Lack of consistency of Hellwig's method

J. A. Milosz[*], Roman Urban[†]

June 13, 2025

Abstract

Hellwig's method was developed in 1969 by Zdzisław Hellwig. It is one of the approaches that can be applied to the important problem of variable selection in linear regression models. This method is still presented in the vast majority of Polish econometrics textbooks, and it is also used in some scientific articles. The main objective of this article is to theoretically demonstrate that Hellwig's method does not satisfy one of the fundamental properties we would expect from a variable selection method, namely, that it is not asymptotically consistent. The article also presents simulations that confirm the theoretical results.

Keywords: Hellwig's method, model selection, linear regression model, econometric modeling

JEL Classification: C18, C52

[*]Student at the Mathematical Institute, University of Wrocław, Wrocław, Poland; e-mail: jakub.aleksander.milosz@outlook.com; ORCID: 0009-0002-9579-6509

[†]Mathematical Institute, University of Wrocław, Wrocław, Poland; e-mail: roman.urban@math.uni.wroc.pl; ORCID: 0000-0001-8931-3342

1 Introduction

The problem of model selection is a very important aspect of econometric modeling. In the context of a linear regression model without an error term, the problem can be stated as follows. Given:

- a set of candidate explanatory variables $\mathcal{X} := \{X_1, X_2, \ldots, X_J\}$ ($J \in \mathbb{N}$) and a sample of size $n \in \mathbb{N}$ drawn from these variables,
- the knowledge that a dependent variable Y is a linear combination of elements from some subset of $\mathcal{X}$ (this will be referred to as the true model),

the objective is to determine precisely which subset of $\mathcal{X}$ determines Y.

To solve the problem above, a model scoring function $S_n : (2^{\mathcal{X}} \setminus \{\emptyset\}) \ni j \longmapsto S_n(j) \in \mathbb{R}$ is defined, which assigns a fixed value to each non-empty subsets $j \subseteq \mathcal{X}$. The subset of $\mathcal{X}$ that maximizes or minimizes (depending on the method used) S_n is then selected. The best-known example of such a function is the information criterion introduced by Akaike (1973). The Akaike's method selects the subset that minimizes $S_n(j) = AIC(j) = -2\ln(\hat{L}(j)) + \#j$, where $\hat{L}(j)$ is the maximum likelihood function of model j, evaluated at the maximum likelihood estimators for that model, and $\#j$ denotes the number of explanatory variables in the model.

An important property of a model selection method is asymptotic consistency. Assume that the method selects model j^*. Then the method is said to be *asymptotically consistent* if and only if, we have:

$$\mathbb{P}\left(j^* \text{ selected by } S_n \text{ is the true model}\right) \longrightarrow 1 \text{ as } n \longrightarrow +\infty.$$

The Bayesian information criterion, introduced by Schwarz (1978), is an example of a method that satisfies this property.

In the late 1960s, Polish economists and econometrician Zdzisaw Hellwig proposed his own version of a selection function (Hellwig 1969), which he continued to analyze over the following decades in Polish academic journals (Hellwig 1974, 1985a, 1985b). He assumed that the data were standardized, meaning that for each $i \in \{1, \ldots, J\}$, $\mathbb{E}[X_i] = 0$ and $\mathrm{Var}[X_i] = 1$ (this assumption will be maintained throughout this article). His method selects the subset of $\mathcal{X}$ that maximizes:

$$S_n(j) = \sum_{i \in I} \frac{\rho^2_{X_i,Y}}{\sum_{m \in I} |\rho_{X_i,X_m}|}, \tag{1}$$

where $I \subseteq \{1, ..., J\}$ is the set of indices of random variables from j, and ρ denotes the sample correlation. Analogously, the selection function is defined not for a sample, but for theoretical random variables:

$$S(j) = \sum_{i \in I} \frac{\mathrm{corr}^2(X_i, Y)}{\sum_{m \in I} |\,\mathrm{corr}(X_i, X_m)|}. \tag{2}$$

Already in the 20th century, a discussion around Hellwig's approach had begun. In the same year, Guzik (1985) and Czerwiński (1985) identified certain limitations in the method's ability to reliably select the true model. In the 21st century, Hellwig's method was critically discussed by Bednarski and Borowicz (2009), and by Serwa (2011). The former attempted to demonstrate the inconsistency of Hellwig's method in the context of the linear regression model; however, the proof provided by the authors is incorrect. The latter succeeded in establishing this inconsistency, but only for linear autoregressive models. In the next section, we prove the inconsistency of the method proposed by Hellwig in the context of a linear regression model without an error term. In the final section, we confirm the theoretical result through simulations and draw conclusions.

2 Lack of consistency

First, note that if j^* is the model selected by S from (2), then the model selected by S_n given in (1) will also be j^* almost surely as $n \longrightarrow +\infty$. Therefore, to prove the lack of consistency, it suffices to show that S chooses a model different from the true one. We will use the following theorem for this:

Theorem 2.1. *For a two-dimensional linear regression model without an error term:*

$$Y = a_1 X_1 + a_2 X_2$$

if $\mathrm{corr}(X_1, X_2) \neq \pm 1$ *and* $\mathrm{corr}(X_1, X_2) \neq 0$*, there exist coefficients* a_1*,* $a_2 \in \mathbb{R}$ *such that selection function* S*, defined in (2), will indicate a one-dimensional model.*

Proof.

According to Hellwig's assumptions: $\mathrm{Var}[X_1] = \mathrm{Var}[X_2] = 1$. From this assumption, we get that: $\mathrm{corr}(X_1, X_2) = \mathrm{Cov}(X_1, X_2) \in (-1, 1) \setminus \{0\}$. We shall identify a set of values a_1, a_2 such that the Hellwig method consistently selects the model containing only the variable X_1, and demonstrate that this set is non-empty, thereby proving the theorem.

First, let us compute how the correlation between X_1 and Y can be expressed in terms of a_1, a_2, and the covariance between X_1 and X_2:

$$\begin{aligned}
\operatorname{corr}(X_1, Y) &= \frac{\operatorname{Cov}(X_1, Y)}{\sqrt{\operatorname{Var}[X_1]\operatorname{Var}[Y]}} \\
&= \frac{\operatorname{Cov}(X_1, a_1X_1 + a_2X_2)}{\sqrt{\operatorname{Cov}(a_1X_1 + a_2X_2, a_1X_1 + a_2X_2)}} \\
&= \frac{a_1 \operatorname{Var}[X_1] + a_2 \operatorname{Cov}(X_1, X_2)}{\sqrt{a_1^2 \operatorname{Var}[X_1] + 2a_1a_2 \operatorname{Cov}(X_1, X_2) + a_2^2 \operatorname{Var}[X_2]}} \\
&= \frac{a_1 + a_2 \operatorname{Cov}(X_1, X_2)}{\sqrt{a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)}}. \qquad (3)
\end{aligned}$$

Analogously, the correlation between X_2 and Y can be expressed as:

$$\operatorname{corr}(X_2, Y) = \frac{a_1 \operatorname{Cov}(X_1, X_2) + a_2}{\sqrt{a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)}}. \qquad (4)$$

Let us now compute the value of the selection function S for the non-empty subsets of the set $\{X_1, X_2\}$. Using equations (3) and fact that $\operatorname{corr}(X_1, X_1) = 1$, for the subset $\{X_1\}$ we have:

$$\begin{aligned}
S(\{X_1\}) &= \frac{\operatorname{corr}^2(X_1, Y)}{|\operatorname{corr}(X_1, X_1)|} \\
&= \left(\frac{a_1 + a_2 \operatorname{Cov}(X_1, X_2)}{\sqrt{a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)}}\right)^2 \\
&= \frac{(a_1 + a_2 \operatorname{Cov}(X_1, X_2))^2}{|a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)|}. \qquad (5)
\end{aligned}$$

Similarly, by applying equation (4) and the fact that $\operatorname{corr}(X_2, X_2) = 1$, we obtain the following for the subset $\{X_2\}$:

$$
\begin{aligned}
S(\{X_2\}) &= \frac{\operatorname{corr}^2(X_2, Y)}{|\operatorname{corr}(X_2, X_2)|} \\
&= \frac{(a_1 \operatorname{Cov}(X_1, X_2) + a_2)^2}{|a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)|}. \qquad (6)
\end{aligned}
$$

Using equations (3) and (4), and the fact that $\operatorname{corr}(X_1, X_1) = \operatorname{corr}(X_2, X_2) = 1$ for the entire set of explanatory variables $\{X_1, X_2\}$, we obtain:

$$
\begin{aligned}
S(\{X_1, X_2\}) &= \frac{\operatorname{corr}^2(X_1, Y)}{|\operatorname{corr}(X_1, X_1)| + |\operatorname{corr}(X_1, X_2)|} + \frac{\operatorname{corr}^2(X_2, Y)}{|\operatorname{corr}(X_2, X_1)| + |\operatorname{corr}(X_2, X_2)|} \\
&= \frac{\left(\frac{a_1 + a_2 \operatorname{Cov}(X_1, X_2)}{\sqrt{a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)}}\right)^2 + \left(\frac{a_1 \operatorname{Cov}(X_1, X_2) + a_2}{\sqrt{a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)}}\right)^2}{1 + |\operatorname{Cov}(X_1, X_2)|} \\
&= \frac{(a_1 + a_2 \operatorname{Cov}(X_1, X_2))^2 + (a_1 \operatorname{Cov}(X_1, X_2) + a_2)^2}{|a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)|(1 + |\operatorname{Cov}(X_1, X_2)|)}. \qquad (7)
\end{aligned}
$$

In order for the Hellwig method to consistently select the model containing only X_1, the following two inequalities must be satisfied:

$$S(\{X_1\}) > S(\{X_2\}), \qquad (8)$$

$$S(\{X_1\}) > S(\{X_1, X_2\}). \qquad (9)$$

Let's express inequality (8) in terms of equations (5) and (6):

$$\frac{(a_1 + a_2 \operatorname{Cov}(X_1, X_2))^2}{|a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)|} > \frac{(a_1 \operatorname{Cov}(X_1, X_2) + a_2)^2}{|a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)|}.$$

Following a series of simplifications, we obtain:

$$a_1^2 + a_2^2 \operatorname{Cov}^2(X_1, X_2) > a_1^2 \operatorname{Cov}^2(X_1, X_2) + a_2^2.$$

This leads to:

$$(1 - \operatorname{Cov}^2(X_1, X_2))(a_1^2 - a_2^2) > 0 \iff \\ (1 - \operatorname{Cov}^2(X_1, X_2))(a_1 - a_2)(a_1 + a_2) > 0. \quad (10)$$

Recalling the assumption that $\operatorname{Cov}(X_1, X_2) \in (-1, 1) \setminus \{0\}$, it follows that $1 - \operatorname{Cov}^2(X_1, X_2) > 0$. Therefore, for the inequality to hold, the coefficients a_1, a_2 must lie in the set $A := \{(a_1, a_2) \in \mathbb{R}^2 : (a_1 - a_2)(a_1 + a_2) > 0\}$.

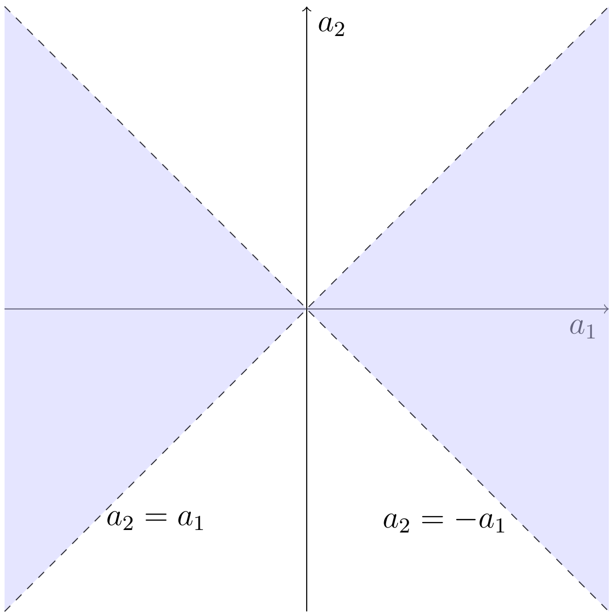

Figure 2.1: Visualization of the set A (in blue).

Let us now express inequality (9) in terms of equations (5) and (7):

$$\frac{(a_1 + a_2 \operatorname{Cov}(X_1, X_2))^2}{|a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)|} \\ > \frac{(a_1 + a_2 \operatorname{Cov}(X_1, X_2))^2 + (a_1 \operatorname{Cov}(X_1, X_2) + a_2)^2}{|a_1^2 + a_2^2 + 2a_1a_2 \operatorname{Cov}(X_1, X_2)|(1 + |\operatorname{Cov}(X_1, X_2)|)}.$$

After multiplying both sides by the denominator on the left-hand side of the inequality above, we obtain:

$$(a_1 + a_2 \operatorname{Cov}(X_1, X_2))^2(1 + |\operatorname{Cov}(X_1, X_2)|) \\ > (a_1 + a_2 \operatorname{Cov}(X_1, X_2))^2 + (a_1 \operatorname{Cov}(X_1, X_2) + a_2)^2.$$

Moving all terms to the left-hand side yields:

$$|\operatorname{Cov}(X_1, X_2)|(a_1 + a_2 \operatorname{Cov}(X_1, X_2))^2 - (a_1 \operatorname{Cov}(X_1, X_2) + a_2)^2 > 0.$$

After multiplying and organizing the terms, we obtain:

$$(|\operatorname{Cov}(X_1, X_2)| - \operatorname{Cov}^2(X_1, X_2))\, a_1^2 + 2 \operatorname{Cov}(X_1, X_2)(|\operatorname{Cov}(X_1, X_2)| - 1)\, a_1a_2 \\ + (|\operatorname{Cov}^3(X_1, X_2)| - 1)\, a_2^2 > 0. \tag{11}$$

Based on the inequality above, let us define the quadratic form $f : \mathbb{R}^2 \to \mathbb{R}$ as follows:

$$f(a_1, a_2) := (|c| - c^2)\, a_1^2 + 2c(|c| - 1)\, a_1a_2 + (|c^3| - 1)\, a_2^2.$$

Where c is a constant from $(-1, 1) \setminus \{0\}$. Let us simplify f so that it takes the form $\alpha(a_1 + \beta\, a_2)^2 + \gamma\, a_2^2$, where α, β, and γ are constants:

$$\begin{aligned}
f(a_1, a_2) &= (|c| - c^2)\left(a_1^2 + \frac{2c(|c|-1)}{|c|(1-|c|)}\, a_1 a_2\right) + (|c^3| - 1)\, a_2^2 \\
&= (|c| - c^2)\left(a_1^2 - \frac{2c}{|c|}\, a_1 a_2 + a_2^2\right) + ((|c^3| - 1) - (|c| - c^2))\, a_2^2 \\
&= (|c| - c^2)\left(a_1^2 - 2\,\mathrm{sng}(c)\, a_1 a_2 + a_2^2\right) + (|c^3| + c^2 - |c| - 1)\, a_2^2 \\
&= (|c| - c^2)(a_1 - \mathrm{sng}(c)\, a_2)^2 + (c^2|c| - c^2 + 2c^2 - 2|c| + |c| - 1)\, a_2^2 \\
&= (|c| - c^2)(a_1 - \mathrm{sng}(c)\, a_2)^2 + (|c| - 1)(c^2 + 2|c| + 1)\, a_2^2 \\
&= (|c| - c^2)(a_1 - \mathrm{sng}(c)\, a_2)^2 - (1 - |c|)(|c| + 1)^2\, a_2^2.
\end{aligned}$$

Let's now determine the set of points (a_1, a_2) for which $f > 0$:

$$\begin{aligned}
f(a_1, a_2) > 0 &\iff \\
(|c| - c^2)(a_1 - \mathrm{sng}(c)\, a_2)^2 - (1 - |c|)(|c| + 1)^2\, a_2^2 > 0 &\iff \\
(a_1 - \mathrm{sng}(c)\, a_2)^2 > \frac{(1 - |c|)(|c| + 1)^2}{|c|(1 - |c|)}\, a_2^2 &\iff \\
(a_1 - \mathrm{sng}(c)\, a_2)^2 > \left(\frac{|c| + 1}{\sqrt{|c|}}\right)^2 a_2^2 &\iff \\
|a_1 - \mathrm{sng}(c)\, a_2| > \frac{|c| + 1}{\sqrt{|c|}}\, |a_2|.&
\end{aligned}$$

From the transformations above, it follows that in order for inequality (11) to hold, the pair (a_1, a_2) must belong to the set $B_{\mathrm{Cov}(X_1,X_2)} := \{(a_1, a_2) \in \mathbb{R}^2 : \sqrt{|\operatorname{Cov}(X_1, X_2)|}\, |a_1 - \mathrm{sng}(\operatorname{Cov}(X_1, X_2))\, a_2| > (|\operatorname{Cov}(X_1, X_2)| + 1)\, |a_2|\}$.

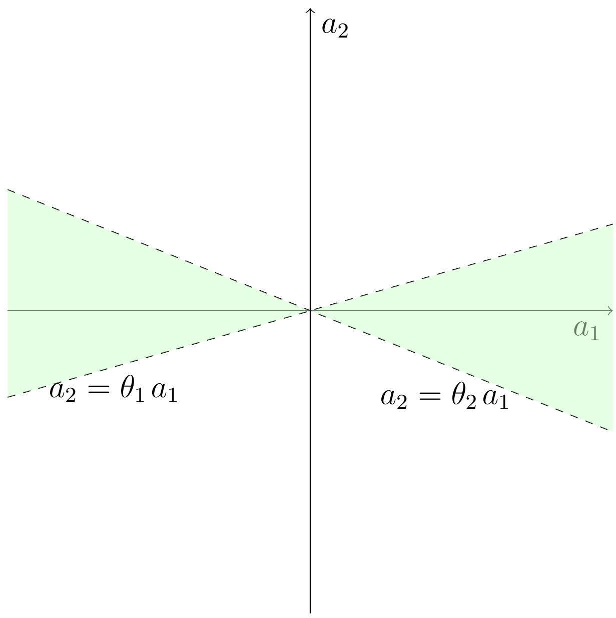

Figure 2.2: Visualization of the set $B_{\mathrm{Cov}(X_1,X_2)}$ for $\mathrm{Cov}(X_1, X_2) = \frac{1}{4}$ (in green). Where $\theta_1 := \frac{\sqrt{|\mathrm{Cov}(X_1,X_2)|}}{|\mathrm{Cov}(X_1,X_2)|+\mathrm{sng}(\mathrm{Cov}(X_1,X_2))\sqrt{|\mathrm{Cov}(X_1,X_2)|+1}}$, $\theta_2 := \frac{\sqrt{|\mathrm{Cov}(X_1,X_2)|}}{-|\mathrm{Cov}(X_1,X_2)|+\mathrm{sng}(\mathrm{Cov}(X_1,X_2))\sqrt{|\mathrm{Cov}(X_1,X_2)|-1}}$.

For the pair a_1, a_2 to satisfy inequalities (10) and (11), it must belong to the set $B_{\mathrm{Cov}(X_1,X_2)} \cap A$. Since $\theta_1 \in (0, 1)$ and $\theta_2 \in (-1, 0)$, it follows that $B_{\mathrm{Cov}(X_1,X_2)} \subset A$, and therefore $B_{\mathrm{Cov}(X_1,X_2)} \cap A = B_{\mathrm{Cov}(X_1,X_2)}$. Observe that for any value of $\mathrm{Cov}(X_1, X_2)$ consistent with the assumption (i.e., $\mathrm{Cov}(X_1, X_2) \in (-1, 1) \setminus \{0\}$), the set is non-empty, which completes the proof. $\square$

As a result of the observations made at the beginning of this section and the theorem just proven, we are now in a position to prove the main theorem of this article.

Theorem 2.2. *For the problem of model selection in the context of a linear regression model, the method proposed by Hellwig (1969, 1974, 1985a, 1985b), referred to as Hellwig's method, is not asymptotically consistent.*

Proof. Assume that:

- $\mathcal{X} = \{X_1, X_2\}$;
- the true model is $Y = a_1 X_1 + a_2 X_2$;
- $\mathrm{corr}(X_1, X_2) \neq \pm 1, \mathrm{corr}(X_1, X_2) \neq 0$;
- and $(a_1, a_2) \in B_{\mathrm{Cov}(X_1, X_2)}$.

Thus, according to theorem 2.1, Hellwig's method will select the subset $\{X_1\}$ instead of the true model $\{X_1, X_2\}$ as the sample size tends to infinity. This demonstrates the lack of asymptotic consistency and completes the proof. $\square$

3 Simulation and conclusion

To conduct the simulation, consider the following example. Let Z_1, Z_2 be independent random variables with distribution $N(0,1)$, and define $X_1 := Z_1$, $X_2 := \frac{1}{4} Z_1 + \frac{\sqrt{15}}{4} Z_2$. Then $\mathrm{corr}(X_1, X_2) = \mathrm{Cov}(X_1, X_2) = \frac{1}{4}$. Assume that the true model is given by:

$$Y = 4X_1 + X_2. \tag{12}$$

We repeated the variable selection procedure for the model above 100,000 times using three methods (Hellwig's, Akaike's, Schwarz's), and sample size $n \in \{20, 100, 1000\}$. The resulting frequencies of selecting each model are presented below.

	$\{X_1\}$	$\{X_2\}$	$\{X_1, X_2\}$ (true model)
Hellwig's method	0.67333	0	0.32667
Akaike's method	0	0	1
Schwarz's method	0	0	1

Table 2.1: Frequencies of selecting each model by each method for sample size $n = 20$.

	$\{X_1\}$	$\{X_2\}$	$\{X_1, X_2\}$ (true model)
Hellwig's method	0.79283	0	0.20717
Akaike's method	0	0	1
Schwarz's method	0	0	1

Table 2.2: Frequencies of selecting each model by each method for sample size $n = 100$.

	$\{X_1\}$	$\{X_2\}$	$\{X_1, X_2\}$ (true model)
Hellwig's method	0.99841	0	0.00159
Akaike's method	0	0	1
Schwarz's method	0	0	1

Table 2.3: Frequencies of selecting each model by each method for sample size $n = 1000$.

The data above confirm the theoretical result shown in the second section. It is clear that for the model (12), Hellwig's method is not asymptotically consistent.

The results presented in this article confirm that the hypothesis posed in Bednarski and Borowicz (2009) regarding Hellwig's method was true. This method does not fulfill its role as a variable selection method for the linear regression model. Therefore, it should not be used in practical or scientific applications.

References

[1] Akaike H., (1973), Information theory and an extension of the maximum likelihood principle, [in:] 2nd International Symposium on Information Theory, [ed.:] B. N. Petrov, F. Csáki, Budapest, 267–281.

[2] Bednarski T., Borowicz F., (2009), On inconsistency of Hellwig's variable choice method in regression models, Discussiones Mathematicae Probability and Statistics 29(1), 41–51.

[3] Czerwiński Z., (1985), O jakości modelu ekonometrycznego (refleksje nad artykułem Zdzisława Hellwiga) (On the quality of an econometric model (reflections on the article by Zdzisław Hellwig)), Przegląd Statystyczny. Statistical Review 32(3), 269–277.

[4] Guzik B., (1985), Metoda Hellwiga w warunkach współliniowości par zmiennych objaśniających (The Hellwig's method in the case of linear dependence of pairs of explanatory variables), Przegląd Statystyczny. Statistical Review 32(1), 33–39.

[5] Hellwig Z., (1969), Problem optymalnego wyboru predyktant (Problem of optimal choice of predictants), Przegląd Statystyczny. Statistical Review 16(3-4), 221–237.

[6] Hellwig Z., (1974), Rozważania nad istotą modelu ekonometrycznego (Reflections on the nature of an econometric model), Ekonomista 74(2), 305–324.

[7] Hellwig Z., (1985a), O jakości modelu ekonometrycznego część I. identyfikacja (Towards the quality of the econometric model part I. identification), Przegląd Statystyczny. Statistical Review 32(1), 3–23.

[8] Hellwig Z., (1985b), O jakości modelu ekonometrycznego część II. estymacja (Towards the quality of the econometric model part II. estimation), Przegląd Statystyczny. Statistical Review 32(4), 287–311.

[9] Serwa D., (2011), Metoda Hellwiga jako kryterium doboru zmiennych do modeli szeregów czasowych (Using Hellwig method to select explanatory variables in time series models), Metody Ilościowe w Badaniach Ekonomicznych. Quantitative Methods in Economics 12(2), 312–321.

[10] Schwarz G. E., (1978), Estimating the dimension of a model, Annals of Statistics 6(2), 461–464.

[11] Shalizi C. R., Kontorovich A., (2010), Almost none of the theory of stochastic processes, available at: `https://www.stat.cmu.edu/~cshalizi/almost-none/v0.1.1/almost-none.pdf`.

A Python code used for the simulation

```
import numpy as np
import statsmodels.api as sm


np.random.seed(0)


hellwig = [0, 0, 0]
aic = [0, 0, 0]
bic = [0, 0, 0]


n = 1000


for _ in range(100000):
    z_1 = [np.random.normal(0, 1) for _ in range(n)]
    z_2 = [np.random.normal(0, 1) for _ in range(n)]

    x_1 = [el_1 for el_1 in z_1]
    x_2 = [1/4 * el_1 + np.sqrt(15)/4 * el_2
           for (el_1, el_2) in zip(z_1, z_2)]
    y = [4 * el_1 + el_2
         for (el_1, el_2) in zip(x_1, x_2)]

    hellwig_s_x_1 = (np.corrcoef(x_1, y)[0, 1])**2
```

```
hellwig_s_x_2 = (np.corrcoef(x_2, y)[0, 1])**2
hellwig_s_x_1_x_2 = (((np.corrcoef(x_1, y)[0, 1])**2 +
                      (np.corrcoef(x_2, y)[0, 1])**2) /
                     (1 + abs(np.corrcoef(x_1, x_2)[0, 1])))
max_hellwig = max(hellwig_s_x_1, hellwig_s_x_2, hellwig_s_x_1_x_2)
for i in range(3):
    if [hellwig_s_x_1, hellwig_s_x_2,
        hellwig_s_x_1_x_2][i] == max_hellwig:
        hellwig[i] = hellwig[i] + 1

model_1 = sm.OLS(y, sm.add_constant(x_1)).fit()
model_2 = sm.OLS(y, sm.add_constant(x_2)).fit()
model_3 = sm.OLS(y, sm.add_constant(
    np.column_stack((x_1, x_2)))).fit()

aic_s_x_1 = model_1.aic
aic_s_x_2 = model_2.aic
aic_s_x_1_x_2 = model_3.aic
min_aic = min(aic_s_x_1, aic_s_x_2, aic_s_x_1_x_2)
for i in range(3):
    if [aic_s_x_1, aic_s_x_2,
        aic_s_x_1_x_2][i] == min_aic:
        aic[i] = aic[i] + 1

bic_s_x_1 = model_1.bic
```

```
    bic_s_x_2 = model_2.bic

    bic_s_x_1_x_2 = model_3.bic

    min_bic = min(bic_s_x_1, bic_s_x_2, bic_s_x_1_x_2)

    for i in range(3):

        if [bic_s_x_1, bic_s_x_2,

            bic_s_x_1_x_2][i] == min_bic:

            bic[i] = bic[i] + 1


print(f'{[el/100000 for el in hellwig]} \n'

      f'{[el/100000 for el in aic]} \n'

      f'{[el/100000 for el in bic]} \n')
```