IS MULTIPLE LINEAR REGRESSION THE PROPER TOOL OF MODELLING A BEHAVIOUR OF REAL SYSTEMS?

Jacek J. Nowak
Katedra Zarządzania SW im. Bogdana Jańskiego w Warszawie
e-mail: jacek.nowak@janski.edu.pl

Abstract: Methodological assumption that multiple linear regression is an adequate tool of modelling the behaviour of real systems is checked. To do this the experiment is organised on the basis of simple "real" system represented as finite discrete automaton. Main result is that in situation of "black box" modelling the approximation of output variables with multiple linear regressions (from several samples and under different conditions) may not fulfil any of criterions of feasible approximation of systems behaviour, also in situations where real relation between input and output variables is strictly linear and only one of variables is omitted.

Keywords: model of real system, multiple linear regression, real system structure, discrete automaton, „black box" modelling, quality of approximation

INTRODUCTION

After almost 40 years the main postulates of W. Leontief on modelling real economic systems are still actual[1]. In present paper we take one of steps on this difficult road of verification of assumptions of modelling real systems.

Multiple linear regression is often used to describe then analyse, forecast or simulate the behaviour of such real systems as economic, environmental, technical and social ones. Most of these models are quantitative models based on statistical data. There still appear failures or unsatisfactory results when applying such quantitative models especially to forecast or simulate behaviour of economic systems. It

---

[1] We mean especially postulate: *"What is really needed, in most cases, is a very difficult and seldom very neat assessment and verification of these assumptions in terms of observed facts."* [W. Leontief, 1971, p. 2].

is often pointed out by experienced practitioners and methodologists[2] that these failures are caused mainly by structural changes in modelled systems[3] or by one or more principal factors from listed below:

       a) omitted or misspecified variables,

       b) multicollinearity (near multicollinearity) between explanatory variables,

       c) using bad data, (e.g. inadequate or incomplete or data with errors),

       d) stochastic nature of relations,

       e) series of data containing time trends,

       f) misspecification of the functional form of relations between variables,

       g) inadequate, incomplete (or lack of) theory of the modelled system.

After the period of significant creativity resulting in many ideas of models taking into account structural changes[4] in recent years the greater attention in researches is paid to testing hypotheses of structural break[5].

But still the more fundamental problem than hypotheses testing is how to construct the adequate model. The opinion is that if there exists no one of above mentioned reasons of failures of models' applications, excluding lack of or incomplete theory, then the regression model with good ratios of its stochastic structure ought to be proper for representing, forecasting or simulating given real system.

But models of real systems, especially econometric ones, are constructed often in situations of "black box" modelling. It means that the data, which modeller possess, are describing only input and output series but are not describing (internal) states of modelled system.

In [Nowak 2000] the following problem was stated: *Is the multiple regression an adequate tool of modelling and forecasting in "black box" situations under the assumption that there exists no one of above mentioned reasons of failures of models' applications (excluding lack of theory or incomplete theory)?*

To answer the above question there was conducted an experiment of modelling simple system under above assumption. We have generated several samples of observations, estimated parameters of linear regressions with two explanatory (input) variables and approximated endogenous (output) variables within and out of samples (using true, observed values of explanatory variables).

Instead of fulfilling the above assumption[6] and very high values of determination coefficients, R^2, the quality of approximation of dependent variables out of samples was very poor. The greater part of regressions produced mostly approxi-

---

[2] See e.g. Mayer (1993), Broemeling (1982), Hendry, Richard (1983), survey in Nowak (1981).

[3] The main reason of these structural changes is the nature of economic (social) reality (cf Marschak, 1950 or Leontief, 1971).

[4] See e.g. survey of such ideas in Nowak (2004).

[5] See e.g. Bai (1999), Elliott and Müller (2006) or Juhl and Xiao (2009).

[6] We could not avoid only some correlation between explanatory variables in some samples instead of generating their values on the base of random tables.

mation errors[7] greater then 10% of really observed values of endogenous variables. The remaining regressions produced nearly half of such infeasible errors. There were several errors greater than 100% (!) of approximated values, some of them greater than 500%. The whole sample, consisted of all generated observations, numbered 34 elements. In addition, great variability of estimates of regression parameters was observed (from one sample to another sample).

There exist several practical criterion s of adequacy (feasibility) of approximations (or forecasts). We can take into account one or combination of at least three following criterions:

1) no error is greater than 10% of observed (approximated) value (very strong condition),
2) no more then 10% of errors is greater than 10% of observed (approximated) values,
3) mean of absolute values of relative errors is not greater than 10%.

Results of [Nowak 2000] showed that none of above criterions of feasibility of approximation was fulfilled by any regression from any sample.

The general conclusion of this investigations was that in situation of "black-box" and instead of lack of principal reasons of failures of modelling multiple linear regression is not the adequate tool of modelling behaviour of system.

The goal of present investigations is to check if the above described results were not casual and conclusion will be valid in situations of changed two types of conditions: A) with another initial states and B) quite different sequence of observations. To check this we must repeat the mentioned experiment with two parts according to A and B.

REPRESENTATION OF REAL SYSTEM

To make our experiment we must choose the point of view, and then the way of representation of real systems. This choice is very important from the methodological as well as from the quality of approximation point of view.

It is widely known that the reaction (output) of every real system does not depend only on impulse (input) acting on this system but also depends on a current state in which this system remains. Therefore we accept the way of representation of real systems as systems of inputs, states and outputs as it is accepted within automata theory in cybernetics or general systems theory.

In economic sciences possibly first who applied such representation of real economic systems were Kornai (1971) and Rey (1971) and (1979).

---

[7] Approximation error we mean difference between observed and theoretical value of any output variable, where theoretical value is calculated from corresponding regression when "true" (observed) values of explanatory variables were used.

To say more exactly we accept as true the working hypothesis that every real system can be fully and explicitly described (on a proper level of generalisation) by the following quintuple

$$<X, Q, Y, \delta, \lambda>, \qquad\qquad (1)$$

where: X - the input set (containing inputs, x's, i.e. actions of the environment on the system), Q - the set of states of the system, Y - the output set (enclosing outputs, y's, i.e. actions of the system on its environment), $\delta: R \rightarrow Q: (x,q) \rightarrow q'$ - the state transition function, where $R \subseteq X \times Q$, $\lambda: R \rightarrow Y: (x,q) \rightarrow y$ - the output function, and X, Q, Y are finite. Formally, the quintuple $<X, Q, Y, \delta, \lambda>$ is also a definition of so called discrete, finite Mealy automaton (or machine).[8]

If the "internal" structure (i. e. transition and output functions) of real system is not known the recognition of the way of performing inputs into outputs becomes the main problem of investigations and the "black box" approach can be applied here as an auxiliary, intermediate tool.

ORGANISATION AND PRINCIPLES OF THE EXPERIMENT

The experiment will consist of two phases as in [Nowak 2000].

In phase I we take some "real" system[9] represented by quintuple (1), i. e. we know everything about its structure (transition and output functions). We fix its structure over the whole experiment – no structural changes. Then we randomly generate[10] sequences of inputs in subsequent steps ("time units")[11] observing resulting sequences of states and of outputs of this system. Then we prepare data that are quantitative characteristics of these sequences of inputs and outputs (input and output variables).

In phase II we arrange the situation of the "black box" modelling, i. e. a modeller, as in many cases in reality, does not know (must forget) the structure of the modelled system which is this "real" system from phase I. Then to describe the behaviour of this system the modeller have only data which are quantitative characteristics of input and output elements in subsequent observations (time units)[12].

---

[8] Theory of automata treats the automaton as only "an information processor" but the author following Kornai (1971) and Rey (1979) does not impose any constraints on the "nature" of elements of the system under consideration, described by the quintuple (1).

[9] Despite of this artificially construction of the system used in experiment its structure has all basic features of real systems according to axioms listed in [Nowak 2003].

[10] Random generation of inputs is applied to avoid time trends or periodic fluctuations in data, and, to minimise probability of multicollinearity (or near multicollinearity) between explanatory variables.

[11] One can interpret these subsequent steps as time units but it is not necessary.

[12] We described in detail the "nature" of economic statistical data or the relation between statistical data and input and output elements of real economic systems in Nowak (1985).

Next we assume that the behaviour of this system (which is now limited to behaviour of input variables) can be adequately represented by linear regression between output and input variables as explanatory ones. Then we estimate parameters of such regression for samples of different sizes, approximate values of output (dependent) variables within and out of these sample (using always "true", observed values of explanatory variables) and check adequacy of these approximations.
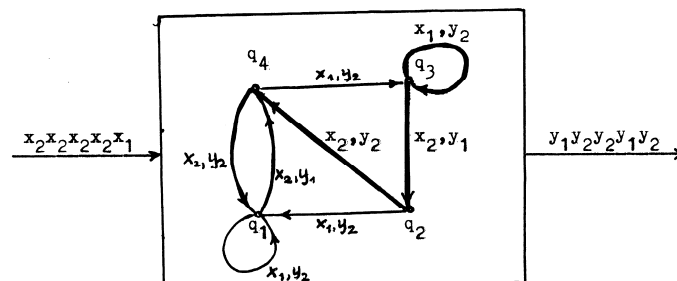
Particular principles of the experiment are as follows. In phase I we take into consideration the same simple system as in [Nowak 2000].

The structure of real system can be presented more conveniently, instead of the quintuple (1) with its algebraic description, as a graph[13], and in general case, as the directed multigraph. So we present the structure of this system in more convenient form as the graph in fig. 1.

Elements of the graph representing the structure of system have following meaning:

-   nodes represent states of system, q's,
-   arrows represent possible transitions of system from state q to state q' (at the arrow-head) - but not vice versa – when input x acts on state q,
-   labels of arrows, pairs (x, y), denote what output is produced when input x acts on state q (the beginning of arrow);
-   first element of label of every arrow shows input x which causes transition from state q to state q' represented by this arrow.

Figure 1. Representation of system G2 with its structure and behaviour in observation no. 2.

Source: Nowak (2000), p. 173.

Elements of a graph representing the structure of system have following meaning:

-   nodes represent states of system, q's,
-   arrows represent possible transitions of system from state q to state q' (at the arrow-head) - but not vice versa – when input x acts on state q,

---

[13] By a graph we mean, in general case, a directed multigraph.

- labels of arrows, pairs (x, y), denote what output is produced when input x acts on state q (the beginning of arrow);
- first element of label of every arrow shows input x which causes transition from state q to state q' represented by this arrow.

Fig. 1 represents an example of the behaviour of system G2, here during observation no. 2 where system's initial state was q_3.

Principles of generating input sequences using random numbers table[14]:
1) watch succeeding digits of random numbers, row by row, from the left to the right, 2) if digit is odd (but other than 9), choose x_1; if digit is even, choose x_2; if digit is 0, omit it; if digit is 9, it means the end of given sequence of inputs.

It is assumed for simplicity that we use only one-element inputs and outputs. Then measuring quantities of input and output elements is simple and values of input and output variables are the frequencies of occurrence of input or output of given kind in given sequence of inputs or outputs.

If we denote any observation in our experiment as quadruple (y_1', y_2', x_1', x_2'), where symbols y_1', y_2', x_1', x_2' denote values of corresponding output and input variables, i. e. quantities (here: frequencies of occurrence) of, respectively, outputs y_1 and y_2 and inputs x_1 and x_2 in corresponding sequences then the quadruple representing observation no. 2 described in fig. 2 can be registered as (y_1', y_2', x_1', x_2') = (2,3,1,4).

To realise part A of our experiment (changed initial state) we generated (applying above principles) new 34 observations using the same sequences of inputs as generated in [Nowak 2000] but starting from other than q_1 initial states.

To realise part B we generated (applying above principles) quite new 34 observations using quite new sequences of inputs[15] and starting from initial state q_1.

In phase II we are in "black box" modelling situation. So we "forget" all information on transition and output functions. But we know all input variables the output variables are dependent of and we have all data concerning these input and output variables. Next we estimate (by least squares method) parameters of linear regressions of the form:

$$y_i' = \beta_0 + \beta_1 x_1' + \beta_2 x_2' + \varepsilon \quad (i=1,2), \tag{2}$$

where β_0, β_1 and β_2 are structural parameters, ε is error term.

We do such estimations for samples of different sizes, which contain observations no. 1-6, 1-12, 13-24, 1-24, 25-29, 30-34 and 1-34. Then we check general goodness of fit, using R^2, and significance of estimated parameters. The most important is checking the quality (feasibility) of approximation of values of output variables, y_i's, the principal element of observed behaviour of modelled system.

---

[14] We used tables published by Martin (1968). In experiment in Nowak (2000) we used random number rows from 1 to 4.
[15] All randomly generated using tables by Martin (1968); rows from 10 to 16.

RESULTS OF NEW EXPERIMENT

Observations we have generated[16] in phase I are contained in table 4 in Appendix. Results of estimation are presented in table 1.

Table 1. Estimates of full regression parameters and adjusted coefficient of determination. R^2.

Full regression for y_1'				Full regression for y_2'			
Initial state: q_3 or q_4; old inputs				Initial state: q_3 or q_4; old inputs.			
Sample	b_0	b_1	b_2	$R^{'2}$	b_0	b_1	b_2	$R^{'2}$
1-6	0,107*	0,071*	0,442	0,893	-0,107*	0,929	0,558	0,992
1-12	0,020*	0,072*	0,478	0,933	-0,020*	0,928	0,522	0,992
13-24	-0,052*	0,174*	0,476	0,948	0,052*	0,826	0,524	0,984
1-24	-0,013*	0,084	0,517	0,946	0,013*	1,017	0,572	0,987
25-29	-1,455*	0,214*	0,623*	0,629	1,455*	0,786*	0,377*	0,684
30-34	0,591*	-0,003*	0,429	0,982	-0,591*	1,003	0,571	0,982
1-34	-0,020*	0,077	0,506	0,939	0,020*	0,923	0,494	0,986
Full regression for y_1'				Full regression for y_2'			
Initial state: q_1 and new inputs				Initial state: q_1 and new inputs			
Próba	b_0	b_1	b_2	$R^{'2}$	b_0	b_1	b_2	$R^{'2}$
1-6	-0,176*	0,054*	0,589	0,989	0,176*	0,946	0,411	0,999
1-12	0,024*	0,172*	0,408	0,900	-0,024*	0,828	0,592	0,985
13-24	-0,043*	0,127*	0,482	0,935	0,043*	0,873	0,518	0,989
1-24	0,000*	0,148	0,453	0,934	0,000*	0,852	0,547	0,989
25-29	0,179*	-0,011*	0,640	0,975	-0,179*	1,011	0,360	0,997
30-34	-0,339*	0,197*	0,551*	0,894	0,339*	0,803	0,449	0,979
1-34	-0,051*	0,131	0,485	0,923	0,051*	0,869	0,515	0,988

* means that the corresponding parameter was not statistically significant[17] (level of significance = 0,10); $R^{'2}$ - adjusted coefficient of determination.
Source: authors calculations.

Because all free terms were statistically insignificant we estimated also parameters of regressions without free term. Instead of insignificance of parameter β_1 in some regressions we did not withdraw variable x_1' from these regressions because they were basic factors the output variables depend of (according to true knowledge of modelled system). It was not surprising that estimates of parameters from different samples were also different. In two small samples we observed also change of sign of estimate of coefficient β_1.

---

[16] Observations generated from initial state q_2 for old input sequences were the same as for initial state q_1.

[17] Under assumptions of normality of distribution and independence of error term (checked by proper statistical tests) and of absence of multicollinearity.

Tables 2 and 3 present results of approximation of output values.

Table 2. Means* of absolute values of relative errors (percent errors).

	Sample							
	1-6	1-12	13-24	1-24	25-29	30-34	1-34	
	Initial state: q_1; old inputs							
Full[1] y_1'	20	21	33	24	148	156	23	n=26
Full[1] y_2'	14	13	13	11	152	96	12	n=34
Wft[2] y_1'	22	21	31	23	96	80	23	n=26
Wft[2] y_2'	11	12	14	12	12	13	11	n=34
	Initial state: q_3 or q_4; old inputs							
Full[1] y_1'	16	16	25	19	42	14	18	n=27
Full[1] y_2'	12	10,4	11	11	37	16	10,2	n=34
Wft[2] y_1'	16	16	24	18	18	16	18	n=27
Wft[2] y_2'	10,4	10,2	12	11	10,4	10,1	10,4	n=34
	Initial state: q_1; new inputs							
Full[1] y_1'	25	23	20	20	29	25	22	n=30
Full[1] y_2'	11	13	11	12	17	11	11	n=33
Wft[2] y_1'	23	23	20	20	27	26	22	n=30
Wft[2] y_2'	12	12	11	12	15	13	11	n=33

[1] Full – full regression [2] Wft – regression without free term. * Calculated without errors when observed value was 0.
Source: authors calculations.

Synthetic evaluation of adequacy of approximation of output variables y's by linear regression models from different samples is presented in table 3 in Appendix.

FINAL REMARKS

Observed variability of regression coefficient cannot be explained with structural changes, because the structure of modelled system was constant over the whole period of our experiment.

According multicollinearity; relatively moderate values coefficients of determination R^2 of regressions between explanatory variables we obtained, from 0,119 (for sample 1-6 and 1-34) through 0,55 to 0,94 (in two small samples, 25-29 and 30-34)[18] for part A of experiment, and most of them from 0,524 to 0,685 (and 0,85 for samples 1-6 and 13-24) in samples for part B, give introductory information that there could be from two to four cases of near multicollinearity. It can be one of reasons why we could obtain some statistically insignificant parameters.

---

[18] For which inputs were generated in special (different than for first 24 observations), not quite random way.

The first basic result of presented experiment is that instead of 1) very good values of adjusted coefficients of determination R'^2 of almost all estimated regressions and 2) almost ideal conditions of modelling such as lack of structural changes, no time trends or errors in data, slight symptoms of multicollinearity (instead of few samples), variables was not misspecified etc. the quality of approximation of output variables with all estimated regressions was poor.

Notice that no one of estimated regressions did not fulfilled in any part of experiment no one of three mentioned in par. 1 criterions of adequacy, also not the most tolerant criterion, i. e. no mean of absolute values of relative errors was less (or equal) 10% (see table 3). Also increasing samples did not lead to substantial improvement of quality of approximations (see table 3).

The next result is that there were not essential differences or improvement in quality of approximation of behaviour of given system when initial states were changed or new input sequences was acting on the system in the same initial state as in first experiment[19]. So this new experiment in changed conditions confirms the results obtained in experiment from 2000yr.

It can be done more and more similar experiments with different sets of observations, new, more complex system, many-element inputs and outputs with "prices" of these elements, new modelling tools or more sophisticated use of regression. But if it is done in situation of "black box" modelling analogous to situation of inadequate, incomplete (or lack of) theory of the modelled system the process of modelling is subject to danger of misspecification of the functional form of relations between variables and, first of all, to omit[20] or to misspecify variables.

If estimates and approximations are not satisfactory there will be supposed that true relation between variables is nonlinear. But what functional form ought to have such nonlinear relation, if we do not know the structure of real system?

In situation in this experiment the true relation between input and output variables was linear (!) as below[21]:

$$y_1' = x_1' + x_2' - y_2' \tag{3}$$

The second basic result of presented experiment is that in situation of "black box" modelling the approximation of output variables with multiple linear regressions (from several samples and under different conditions) may not fulfil any of criterions of feasible approximation of systems behaviour, also in situations where real relation between input and output variables is strictly linear and only one of variables is omitted.

---

[19] Only some improvement was in regressions from 5-element samples (25-29 and 30-34), but still with infeasible quality of approximation.

[20] We mean variables representing states of system.

[21] Because of assuming very simple system and one-element inputs and outputs. But relation (3) in general may be no longer true under other assumptions than applied in our simple situation.

But what could be the reason to modeller to include into regression model the other output variable as explanatory variable if he had true information that outputs of type y_1 (therefore output variable y_1') are results caused only by factors named inputs x_1 and x_2 and internal (nonobserved) states?

And what would be the results of assuming and estimating any nonlinear relation between variables if true relation is linear?

REFERENCES

Broemeling L., ed. (1982), *Structural Change in Econometrics*, "J. of Econometrics" vol.19, no.1, pp. 1-182; "Annals of Applied Econometrics" 1982-2.

Elliot G., Müller U.K. (2006), *Efficient tests for general persistent variation in regression coefficients*, "Review of Economic Studies", Vol. 73, pp. 907-940.

Hendry F., Richard J.-F.(1983) *The Econometric Analysis of Economic Time Series*, "Intern. Statist. Review" No. 2, pp. 111-148 (discussion pp. 148-163).

Juhl T. and Z. Xiao (2009), *Tests for changing mean with monotonic power*, "Journal of Econometrics", Vol. 148, pp. 14-24.

Kornai J. (1971), *Anti-Equilibrium*. North-Holland, Amsterdam.

Leontief W. (1971), *Theoretical Assumptions and Nonobserved Facts*, "American Economic Review" Vol. LXI, no.1, pp. 1-7.

Martin F.F. (1968), *Computer Modeling and Simulation*. Wiley, New York.

Marschak j. (1950), *Statistical Inference in Economics: An Introduction*, in: "Statistical Inference In Dynamic Economic Models" (ed. T.c. Koopmans). Wiley.

Mayer T. (1993) *Truth versus Precision on Economics*, Edward Elgar Publ.

Nowak J.J. (2004), *On econometric approaches to building and verification of simulation model in situations of structural changes* (in Polish), in: "Symulacja systemów gospodarczych. Część II" (red. A. Balcerak i W. Kwaśnicki). Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2004 (Prace Naukowe Instytutu Organizacji i Zarządzania Politechniki Wrocławskiej nr 75), pp. 47-63.

Nowak J.J. (2003), *Minimal structures of real systems and their models in discrete simulation* (in Polish), in: "Symulacja systemów gospodarczych" (red. A. Balcerak, E. Radosiński, B. Mielczarek), Prace Naukowe Instytutu Organizacji i Zarządzania Politechniki Wrocławskiej nr 74, Wyd. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2003; pp. 143-161.

Nowak J.J. (2000), *On essential limitation of multiple regression as simulation* (in Polish), in: "Symulacja Systemów Gospodarczych. Prace Szkoły Antałówka'2000". Wyd. Wyższej Szkoły Przedsiębiorczości i Zarządzania im. L. Koźmińskiego oraz Politechniki Wrocławskiej, Warszawa-Wrocław, pp. 169-189.

Nowak J.J. (1985), *On Some Foundations of Economic Model Building*, in: "Summer MIZAR Workshop, 16.06-15-09, Louvain-la-Neuve 1985", "Cahiers des Mathematiques Appliquees aux Sciences Humaines", No. 1, Bruxelles.

Nowak J. (1981), *Evaluation of stability of econometric model and constancy of elements of its structure with application to forecasting* (unpublished doctoral thesis written in Academy of Economics in Cracow, Dept. of Econometrics and Economic Cybernetics – in Polish). Kraków-Warszawa.

Rey-Werszowec K. (1979), *Tentative de construction d'echelons de structures economiques* (paper at the "Colloque Structures Economiques et Econometrie", Lyon, France).

Rey K. (1971), La planification sur le chemin de l'automatisation, "La Pologne", No 7 (203) Juillet, pp. 7 and 21.

Czy liniowa regresja wieloraka jest odpowiednim narzędziem modelowania zachowania systemów rzeczywistych?

Streszczenie: W pracy sprawdzono metodologiczne założenie, że liniowa regresja wieloraka jest odpowiednim narzędziem modelowania zachowania systemów rzeczywistych. W tym celu przeprowadzono eksperyment, w którym wzięto pod uwagę prosty system "rzeczywisty" reprezentowany z pomocą dyskretnego automatu skończonego. Główny wynik tego eksperymentu mówi, że aproksymacja zmiennych wyjściowych z pomocą liniowej regresji wielorakiej (w sytuacji modelowania "czarnej skrzynki", na podstawie szeregu prób i w różnych warunkach) może nie spełniać żadnego z kryteriów dopuszczalnej aproksymacji zachowania modelowanego systemu, również w sytuacjach, gdy rzeczywista relacja pomiędzy zmiennyymi wyjściowymi a zmiennymi wejściowymi jest dokładnie liniowa oraz gdy tylko jedna ze zmiennych jest pominięta.

Słowa kluczowe: automaty dyskretne, model systemu rzeczywistego, liniowa regresja wieloraka

APPENDIX

TABLE 3. Quality of approximations of output variables with linear regression models.

Sample (Numbers of observations)	Regression for y_1'		Regression for y_2'	
	Full	Without free term	Full	Without free term
Initial state q_3 or q_4 and old inputs				
1-6	- 14	- 20	+ 16	+ 12
1-12	- 15	-19	+ 9	+ 10
13-24	-17 $_1$	-20 $_2$	+ 14	- 23
1-24	- 20 $_2$	- 16 $_1$	+ 8	+ 13
25-29	- 20 $_3$	- 26 $_{10}$	- 21 $_4$	+ 13
30-34	+ 13	- 26 $_2$	+ 15 $_1$	+ 15
1-34	- 17	- 17	+ 10	+ 12
Initial state q_1 and new inputs				
1-6	- 18 $_1$	- 17	+ 11	+ 10
1-12	- 23	- 23	+ 13	+ 13
13-24	- 17	- 17	+ 12	+ 12
1-24	- 17	- 17	+ 12	+ 12
25-29	- 21 $_3$	- 21	- 18	+ 15
30-34	- 20 $_1$	- 20 $_1$	+10	+ 15
1-34	- 17	- 17	+ 12	+ 12

Meaning of signs and numbers in the table:

"- n" means that majority of approximation errors were greater than 10% of really observed values of endogenous variables, and n is the number of such errors,

"+ n" means that most of approximation errors were not greater than 10%, and n is the number of errors greater than 10% of really observed values,

number appearing as subscript means number of errors greater than 100%,

 "full" means linear regression with free term and two explanatory (input) variables.

Source: authors calculations.

TABLE 4. Observations for system G2 generated in new conditions of experiment.

No. of obser-vation	Initial state: q_3 or q_4				Initial state: q_1			
	"Old" input sequences				New input sequences			
	Quantities of inputs		Quantities of outputs		Quantities of inputs		Quantities of outputs	
	$y_1{}'$	$y_2{}'$	$x_1{}'$	$x_2{}'$	$y_1{}'$	$y_2{}'$	$x_1{}'$	$x_2{}'$
1	4	15	12	7	4	13	11	6
2	2	3	1	4	2	4	3	3
3	0	2	2	0	5	13	10	8
4	1	2	2	1	1	4	3	2
5	1	3	1	3	1	1	0	2
6	2	3	2	3	1	2	1	2
7	0	1	0	1	4	10	8	6
8	4	9	6	7	0	2	2	0
9	1	3	3	1	4	6	5	5
10	0	2	2	0	1	4	4	1
11	1	1	1	1	3	8	8	3
12	0	1	1	0	3	9	5	7
13	1	2	1	2	2	2	0	4
14	2	4	2	4	4	7	4	7
15	0	1	1	0	5	9	6	8
16	2	2	2	2	0	2	0	2
17	6	13	7	12	1	0	1	0
18	2	5	4	3	0	1	0	1
19	2	2	1	3	3	2	2	3
20	0	4	4	0	1	5	3	3
21	2	3	1	4	10	26	20	16
22	5	9	6	8	1	3	2	2
23	0	1	1	0	1	3	1	3
24	7	12	9	10	5	9	5	9
25	3	9	6	6	0	2	2	0
26	4	6	1	9	3	13	11	5
27	1	10	10	1	4	15	13	6
28	5	7	3	9	4	7	5	6
29	4	7	6	5	1	5	5	1
30	1	9	9	1	1	2	2	1
31	3	9	6	6	6	15	14	7
32	3	8	5	6	5	8	7	6
33	5	6	1	10	2	3	1	4
34	2	10	9	3	1	4	2	3

Source: authors calculations.