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

Blazejowski, 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*

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 for helpful comments.

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

Summary

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 to compare model reduction strategies. Allowing for the investigation of other specifications, we confirm the same set of significant determinants but find that Hendrys' model is not the most probable.

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.

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

*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 2018b and PcGive version 14.2 with Ox Professional version 7.20 on a PC with 8 x Intel(R) Core(TM) i7-8550U CPU@1.80GHz, 16.0 Gb of RAM running under Debian GNU/Linux "buster/sid" 64 bit.

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, first of all, by having less restrictive prior assumptions of parameters. Sala-i-Martin et al. (2004) showed that the BMA approach may be understood as limiting case of 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 1.0 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, section 4 presents the BACE results together with similarities in the inconclusive inference on the relevance of excess labor demand (replication in a wide sense), and section 5 concludes.

2 Data

In our paper and replication files, we used the same data definitions as in Hendry (2001) 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 $c_t^* = c_t - p_t + 0.006 \times (trend - 69.5) + 2.37$. This exception was due to an inconsistency between the JAE archive and Hendrys' paper, where it was defined as $c_t^* = (c_t - p_t)^*$.

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 (`Model: GUM52`; residual standard deviation $\hat{\sigma} = 1.21\%$, Schwartz Criterion $SC = -7.3$). After excluding indicators from the initial model, we also received exact results (`Model: GUMnoIndicators`; $\hat{\sigma} = 2.5\%$, $SC = -6.63$). In the next step, we added dummy variables I_b, I_l, I_m

[3] The BACE 1.0 package is available at http://ricardo.ecn.wfu.edu/gretl/cgi-bin/gretldata.cgi?opt=SHOW_FUNCS.

concerning outliers in particular years to (`Model: GUMnoIndicators`), and we obtained the same results as in the paper (`Model: 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 (`Model: 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 (`Model: GUMfinal`; $\hat{\sigma} = 1.15\%$, $SC = -8.33$). At this point, we had the following specification:

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

After the reduction of model (1) at a 1% significance level[4], we obtained model (6) in Hendry (2001) and (`Model: FinalModel`; $\hat{\sigma} = 1.14\%$, $SC = -8.66$) in our notation. Detailed results are available in the table 1.

Table 1: Comparison of Hendrys' and the replication results

	Hendry (2001)	Replication[a]
y^d_{t-1}	0.180 (0.032)	0.184 (0.032)
Δm_{t-1}	0.187 (0.028)	0.182 (0.028)
S_{t-1}	−0.834 (0.088)	−0.834 (0.087)
$\Delta R_{s,t-1}$	0.618 (0.106)	0.619 (0.106)
π^*_{t-1}	−0.186 (0.024)	−0.186 (0.024)
$\Delta p_{e,t}$	0.265 (0.025)	0.265 (0.025)
$I_{d,t}$	0.038 (0.002)	0.038 (0.002)
$\Delta p_{o,t-1}$	0.041 (0.010)	0.041 (0.010)
Δp_{t-1}	0.267 (0.027)	0.268 (0.026)
R^2	0.975	0.975
$\hat{\sigma}$	1.14%	1.14%
SC	−8.66	−8.66

[a]Replication was performed in two programs: gretl and PcGive/Autometrics.
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[5]. We repeated this automatic model selection procedure using Autometrics for model (1), and we obtained exactly the same estimates as in `Model: GUMfinal` in gretl (i.e., with slight differences compared to model (6) in Hendrys' paper). We suppose that these differences are due to different computer architectures (64 bit and 32 bit).

[4]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).

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

4 BACE results

To verify the correctness of Hendrys' variable selection strategy, we used the BACE approach (replication in a wide sense). This procedure enables searching the whole model space and selecting the most probable regressions. The BACE analysis was performed for the set of $k = 20$ variables (including the intercept) defined in `Model: GUMfinal` (model (1)), and so the total number of possible models was $2^k = 1,048,576$. The following parameters for the MC^3 algorithm were set:

- Total number of Monte Carlo iterations: 1,000,000, including 25% burn-in draws,
- Model prior distribution: binomial with prior average model size equal to $k/2 = 10$, which means that all possible specifications are equally probable,
- Significance level for the initial model $\alpha = 60\%$, which means that we dropped the most statistically insignificant variables in the initial model at the beginning of the procedure.

The BACE approach 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 Δp_t (see Doppelhofer & Weeks, 2009; Koop, Poirier, & Tobias, 2007). PIP is calculated as the frequency of appearance of a given variable in all considered models. The BACE results, obtained after 1,000,000 Monte Carlo iterations, are presented in table 2.

Table 2: Posterior inclusion probabilities and posterior estimates of regression coefficients obtained by BACE

	PIP	Avg. Mean	Avg. Std. Dev.
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 2, 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 $PIP \geq 0.987$, 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

$PIP \leq 0.195$. Our results were consistent with Hendrys' paper because the highly probable determinants according to the BACE approach were the same as in model (6). This result confirms that the "conservative" model reduction strategy was relevant in the case of modeling UK inflation. Moreover, our results confirmed the inconclusive inference on the relevance of U^d_{t-1}, i.e., the excess labor demand (with $PIP = 0.61$, it could not be classified as a highly probable determinant).

In addition to the posterior characteristics presented in table 2, the BACE approach allows models to be ranked according to their posterior probabilities. Table 3 presents the coefficient estimates and model statistics for the top 10 models. The total probability of these models was 50.8%.

Table 3: 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}
$P(M_j)$	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 $P(M_j)$ denotes the posterior model probability of model j.

The most probable model M_1 had the posterior probability 0.2. The second probable model M_2, with probability 0.095, was model (6) in Hendry (2001) and `FinalModel` in our notation. In addition, M_1 fit the data better then M_2 based on the following statistics: $\overline{R}^2_{M_1} > \overline{R}^2_{M_2}$, $\hat{\sigma}_{M_1} < \hat{\sigma}_{M_2}$ and $SC_{M_1} < SC_{M_2}$. 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 omits 80% of the total information contained in the entire model space. As a consequence, the average coefficient estimates presented in table 2 were different than coefficient estimates for the `FinalModel` in table 1. The greatest differences were noticed for the following variables: Δp_{t-1}, y^d_{t-1}, Δm_{t-1} and $\Delta p_{o,t-1}$.

Taking our results into consideration, we confirmed the simplification problem about the relevance of U^d_{t-1}, as in

Hendrys' paper. If we use all available information contained in the whole model space, U^d_{t-1} will be classified as a medium determinant variable with $PIP = 0.61$. 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}. Furthermore, setting the target size to "medium" (2.5% significance level) in Autometrics also leads to M_1. The posterior probabilities of the other models $P(M_3),\ldots,P(M_{10})$ were less than 0.038. These models differ from the two best models only by the least probable variables, and they did not contribute substantial information in this case.

5 Conclusions

Replication of Hendrys' model for UK inflation in a narrow sense was performed in two programs (gretl and PcGive/Autometrics) and brought exactly the same results, although they were slightly different than the original. In the replication in a wide sense, we used BACE as an automatic model reduction strategy. Taking into account the whole model space, we obtained the same set of determinants as in Hendrys' paper, although his `FinalModel` was the second one in the ranking, and it was more than two times less probable then the most likely model containing the additional variable U^d_{t-1}. Hendrys' model omitted over 90% of the information contained in all possible models, which leads to different coefficients estimates. Referring directly to the findings in Hendry (2001), inference based on just one model may lead to slightly different conclusions than inference based on the whole model space.

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.

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). Emerald Group Publishing Limited. doi: 10.1016/S1574-8715(07)00201-1

Cottrell, A., & Lucchetti, R. (2018, August). Gretl User's Guide [Computer software manual]. Retrieved from http://ricardo.ecn.wfu.edu/pub/gretl/manual/PDF/gretl-guide-a4.pdf

Doornik, J. A. (2009, May). Autometrics. In J. Castle & N. Shephard (Eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry.* Oxford University Press. doi: 10.1093/acprof:oso/9780199237197.001.0001

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

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

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

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

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

Koop, G., Poirier, D. J., & Tobias, J. L. (2007). *Bayesian Econometric Methods*. New York: Cambridge University Press. doi: 10.1017/CBO9780511802447

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

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