Model simplification and variable selection: A Replication of the UK inflation model by Hendry (2001)

Błażejowski, Marcin and Kufel, Paweł and Kwiatkowski, Jacek

WSB University in Torun, Nicolas Copernicus University in Torun

30 August 2018

Model simplification and variable selection: A Replication of the UK inflation model by Hendry (2001)

Marcin Błażejowski*†, Paweł Kufel†, Jacek Kwiatkowski‡

†*WSB University in Torun*

‡*Nicolas Copernicus University in Torun*

September 6, 2019

Acknowledgments: Financial support from the National Center of Science, Poland (contract/grant number UMO-2016/21/B/HS4/01970), is gratefully acknowledged. We are grateful to Riccardo Lucchetti, Magdalena Osińska, Mark Steel and two reviewers for helpful suggestions and comments.

Conflict of Interest Statement: The authors declare no conflict of interest.

Abstract

In this paper, we revisit the well-known UK inflation model by Hendry (*Journal of Applied Econometrics* 2001, 16:255-275. doi: 10.1002/jae.615). We replicate the results in a narrow sense using the gretl and PcGive programs. In a wide sense, we extend the study of model uncertainty using the Bayesian averaging of classical estimates (BACE) approach as automatic model reduction strategy. We consider three different specifications to compare BACE variable selection with Hendrys' reduction. We find that BACE method can recover the path of non-trivial reduction strategy.

Keywords: BACE, gretl, model uncertainty, reduction strategy

1 Introduction

This paper concerns a replication of a model of UK inflation, 1875–1991, by Hendry (2001) based on data provided by JAE services at (http://qed.econ.queensu.ca/jae/2001-v16.3/hendry). To replicate Hendrys' procedure for modeling inflation in the UK in a narrow sense, we used the gretl[1] (see Cottrell & Lucchetti, 2018) and PcGive/Autometrics (see Doornik, 2009) program[2]. Our extension, in a wide sense, of Hendrys' work employed the Bayesian averaging of classical estimates (BACE) approach proposed by Sala-i-Martin, Doppelhofer, and Miller (2004) to compare model reduction strategies and the variable selection procedure.

*Correspondence to: Wyższa Szkoła Bankowa w Toruniu, Marcin Błażejowski, ul. Młodzieżowa 31a, 87-100 Toruń. Email: marcin.blazejowski@wsb.torun.pl. Phone: +48566609245.

[1]Gretl is an open-source software for econometric analysis available at http://gretl.sf.net.

[2]We used gretl version 2019a and PcGive version 14.2 with Ox Professional version 7.20 on a PC machine running under Debian GNU/Linux 64 bit.

When we consider the large number of variables, it is difficult to decide which model is the most appropriate for analyzing the dependencies, i.e., to find the optimal set of variables in terms of goodness of fit measures. Using BACE, we can obtain the most probable set of determinants along with posterior parameter estimates based on the whole model space instead of making decisions based only on a single model. This approach is an alternative to the earlier and familiar Bayesian model averaging (BMA) (see Fernández, Ley, & Steel, 2001; Ley & Steel, 2012), from which it differs by having non-informative prior assumptions of regression parameters. Discussion on the effect of prior assumptions in BMA is presented in Ley and Steel (2009). Sala-i-Martin et al. (2004) showed that the BACE approach may be understood as Bayesian analysis in the situation where prior information is ”dominated” by the data. The parameter estimates are averaged across all possible combinations of models obtained by means of OLS. In our case, the BACE analysis was performed in the BACE 2.0 package for the gretl program[3] (see Błażejowski & Kwiatkowski, 2018).

The remainder of the paper is structured as follows. In section 2, we discuss issues related to data transformation, section 3 presents the research scenario for the replication in a narrow sense. In section 4 we consider model uncertainty using BACE selection strategy (replication in a wide sense), section 5 concludes.

2 Data

In our paper and replication files, we used the same data sample period (1875–1991) and definitions as in Hendry (2001)[4] with the following exceptions:

1. *Profit markup* (π_t^*) was taken directly from the `jaedfh4.dat` file (part of the `dfhdata.zip` archive); this variable exists as ”`pistarn`” in the JAE archive.

2. *Short-long spread* $(R_{s,t} - R_{l,t} + 0.006)$ was named S_t, similar to Clements and Hendry (2008, pp. 11).

3. *Excess demand* $\left(y_t^d\right)$ was taken directly from the `jaedfhm.dat` file (part of the `dfhdata.zip` archive); this variable exists as ”`gdpd`” in the JAE archive.

4. *The real exchange rate* was defined as $e_{r,t} = p_t - p_{£,t} - 0.52$. We found an inaccuracy in the paper by Hendry (2001) and data definitions in the JAE archive. The calculation of $e_{r,t} = p_t - p_{£,t} + 0.52$ (equation (3) in Hendry (2001, pp. 263)) is misleading with the form of calculating $e_{r,t}$ in the formula for ”`pistarn`” (`readme.h.txt` file) and refers to subtracting (not adding) the intercept value (0.52).

5. According to formulas in the JAE archive (`readme.h.txt` file), the variable *Unit labor costs in constant prices* was defined as $(c-p)_t^* = c_t - p_t + 0.006 \times (trend - 69.5) + 2.37$.

3 Research scenario

To replicate the Hendry (2001) results in a narrow sense, we proceed as follows. In the case of the initial model for all 52 variables, we received identical output to that in the original model (`GUM52`; residual standard deviation $\hat{\sigma} = 1.21\%$,

[3] The BACE 1.2 package is available at http://ricardo.ecn.wfu.edu/gretl/cgi-bin/gretldata.cgi?opt=SHOW_FUNCS.
[4] Description of variables are available in table 6 in Appendix.

Schwartz Criterion $SC = -7.3$). After excluding indicators from the initial model, we also received identical results (`GUMnoIndicators`; $\hat{\sigma} = 2.5\%$, $SC = -6.63$). In the next step, we added dummy variables I_b, I_l, I_m concerning outliers in particular years to (`GUMnoIndicators`), and we obtained the same results as in the paper (`GUMfirstReduction`; $\hat{\sigma} = 1.16\%$, $SC = -8.08$). In the next step, the dummy variables I_b, I_l, I_m were substituted by one overall index, I_d, and once again, we obtained the same results (`GUMsecondReduction`; $\hat{\sigma} = 1.15\%$, $SC = -8.16$). Finally, we expressed the general model in terms of π^*_{t-1} with indicators restricted to I_d (`GUMfinal`; $\hat{\sigma} = 1.15\%$, $SC = -8.33$). At this point, we had the following specification:

$$\Delta p_t = f(\Delta p_{t-1}, y^d_{t-1}, m^d_{t-1}, n^d_{t-1}, U^d_{t-1}, S_{t-1}, R_{l,t-1}, \Delta p_{e,t}, \Delta p_{e,t-1}, \Delta U_{r,t-1},$$
$$\Delta w_{t-1}, \Delta c_{t-1}, \Delta m_{t-1}, \Delta n_{t-1}, \Delta R_{s,t-1}, \Delta R_{l,t-1}, \Delta p_{o,t-1}, I_{d,t}, \pi^*_{t-1}; \varepsilon_t). \quad (1)$$

After the reduction of model (1) at a 1% significance level[5], we obtained Hendrys' model (6) (Hendry, 2001, pp. 267) and (`FinalModel`; $\hat{\sigma} = 1.14\%$, $SC = -8.66$) in our notation. Detailed results are available in the table 1.

Table 1: Comparison of Hendrys' estimates and the replication results

	Hendrys' model (6)	Replication in gretl	Replication in Autometrics
y^d_{t-1}	0.180 (0.032)	0.184 (0.032)	0.184 (0.032)
Δm_{t-1}	0.187 (0.028)	0.182 (0.028)	0.182 (0.028)
S_{t-1}	−0.834 (0.088)	−0.834 (0.087)	−0.834 (0.087)
$\Delta R_{s,t-1}$	0.618 (0.106)	0.619 (0.106)	0.619 (0.106)
π^*_{t-1}	−0.186 (0.024)	−0.186 (0.024)	−0.186 (0.024)
$\Delta p_{e,t}$	0.265 (0.025)	0.265 (0.025)	0.265 (0.025)
$I_{d,t}$	0.038 (0.002)	0.038 (0.002)	0.038 (0.002)
$\Delta p_{o,t-1}$	0.041 (0.010)	0.041 (0.010)	0.041 (0.010)
Δp_{t-1}	0.267 (0.027)	0.268 (0.026)	0.268 (0.026)
R^2	0.975	0.975	0.975
$\hat{\sigma}$	1.14%	1.14%	1.14%
SC	−8.66	−8.66	−8.66

Standard errors in parentheses.

According to results in table 1, we found minor differences in the coefficient estimates for four variables—y^d_{t-1}, Δm_{t-1}, $\Delta R_{s,t-1}$, and Δp_{t-1}—and two differences in standard errors for S_{t-1} and Δp_{t-1}. The remaining coefficients and the model statistics were identical. In his paper, Hendry used the PcGets automatic model selection procedure with a 1% significance level for the model (1) to check the correctness of the simplification[6]. We repeated this automatic model selection procedure using Autometrics for model (1), and we obtained the same estimates as in `GUMfinal` in gretl (i.e., with slight differences compared to model (6) in Hendrys' paper). We suppose that these differences are due to the precision of the computers used for calculations, the 32 or 64 bits software or precision involved in data storage.

[5] Hendry and Krolzig (2001) classified simplification at a 1% significance level as a "conservative" strategy and at a 5% significance level as a "liberal" strategy. Currently these strategies are renamed the "small" and "standard" target size (see Doornik & Hendry, 2013).

[6] See subsection 4.3 in Hendry (2001).

4 BACE analysis

In order to perform replication in a wide sense, we analyzed model selection procedure employing the BACE approach. This procedure enables searching the whole model space and selecting the most probable regressions. The BACE also enables calculations of the averages of the posterior means and standard deviations of parameters as well as posterior inclusion probabilities (PIP). The posterior inclusion probability is the probability that, conditional on the data but unconditional with respect to the model space, the independent variable is relevant in explaining dependent variable. PIP is calculated as the frequency of appearance in Marcov chain of a given variable in all considered models (see Doppelhofer & Weeks, 2009; Koop, Poirier, & Tobias, 2007). We assumed three scenarios for BACE analysis:

1. `GUM52`—we used all 52 available explanatory variables (including trend and outliers).
2. `GUMId`—we used I_d composite indicator which captures the 22 large outliers and elements of π^*.
3. `GUMfinal`—we used I_d and π^* variables.

4.1 GUM52 results

We formulated the following initial inflation model (`GUM52`) for the UK in 1875-1991:

$$\begin{aligned}\Delta p_t = &f\Big(\Delta p_{t-1}, y^d_{t-1}, m^d_{t-1}, n^d_{t-1}, U^d_{t-1}, e_{r,t-1}, c_{t-1}, S_{t-1}, p_{t-1}, p_{o,t-1}, R_{l,t-1}, \Delta p_{e,t-1}, \Delta p_{e,t},\\ &\Delta U_{r,t-1}, \Delta w_{t-1}, \Delta c_{t-1}, \Delta m_{t-1}, \Delta n_{t-1}, \Delta R_{s,t-1}, \Delta R_{l,t-1}, \Delta p_{o,t-1}, trend, outliers;\ \varepsilon_t\Big). \qquad (2)\end{aligned}$$

Trend, individual-year indicators (outliers) and the rest of the variables, are defined identically as in Hendry (2001). The whole model space in `GUM52` (including intercept) was equal to $2^{52} = 4,503,599,627,370,496$. The total number of Monte Carlo iterations was 5,000,000 (including 25% burn-in draws). Model prior was set to uniform, which means that all possible specifications are equally probable (or, in other words, we do not prefer any of 2^{52} possible specifications). Additionally we performed jointness analysis based on J_{LS} measure introduced in Ley and Steel (2007), because it can identify independence, substitutes and complements between variables in the regression model (see Doppelhofer & Weeks, 2009).

Results of BACE analysis for the `GUM52` are in table 2. We can draw the following conclusions:

1. We found that 6 out of 7 variables from Hendrys' final model (excluding I_d and π^*) are highly probable (PIP $\geq 2/3$). The 7th variable, i.e. y^d_{t-1} (GDP excess demand), is lowly probable (PIP $< 1/3$). Instead of y^d_{t-1} we obtained Δn_{t-1} (nominal National Debt), which is classified as medium probable variable ($1/3 \leq$ PIP $< 2/3$). We can find a possible explanation of this discrepancy using jointness analysis. According to jointness results, variable y^d_{t-1} is a strong substitute for Δn_{t-1} with value of $J_{LS}(y^d_{t-1}, \Delta n_{t-1}) = -2.0431$ (Doppelhofer & Weeks, 2009), hence we can say that these two variables swap between each other.

2. Individual-year indicators can be divided into three groups:

Table 2: BACE posterior inclusion probabilities and posterior estimates of regression coefficients for the `GUM52`

	PIP	Avg. Mean	Avg. Std. Dev.		PIP	Avg. Mean	Avg. Std. Dev.
I_{1921}	1.0000	-0.1522	0.0194	Δn_{t-1}	0.5937	0.0357	0.0378
I_{1915}	1.0000	0.1046	0.0143	I_{1916}	0.5621	0.0223	0.0249
I_{1922}	0.9999	-0.1224	0.0183	I_{1944}	0.3638	-0.0085	0.0144
I_{1975}	0.9978	0.0729	0.0178	I_{1918}	0.3142	0.0052	0.0213
I_{1917}	0.9959	0.1006	0.0279	I_{1914}	0.3017	-0.0060	0.0117
I_{1881}	0.9947	-0.0494	0.0146	$p_{o,t-1}$	0.2861	0.0011	0.0025
Δp_{t-1}	0.9943	0.2964	0.0641	p_{t-1}	0.2427	-0.0008	0.0145
$\Delta p_{e,t}$	0.9938	0.1741	0.0447	c_{t-1}	0.2414	0.0013	0.0130
I_{1880}	0.9938	0.0499	0.0151	$const$	0.2258	0.0024	0.0241
Δm_{t-1}	0.9919	0.2419	0.0590	I_{1901}	0.2224	-0.0039	0.0096
S_{t-1}	0.9912	-0.6809	0.1572	$I_{2,t}$	0.2174	0.0016	0.0047
I_{1980}	0.9899	0.0499	0.0159	$\Delta U_{r,t-1}$	0.2041	-0.0309	0.0865
I_{1973}	0.9891	-0.0579	0.0162	$trend$	0.1922	<0.0001	0.0001
I_{1971}	0.9887	0.0447	0.0143	$R_{l,t-1}$	0.1668	0.0099	0.0501
I_{1943}	0.9852	-0.0538	0.0171	n^d_{t-1}	0.1636	-0.0006	0.0022
I_{1945}	0.9759	-0.0472	0.0162	Δc_{t-1}	0.1511	0.0095	0.0371
$\Delta R_{s,t-1}$	0.9743	0.5235	0.1778	m^d_{t-1}	0.1442	-0.0010	0.0085
I_{1900}	0.9710	0.0401	0.0149	y^d_{t-1}	0.1336	0.0056	0.0275
I_{1919}	0.9578	0.0532	0.0222	$\Delta p_{e,t-1}$	0.1232	0.0037	0.0192
$\Delta p_{o,t-1}$	0.9308	0.0417	0.0199	I_{1941}	0.1223	-0.0012	0.0067
I_{1920}	0.9248	0.0469	0.0221	Δw_{t-1}	0.1126	0.0046	0.0337
I_{1940}	0.8781	0.0375	0.0206	I_{1974}	0.1119	0.0004	0.0089
$e_{r,t-1}$	0.8070	-0.0361	0.0238	U^d_{t-1}	0.1070	0.0004	0.0188
I_{1970}	0.8030	0.0261	0.0177	$\Delta R_{l,t-1}$	0.1027	0.0115	0.1362
I_{1942}	0.7439	-0.0266	0.0206	I_{1946}	0.0979	0.0005	0.0049
I_{1939}	0.6401	0.0193	0.0184	I_{1979}	0.0943	0.0003	0.0050

(a) highly probable: 1880, 1881, 1900, 1915, 1917, 1919, 1920, 1921, 1922, 1940, 1942, 1943, 1945, 1970, 1971, 1973, 1975, 1980 (18 variables).

(b) medium probable: 1916, 1939, 1944 (3 variables).

(c) lowly probable: 1901, 1914, 1918, 1941, 1946, 1974, 1979 (7 variables).

The results coincided with a set of variables included into composite indicator I_d which captures the 22 large outliers. The only exception is the indicator for 1918, which according to our analysis is lowly probable ($\text{PIP}_{I_{1918}} \approx 0.31$).

3. Although the π^* is not directly present in the `GUM52`, variable e_r (being an element of π^*) is also highly probable ($\text{PIP}_{e_r} \approx 0.81$).

In our opinion BACE variable selection results for `GUM52` are close to Hendrys' final selection. It is worth to mention that these two selections can not be compared directly, because in `GUM52` we do not have two important components, i.e. I_d and π^*, since we considered their elements separately. Having the above in mind we are convinced that in order to maintain the comparability of results we should replace individual-year indicators by the I_d composite indicator and include elements of π^*[7].

[7] Profit markup is defined as follows: $\pi^*_t = 0.25e_{r,t} - 0.675(c-p)^*_t - 0.075(p_o - p)_t + 0.11I_{2,t} + 0.25$ where $(c-p)^*_t = c_t - p_t + 0.006 \times (trend - 69.5) + 2.37$.

4.2 GUMId results

Since BACE results for `GUM52` exhibit the important role of individual-year indicators, we decided to include the I_d composite indicator as defined in Hendry (2001, pp. 266). Additionally, we reconstructed variables used in π^* according to Hendry (2001, pp. 267) to make results more comparable with `GUMfinal`. As a result we replaced variables c and p_o by their equivalents in constant prices, i.e. $(c-p)^*$ and (p_o-p) respectively and we included lagged I_2, so we got the following model specification (`GUMId`):

$$\begin{aligned}\Delta p_t =& f\Big(\Delta p_{t-1}, y^d_{t-1}, m^d_{t-1}, n^d_{t-1}, U^d_{t-1}, e_{r,t-1}, S_{t-1}, R_{l,t-1}, \Delta p_{e,t}, \Delta p_{e,t-1}, \Delta U_{r,t-1}, \Delta w_{t-1}, p_{t-1},\\ & \Delta c_{t-1}, \Delta m_{t-1}, \Delta n_{t-1}, \Delta R_{s,t-1}, \Delta R_{l,t-1}, \Delta p_{o,t-1}, I_{2,t-1}, (c-p)^*_{t-1}, (p_o-p)_{t-1}, I_{d,t};\ \varepsilon_t\Big).\end{aligned} \tag{3}$$

The whole model space in `GUMId` (including intercept) was equal to $2^{24} = 16,777,216$. Model prior was set to uniform and the total number of Monte Carlo iterations was 1,000,000 (including 25% burn-in draws). The posterior results are in table 3.

Table 3: BACE posterior inclusion probabilities and posterior estimates of regression coefficients for the `GUMId`

		PIP	Avg. Mean	Avg. Std. Dev.
	$I_{d,t}$	1.0000	0.0377	0.0016
Hendrys' selection	$\Delta p_{e,t}$	1.0000	0.2454	0.0271
Hendrys' selection	S_{t-1}	1.0000	-0.8516	0.1218
Hendrys' selection	Δp_{t-1}	1.0000	0.2473	0.0382
Hendrys' selection	$\Delta R_{s,t-1}$	0.9999	0.6263	0.1262
Hendrys' selection	Δm_{t-1}	0.9997	0.1878	0.0384
Hendrys' selection	$\Delta p_{o,t-1}$	0.9975	0.0480	0.0121
Hendrys' selection	y^d_{t-1}	0.9884	0.1456	0.0466
π^*_{t-1}	$(c-p)^*_{t-1}$	0.9827	0.0882	0.0262
π^*_{t-1}	$(p_o-p)_{t-1}$	0.9460	0.0132	0.0066
π^*_{t-1}	$const$	0.8366	-0.0479	0.0297
π^*_{t-1}	$e_{r,t-1}$	0.6377	-0.0226	0.0221
π^*_{t-1}	$I_{2,t-1}$	0.3387	-0.0045	0.0079
	U^d_{t-1}	0.2621	-0.0143	0.0303
	p_{t-1}	0.1542	0.0002	0.0021
	Δn_{t-1}	0.1397	0.0023	0.0087
	Δc_{t-1}	0.1236	0.0042	0.0203
	$\Delta R_{l,t-1}$	0.1163	0.0223	0.1067
	$R_{l,t-1}$	0.1101	0.0047	0.0321
	n^d_{t-1}	0.1095	0.0002	0.0012
	$\Delta p_{e,t-1}$	0.1019	0.0019	0.0113
	m^d_{t-1}	0.1004	0.0004	0.0047
	Δw_{t-1}	0.0982	-0.0025	0.0219
	$\Delta U_{r,t-1}$	0.0925	-0.0025	0.0238

According to results presented in table 3, BACE procedure gives the same set of UK inflation factors as in Hendry (2001). Variables selected in model (6) in Hendry (2001) are top 8 most probable variables with PIP > 0.988. The next 5 most probable variables indicated by BACE procedure are elements of π^*.

4.3 GUMfinal results

In the next step we considered scenario using I_d and π^* directly. The BACE analysis was performed for the set of 20 variables (including the intercept) defined in `GUMfinal` (model (1)), so the total number of possible models was $2^{20} = 1,048,576$. We assumed that all possible specifications were equally probable. The BACE results, obtained after 1,000,000 Monte Carlo iterations, are presented in table 4.

Table 4: BACE posterior inclusion probabilities and posterior estimates of regression coefficients for the `GUMfinal`

		PIP	Avg. Mean	Avg. Std. Dev.
Hendrys' model (6)	π^*_{t-1}	1.000000	-0.186844	0.025828
	$I_{d,t}$	1.000000	0.037903	0.001573
	$\Delta p_{e,t}$	1.000000	0.264119	0.025146
	S_{t-1}	1.000000	-0.856166	0.090581
	Δp_{t-1}	1.000000	0.279046	0.033585
	y^d_{t-1}	0.999996	0.193686	0.036891
	$\Delta R_{s,t-1}$	0.999949	0.609606	0.114351
	Δm_{t-1}	0.999936	0.173201	0.029831
	$\Delta p_{o,t-1}$	0.987283	0.038862	0.011714
	U^d_{t-1}	0.610013	-0.041815	0.040875
	n^d_{t-1}	0.194672	0.000631	0.001692
	$R_{l,t-1}$	0.151855	0.006635	0.022907
	$\Delta R_{l,t-1}$	0.126007	0.026201	0.111685
	$\Delta p_{e,t-1}$	0.105244	0.002085	0.011372
	m^d_{t-1}	0.104247	-0.000513	0.004491
	$\Delta U_{r,t-1}$	0.097311	-0.002368	0.024480
	const	0.095136	0.000021	0.000643
	Δc_{t-1}	0.090481	-0.000170	0.010168
	Δn_{t-1}	0.089092	0.000461	0.004619
	Δw_{t-1}	0.085100	-0.000306	0.012303

According to the results in table 4, the set of variables used in the BACE analysis can be divided into 3 groups: highly probable determinants $(\pi^*_{t-1}, I_{d,t}, \Delta p_{e,t}, S_{t-1}, \Delta p_{t-1}, y^d_{t-1}, \Delta R_{s,t-1}, \Delta m_{t-1}, \Delta p_{o,t-1})$ with PIP ≥ 0.987, medium probable (U^d_{t-1}) with PIP $= 0.61$ and lowly probable $(n^d_{t-1}, R_{l,t-1}, \Delta R_{l,t-1}, \Delta p_{e,t-1}, m^d_{t-1}, \Delta U_{r,t-1}, const, \Delta c_{t-1}, \Delta n_{t-1}, \Delta w_{t-1})$ with PIP ≤ 0.195. Our results are consistent with Hendrys' paper because the highly probable determinants according to the BACE approach are the same as in model (6).

In addition, BACE approach allows models to be ranked according to their posterior probabilities. Table 5 presents the posterior probability, coefficient estimates and model statistics for the top 10 models. The most probable model M_1 has the posterior probability 0.200. The second probable model M_2, with probability 0.095, is model (6) in Hendry (2001) and `FinalModel` in our notation. These two best models differ only by the variable U^d_{t-1}, i.e., the excess labor demand. Although the posterior probability of the highest ranked model M_1 was more than twice as large as that for the second model M_2, an inference based only on M_1 leaves out 80% of the posterior probability mass. The "conservative" model reduction strategy dropped U^d_{t-1}, leading to M_2, while the "liberal" strategy leads to M_1, which includes U^d_{t-1}. This result confirms that the "conservative" model reduction strategy is relevant in the case of modeling UK inflation. Variable U^d_{t-1} with PIP $= 0.61$ can not be classified as a highly probable determinant, which confirms that this variable should be dropped from `FinalModel`. The posterior probabilities of the other models $P(M_3 \mid y), \ldots, P(M_{10} \mid y)$ are less than 0.038. These models differ from the two best models only by the least probable variables, and they do not contribute substantial

Table 5: Coefficient estimates and model statistics for top 10 models

Model	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
$P(M_j \mid y)$	0.200	0.095	0.037	0.035	0.026	0.024	0.023	0.023	0.022	0.021
π^*_{t-1}	−0.187 (***)	−0.186 (***)	−0.194 (***)	−0.196 (***)	−0.168 (***)	−0.177 (***)	−0.190 (***)	−0.185 (***)	−0.188 (***)	−0.188 (***)
$I_{d,t}$	0.038 (***)	0.038 (***)	0.038 (***)	0.038 (***)	0.037 (***)	0.038 (***)	0.038 (***)	0.038 (***)	0.038 (***)	0.038 (***)
$\Delta p_{e,t}$	0.265 (***)	0.265 (***)	0.263 (***)	0.262 (***)	0.262 (***)	0.263 (***)	0.262 (***)	0.264 (***)	0.264 (***)	0.268 (***)
S_{t-1}	−0.857 (***)	−0.834 (***)	−0.882 (***)	−0.872 (***)	−0.833 (***)	−0.854 (***)	−0.860 (***)	−0.874 (***)	−0.850 (***)	−0.857 (***)
Δp_{t-1}	0.287 (***)	0.268 (***)	0.288 (***)	0.272 (***)	0.264 (***)	0.283 (***)	0.279 (***)	0.287 (***)	0.291 (***)	0.286 (***)
y^d_{t-1}	0.188 (***)	0.184 (***)	0.216 (***)	0.223 (***)	0.191 (***)	0.192 (***)	0.188 (***)	0.177 (***)	0.187 (***)	0.189 (***)
$\Delta R_{s,t-1}$	0.625 (***)	0.619 (***)	0.572 (***)	0.547 (***)	0.635 (***)	0.633 (***)	0.611 (***)	0.652 (***)	0.601 (***)	0.605 (***)
Δm_{t-1}	0.178 (***)	0.182 (***)	0.167 (***)	0.167 (***)	0.162 (***)	0.168 (***)	0.174 (***)	0.176 (***)	0.176 (***)	0.177 (***)
$\Delta p_{o,t-1}$	0.037 (***)	0.041 (***)	0.041 (***)	0.045 (***)	0.042 (***)	0.038 (***)	0.037 (***)	0.038 (***)	0.037 (***)	0.036 (***)
U^d_{t-1}	−0.069 (**)		−0.062 (**)			−0.062 (**)	−0.069 (**)	−0.075 (**)	−0.076 (**)	−0.064 (**)
n^d_{t-1}			0.003 (–)	0.004 (–)						
$R_{l,t-1}$					0.051 (–)	0.028 (–)				
$\Delta R_{l,t-1}$										0.118 (–)
$\Delta p_{e,t-1}$							0.020 (–)			
m^d_{t-1}								−0.006 (–)		
$\Delta U_{r,t-1}$									−0.043 (–)	
R^2	0.976	0.975	0.977	0.975	0.976	0.977	0.977	0.977	0.977	0.977
$\overline{R}^2$	0.974	0.973	0.975	0.974	0.974	0.974	0.974	0.974	0.974	0.974
$\hat{\sigma}$	1.11%	1.14%	1.11%	1.13%	1.11%	1.11%	1.12%	1.12%	1.12%	1.12%
SC	−8.67	−8.66	−8.65	−8.64	−8.64	−8.64	−8.64	−8.64	−8.64	−8.64

(***) significance at 1%, (**) significance at 5%, (–) insignificance at 10%, $\overline{R}^2$ stands for the adjusted R^2, and $P(M_j \mid y)$ denotes the posterior model probability of model M_j.

information in this case.

5 Conclusions

Replication of Hendrys' model (6) for UK inflation in a narrow sense was performed in two alternative programs (gretl and PcGive/Autometrics) and brought the same results, in line with the original paper. In our opinion slight differences comes from the precision of the computers architecture used for calculations or changes in data storage. In the replication in a wide sense, we used BACE as an automatic variable selection procedure. Taking into account the whole model space, we obtained the same set of determinants as in Hendrys' paper. All considered specifications (`GUM52`, `GUMId` and `GUMfinal`) indicate almost the same set of significant variables. The only difference was for specification with initial set of all 52 variables (`GUM52`), where the nominal National Debt was replaced by GDP excess demand, but referring directly to the findings from jointness results these variables were strong substitute. BACE selection strategy for two other specifications (`GUMId` and `GUMfinal`) gave the same set of relevant variables in explanation of UK inflation as in Hendry (2001). According to our research, BACE selection procedure recovered the path of reduction strategy used in the original paper. Moreover we can say that BACE, which takes into account model uncertainty, can properly select significant factors from the wide range of possible determinants.

References

Błażejowski, M., & Kwiatkowski, J. (2018). *Bayesian Averaging of Classical Estimates (BACE) for gretl* (gretl working papers No. 6). Universita' Politecnica delle Marche (I), Dipartimento di Scienze Economiche e Sociali. (Cit. on p. 2).

Clements, M. P., & Hendry, D. F. (2008). Chapter 1 Forecasting Annual UK Inflation Using an Econometric Model over 1875–1991. In *Forecasting in the presence of structural breaks and model uncertainty* (pp. 3–39). doi:10.1016/S1574-8715(07)00201-1. (Cit. on p. 2)

Cottrell, A., & Lucchetti, R. (2018). *Gretl User's Guide*. (Cit. on p. 1).

Doornik, J. A. (2009). Autometrics. In J. Castle & N. Shephard (Eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry* (9780199237197). OUP Catalogue. doi:10.1093/acprof:oso/9780199237197.001.0001. (Cit. on p. 1)

Doornik, J. A., & Hendry, D. F. (2013). *Empirical Econometric Modelling – PcGive 14*. London: Timberlake Consultants Ltd. (Cit. on p. 3).

Doppelhofer, G., & Weeks, M. (2009). Jointness of growth determinants. *Journal of Applied Econometrics*, *24*, 209–244. doi:10.1002/jae.1046. (Cit. on p. 4)

Fernández, C., Ley, E., & Steel, M. F. J. (2001). Benchmark Priors for Bayesian Model Averaging. *Journal of Econometrics*, *100*, 381–427. doi:10.1016/S0304-4076(00)00076-2. (Cit. on p. 2)

Hendry, D. F. (2001). Modelling UK Inflation, 1875–1991. *Journal of Applied Econometrics*, *16*, 255–275. doi:10.1002/jae.615. (Cit. on pp. 1–4, 6–8)

Hendry, D. F., & Krolzig, H.-M. (2001). *Automatic Econometric Model Selection Using PcGets*. London: Timberlake Consultants Press. (Cit. on p. 3).

Koop, G., Poirier, D. J., & Tobias, J. L. (2007). *Bayesian Econometric Methods*. Cambridge Books. doi:10.1017/CBO9780511802447. (Cit. on p. 4)

Ley, E., & Steel, M. F. (2007). Jointness in Bayesian Variable Selection with Applications to Growth Regression. *Journal of Macroeconomics*, *29*(3), 476–493. doi:10.1016/j.jmacro.2006.12.002. (Cit. on p. 4)

Ley, E., & Steel, M. F. (2009). On the effect of prior assumptions in bayesian model averaging with applications to growth regression. *Journal of Applied Econometrics*, *24*(4), 651–674. doi:10.1002/jae.1057. (Cit. on p. 2)

Ley, E., & Steel, M. F. (2012). Mixtures of g-priors for Bayesian model averaging with economic applications. *Journal of Econometrics*, *171*, 251–266. doi:10.1016/j.jeconom.2012.06. (Cit. on p. 2)

Sala-i-Martin, X., Doppelhofer, G., & Miller, R. I. (2004). Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *The American Economic Review*, *94*, 813–835. doi:10.1257/0002828042002570. (Cit. on pp. 1, 2)

Appendix

Table 6: List of variables and their definitions used in paper

Variable	Definition
Y_t	real GDP, £ million, 1985 prices
P_t	implicit deflator of GDP (1985=1)
M_t	nominal broad money, £ million
$R_{s,t}$	three-month treasury bill rate, fraction p.a.
$R_{l,t}$	long term bond interest rate, fraction p.a.
$R_{n,t}$	opportunity-cost of money measure
N_t	nominal National Debt, £ million
U_t	unemployment
$Wpop_t$	working population
$U_{r,t}$	unemployment rate, fraction
L_t	employment
K_t	Gross capital stock
W_t	wages
H_t	normal hours (from 1920)
$P_{e,t}$	world prices, (1985=1)
E_t	annual-average effective exchange rate
$P_{nni,t}$	deflator of net national income, (1985=1)
$P_{cpi,t}$	consumer price index, (1985=1)
$P_{o\$,t}$	commodity price index, $
m_t^d	money excess demand
y_t^d	GDP excess demand
s_t	short–long spread
n_t^d	excess demand for debt
$e_{r,t}$	real exchange rate
π_t^*	profit markup
U_t^d	excess demand for labour
$p_{o,t}$	commodity prices in Sterling