

Supplementary material to the paper
“Forecasting using a large number of predictors:
is Bayesian regression a valid alternative
to principal components?”

Christine De Mol, Université Libre de Bruxelles, ECARES
Domenico Giannone*, European Central Bank, ECARES and CEPR
Lucrezia Reichlin, European Central Bank, ECARES and CEPR

March 5, 2008

*Corresponding author. European Central Bank, Directorate General Research, Kaiserstrasse
29, D-60311 Frankfurt am Main, Germany. Tel.: +49 69 1344 7849.
E-mail address: Domenico.Giannone@ecb.int

Table 1: Principal component forecasts

Industrial Production							
	Number of Principal Components						
	1	3	6	10	25	50	75
MSFE $p=0$	0.91	0.62	0.56	0.54	0.65	0.93	1.56
MSFE $p=6$	1.08	0.67	0.54	0.57	1.10	1.90	3.15
Variance* $p=0$	0.23	0.70	0.79	0.97	1.28	1.43	1.78
Variance* $p=6$	0.25	0.84	1.00	1.13	1.77	3.29	4.54

Consumer Price Index							
	Number of Principal Components						
	1	3	6	10	25	50	75
MSFE $p=0$	0.57	0.55	0.57	0.69	0.83	1.17	1.69
MSFE $p=6$	0.59	0.69	0.80	0.81	1.44	2.43	3.18
Variance* $p=0$	0.36	0.55	0.61	0.63	0.69	0.89	1.69
Variance* $p=6$	0.51	0.69	0.92	0.96	1.37	2.80	3.65

Evaluation sample: 1971-2002. MSFE are relative to a the Naive, Random Walk, forecast.
 *The variance of the forecast relative to the variance of the series.

Using lagged regressors

Empirical results including lagged predictors when computing the forecasts are not reported in the paper. In this additional section we look at the out-of-sample accuracy of the different forecasting methods when also lagged predictors are included. We consider the case $p = 6$. Qualitative results are robust across different choices of p . Results are reported in Table 1 (principal component regression), Table 2 (Bayesian regression with Gaussian prior) and Table 3 (Bayesian regression with double-exponential prior). We report the MSFE for the entire evaluation sample (1971-2002) for $p = 6$ and also for $p = 0$. The design of the exercise is the same as in Section 3 of the paper with the exception of the Bayesian predictions with double-exponential prior. Instead of fixing the number of non-zero coefficients in the entire evaluation sample, we select the prior ν that delivers a given number (k) of non zero coefficients in the initial sample 1959 – 1970. Then the prior ν is kept fixed at each estimation step over the evaluation sample. For this reason the results obtained here for $p = 0$ are not the same as those reported for Lasso regression in Section 3 of the paper.

Results suggest that including lagged regressors does not improve forecast accuracy. The forecasts based on PC become more volatile when using lags of the predictors. For Bayesian regression, with Gaussian and double-exponential prior, the volatility of the forecasts decreases when including lagged regressors. This suggests that with the dataset at hand Bayesian regression methods are more appropriate when using lagged predictors.

Table 2: Bayesian forecasts with Gaussian prior

Industrial Production									
	In-sample Residual variance								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
MSFE p=0	0.96	0.70	0.60	0.56	0.56	0.58	0.64	0.72	0.83
MSFE p=6	0.72	0.64	0.61	0.60	0.61	0.65	0.70	0.78	0.88
Variance* p=0	0.71	0.63	0.57	0.49	0.39	0.29	0.19	0.12	0.07
Variance* p=6	0.46	0.43	0.39	0.34	0.27	0.21	0.14	0.09	0.06

Consumer Price Index									
	In-sample Residual variance								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
MSFE p=0	0.88	0.72	0.66	0.63	0.62	0.63	0.66	0.73	0.84
MSFE p=6	0.74	0.70	0.67	0.66	0.65	0.65	0.67	0.70	0.80
Variance* p=0	0.41	0.35	0.32	0.28	0.24	0.19	0.13	0.08	0.05
Variance* p=6	0.32	0.33	0.31	0.29	0.26	0.22	0.17	0.11	0.06

Evaluation sample: 1971-2002. MSFE are relative to the Naive, Random Walk, forecast.
 *The variance of the forecast relative to the variance of the series.

Table 3: Lasso forecasts

Industrial Production							
	Number of non-zero coefficients						
	1	3	5	10	25	50	75
MSFE p=0	0.86	0.69	0.64	0.60	0.64	0.77	1.10
MSFE p=6	0.73	0.72	0.72	0.70	0.57	0.57	0.58
Variance* p=0	0.07	0.16	0.24	0.40	0.53	0.65	0.79
Variance* p=6	0.11	0.12	0.12	0.14	0.38	0.44	0.51

Consumer Price Index							
	Number of non-zero coefficients						
	1	3	5	10	25	50	75
MSFE p=0	0.90	0.76	0.62	0.59	0.68	0.86	1.06
MSFE p=6	0.86	0.75	0.71	0.66	0.66	0.67	0.68
Variance* p=0	0.05	0.09	0.18	0.26	0.33	0.39	0.50
Variance* p=6	0.05	0.08	0.10	0.18	0.25	0.34	0.38

Evaluation sample: 1971-2002. MSFE are relative to the Naive, Random Walk, forecast.
 *The variance of the forecast relative to the variance of the series.

Description of the dataset

Table 4: Data transformation

	Definition	Transformation
1	$X_{it} = Z_{it}$	no transformation
2	$X_{it} = \Delta Z_{it}$	monthly difference
4	$X_{it} = \ln Z_{it}$	log
5	$X_{it} = \Delta \ln Z_{it} \times 100$	monthly growth rate
6	$X_{it} = \Delta \ln \frac{Z_{it}}{Z_{it-12}} \times 100$	monthly difference of yearly growth rate

Table 5: Data description

Code	Description	Transf.	Lasso Selection*	
			IP	CPI
a0m052	Personal income (AR, bil. chain 2000 \$)	5		
A0M051	Personal income less transfer payments (AR, bil. chain 2000 \$)	5		II
A0M224 R	Real Consumption (AC) A0m224gmdc	5		
A0M057	Manufacturing and trade sales (mil. Chain 1996 \$)	5		
A0M059	Sales of retail stores (mil. Chain 2000 \$)	5		
IPS10	INDUSTRIAL PRODUCTION INDEX - TOTAL INDEX	5		
IPS11	INDUSTRIAL PRODUCTION INDEX - PRODUCTS, TOTAL	5		
IPS299	INDUSTRIAL PRODUCTION INDEX - FINAL PRODUCTS	5		
IPS12	INDUSTRIAL PRODUCTION INDEX - CONSUMER GOODS	5		
IPS13	INDUSTRIAL PRODUCTION INDEX - DURABLE CONSUMER GOODS	5		
IPS18	INDUSTRIAL PRODUCTION INDEX - NONDURABLE CONSUMER GOODS	5		
IPS25	INDUSTRIAL PRODUCTION INDEX - BUSINESS EQUIPMENT	5		
IPS32	INDUSTRIAL PRODUCTION INDEX - MATERIALS	5	II	
IPS34	INDUSTRIAL PRODUCTION INDEX - DURABLE GOODS MATERIALS	5		II
IPS38	INDUSTRIAL PRODUCTION INDEX - NONDURABLE GOODS MATERIALS	5	II	
IPS43	INDUSTRIAL PRODUCTION INDEX - MANUFACTURING (SIC)	5		
IPS307	INDUSTRIAL PRODUCTION INDEX - RESIDENTIAL UTILITIES	5		
IPS306	INDUSTRIAL PRODUCTION INDEX - FUELS	5		
PMP	NAPM PRODUCTION INDEX (PERCENT)	1		
A0m082	Capacity Utilization (Mfg)	2		
LHEL	INDEX OF HELP-WANTED ADVERTISING IN NEWSPAPERS (1967=100;SA)	2		
LHELX	EMPLOYMENT: RATIO; HELP-WANTED ADS:NO. UNEMPLOYED CLF	2		
LHEM	CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS.,SA)	5		
LHNAG	CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC.INDUSTRIES (THOUS.,SA)	5		
LHUR	UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS & OVER (%;SA)	2		
LHU680	UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA)	2		
LHU5	UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)	5		
LHU14	UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA)	5		
LHU15	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS.,SA)	5		
LHU26	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 TO 26 WKS (THOUS.,SA)	5		
LHU27	UNEMPLOY.BY DURATION: PERSONS UNEMPL.27 WKS + (THOUS.,SA)	5		
A0M005	Average weekly initial claims, unemploy. insurance (thous.)	5		
CES002	EMPLOYEES ON NONFARM PAYROLLS - TOTAL PRIVATE	5		
CES003	EMPLOYEES ON NONFARM PAYROLLS - GOODS-PRODUCING	5		
CES006	EMPLOYEES ON NONFARM PAYROLLS - MINING	5		II
CES011	EMPLOYEES ON NONFARM PAYROLLS - CONSTRUCTION	5		
CES015	EMPLOYEES ON NONFARM PAYROLLS - MANUFACTURING	5		
CES017	EMPLOYEES ON NONFARM PAYROLLS - DURABLE GOODS	5		
CES033	EMPLOYEES ON NONFARM PAYROLLS - NONDURABLE GOODS	5		
CES046	EMPLOYEES ON NONFARM PAYROLLS - SERVICE-PROVIDING	5		
CES048	EMPLOYEES ON NONFARM PAYROLLS - TRADE, TRANSPORTATION, AND UTILITIES	5		
CES049	EMPLOYEES ON NONFARM PAYROLLS - WHOLESALE TRADE	5	II	
CES053	EMPLOYEES ON NONFARM PAYROLLS - RETAIL TRADE	5		
CES088	EMPLOYEES ON NONFARM PAYROLLS - FINANCIAL ACTIVITIES	5	I	I
CES140	EMPLOYEES ON NONFARM PAYROLLS - GOVERNMENT	5		
A0M048	Employee hours in nonag. establishments (AR, bil. hours)	5		
CES151	AVG WEEKLY HOURS OF PROD. OR NONSUPERV. WORKERS ON PRIVATE NONFARM	1		
CES155	AVG WEEKLY HOURS OF PROD. OR NONSUPERV. WORKERS ON PRIVATE NONFARM	2		
aom001	Average weekly hours, mfg. (hours)	1		
PMEMP	NAPM EMPLOYMENT INDEX (PERCENT)	1		I
HSPR	HOUSING STARTS:NONFARM(1947-58);TOTAL FARM&NONFARM(1959-)(THOUS.,SA	4		
HSNE	HOUSING STARTS:NORTHEAST (THOUS.U.)S.A.	4	II	
HSMW	HOUSING STARTS:MIDWEST(THOUS.U.)S.A.	4		II
HSSOU	HOUSING STARTS:SOUTH (THOUS.U.)S.A.	4		
HSWST	HOUSING STARTS:WEST (THOUS.U.)S.A.	4		I
HSBR	HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS.,SAAR)	4	I	
HSBNE	HOUSES AUTHORIZED BY BUILD. PERMITS:NORTHEAST(THOU.U.)S.A	4	II	I
HSBMW	HOUSES AUTHORIZED BY BUILD. PERMITS:MIDWEST(THOU.U.)S.A.	4		I-II
HSBSOU	HOUSES AUTHORIZED BY BUILD. PERMITS:SOUTH(THOU.U.)S.A.	4		I
HSBWST	HOUSES AUTHORIZED BY BUILD. PERMITS:WEST(THOU.U.)S.A.	4		

Table 5: Data description

Code	Description	Transf.	Lasso Selection*	
			IP	CPI
FMI	PURCHASING MANAGERS' INDEX (SA)	1		
PMNO	NAPM NEW ORDERS INDEX (PERCENT)	1	I	
PMDEL	NAPM VENDOR DELIVERIES INDEX (PERCENT)	1		
PMNV	NAPM INVENTORIES INDEX (PERCENT)	1		II
A0M008	Mfrs' new orders, consumer goods and materials (bil. chain 1982 \$)	5		
A0M007	Mfrs' new orders, durable goods industries (bil. chain 2000 \$)	5		
A0M027	Mfrs' new orders, nondefense capital goods (mil. chain 1982 \$)	5		
A1M092	Mfrs' unfilled orders, durable goods indus. (bil. chain 2000 \$)	5	I	
A0M070	Manufacturing and trade inventories (bil. chain 2000 \$)	5	I	
A0M077	Ratio, mfg. and trade inventories to sales (based on chain 2000 \$)	2		
FM1	MONEY STOCK: M1(CURR,TRAV,CKS,DEM DEP,OTHER CK'ABLE DEP)(BIL\$,SA)	6		
FM2	MONEY STOCK:M2(M1+O'NITE RPS,EURO\$,GP&BD MMMFS&SAV&SM TIME DEP(BIL\$,	6		I
FM3	MONEY STOCK: M3(M2+LG TIME DEP,TERM RP'S&INST ONLY MMMFS)(BIL\$,SA)	6		
FM2DQ	MONEY SUPPLY - M2 IN 1996 DOLLARS (BCI)	5	I	I
FMFBA	MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES(MIL\$,SA)	6		
FMRRRA	DEPOSITORY INST RESERVES:TOTAL,ADJ FOR RESERVE REQ CHGS(MIL\$,SA)	6		
FMRNBA	DEPOSITORY INST RESERVES:NONBORROWED,ADJ RES REQ CHGS(MIL\$,SA)	6		
FCLNQ	COMMERCIAL & INDUSTRIAL LOANS OUSTANDING IN 1996 DOLLARS (BCI)	6		
FCLBMC	WKLY RP LG COM'L BANKS:NET CHANGE COM'L & INDUS LOANS(BIL\$,SAAR)	1		II
CCINRV	CONSUMER CREDIT OUTSTANDING - NONREVOLVING(G19)	6		
A0M095	Ratio, consumer installment credit to personal income (pct.)	2		
FSPCOM	S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)	5		
FSPIN	S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)	5	II	
FSDXP	S&P'S COMPOSITE COMMON STOCK: DIVIDEND YIELD (% PER ANNUM)	2	I	
FSPXE	S&P'S COMPOSITE COMMON STOCK: PRICE-EARNINGS RATIO (%NSA)	5		
FYFF	INTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM,NSA)	2		
CP90	Cmmercial Paper Rate (AC)	2		
FYGM3	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA)	2		
FYGM6	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.(% PER ANN,NSA)	2		
FYGT1	INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA)	2		
FYGT5	INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.(% PER ANN,NSA)	2		
FYGT10	INTEREST RATE: U.S.TREASURY CONST MATURITIES,10-YR.(% PER ANN,NSA)	2		
FYAAAC	BOND YIELD: MOODY'S AAA CORPORATE (% PER ANNUM)	2		
FYBAAC	BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM)	2		II
scp90	cp90-fyff	1		II
sfygm3	fygm3-fyff	1	I	
sFYGM6	fygm6-fyff	1		
sFYGT1	fygt1-fyff	1		
sFYGT5	fygt5-fyff	1		
sFYGT10	fygt10-fyff	1	II	
sFYAAAC	fyaaac-fyff	1		
sFYBAAC	fybaac-fyff	1		
EXRUS	UNITED STATES;EFFECTIVE EXCHANGE RATE(MERM)(INDEX NO.)	5		
EXRSW	FOREIGN EXCHANGE RATE: SWITZERLAND (SWISS FRANC PER U.S.\$)	5		
EXRJAN	FOREIGN EXCHANGE RATE: JAPAN (YEN PER U.S.\$)	5		
EXRUK	FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)	5		
EXRCAN	FOREIGN EXCHANGE RATE: CANADA (CANADIAN \$ PER U.S.\$)	5		
PWFSA	PRODUCER PRICE INDEX: FINISHED GOODS (82=100,SA)	6		
PWFCSA	PRODUCER PRICE INDEX:FINISHED CONSUMER GOODS (82=100,SA)	6		
PWMSA	PRODUCER PRICE INDEX:INTERMED MAT.SUPPLIES & COMPONENTS(82=100,SA)	6		
PWCMSA	PRODUCER PRICE INDEX:CRUDE MATERIALS (82=100,SA)	6		
PSM99Q	INDEX OF SENSITIVE MATERIALS PRICES (1990=100)(BCI-99A)	6	I	
PMCP	NAPM COMMODITY PRICES INDEX (PERCENT)	1	II	II
PUNEW	CPI-U: ALL ITEMS (82-84=100,SA)	6		I
PU83	CPI-U: APPAREL & UPKEEP (82-84=100,SA)	6		
PU84	CPI-U: TRANSPORTATION (82-84=100,SA)	6	I	
PU85	CPI-U: MEDICAL CARE (82-84=100,SA)	6	II	
PUC	CPI-U: COMMODITIES (82-84=100,SA)	6		
PUCD	CPI-U: DURABLES (82-84=100,SA)	6		
PUS	CPI-U: SERVICES (82-84=100,SA)	6		
PUXF	CPI-U: ALL ITEMS LESS FOOD (82-84=100,SA)	6		
PUXHS	CPI-U: ALL ITEMS LESS SHELTER (82-84=100,SA)	6		
PUXM	CPI-U: ALL ITEMS LESS MIDICAL CARE (82-84=100,SA)	6		
GMDC	PCE,IMPL PR DEFL:PCE (1987=100)	6		
GMDCD	PCE,IMPL PR DEFL:PCE; DURABLES (1987=100)	6		
GMDCN	PCE,IMPL PR DEFL:PCE; NONDURABLES (1996=100)	6		
GMDCS	PCE,IMPL PR DEFL:PCE; SERVICES (1987=100)	6		
CES275	AVG HOURLY EARNINGS OF PROD. OR NONSUPERV. WORKERS ON PRIVATE NONFARM	6		I
CES277	AVG HOURLY EARNINGS OF PROD. OR NONSUPERV WORKERS ON PRIVATE NONFARM	6	II	
CES278	AVG HOURLY EARNINGS OF PROD. OR NONSUPERV. WORKERS ON PRIVATE NONFARM	6		
HHSNTN	U. OF MICH. INDEX OF CONSUMER EXPECTATIONS(BCD-83)	2		

*We indicate when, for forecasting IP or CPI, the variable has been selected by Lasso regression at the beginning (I), 1970 : 1, and/or and the end (II),2001 : 12, of the out-of-sample evaluation period.