Treatment effect estimation in high-dimension: An inference-based approach

Ulrich Aiounou,* Emmanuel Flachaire,* Sullivan Hué,* Sébastien Laurent*†

April 29, 2024

Abstract

Post-Lasso and Post-Double-Lasso are becoming the most popular methods for estimating average treatment effects from linear regression models with many covariates. However, these methods can suffer from substantial omitted variable bias in finite sample. We propose a new method called Post-Double-Autometrics, which is based on Autometrics, and show that this new method outperforms Post-Double-Lasso when explanatory variables are weakly correlated with the endogenous variable but correlated with the treatment variable.

JEL Classification: C21, C52, C55.
Keywords: Treatment effect, High dimension, Lasso, Autometrics.

---

*Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS
†Aix-Marseille Graduate School of Management – IAE, France.
The authors acknowledge research support by the French National Research Agency Grants ANR-17-EURE-0020 and ANR-21-CE26-0007-01 (project MLEforRisk) and by the Excellence Initiative of Aix-Marseille University - A*MIDEX.

1 Introduction

Post-Lasso (Belloni and Chernozhukov, 2013) and Post-Double-Lasso (Belloni et al., 2014) are becoming the most popular methods for estimating average treatment effects in a linear regression model with many covariates. These methods allow to use many covariates, to avoid omitted variable bias, while controlling overfitting issue and loss of accuracy with a Lasso variable selection method.

Wüthrich and Zhu (2023) recently showed that these methods can suffer from substantial omitted variable bias in finite sample, even in simple settings favorable to the Lasso. They also show that the performance of these methods is very sensitive to the choice of the regularization parameter of the Lasso, and no specific choice of this parameter systematically outperforms the others.

In this paper, we consider an alternative variable selection method, Autometrics (Doornik et al., 2009), which is based on statistical inference. This algorithm performs automatic selection of variables based on "Hendry's theory of reduction" (see Chapter 6 of Hendry and Doornik, 2014) and an automatic general-to-specific model selection (Hendry, 2000).

Autometrics is known to have very good properties in the sense that, unlike Lasso, its main hyperparameter (i.e., the target size α) allows to determine approximatively the frequency of irrelevant explanatory variables included in the terminal model. Importantly, like the hyperparameter of Lasso (i.e., λ), the target size also influences the frequency of relevant variables selected included in the terminal model. However, we show in this paper that there are cases where Autometrics fails to keep some relevant variables, which translates into a biased estimate of the average treatment effect when these variables are correlated with the treatment variable.

To overcome this problem, we propose a new method called Post-Double-Autometrics, that extends Post-Double-Lasso by using Autometrics instead of Lasso to do variable selection. We show through Monte Carlo experiments and an empirical illustration that the proposed method outperforms Post-Double-Lasso in some situations. Figure 1 provides a numerical illustration of this phenomenon in a simple simulation experiment, with $n = 500$ observations and $p = 200$ covariates. The Post-Double-Lasso estimator of the treatment effect is biased, shifted much to the right compared to the true value. To the opposite, the Post-Double-Autometrics estimator is approximately unbiased and behaves almost like the oracle estimator. Our simulation results also show that Post-Double-Autometrics provides the best variable selection method.

The reminder of the paper is structured as follows. Section 2 presents the pitfalls of Lasso-based methods and introduce our new approach. Section 3 is devoted to Monte Carlo experiments and Section 4 to an empirical illustration. The paper concludes with Section 5.

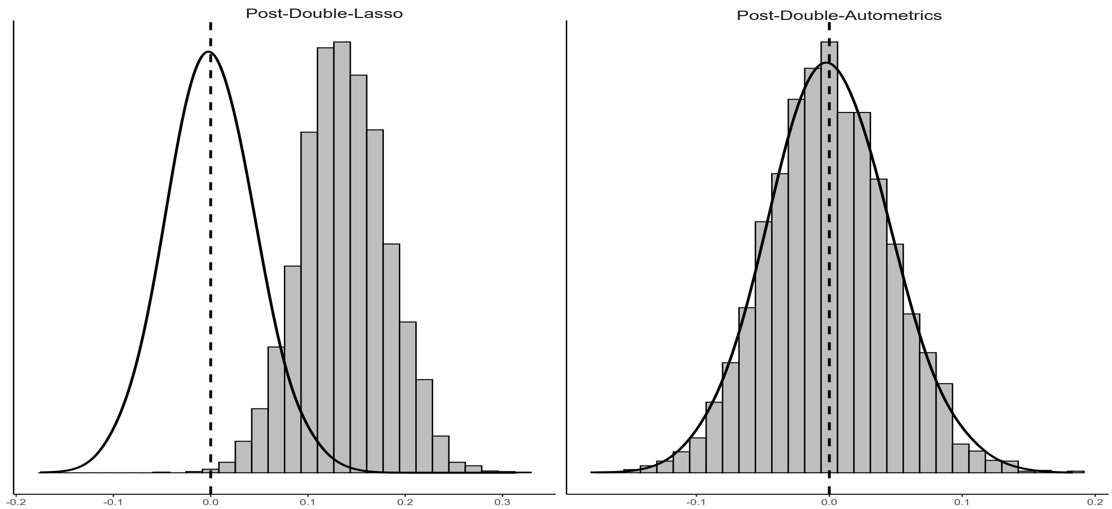

Figure 1: Empirical distribution (histogram) of the Post-Double-Lasso (left panel) and Post-Double-Autometrics estimators (right panel) of the treatment effect. In both panels, the distribution of the oracle estimator of the treatment effect corresponds to the solid line while the true value corresponds to the dotted vertical line.

2 Treatment effect estimation in high dimensions

2.1 Pitfalls of Post-Lasso and Post-Double-Lasso

Following the literature, we consider the following standard linear regression model

$$Y_i = D_i\delta + X_i\beta + \epsilon_i, \tag{1}$$

$$D_i = X_i\gamma + \eta_i, \tag{2}$$

for $i = 1, \ldots, n$, where Y_i is the outcome, D_i is a dummy corresponding to the treatment variable, X_i is a p-vector of control variables $X = (X_1, \ldots, X_p) \in \mathbb{R}^p$, and ϵ_i and η_i are error terms. The parameter of interest δ is the average treatment effect.

In observational studies, the unconfoundedness condition is crucial to obtain unbiased estimation of the average treatment effect δ. This condition requires that X contains all the confounder variables, which explain both the outcome Y and the treatment D. Otherwise, the estimation may suffer from substantial omitted variable bias (OVB) or counfounding bias. In practice, it is difficult to account for all counfounder variables. Here, we consider the strategy of including many control variables and using a variable selection method to select the relevant control variables.

An early approach developed in the literature is Lasso, which performs estimation and variable selection simultaneously (Tibshirani, 1996). Considering the linear regression in (1), Lasso solves the following penalized regression problem

$$\hat{\theta} = \arg\min_{\beta} \left[\sum_{i=1}^{n} (Y_i - Z_i\theta)^2 + \lambda \sum_{j=1}^{p+1} |\theta_j| \right], \tag{3}$$

where $Z_i = (D_i, X_i)$ and $\theta = (\delta, \beta)$. Estimation is feasible with high dimensional data, when the number of covariates p is large and can exceed the number of observations n. The penalization restricts the magnitude of the coefficients, making some of them exactly zero with a sufficiently large λ. Several approaches have been proposed in the literature to set λ. For example, Bickel et al. (2009) proposed to set $\lambda^{bya} = 2\sigma\sqrt{2n^{-1}(1+\tau)\log p}$, where σ is the standard error of the error term, and $\tau > 0$. Belloni et al. (2012) proposed a penalization choice, denoted as λ^{bcch}, to take into account heteroscedastic and clustered errors. Another popular approach is to set λ by cross-validation, with the value minimizing the prediction error, λ^{min}, or the value associated to the most parsimonious model within a 1-standard-error interval, λ^{1se}.

The penalization shrinks all penalized coefficients towards zero, introducing some bias in the estimates. Since the interest is on the estimation of the parameter δ, it is recommended to use the Post-Lasso (Belloni and Chernozhukov, 2013) described below:

1. use Lasso on (1) to select a subset of variables, denoted X^*;

2. estimate δ from the OLS regression of Y on D and X^*.

Belloni et al. (2014) show that Lasso may fail to select some relevant variables in the first step, which may lead to a biased estimate of the parameter of interest δ. This typically occurs when some control variables in X are correlated with D, but only have a small effect on Y. To solve this issues, they proposed the Post-Double-Lasso:

1. use Lasso on a restricted version of (1) in which D is ignored to select a subset of variables X^*;

2. use Lasso on (2) to select a subset of variables X^{**};

3. estimate δ from the OLS regression of Y on D and the union of X^* and X^{**}.

This method gives control variables omitted in the first step a second chance to be recovered in the second step, especially those correlated to D and likely to lead to an omitted variable bias.

Although the Post-Double-Lasso is a substantial improvement over the Post-Lasso, Wüthrich and Zhu (2023) show that it can still lead to OVB issue, when the coefficients of relevant variables are not large enough for the Lasso to select them but large enough to cause bias in the estimation of δ. Moreover, they show that the extent of the bias depends on the choice of the regularization parameter, and that there is no single choice of λ among popular ones systematically leading to better results.

2.2 Post-Double-Autometrics: An inference-based method

To address the OVB issue, we propose to use another variable selection method: Autometrics (Doornik et al., 2009). Autometrics is a popular variables selection model among econometricians because it relies on statistical inference to select the relevant variables. There exists at least six different software implementations of automatic general-to-specific model selection (Matlab, OxMetrics, Scilab, STATA, EViews and R). Autometrics is one of them and is implemented in OxMetrics. See Sucarrat (2020) for more details on general-to-specific modelling and its implementations.

More specifically, Autometrics is an algorithm performing automatic variables selection with the "general-to-specific" model selection approach (Hendry, 2000). This method starts from a generalized unrestricted model (GUM) which includes all potential relationships between the outcome Y and the control variables X, such as dynamic effects, breaks, or non-linearities@@Préciser que ce ne sont pas des non-linéarités comme avec du ML.. The algorithm then performs a battery of tests to remove insignificant variables from X and to find a congruent parsimonious model.[1]

Autometrics has two main advantages. Firstly, unlike the Lasso performing variable selection based on the magnitude of coefficients (Wüthrich and Zhu, 2023), Autometrics relies on statistical inference. It implies that Autometrics-based methods will select variables as soon as their effect on Y or D is significant, independently of the coefficients magnitude. Secondly, the tuning parameter of Autometrics is the target nominal size α. It allows the user to control the expected proportion of irrelevant variables included in the final model, something not possible with Lasso.[2]

A standard implementation of Autometrics is to use this variable selection method directly on (1) by forcing the presence of D in the terminal model. Importantly, Autometrics is expected to miss relevant variables with a low degree of non-centrality and if these variables are strongly correlated with D, this can lead to a biased estimate of δ.

[1]For more details on Autometrics, see Doornik et al. (2009) and Castle et al. (2023).

[2]Typical values of this parameter are $\alpha = 0.05$, or $\alpha = 0.01$ for a more conservative user.

Therefore, in the same way as Belloni et al. (2014), we propose the Post-Double-Autometrics method to estimate the average treatment effect in presence of many control variables. Our method is similar to the Post-Double-Lasso described above, with Lasso being replaced by Autometrics in the first two steps. By giving control variables omitted in the first step a second chance to be recovered in the second step, this method should improve over the usual implementation of Autometrics in some cases.

Post-Double-Autometrics shares similarities with the indicator saturation method proposed by Hendry et al. (2006). Their approach consists in splitting the initial sample in two blocks, and applying a general-to-specific variable selection method on each block. A final variable selection is then applied on the union of variables selected in each block. However, the dependent variable considered in both blocks of the indicator saturation method is the same (Y). Moreover, the main objective of the indicator saturation method is to detect outliers or breaks and not estimating an average treatment effect as in the Post-Double-Autometrics.

3 Simulations

In this section, we use Monte Carlo simulations to illustrate the finite sample properties of the Post-Double-Autometrics method, and compare it to those of the benchmark Post-Double-Lasso. We also include in the comparison Post-Lasso and Autometrics. The Monte Carlo simulation is performed on 5,000 replications, and the experiment is inspired from Wüthrich and Zhu (2023).

The data generating process (DGP) is corresponds to Equations (1)-(2), where $X_i \overset{i.i.d}{\sim} \mathcal{N}(0_p, \Sigma_X)$, $\epsilon_i \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$, and $\eta_i \overset{i.i.d}{\sim} \mathcal{N}\left(0, \sigma_\eta^2\right)$. We set $\Sigma_X = I_p$, $\sigma_\epsilon^2 = \sigma_\eta^2 = 1$, $n = 500$, $p = 200$, and $\delta = 0$, which is our parameter of interest. To assess the variable selection properties of Post-Double-Autometrics, we set some β and γ parameters to zero, while all others have non-zero values. The non-zero coefficients are $\beta_j = \gamma_j$ for $j = 1, \ldots, 5$, and are set using a measure of non-centrality as follows: QUID CAS X CORRELES (POSITIVEMENT ET NEGATIVEMENT)?Quel message en tirer par rapport au setup principal ?

$$\beta_j = \psi_j^Y \sqrt{\sigma_\epsilon^2 \left[\Sigma_X^{-1}\right]_{jj}}, \tag{4}$$

$$\gamma_j = \psi_j^D \sqrt{\sigma_\eta^2 \left[\Sigma_X^{-1}\right]_{jj}}, \tag{5}$$

where ψ_j^Y and ψ_j^D are the non-centrality parameters of the relevant control variables X_j in (1)-(2). Instead of choosing arbitrary values for β_j and γ_j, non-centrality parameters allow to calibrate their statistical significance. DANS CE CAS-CI il me semble que $\beta_j = \psi_j^Y$ donc l'argument tombe. PRQ NE PAS AUSSI TESTER AVEC LES DONNEES DE L'APPLICATION ?

Specifically, we choose $\psi_j^Y = \psi_j^D = \{1, 4, 6, 10\}$ for $j = 1, \ldots, 5$ so that these variables have expected t-statistics equal to $\{1, 4, 6, 10\}$ for the null hypotheses $H_0 : \beta_j = 0$ and $H_0 : \gamma_j = 0$. These values also allow to match approximately the population $R^2 = \{0.01, 0.1, 0.3, 0.5\}$ in (1) and the first step of the Post-Double-Lasso considered in Belloni et al. (2014) and Wüthrich and Zhu (2023).

To study the finite sample performance of the different methods mentioned above, we report the average bias and root mean squared error (RMSE) of the treatment effect δ as well as the proportion of retained relevant (potency) and irrelevant (gauge) variables.

3.1 Main results

Table 1 shows bias and RMSE of the treatment effect δ as well as the potency and gauge, for $\psi^Y = \psi^D = 4$, $n = 500$ and $p = 200$. Note that for the Post-Double-Lasso and Post-Double-Autometrics the gauge and potency are computed on the final model, i.e., on the union of X^* and X^{**}. The results correspond to average values of each criterion over 5,000 replications. We consider several choice of λ and α. Values in bold are those corresponding to Figure 1. The results show that:

- **Post-Double-Autometrics outperforms Post-Double-Lasso**

Post-Double-Autometrics with $\alpha = \{0.01, 0.05\}$ outperforms other estimation methods. They provide the best treatment effect estimation (lowest bias and RMSE) and the best variable selection (high potency).

- **Double-selection improves on single-selection**

Both double selection methods provide substantial improvements over single selection methods. Potency is always increased in the double selection approaches, while bias and RMSE are reduced.

- **Autometrics-based methods allow to control the gauge**

Methods based on Autometrics control the proportion of irrelevant selected variables, while selecting most relevant variables. In the last column, the gauge is approximately equal to α for the standard Autometrics method.

The Post-Double-Autometrics method applies Autometrics successively to two different models and the expected gauge of each model is therefore α. If the number of redundant variables is very large (like in our simulation), approximatively $2\alpha\%$ of redundant variables are expected to be retained when considering the union between X^* and X^{**}, which is exactly what we observe in Table 1.

- **Lasso-based methods are very sensitive to the choice of λ**

Methods based on Lasso can perform well or badly depending on the choice of the λ penalty term. With λ^{bya} most of the relevant variables are missed (potency

Model $(\psi^Y = \psi^D = 4)$	Bias	RMSE	Potency	Gauge
Post-Lasso, λ^{bya}	0.1388	0.1458	0.0006	0.0000
Post-Lasso, λ^{bcch}	0.0806	0.0959	0.3449	0.0000
Post-Lasso, λ^{min}	0.0124	0.0526	0.9361	0.0774
Post-Lasso, λ^{1se}	0.0795	0.1094	0.3811	0.0038
Post-Double-Lasso, λ^{bya}	**0.1372**	**0.1443**	**0.0076**	**0.0000**
Post-Double-Lasso, λ^{bcch}	0.0501	0.0717	0.5552	0.0001
Post-Double-Lasso, λ^{min}	0.0007	0.0469	0.9986	0.1413
Post-Double-Lasso, λ^{1se}	0.0200	0.0582	0.8079	0.0076
Autometrics, $\alpha = 0.05$	0.0074	0.0552	0.9176	0.0458
Autometrics, $\alpha = 0.01$	0.0236	0.0708	0.7882	0.0110
Post-Double-Autometrics, $\alpha = 0.05$	**0.0001**	**0.0463**	**0.9953**	**0.0910**
Post-Double-Autometrics, $\alpha = 0.01$	**0.0016**	**0.0460**	**0.9771**	**0.0229**

Table 1: Bias and RMSE of treatment effect δ, proportion of relevant (potency) and irrelevant (gauge) selected variables, for $\psi^Y = \psi^D = 4$, $n = 500$, $p = 200$.

close to 0) and the treatment effect is poorly estimated (large bias and RMSE). With λ^{min}, most of the relevant variables are correctly selected (potency close to 1), the bias and RMSE are quite low, but there is over-selection (higher gauge).

3.2 Additional simulation results

We study the robustness of our previous results with additional simulation experiments. By increasing (resp. decreasing) the non-centrality parameter of relevant control variables, i.e., ψ^Y and ψ^D, they become easier (resp. more difficult) to select. By decreasing the number of variables p, we change the dimension of the model.

SL: JE NE COMPREND PAS PRQ AUTOMETRICS ET POST-DOUBLE-AUTOMETRICS A UNE GAUGE SI FAIBLE. DIMINUER $\psi^Y = \psi^D$ A 1 NE DEVRAIT PAS CHANGER LA GAUGE. SL: MAINTENANT ON COMPREND MIEUX PRQ MAIS NE DEVRAIT-ON PAS AUSSI PRENDRE DES VALEURS COMME 2 OU 3?

Table 2 reports the simulation results for $\psi^Y = \psi^D = 1$. The results show that all methods perform well in the treatment effect estimation (low bias and RMSE), and poorly in variable selection. They fail to select relevant control variables (potency close to 0), but these variables do not have enough effect on Y and D to cause a bias in the estimate of δ. @@ Peut-on dire que les relevant variables sont

Model $(\psi^Y = \psi^D = 1)$	Bias	RMSE	Potency	Gauge
Post-Lasso, λ^{bya}	0.0099	0.0459	0.0000	0.0000
Post-Lasso, λ^{bcch}	0.0099	0.0459	0.0004	0.0000
Post-Lasso, λ^{min}	0.0086	0.0465	0.0549	0.0144
Post-Lasso, λ^{1se}	0.0099	0.0459	0.0004	0.0001
Post-Double-Lasso, λ^{bya}	0.0099	0.0459	0.0000	0.0000
Post-Double-Lasso, λ^{bcch}	0.0099	0.0459	0.0010	0.0001
Post-Double-Lasso, λ^{min}	0.0070	0.0458	0.1098	0.0289
Post-Double-Lasso, λ^{1se}	0.0099	0.0459	0.0009	0.0001
Autometrics, $\alpha = 0.05$	0.0080	0.0470	0.0716	0.0243
Autometrics, $\alpha = 0.01$	0.0094	0.0460	0.0148	0.0033
Post-Double-Autometrics, $\alpha = 0.05$	0.0064	0.0458	0.1377	0.0470
Post-Double-Autometrics, $\alpha = 0.01$	0.0091	0.0457	0.0280	0.0063

Table 2: Bias and RMSE of the treatment effect δ, proportion of relevant (potency) and irrelevant (gauge) selected variables, for $\psi^Y = \psi^D = 1$, $n = 500$, $p = 200$.

vraiment "relevant" et que la selection est mauvaise? (avec un t-stat theorique=1, elles ne sont pas statistiquement significatives)

Table 3 shows simulation results for $\psi^Y = \psi^D = 10$. The results indicate that all double selection methods perform well both in the treatment effect estimation (low bias and RMSE) and in variable selection (high potency, low gauge). The methods select almost all relevant control variables, regardless of the choice of λ and α. In this framework, there is no clear improvement of Post-Double-Autometrics on standard Autometrics. EF: rajouter bcch et 1se fonctionnent mieux (gauge) pour des psi eleves

Table 4 shows simulation results for $\psi^Y = 4$ and $\psi^D = 10$. The double selection methods outperform single selection methods: potency is always much higher and bias is substantially reduced. This is true for methods based on Lasso and Autometrics. In this framework, relevant control variables are less easy to detect in the outcome regression (1), and more easy to detect in treatment regression (2).

Table 5 shows simulation results for $\psi^Y = \psi^D = 6$ and $p = 50$. In the previous results, Post-Double-Lasso with λ^{min} perform better than the others. Here, the gauge is very large, with only 37.93% of irrelevant variables selected on average. @@ Cette table n'est pas tellement convaincante pour dire que λ^{min} fonctionne mal. Dans toutes nos simus, λ^{min} fonctionne plutôt bien pour estimer δ. Peut-on trouver un cas ou il marche mal ? =>Rajouter un cas avec $p >> n$? On peut retirer le cas $p = 50$ et $n = 500$

Model $(\psi^Y = \psi^D = 10)$	Bias	RMSE	Potency	Gauge
Post-Lasso, λ^{bya}	0.4806	0.4840	0.0548	0.0000
Post-Lasso, λ^{bcch}	0.0007	0.0466	0.9991	0.0001
Post-Lasso, λ^{min}	0.0213	0.0531	1	0.1301
Post-Lasso, λ^{1se}	0.0105	0.0482	0.9998	0.0230
Post-Double-Lasso, λ^{bya}	0.0004	0.0453	0.9994	0.0000
Post-Double-Lasso, λ^{bcch}	0.0001	0.0449	1	0.0002
Post-Double-Lasso, λ^{min}	0.0008	0.0470	1	0.1514
Post-Double-Lasso, λ^{1se}	0.0001	0.0453	1	0.0117
Autometrics, $\alpha = 0.05$	0.0000	0.0476	1	0.0471
Autometrics, $\alpha = 0.01$	0.0000	0.0458	1	0.0116
Post-Double-Autometrics, $\alpha = 0.05$	-0.0003	0.0463	1	0.0922
Post-Double-Autometrics, $\alpha = 0.01$	0.0000	0.0451	1	0.0233

Table 3: Bias and RMSE of the treatment effect δ, proportion of relevant (potency) and irrelevant (gauge) selected variables, for $\psi^Y = \psi^D = 10$, $n = 500$, $p = 200$.

Model $(\psi^Y = 4, \psi^D = 10)$	Bias	RMSE	Potency	Gauge
Post-Lasso, λ^{bya}	0.2011	0.2037	0.0000	0.0000
Post-Lasso, λ^{bcch}	0.1675	0.1743	0.1610	0.0000
Post-Lasso, λ^{min}	0.1179	0.1461	0.4177	0.0490
Post-Lasso, λ^{1se}	0.1959	0.2004	0.0234	0.0010
Post-Double-Lasso, λ^{bya}	0.0016	0.0457	0.9940	0.0000
Post-Double-Lasso, λ^{bcch}	0.0000	0.0449	1	0.0001
Post-Double-Lasso, λ^{min}	0.0007	0.0471	1	0.1465
Post-Double-Lasso, λ^{1se}	0.0002	0.0454	1	0.0096
Autometrics, $\alpha = 0.05$	0.0428	0.1017	0.7517	0.0416
Autometrics, $\alpha = 0.01$	0.0955	0.1435	0.4858	0.0085
Post-Double-Autometrics, $\alpha = 0.05$	-0.0002	0.0463	1	0.0915
Post-Double-Autometrics, $\alpha = 0.01$	0.0000	0.0452	1	0.0231

Table 4: Bias and RMSE of the treatment effect δ, proportion of relevant (potency) and irrelevant (gauge) selected variables, for $\psi^Y = 4$, $\psi^D = 10$, $n = 500$, $p = 200$

Model $(\psi^Y = \psi^D = 4)$	Bias	RMSE	Potency	Gauge
Post-Lasso, λ^{bya}	0.1369	0.1442	0.0075	0.0000
Post-Lasso, λ^{bcch}	0.0585	0.0787	0.4944	0.0002
Post-Lasso, λ^{min}	0.0017	0.0470	0.9866	0.2260
Post-Lasso, λ^{1se}	0.0496	0.0882	0.5833	0.0134
Post-Double-Lasso, λ^{bya}	0.1221	0.1307	0.0824	0.0000
Post-Double-Lasso, λ^{bcch}	0.0263	0.0559	0.7336	0.0005
Post-Double-Lasso, λ^{min}	-0.0008	0.0462	0.9999	0.3793
Post-Double-Lasso, λ^{1se}	0.0064	0.0485	0.9181	0.0233
Autometrics, $\alpha = 0.05$	0.0008	0.0466	0.9721	0.0545
Autometrics, $\alpha = 0.01$	0.0075	0.0501	0.9021	0.0118
Post-Double-Autometrics, $\alpha = 0.05$	-0.0009	0.0454	0.9989	0.1065
Post-Double-Autometrics, $\alpha = 0.01$	-0.0004	0.0451	0.9922	0.0236

Table 5: Bias and RMSE of the treatment effect δ, proportion of relevant (potency) and irrelevant (gauge) selected variables, for $\psi^Y = \psi^D = 4$, $n = 500$ and $p = 50$.

Additional results are in Appendix A. @@ Il faut discuter de ce qu'il est utile de garder ?

@@ Message principal: Resultats stables avec Autometrics. Avec lasso, λ_{bcch} est le meilleur pour ψ eleve, mais sinon c'est λ_{min} pour valeurs intermediaires

4 Empirical Application

In this section, we revisit an international economic growth example with our Autometrics-based approaches. In the empirical growth literature, estimating the effect of an initial level of GDP per capita on the growth rates of GDP per capita is a key issue. Indeed, an important finding of the Solow-Swan-Ramsey growth model is the hypothesis of convergence, stating that poorer countries should catch up with richer ones over time as they grow faster. Following this hypothesis, the effect of the initial level of GDP on its growth rate, for a given country, should be negative. To assess whether or not this assumptions holds empirically, we use the data of Barro and Lee (1994). This dataset consists of a panel of 138 countries for the period of 1960 to 1985, and includes $n = 90$ observations and $p = 60$ control variables (X), such as variables measuring the education, trade openness or mortality rate. The dependent variable (Y) corresponds to the national growth rates in GDP per capita, and the variable of interest (D) to the real GDP per capita in 1965. See Table 14 for a complete description of the variables included in the dataset.

Many contributions in the literature investigated the convergence hypothesis. Using a simple bivariate regression model of growth rates on the initial level of GDP, Barro and Sala-i Martin (1995) showed that the convergence hypothesis is rejected. However, this estimate is biased as the model does not take into account other countries' characteristics potentially affecting the growth rate, the initial level of GDP, or both. The literature thus focused on estimating this effect conditional on some characteristics (Levine and Renelt, 1992; Sala-I-Martin, 1997; Sala-i Martin et al., 2004). Given that the number of control variables is close to the number of observations, covariate selection is critical and the findings of these articles were severely criticized as they relied on ad-hoc procedures. More recently, Belloni et al. (2011) and Belloni and Chernozhukov (2011) shed light on the convergence hypothesis using Lasso-based approaches. The authors find a negative and statistically significant effect of the initial level of GDP on its growth rate, supporting the convergence hypothesis of the Solow-Swan-Ramsey growth model. We revisit this application using the different methods presented in Section 2 to show whether or not our results are supporting those of the literature, and ultimately if the convergence hypothesis is verified.

Table 6 displays the results obtained for Post-Lasso and Post-Double-Lasso with $\lambda = \{\lambda^{bya}, \lambda^{bcch}, \lambda^{min}, \lambda^{1se}\}$, Autometrics and Post-Double-Autometrics with $\alpha = \{0.05, 0.01\}$, as well as OLS regressions with and without control variables. We also report the results of an OLS regression involving all control variables with the correction of Cattaneo et al. (2018) to control for high-dimension and heteroscedasticity. The results of the OLS methods whithout control variables suggest that the convergence hypothesis is rejected as the coefficient is not significant, which is consistent with the finding of Barro and Sala-i Martin (1995). This conclusion is robust to the inclusion of all control variables in the model, with or without the standard error correction. Given that these last two models involve many control variables, the estimate is not accurate as it can be seen from the relatively large standard errors and confidence intervals. More reliable results can thus be obtained from the methods based on covariate selection presented in Section 2.

As expected from Post-Lasso and Autometrics, results are inconsistent across λ and α values as both methods suffer from OVB. More interestingly, the results of the Post-Double-Lasso are also inconsistent across λ. Indeed, while the Post-Double-Lasso with λ^{bya}, λ^{1se} and λ^{bcch} provides negative and significant coefficients, it leads to an insignificant coefficient with λ^{min}. The conclusion regarding the convergence hypothesis thus depends on the choice of λ, which is a major issue as the optimal choice of λ is unknown in practice. In contrast, Post-Double-

Model	$\hat{\delta}$	Robust s.e.	90% CI	k^*
OLS without control variables	0.001	0.005	$[-0.007; 0.010]$	0
OLS with all control variables	-0.009	0.032	$[-0.063; 0.044]$	60
OLS, HCK standard errors	-0.009	0.035	$[-0.068; 0.049]$	60
Post-Lasso, λ^{bya}	0.001	0.005	$[-0.007; 0.010]$	0
Post-Lasso, λ^{min}	-0.051	0.014	$[-0.074; -0.028]$	20
Post-Lasso, λ^{1se}	-0.035	0.012	$[-0.054; -0.015]$	10
Post-Lasso, λ^{bcch}	-0.017	0.009	$[-0.033; -0.002]$	7
Post-Double-Lasso, λ^{bya}	-0.042	0.017	$[-0.069; -0.014]$	7
Post-Double-Lasso, λ^{min}	-0.031	0.021	$[-0.065; 0.003]$	23
Post-Double-Lasso, λ^{1se}	-0.056	0.019	$[-0.088; -0.025]$	13
Post-Double-Lasso, λ^{bcch}	-0.050	0.016	$[-0.076; -0.024]$	7
Autometrics, $\alpha = 0.05$	-0.005	0.012	$[-0.025; 0.014]$	12
Autometrics, $\alpha = 0.01$	-0.043	0.011	$[-0.060; -0.025]$	3
Post-Double-Autometrics, $\alpha = 0.05$	-0.026	0.019	$[-0.057; 0.005]$	14
Post-Double-Autometrics, $\alpha = 0.01$	-0.014	0.020	$[-0.046; 0.018]$	11

Table 6: Application on the Growth data: Estimate, robust standard error, 90% confidence interval of the treatment effect δ, and number of selected variables k^*.

Autometrics provides consistent results for both values of α and suggests that the convergence hypothesis is not verified. This conclusion does not support the previous findings of the literature, but is consistent with those of OLS methods.[3] Moreover, the results obtained are more accurate than those of OLS methods with all control variables as it can be seen from the smaller standard errors and confidence intervals.

The number of covariates selected by the Post-Double-Lasso also provides some information regarding its inconsistency. Indeed, while the Post-Double-Lasso with λ^{min} selects at least 10 more control variables than with the other three choices, it is the only example for which the Post-Double-Lasso rejects the hypothesis of convergence. This finding thus suggest that decreasing the penalization parameter increases the number of control variables selected, leading to a modification of the conclusion of the model.[4] It echoes with the recommendation of Wüthrich and Zhu (2023) to increase or decrease the penalization parameter to check the robustness of Post-Double-Lasso results.

[3] Note that our results are also consistent with those of the Post-Double-Lasso with λ^{min}, which has been shown to be the overall best choice of λ in our previous section across several Monte Carlo simulation experiments.

[4] In a similar fashion, Belloni and Chernozhukov (2011) reduce the penalization parameter and show that the number of control variables selected increases. However, their results remain qualitatively similar.

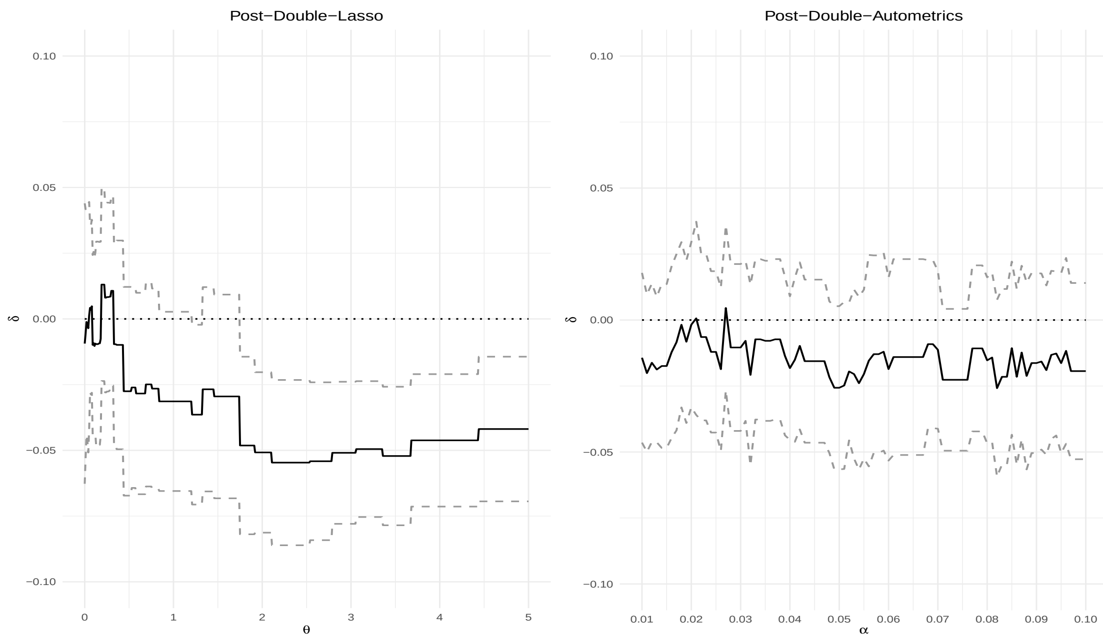

Figure 2: Estimate (solid line) and 90% confidence interval (dashed line) of δ with the Post-Double-Lasso and penalty term $\lambda = \theta\lambda^{min}$ for different values of θ (Left), and with Post-Double-Autometrics for different values of α (Right)

To illustrate the (lack of) robustness of the Post-Double-Lasso, Figure 2 (Left) displays the estimate and 90% confidence interval of δ with λ^{min} multiplied by $\theta \in (0, 5)$.[5] The results suggest that Post-Double-Lasso conclusions are not robust to an increase or a decrease of the penalization parameter as the convergence hypothesis depends on θ. Indeed, while the coefficient is not significant for $\theta < 1.75$, it becomes negative and significant for $\theta \geqslant 1.75$. Qualitatively similar results are observed for λ^{bya}, λ^{1se} and λ^{bcch} in Figures 7-9 in Appendix B. Therefore, the conclusion of the Post-Double-Lasso regarding the convergence hypothesis depends on the choice λ, but also on the value of θ, as a coefficient previously significant can become insignificant when λ is multiplied by θ, and vice-versa.

On the other hand, the results of Post-Double-Autometrics are robust to the choice of α. Figure 2 (Right) displays the estimate and 90% confidence interval of δ for several plausible values of α. The results show that the estimates and confidence intervals are stable across all values of α, and that the convergence hypothesis is never satisfied. In summary, our results do not support the convergence hypothesis of the Solow-Swan-Ramsey growth model, unlike previous findings of the literature, and we recommend to rely on Post-Double-Autometrics rather than Post-Double-

[5]Note that when $\theta = 0$, the results are identical to those of the OLS method with all control variables, and when $\theta = 1$ the results are those displayed in Table 6 .

Lasso for its robustness.

Finally, Table 7 displays the control variables selected by Post-Double-Autometrics and Post-Double-Lasso. Among all control variables, the most important are the black market premium (*bmp1l*), which is a proxy of trade openness, the life expectancy at birth (*lifee065*), the percentage of workers within the population (*worker65*), and the ratio of export to GDP (*ex1*) as they are selected by several methods. Interestingly, the variable measuring the ratio of import to GDP (*im1*) is only selected by the three methods for which the convergence hypothesis is not verified, i.e., Post-Double-Autometrics with $\alpha = \{0.01, 0.05\}$ and Post-Double-Lasso with λ^{min}. This variable is thus probably a major factor to the significance or not of the coefficient of the initial GDP.

To confirm it, Table 8 reports the t-statistics of *im1* in the two steps of Post-Double-Autometrics with $\alpha = \{0.01, 0.05\}$. The results confirm that this variable is relevant, as it is selected in both steps, but also that t-statistics are barely high enough for it.[6] This finding is reminiscent of previous Monte Carlo simulation experiments in which non-centrality coefficients of relevant variables are small in both equations of Y and D. In such setup, we showed that Post-Double-Autometrics accurately selects relevant variables unlike Post-Double-Lasso, as in this application.

These three methods also select other variables in common, such as variables related to the population not working (*pop1565* and *pop6565*) or to education (*nof65*, *seccm65* and *sf65*). The results thus suggest that the most important control variables allowing to reject the convergence hypothesis are mainly related to trade openness, infant life expectancy, proportion of worker in the population, exportation and importation (relative to the GDP), and education.[7] Finally, Post-Double-Lasso with λ^{min} also selects several additional variables compared to Post-Double-Autometrics. These ones are mainly related to education, but also to tariff restriction, political instablity, exchange rate and government expenditures. However, given that Post-Double-Lasso with λ^{min} selects on average many irrelevant variables (see previous section), unlike Post-Double-Autometrics, most of these variables are probably irrelevant and could be dropped from the analysis.

5 Conclusion

@@ To be completed

[6] The correlation of *im1* with Y (0.180) and D (0.364) support these results.

[7] Note that two variables related to education (*pf65*) and infant life expectancy (*mort1*) are only selected by both Post-Double-Autometrics methods.

	PDA, $\alpha = 0.05$	PDA, $\alpha = 0.01$	PDL, λ^{bya}	PDL, λ^{min}	PDL, λ^{1se}	PDL, λ^{bcch}
bmp1l	✓	✓		✓	✓	✓
pf65	✓	✓				
lifee065	✓	✓	✓	✓	✓	✓
mort1	✓	✓				
worker65	✓	✓		✓	✓	
ex1	✓	✓		✓	✓	
im1	✓	✓		✓		
mort65	✓					
gpop1	✓					
no65	✓					
nof65	✓			✓	✓	
pop1565	✓			✓		
secc65	✓					
seccm65	✓			✓		
sf65		✓	✓	✓		✓
pop6565		✓	✓	✓	✓	✓
syr65		✓				
syrf65		✓				
freetar			✓	✓	✓	✓
hm65			✓	✓	✓	✓
hf65				✓		
pm65				✓		
geerec1				✓		
gde1				✓		
govsh41				✓	✓	
gvxdxe41					✓	
humanf65			✓	✓	✓	✓
pinstab1				✓		
seccf65				✓		
teapri65			✓	✓	✓	
teasec65				✓	✓	
xr65				✓		

Note: PDA and PDL refer to Post-Double-Autometrics and Post-Double-Lasso methods, respectively. The symbol "✓" indicates that a variable is selected by the corresponding method.

Table 7: Application on the Growth data: Variables selected by Post-Double-Autometrics and Post-Double Lasso methods

Model	PDA, $\alpha = 0.01$	PDA, $\alpha = 0.05$
Step 1	2.134	2.634
Step 2	-3.846	-2.507

Note: PDA refers to Post-Double-Autometrics.

Table 8: Application on the Growth data: t-statistics of the variable *im1* in Post-Double-Autometrics

References

Barro, R. and Sala-i Martin, X. (1995). *Economic Growth*. Advanced Series in Economics Series. McGraw-Hill.

Barro, R. J. and Lee, J.-W. (1994). Data set for a panel of 138 countries.

Belloni, A., Chen, D., Chernozhukov, V., and Hansen, C. (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6):2369–2429.

Belloni, A. and Chernozhukov, V. (2011). L1-penalized quantile regression in high-dimensional sparse models.

Belloni, A. and Chernozhukov, V. (2013). Least squares after model selection in high-dimensional sparse models. *Bernouilli*, 19(2):521–547.

Belloni, A., Chernozhukov, V., and Hansen, C. (2011). Inference for high-dimensional sparse econometric models. *arXiv preprint arXiv:1201.0220*.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, 81(2):608–650.

Bickel, P. J., Ritov, Y., and Tsybakov, A. B. (2009). Simultaneous analysis of Lasso and Dantzig selector. *The Annals of Statistics*, 37(4):1705 – 1732.

Castle, J., Doornik, J., and Hendry, D. (2023). Robust discovery of regression models. *Econometrics and Statistics*, 26:31–51.

Cattaneo, M. D., Jansson, M., and Newey, W. K. (2018). Inference in linear regression models with many covariates and heteroscedasticity. *Journal of the American Statistical Association*, 113(523):1350–1361.

Doornik, J. A. et al. (2009). Autometrics. *Castle, and Shephard (2009)*, pages 88–121.

Hendry, D., Johansen, S., and Santos, C. (2006). Selecting a regression saturated by indicators. *Unpublished paper, Economics Department, University of Oxford.*

Hendry, D. F. (2000). *Econometrics: alchemy or science?: essays in econometric methodology*. OUP Oxford.

Hendry, D. F. and Doornik, J. A. (2014). *Empirical model discovery and theory evaluation: automatic selection methods in econometrics*. MIT Press.

Levine, R. and Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *The American economic review*, pages 942–963.

Sala-i Martin, X., Doppelhofer, G., and Miller, R. I. (2004). Determinants of long-term growth: A bayesian averaging of classical estimates (bace) approach. *American economic review*, 94(4):813–835.

Sala-I-Martin, X. X. (1997). I just ran two million regressions. *The American Economic Review*, 87(2):178–183.

Sucarrat, G. (2020). User-specified general-to-specific and indicator saturation methods. *R Journal*, 12(2).

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288.

Wüthrich, K. and Zhu, Y. (2023). Omitted variable bias of lasso-based inference methods: A finite sample analysis. *Review of Economics and Statistics*, 105(4):982–997.

A Additional figures and tables for the Monte Carlo simulation study

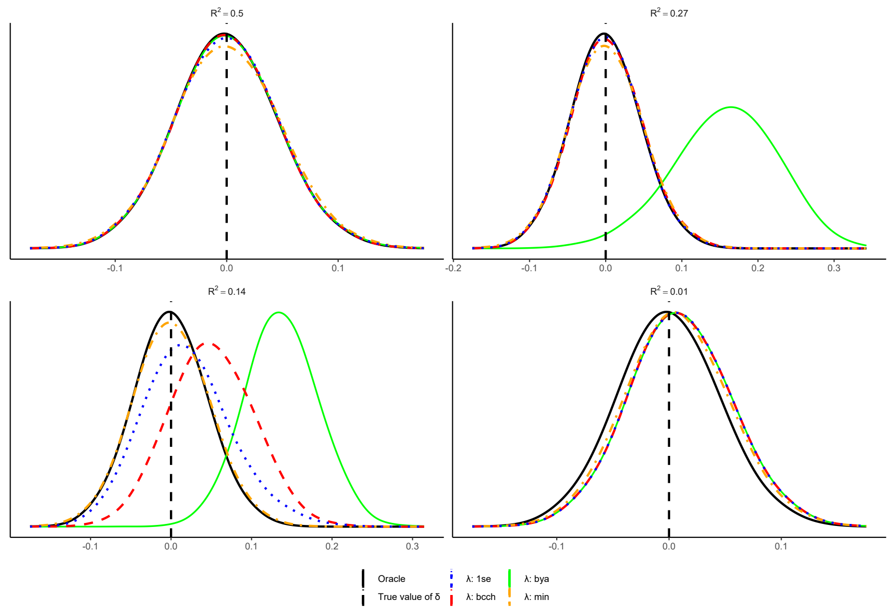

Figure 3: Distributions of the treatment effect $\hat{\delta}$ with Post-Double-Lasso and of the oracle estimator, for $\psi^Y = \psi^D = 4$, $n = 500$ and $p = 200$

@@ A revoir

Figures **??** and **??** display the finite sample distributions of the treatment effect estimator $\hat{\delta}$ obtained, respectively, from Post-Double-Lasso with λ^{bya} and Post-Double-Autometrics with $\alpha = 0.05$.[8] Densities of the oracle estimator are represented in black curves. Results show that Post-Double-Lasso with λ^{bya} does not perform well with intermediate values of the non-centrality parameters $\{4, 6\}$. These results are consistent with those of Wüthrich and Zhu (2023) and illustrate the inability of the Lasso to select relevant control variables when their coefficients are not large enough. In contrast, Post-Double-Autometrics with $\alpha = 0.05$ always perform well, with distributions close to those of the oracle estimator. These results show that Post-Double-Autometrics consistently estimate treatment effect regardless of the coefficients magnitude of control variables, unlike Post-Double-Lasso.

We study the robustness of our previous results by investigating the effect of the choice of λ and α. For Post-Double-Lasso, we consider $\lambda = \{\lambda^{bya}, \lambda^{bcch}, \lambda^{min}, \lambda^{1se}\}$,

[8] We set $\tau = 0.1$ in the Monte Carlo simulation and empirical illustration.

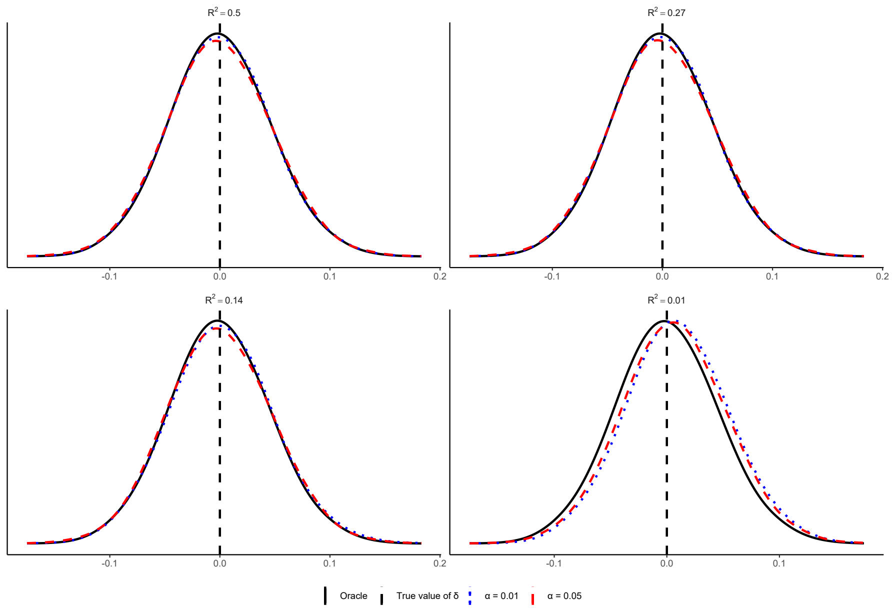

Figure 4: Distributions of the treatment effect $\hat{\delta}$ with Post-Double-Autometrics and of the oracle estimator, for $\psi^Y = \psi^D = 4$, $n = 500$ and $p = 200$

and $\alpha = \{0.01, 0.05\}$ for Post-Double-Autometrics. For each method, we measure the bias, average of the standard errors, and RMSE to evaluate the estimation accuracy of δ, and the potency and gauge, i.e., the percentage of relevant and irrelevant control variables selected in the final model, to assess the consistency of variables selection.

Table **??** reports the average value of each indicator across 5,000 replications. Results show that every method leads to unbiased estimate of δ for very small or large values of control variables coefficients. When $\psi^Y = \psi^D = 1$, Post-Double-Lasso and Post-Double-Autometrics methods select very few relevant control variables (potency close to 0), but these variables do not have enough effect on Y and D to cause a bias in the estimate of δ. For $\psi^Y = \psi^D = 10$, each method selects almost all relevant control variables (potency close to 1), regardless of the choice of λ and α.

Results also suggest that the variable selection inconsistency of Post-Double-Lasso highlighted in Figure **??** for intermediate values of ψ^Y and ψ^D should be tempered. The consistency of Post-Double-Lasso depends on the choice of the regularization parameter λ. Post-Double-Lasso with λ^{min} consistently estimate the treatment effect regardless of the coefficients magnitude of control variables. For intermediate values of non-centrality parameters, the bias is close to 0 and leads to similar values of the RMSE and average standard errors. Similar results are observed for λ^{bcch} or λ^{1se}, even though the estimate of δ is slightly biased (0.0501

Model $(\psi^Y = \psi^D = 6)$	Bias	RMSE	Potency	Gauge
Post-Lasso, λ^{bya}	0.2625	0.2666	0.0125	0.0000
Post-Lasso, λ^{bcch}	0.0253	0.0685	0.8938	0.0001
Post-Lasso, λ^{min}	0.0121	0.0495	0.9996	0.0979
Post-Lasso, λ^{1se}	0.0217	0.0722	0.9230	0.0101
Post-Double-Lasso, λ^{bya}	0.1568	0.1707	0.3920	0.0000
Post-Double-Lasso, λ^{bcch}	0.0017	0.0457	0.9909	0.0002
Post-Double-Lasso, λ^{min}	0.0008	0.0470	1	0.1510
Post-Double-Lasso, λ^{1se}	0.0001	0.0453	0.9998	0.0116
Autometrics, $\alpha = 0.05$	0.0002	0.0480	0.9988	0.0471
Autometrics, $\alpha = 0.01$	0.0012	0.0493	0.9945	0.0116
Post-Double-Autometrics, $\alpha = 0.05$	-0.0002	0.0462	1	0.0921
Post-Double-Autometrics, $\alpha = 0.01$	0.0000	0.0451	1	0.0233

Table 9: Bias and root-MSE of treatment effect estimator $\hat{\delta}$, and proportion of relevant (potency) and irrelevant (gauge) selected variables, for $\psi^Y = \psi^D = 6$, $n = 500$ and $p = 200$

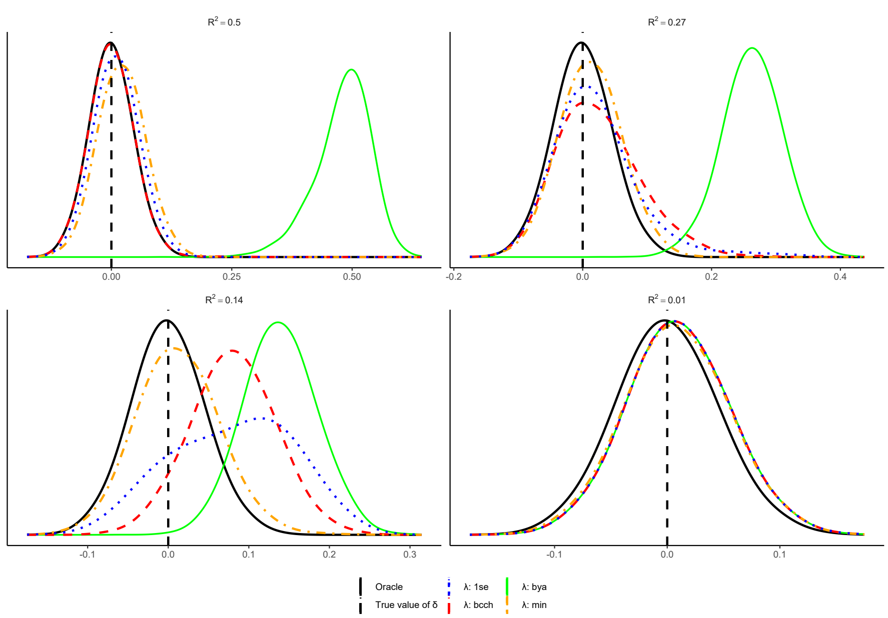

Figure 5: Distributions of the treatment effect $\hat{\delta}$ with Post-Lasso and of the oracle estimator, for $n = 500$ and $p = 200$

Model	Bias	RMSE	Potency	Gauge
		$\psi^Y = \psi^D = 1$		
Post-Lasso, λ^{bya}	0.0099	0.0459	0.0000	0.0000
Post-Lasso, λ^{bcch}	0.0099	0.0459	0.0004	0.0000
Post-Lasso, λ^{min}	0.0086	0.0465	0.0549	0.0144
Post-Lasso, λ^{1se}	0.0099	0.0459	0.0004	0.0001
Post-Double-Lasso, λ^{bya}	0.0099	0.0459	0.0000	0.0000
Post-Double-Lasso, λ^{bcch}	0.0099	0.0459	0.0010	0.0001
Post-Double-Lasso, λ^{min}	0.0070	0.0458	0.1098	0.0289
Post-Double-Lasso, λ^{1se}	0.0099	0.0459	0.0009	0.0001
Autometrics, $\alpha = 0.05$	0.0080	0.0470	0.0716	0.0243
Autometrics, $\alpha = 0.01$	0.0094	0.0460	0.0148	0.0033
Post-Double-Autometrics, $\alpha = 0.05$	0.0064	0.0458	0.1377	0.0470
Post-Double-Autometrics, $\alpha = 0.01$	0.0091	0.0457	0.0280	0.0063
		$\psi^Y = \psi^D = 10$		
Post-Lasso, λ^{bya}	0.4806	0.4840	0.0548	0.0000
Post-Lasso, λ^{bcch}	0.0007	0.0466	0.9991	0.0001
Post-Lasso, λ^{min}	0.0213	0.0531	1	0.1301
Post-Lasso, λ^{1se}	0.0105	0.0482	0.9998	0.0230
Post-Double-Lasso, λ^{bya}	0.0004	0.0453	0.9994	0.0000
Post-Double-Lasso, λ^{bcch}	0.0001	0.0449	1	0.0002
Post-Double-Lasso, λ^{min}	0.0008	0.0470	1	0.1514
Post-Double-Lasso, λ^{1se}	0.0001	0.0453	1	0.0117
Autometrics, $\alpha = 0.05$	0.0000	0.0476	1	0.0471
Autometrics, $\alpha = 0.01$	0.0000	0.0458	1	0.0116
Post-Double-Autometrics, $\alpha = 0.05$	-0.0003	0.0463	1	0.0922
Post-Double-Autometrics, $\alpha = 0.01$	0.0000	0.0451	1	0.0233

Table 10: Bias and root-MSE of treatment effect estimator $\hat{\delta}$, and proportion of relevant (potency) and irrelevant (gauge) selected variables, for $\psi^Y = \psi^D = \{1, 10\}$, $n = 500$ and $p = 200$

Model	Bias	RMSE	Potency	Gauge
$\psi^Y = 4, \psi^D = 6$				
Post-Lasso, λ^{bya}	0.1777	0.1823	0.0001	0.0000
Post-Lasso, λ^{bcch}	0.1135	0.1269	0.3113	0.0000
Post-Lasso, λ^{min}	0.0299	0.0705	0.8490	0.0769
Post-Lasso, λ^{1se}	0.1358	0.1569	0.2085	0.0032
Post-Double-Lasso, λ^{bya}	0.1352	0.1446	0.2364	0.0000
Post-Double-Lasso, λ^{bcch}	0.0073	0.0474	0.9481	0.0001
Post-Double-Lasso, λ^{min}	0.0007	0.0471	1	0.1464
Post-Double-Lasso, λ^{1se}	0.0007	0.0455	0.9956	0.0096
Autometrics, $\alpha = 0.05$	0.0154	0.0675	0.8831	0.0451
Autometrics, $\alpha = 0.01$	0.0444	0.0971	0.7087	0.0103
Post-Double-Autometrics, $\alpha = 0.05$	-0.0002	0.0463	1	0.0915
Post-Double-Autometrics, $\alpha = 0.01$	0.0000	0.0452	0.9999	0.0231
$\psi^Y = 4, \psi^D = 10$				
Post-Lasso, λ^{bya}	0.2011	0.2037	0.0000	0.0000
Post-Lasso, λ^{bcch}	0.1675	0.1743	0.1610	0.0000
Post-Lasso, λ^{min}	0.1179	0.1461	0.4177	0.0490
Post-Lasso, λ^{1se}	0.1959	0.2004	0.0234	0.0010
Post-Double-Lasso, λ^{bya}	0.0016	0.0457	0.9940	0.0000
Post-Double-Lasso, λ^{bcch}	0.0000	0.0449	1	0.0001
Post-Double-Lasso, λ^{min}	0.0007	0.0471	1	0.1465
Post-Double-Lasso, λ^{1se}	0.0002	0.0454	1	0.0096
Autometrics, $\alpha = 0.05$	0.0428	0.1017	0.7517	0.0416
Autometrics, $\alpha = 0.01$	0.0955	0.1435	0.4858	0.0085
Post-Double-Autometrics, $\alpha = 0.05$	-0.0002	0.0463	1	0.0915
Post-Double-Autometrics, $\alpha = 0.01$	0.0000	0.0452	1	0.0231

Table 11: Bias and root-MSE of treatment effect estimator $\hat{\delta}$, and proportion of relevant (potency) and irrelevant (gauge) selected variables, for $\psi^Y = 4$, $\psi^D = \{6, 10\}$, $n = 500$ and $p = 200$

Model	Bias	RMSE	Potency	Gauge
		$\psi^Y = \psi^D = 1$		
Post-Lasso, λ^{bya}	0.0093	0.0459	0.0000	0.0000
Post-Lasso, λ^{bcch}	0.0092	0.0458	0.0020	0.0001
Post-Lasso, λ^{min}	0.0064	0.0460	0.1221	0.0472
Post-Lasso, λ^{1se}	0.0093	0.0459	0.0008	0.0002
Post-Double-Lasso, λ^{bya}	0.0093	0.0459	0.0000	0.0000
Post-Double-Lasso, λ^{bcch}	0.0091	0.0458	0.0043	0.0002
Post-Double-Lasso, λ^{min}	0.0041	0.0453	0.2343	0.0948
Post-Double-Lasso, λ^{1se}	0.0093	0.0459	0.0012	0.0003
Autometrics, $\alpha = 0.05$	0.0063	0.0459	0.1166	0.0375
Autometrics, $\alpha = 0.01$	0.0084	0.0459	0.0289	0.0061
Post-Double-Autometrics, $\alpha = 0.05$	0.0041	0.0453	0.2210	0.0742
Post-Double-Autometrics, $\alpha = 0.01$	0.0076	0.0455	0.0571	0.0115
		$\psi^Y = \psi^D = 10$		
Post-Lasso, λ^{bya}	0.4240	0.4326	0.2012	0.0000
Post-Lasso, λ^{bcch}	-0.0006	0.0452	0.9998	0.0004
Post-Lasso, λ^{min}	0.0033	0.0468	1.0000	0.3239
Post-Lasso, λ^{1se}	0.0028	0.0455	1.0000	0.0552
Post-Double-Lasso, λ^{bya}	-0.0007	0.0451	0.9999	0.0000
Post-Double-Lasso, λ^{bcch}	-0.0008	0.0450	1.0000	0.0007
Post-Double-Lasso, λ^{min}	-0.0009	0.0462	1.0000	0.3871
Post-Double-Lasso, λ^{1se}	-0.0008	0.0451	1.0000	0.0291
Autometrics, $\alpha = 0.05$	-0.0010	0.0456	1.0000	0.0543
Autometrics, $\alpha = 0.01$	-0.0008	0.0452	1.0000	0.0116
Post-Double-Autometrics, $\alpha = 0.05$	-0.0009	0.0454	1.0000	0.1063
Post-Double-Autometrics, $\alpha = 0.01$	-0.0008	0.0451	1.0000	0.0233

Table 12: Bias and root-MSE of treatment effect estimator $\hat{\delta}$, and proportion of relevant (potency) and irrelevant (gauge) selected variables, for $\psi^Y = \psi^D = \{1, 10\}$, $n = 500$ and $p = 50$

Model	Bias	RMSE	Potency	Gauge
		$\psi^Y = \psi^D = 4$		
Post-Lasso, λ^{bya}	0.1369	0.1442	0.0075	0.0000
Post-Lasso, λ^{bcch}	0.0585	0.0787	0.4944	0.0002
Post-Lasso, λ^{min}	0.0017	0.0470	0.9866	0.2260
Post-Lasso, λ^{1se}	0.0496	0.0882	0.5833	0.0134
Post-Double-Lasso, λ^{bya}	0.1221	0.1307	0.0824	0.0000
Post-Double-Lasso, λ^{bcch}	0.0263	0.0559	0.7336	0.0005
Post-Double-Lasso, λ^{min}	-0.0008	0.0462	0.9999	0.3793
Post-Double-Lasso, λ^{1se}	0.0064	0.0485	0.9181	0.0233
Autometrics, $\alpha = 0.05$	0.0008	0.0466	0.9721	0.0545
Autometrics, $\alpha = 0.01$	0.0075	0.0501	0.9021	0.0118
Post-Double-Autometrics, $\alpha = 0.05$	-0.0009	0.0454	0.9989	0.1065
Post-Double-Autometrics, $\alpha = 0.01$	-0.0004	0.0451	0.9922	0.0236
		$\psi^Y = \psi^D = 6$		
Post-Lasso, λ^{bya}	0.2438	0.2503	0.0754	0.0000
Post-Lasso, λ^{bcch}	0.0111	0.0562	0.9475	0.0004
Post-Lasso, λ^{min}	0.0018	0.0464	1.0000	0.2589
Post-Lasso, λ^{1se}	0.0061	0.0536	0.9752	0.0271
Post-Double-Lasso, λ^{bya}	0.0584	0.0866	0.7610	0.0000
Post-Double-Lasso, λ^{bcch}	-0.0005	0.0451	0.9982	0.0007
Post-Double-Lasso, λ^{min}	-0.0008	0.0462	1.0000	0.3874
Post-Double-Lasso, λ^{1se}	-0.0008	0.0451	0.9999	0.0291
Autometrics, $\alpha = 0.05$	-0.0010	0.0456	1.0000	0.0544
Autometrics, $\alpha = 0.01$	-0.0007	0.0453	0.9993	0.0117
Post-Double-Autometrics, $\alpha = 0.05$	-0.0009	0.0453	1.0000	0.1064
Post-Double-Autometrics, $\alpha = 0.01$	-0.0008	0.0450	1.0000	0.0233

Table 13: Bias and root-MSE of treatment effect estimator $\hat{\delta}$, and proportion of relevant (potency) and irrelevant (gauge) selected variables, for $\psi^Y = \psi^D = \{4, 6\}$, $n = 500$ and $p = 50$

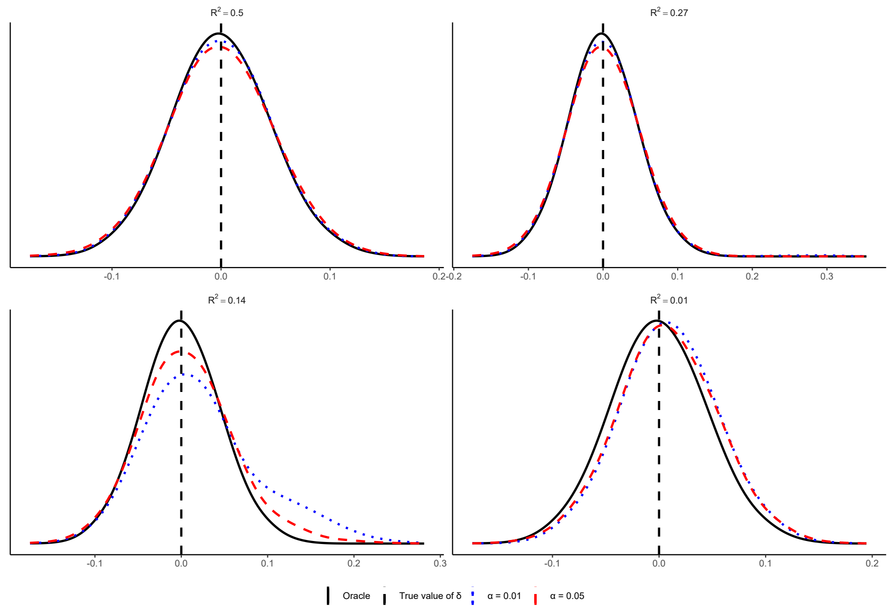

Figure 6: Distributions of the treatment effect $\hat{\delta}$ with Autometrics and of the oracle estimator, for $n = 500$ and $p = 200$

and 0.0200) for $\psi^Y = \psi^D = 4$. In constrats, Post-Double-Lasso with λ^{bya} leads to biased estimates of δ (0.1372 and 0.1568) as illustrated in Figure **??**, and the RMSE is thus larger than the average of the standard errors. These results come from the inconsistent selection of relevant control variables. For $\psi^Y = \psi^D = 6$, Post-Double-Lasso with λ^{bcch}, λ^{min}, and λ^{1se} leads to potency values close to 1, but does not event select half of the relevant control variables with λ^{bya}.

Results also highlight that Post-Double-Autometrics leads to unbiased estimate of δ, regardless of the choice α and of the coefficients magnitude of control variables. Moreover, Post-Double-Autometrics allows to reach high potency values while controling the gauge level, unlike Post-Double-Lasso. When $\psi^Y = \psi^D = 4$, Post-Double-Lasso with λ^{min} reaches potency values close to 1 at the cost of a relatively large number of irrelevant variables selected in the final model. Whereas with λ^{bcch} and λ^{1se}, the gauge of Post-Double-Lasso is relatively low, but the estimate of δ is slightly biased due to a lower potency. In contrasts, the potency of Post-Double-Autometrics is close to 1 for both values of α, and the gauge level remains close to 2α. Therefore, Post-Double-Autometrics allows to consistently estimate treatment effects while controlling the propotion of selected irrelevant variables, regardless of the choice of α, unlike Post-Double-Lasso.

B Additional tables and figures for the empirical illustration

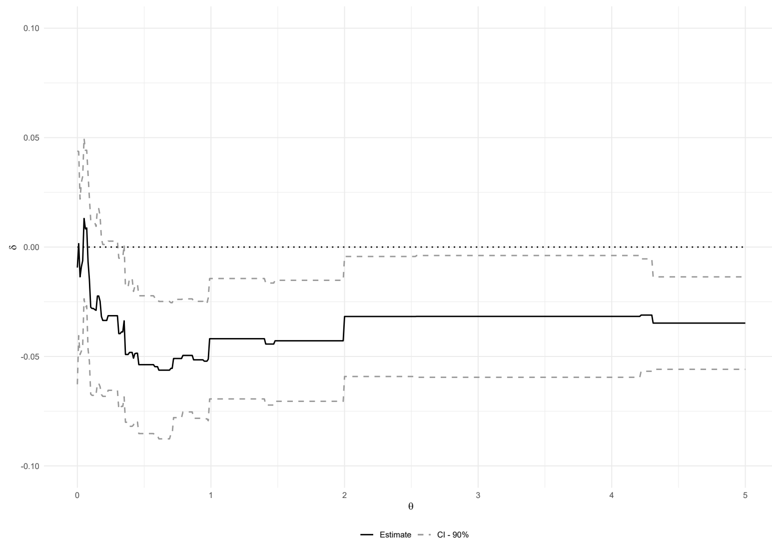

Figure 7: Estimate and 90% confidence interval of δ with the Post-Double-Lasso, λ^{bya}, depending on θ

Status	Name	Description
Dependent variable	Outcome	National growth rates in GDP per capita for the periods 1965-1975
Variable of interest	gdpsh465	Real GDP per capita (1980 international prices) in 1965
Control variable	bmp1l	Black market premium Log (1+BMP)
Control variable	freeop	Measure of "Free trade openness
Control variable	freetar	Measure of tariff restriction
Control variable	h65	Total gross enrollment ratio for higher education in 1965
Control variable	hm65	Male gross enrollment ratio for higher education in 1965
Control variable	hf65	Female gross enrollment ratio for higher education in 1965
Control variable	p65	Total gross enrollment ratio for primary education in 1965
Control variable	pm65	Male gross enrollment ratio for primary education in 1965
Control variable	pf65	Female gross enrollment ratio for primary education in 1965
Control variable	s65	Total gross enrollment ratio for secondary education in 1965
Control variable	sm65	Male gross enrollment ratio for secondary education in 1965
Control variable	sf65	Female gross enrollment ratio for secondary education in 1965
Control variable	fert65	Total fertility rate (children per woman) in 1965
Control variable	mort65	Infant Mortality Rate in 1965
Control variable	lifee065	Life expectancy at age 0 in 1965
Control variable	gpop1	Growth rate of population
Control variable	fert1	Total fertility rate (children per woman)
Control variable	mort1	Infant Mortality Rate (ages 0-1)
Control variable	invsh41	Ratio of real domestic investment (private plus public) to real GDP
Control variable	geetot1	Ratio of total nominal government expenditure on education to nominal GDP
Control variable	geerec1	Ratio of recurring nominal government expenditure on education to nominal GDP
Control variable	gde1	Ratio of nominal government expenditure on defense to nominal GDP
Control variable	govwb1	Ratio of nominal government "consumption" expenditure to nominal GDP (using current local currency)
Control variable	govsh41	Ratio of real government "consumption" expenditure to real GDP (Period average)
Control variable	gvxdxe41	Ratio of real government "consumption" expenditure net of spending on defense and on education to real GDP
Control variable	high65	Percentage of "higher school attained" in the total pop in 1965
Control variable	highm65	Percentage of "higher school attained" in the male pop in 1965
Control variable	highf65	Percentage of "higher school attained" in the female pop in 1965
Control variable	highc65	Percentage of "higher school complete" in the total pop
Control variable	highcm65	Percentage of "higher school complete" in the male pop
Control variable	highcf65	Percentage of "higher school complete" in the female pop
Control variable	human65	Average schooling years in the total population over age 25 in 1965
Control variable	humanm65	Average schooling years in the male population over age 25 in 1965
Control variable	humanf65	Average schooling years in the female population over age 25 in 1965
Control variable	hyr65	Average years of higher schooling in the total population over age 25
Control variable	hyrm65	Average years of higher schooling in the male population over age 25
Control variable	hyrf65	Average years of higher schooling in the female population over age 25
Control variable	no65	Percentage of "no schooling" in the total population
Control variable	nom65	Percentage of "no schooling" in the male population
Control variable	nof65	Percentage of "no schooling" in the female population
Control variable	pinstab1	Measure of political instability
Control variable	pop65	Total Population in 1965
Control variable	worker65	Ratio of total Workers to population
Control variable	pop1565	Population Proportion under 15 in 1965
Control variable	pop6565	Population Proportion over 65 in 1965
Control variable	sec65	Percentage of "secondary school attained" in the total pop in 1965
Control variable	secm65	Percentage of "secondary school attained" in male pop in 1965
Control variable	secf65	Percentage of "secondary school attained" in female pop in 1965
Control variable	secc65	Percentage of "secondary school complete" in the total pop in 1965
Control variable	seccm65	Percentage of "secondary school complete" in the male pop in 1965
Control variable	seccf65	Percentage of "secondary school complete" in female pop in 1965
Control variable	syr65	Average years of secondary schooling in the total population over age 25 in 1965
Control variable	syrm65	Average years of secondary schooling in the male population over age 25 in 1965
Control variable	syrf65	Average years of secondary schooling in the female population over age 25 in 1965
Control variable	teapri65	Pupil/Teacher Ratio in primary school
Control variable	teasec65	Pupil/Teacher Ratio in secondary school
Control variable	ex1	Ratio of export to GDP (in current international prices)
Control variable	im1	Ratio of import to GDP (in current international prices)
Control variable	xr65	Exchange rate (domestic currency per U.S. dollar) in 1965
Control variable	tot1	Terms of trade shock (growth rate of export prices minus growth rate of import prices)

Table 14: Application on the Growth data: Status, name and description of variables

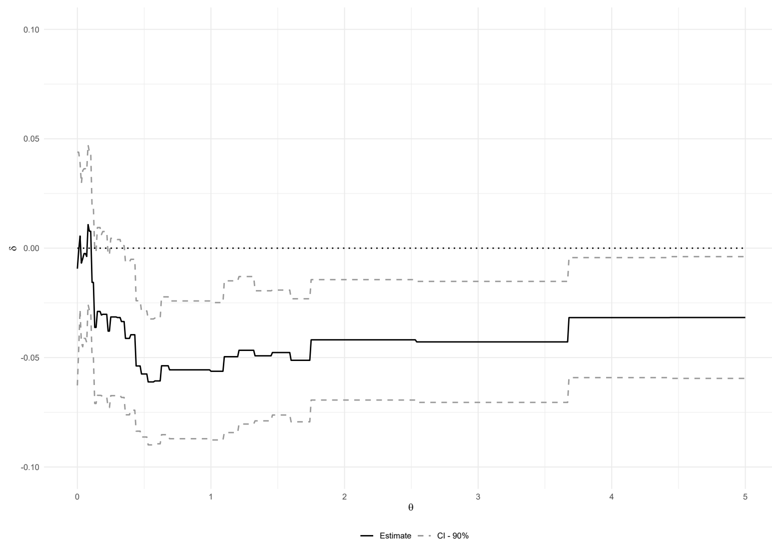

Figure 8: Estimate and 90% confidence interval of δ with the Post-Double-Lasso, λ^{1se}, depending on θ

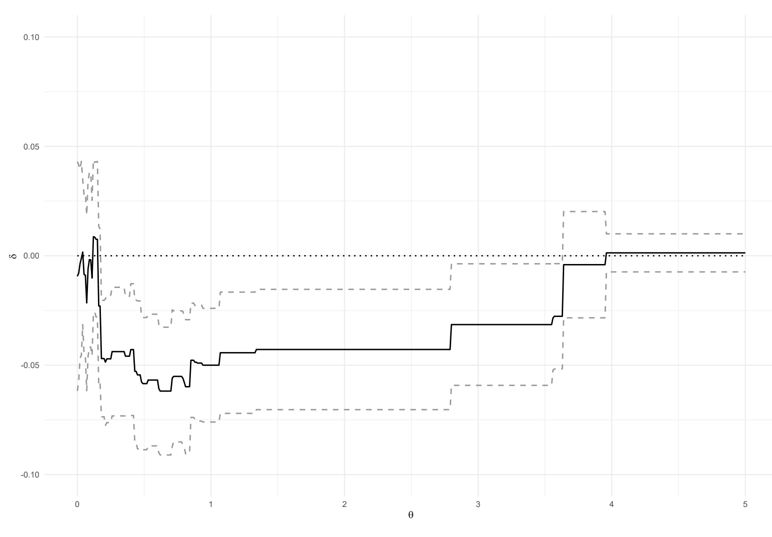

Figure 9: Estimate and 90% confidence interval of δ with the Post-Double-Lasso, λ^{bcch}, depending on θ