# Online Appendix for "$L_2$Boosting for Economic Applications"

YE LUO AND MARTIN SPINDLER

**Abstract.** In this supplement additional material, in particular simulation results and applications, for the paper "$L_2$Boosting for Economic Applications"is presented.

**Key words:** $L_2$Boosting, High-dimensional, instrumental variables, treatment effects, post-selection inference.

In this supplement additional material for the article "$L_2$Boosting for Economic Applications"is presented. First, a brief literature review of using boosting in Economics and Finance is given. The main part shows – by simulations and applications – how boosting can be used for estimation of treatment effects in a setting with very many control variables and with very many potential instrumental variables.

## 1. A Brief Review of the Literature

In this section we give a very brief review of applications of boosting in Economics and Finance. As the strength of machine learning is in prediction and model selection, boosting has been mainly used in these domains. Although boosting has been shown to be a useful approach in many statistical applications, it has been more or less ignored in empirical economics and finance. Some of the few exceptions include the following applications.

Boosting has been used for modeling and predicting volatility,amongst others, by Mittnik, Robinzonov and Spindler (2015), Audrino and Bühlmann (2003) and Audrino and Bühlmann (2009). Audrino and Bühlmann (2009) model stock–index volatility in a GARCH framework and employ boosting for componentwise knot selection in bivariate–spline estimation. Mittnik, Robinzonov and Spindler (2015) also employ a GARCH framework for modeling volatility but allow for a large set of macroeconomic variables which drive volatility. Their data set consists of monthly data with 253 months in total and 40 macroeconomic variables leading to more than 80 predictors (allowing lags). They employ boosting for model estimation and variable selection.

Luo: University of Florida. Spindler: University of Hamburg, Hamburg Business School, Moorweidenstr. 18, 20148 Hamburg, Germany, martin.spindler@uni-hamburg.de.

TABLE 1. Simulation results.

|      | post-Lasso | BA    | post-BA | oBA   |
|------|------------|-------|---------|-------|
| bias | 0.194      | 0.142 | 0.142   | 0.141 |
| RP   | 0.032      | 0.060 | 0.064   | 0.056 |

Bai and Ng (2009) use boosting to select the predictors in factor-augmented autoregressions. Ng (2013) classifies and predicts recessions with boosting.

## 2. IV ESTIMATION WITH MANY INSTRUMENTS

In this section we demonstrate how boosting can be used for IV estimation in a setting with very many instruments.

2.1. **Simulation.** The simulations are based on a simple instrumental variables model data-generating process (DGP):

$$y_i = \beta d_i + e_i, \tag{1}$$

$$d_i = z_i \Pi + v_i, \tag{2}$$

$$(e_i, v_i) \sim N\left(0, \begin{pmatrix} \sigma_e^2 & \sigma_{ev} \\ \sigma_{ev} & \sigma_v^2 \end{pmatrix}\right) \; i.i.d., \tag{3}$$

where $\beta = 1$ is the parameter of interest. The regressors $Z_i = (z_{i1}, \ldots, z_{i100})'$ are normally distributed $N(0, \Sigma_Z)$ with $\mathbb{E}[z_{ih}^2] = \sigma_z^2$ and $Corr(z_{ih}, z_{ij}) = 0.5^{|j-h|}$. $\sigma_z^2$ and $\sigma_e^2$ are set to unit, $Corr(e, v) = 0.6$. $\sigma_v^2 = 1 - \Pi'\Sigma_z \Sigma$ so that the the unconditional variance of the endogenous variable equals 1. The first stage coefficients are set according to $\Pi = C\tilde{\Pi}$. For $\tilde{\Pi}$ we use a sparse design, i.e., $\tilde{\Pi} = (1, \ldots, 1, 0, \ldots, 0)$ with $s$ coordinates equal to one and all other $p - s$ equal to zero. $C$ is set in such a way that we generate target values for the concentration parameter $\mu = \frac{n\Pi'\Sigma_z\Pi}{\sigma_v^2}$ which determines the behavior of IV estimators. We set the sample size equal $n$ to 100, $s = 5$, $p = 100$ and the concentration parameter equal to 180. We estimate the first stage and calculate the first stage predictions with $L_2$Boosting and its variants. The simulation results in Table 1 reveal that boosting has a smaller bias than post-Lasso in this setting. While post-Lasso produces rejection rates below the nominal 5% level, boosting is slightly above.

2.2. **Application.** We consider IV estimation of the effects of federal appellate court decisions regards in eminent domain on macroeconomic outcomes, here in particular the log of the GDP.[1] The structural model is given by

$$y_{ct} = \alpha_c + \alpha_t + \gamma_c t + \beta TakingsLaw_{ct} + W_{ct}'\delta + \varepsilon_{ct}, \tag{4}$$

where $y_{ct}$ is the economic outcome, here log of GDP, for circuit $c$ at time $t$, $TakingsLaw_{ct}$ number of pro-plaintiff appellate takings decisions in circuit $c$ and time $t$, $W_{ct}$ judicial pool characteristics, a dummy for whether there were no cases in that circuit-year,

---

[1]We refer to Belloni et al. (2012) for more information on this application.

TABLE 2. Effect of Federal Appellate Takings Law Decisions on Economic Outcomes.

|  | post-Lasso | BA | post-BA | oBA |
|---|---|---|---|---|
| $\hat{\beta}$ | 0.005 | 0.005 | 0.004 | 0.008 |
| se | 0.012 | 0.007 | 0.006 | 0.006 |

TABLE 3. Simulation results.

|  | post-Lasso | BA | post-BA | oBA |
|---|---|---|---|---|
| bias | 0.082 | 0.121 | 0.136 | 0.121 |
| RP | 0.002 | 0.042 | 0.054 | 0.042 |

and the number of takings appellate decisions; $\alpha_c$, $\alpha_t$ and $\gamma_c t$ denote circuit-specific, time-specific and circuit-specific time trends. The parameter of interest, $\beta$, represents the effect of an additional decision upholding individual property interpreted as more protective individual property rights. The sample size is 312. The analysis of the causal effect of takings law is complicated by potential endogeneity between taking law decisions and economic variables. We employ an instrumental variables strategy that relies on the random assignment of judges to federal appellate panels and uses characteristics of federal circuit court judges (e.g. gender ,race, religion, political affiliation, etc.) as instruments. This gives 138 instruments. We estimate the effect $\beta$ by doing the selection of IVs and estimation the first-stage predicted values $Taki\hat{n}gsLaw_{ct}$ by employing the boosting algorithms introduced before. The results are given in Table 2. The boosting estimates agree with the Lasso estimate but give smaller standard errors. The economic conclusions remain unchanged.

## 3. INFERENCE ON TREATMENT EFFECTS AFTER SELECTION AMONG HIGH-DIMENSIONAL CONTROLS

3.1. **Simulation.** Here we consider the following data-generating process:

$$y_i = d_i \alpha_0 + x_i' \theta_g + \xi_i \tag{5}$$

$$d_i = x_i' \theta_m + \nu_i, \tag{6}$$

where $(\xi_i, \nu_i)' \sim N(0, I_2)$ with $I_2$ the $2 \times 2$ identity matrix, $p = 200$, $x_i \sim N(0, \Sigma)$ with $\Sigma_{kj} 0 = .5^{|j-k|}$. The parameter of interest, $\alpha_0$, is set equal to 0.5 and the sample size is $n = 100$. We consider a design with a decaying sequence of $\theta_m$ and $\theta_g$, namely $1/j^2$ for $j = 1, \ldots, p$. The results in Table 3 show that post-Lasso has a smaller bias than the boosting algorithms, but too small rejection rates (RP) compared to the nominal 5% level. Boosting has rejection rates close to the nominal level.

3.2. **Application.** In Macroeconomics an important questions is how the rates ($Y$) at which economies of different countries grow are related to the initial wealth levels in

each country ($D$) controlling for country's institutional, educational, and other similar characteristics ($W$). The relationship is captured by $beta_1$, the "speed of convergence/divergence", it measures the speed at which poor countries catch up $beta_1 < 0$ or fall behind $beta_1 > 0$ rich countries, after controlling for $W$. Hence the model is given as

$$Y = \beta_1 D + \beta_2' W + \epsilon. \tag{7}$$

For the analysis we use the Barro-Lee data set with 90 countries (observations) and about 60 controls. We estimate the parameter of interest by the double selection method employing both Lasso and $L_2$Boosting for the two selection steps. The double selection method implicitly constructs an orthogonal moment condition. The results are given in Table 4. Here again, the boosting estimates agree with the Lasso estimate and confirm the convergence hypothesis.

TABLE 4. Effect of Initial GDP level om Growth.

|  | post-Lasso | BA | post-BA | oBA |
|---|---|---|---|---|
| $\hat{\beta}$ | $-0.040$ | $-0.042$ | $-0.042$ | $-0.041$ |
| se | 0.015 | 0.012 | 0.012 | 0.013 |

REFERENCES

**Audrino, Francesco, and Peter Bühlmann.** 2003. "Volatility Estimation with Functional Gradient Descent for Very Highdimensional Financial Time Series." *Journal of Computational Finance*, 6(3): 65–89.

**Audrino, Francesco, and Peter Bühlmann.** 2009. "Splines for Financial Volatility." *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 71(3): 655–670.

**Bai, Jushan, and Serena Ng.** 2009. "Boosting Diffusion Indices." *Journal of Applied Econometrics*, 24(4): 607–629.

**Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Christian Hansen.** 2012. "Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain." *Econometrica*, 80: 2369–2429. Arxiv, 2010.

**Mittnik, Stefan, Nikolay Robinzonov, and Martin Spindler.** 2015. "Stock market volatility: Identifying major drivers and the nature of their impact." *Journal of Banking & Finance*, 58: 1 – 14.

**Ng, Serena.** 2013. "BOOSTING RECESSIONS."