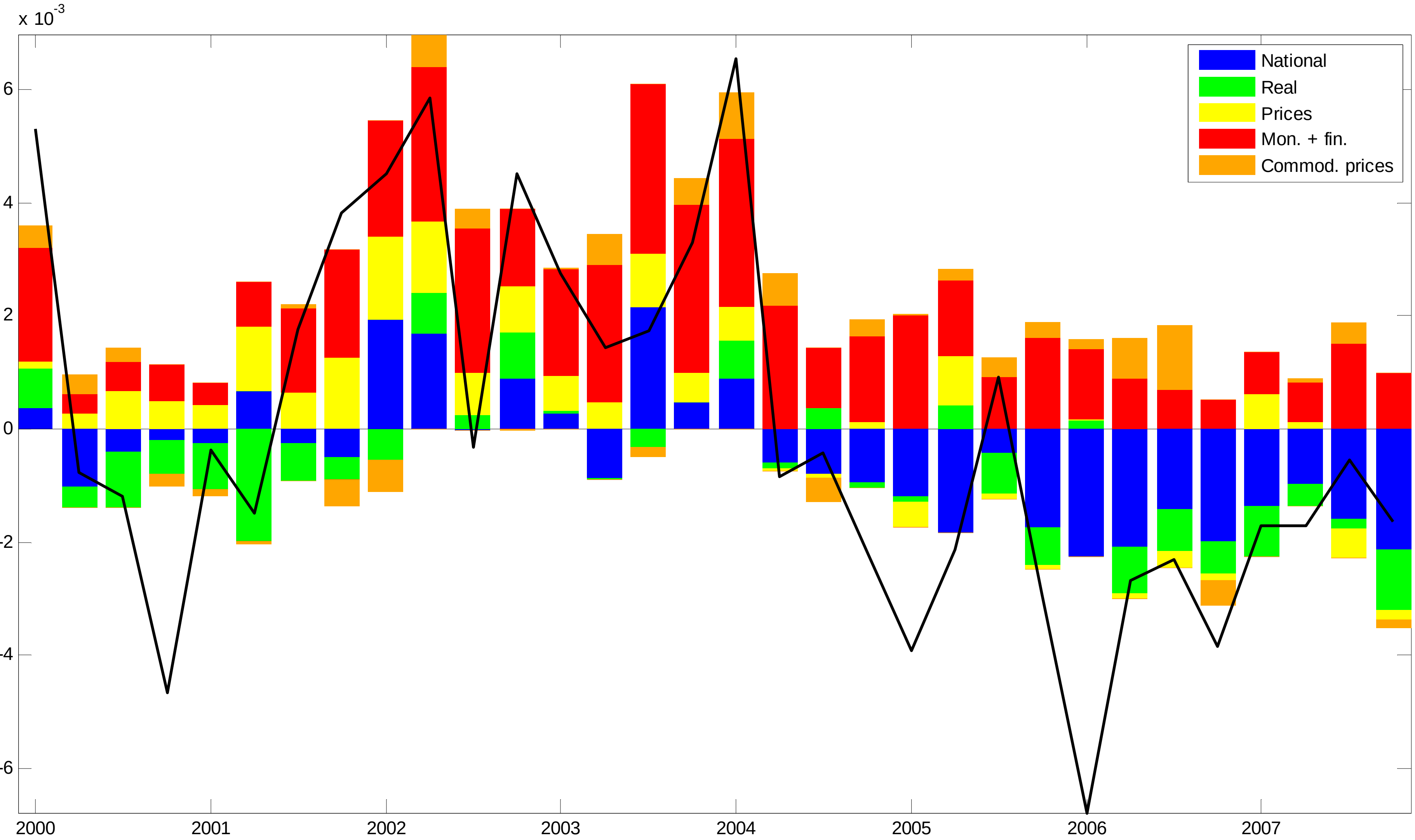
Contribution to NZL's GDP growth forecast

- Regional breakdown



Contribution to NZL's GDP growth forecast

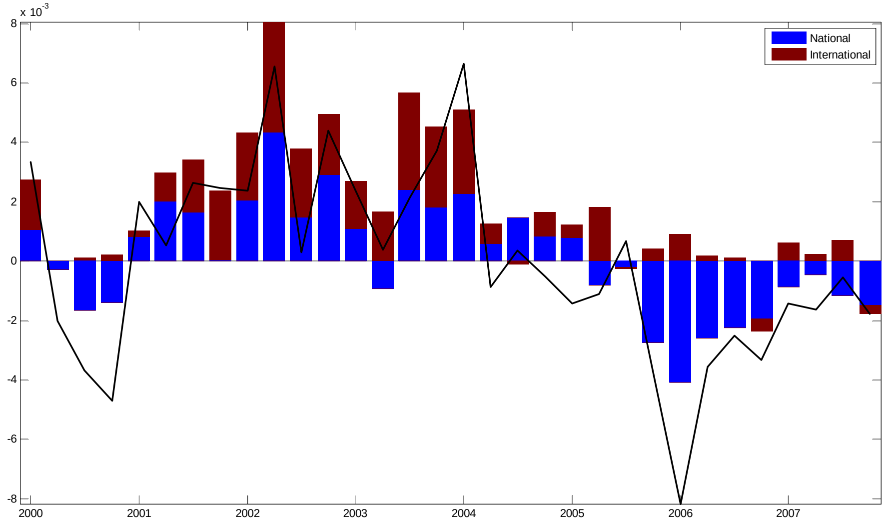
- Economic breakdown



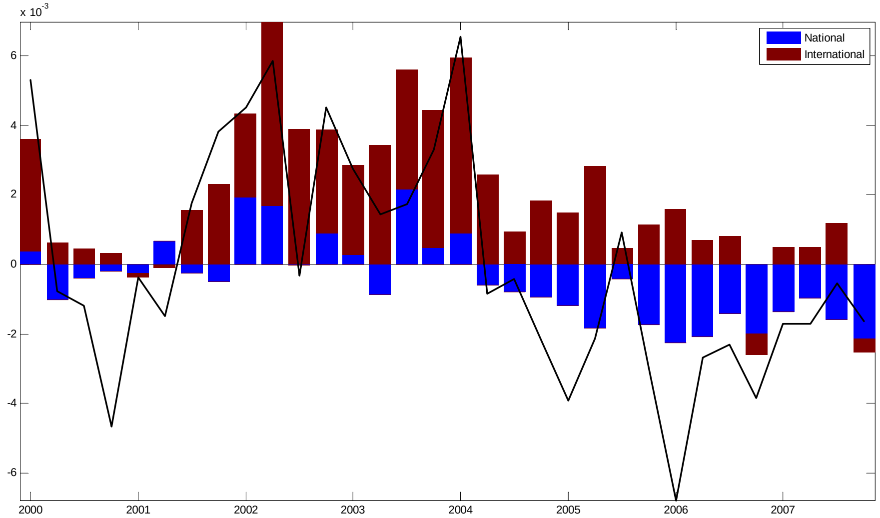
Do national surveys capture international information?

- National vs. international

with surveys



without surveys



- Correlation btw. survey factors and nat./int. factors: 0.43/0.49.

- Surveys seem to reflect mainly information on North America and foreign and monetary/financial developments

Contributions

- Systematic assessment of the value of exploiting international information in producing national forecasts

- Use of new data-rich approaches

- Link between forecasting and structural interpretation

Main findings

- International information can help improving national forecasts, with particularly large gains for long horizons.

- Information from large international datasets is worth exploiting.

- Performance of data-rich methods differs widely. RR, EN and PLS work particularly well in our international context.

- Histor. decomp. suggest interpretations of drivers of forecasts that accord with narratives about eco. circumstances of the time.

- National surveys capture international information.