

Supplementary Materials. Comparison of predictive performance of different models

On the full sample set, we divide the test set and the prediction set. On the test set, we use the method of 5-fold cross-validation to make hyperparameter selection for different models. Then, on the prediction set, the out-of-sample prediction performance of different models is compared. In most cases, the predictive power of the random forest model is the best, followed by the XGBOOST method (in Table S1, we made **bold** display for the models with better prediction performance). This result is consistent with the research conclusion of Grinsztajn et al. (2022).

For details, please refer to Tables S1 and S2

Table S1 Comparison of out-of-sample prediction performance of different machine learning models

Methods		NDVI	Δy_{t+1}	Δy_{t+2}	Δy_{t+3}	Δy_{t+4}	Δy_{t+5}	Δy_{t+6}	Δy_{t+7}	Δy_{t+8}
RF	MSE	0.0103	0.0151	0.0194	0.0180	0.0159	0.0178	0.0168	0.0170	0.0139
	R^2	0.9715	0.2808	0.5862	0.7373	0.8003	0.8162	0.8495	0.8627	0.8980
MLP	MSE	0.0125	0.0210	0.0469	0.0687	0.0749	0.0806	0.0916	0.0916	0.0970
	R^2	0.9656	-0.0040	-0.0003	-0.0018	0.0585	0.1704	0.1819	0.2604	0.2863
SVR	MSE	0.0423	0.0203	0.0446	0.0561	0.0519	0.0544	0.0546	0.0503	0.0497
	R^2	0.8832	0.0303	0.0494	0.1814	0.3476	0.4398	0.5120	0.5941	0.6344
XGBOOST	MSE	0.0222	0.0134	0.0200	0.0206	0.0216	0.0241	0.0221	0.0223	0.0192
	R^2	0.9387	0.3602	0.5740	0.6992	0.7284	0.7516	0.8028	0.8203	0.8589
OLS	MSE	0.0110	0.0190	0.0271	0.0267	0.0244	0.0246	0.0279	0.0290	0.0311
	R^2	0.9696	0.0911	0.4222	0.6100	0.6938	0.7466	0.7509	0.7661	0.7710

Note: Δy_{t+h} represents the change value of per capita disposable income of rural residents in period h after the implementation of the policy compared with period t. We use **bold** to highlight the model with the best out-of-sample predictive performance in a column.

Table S2 Hyperparameter selection of different models

[illegible]

MLP	activation	ReLU	tanh	tanh	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU
	hidden_layer_sizes	(50,)	(50,)	(50,)	(100,)	(100,)	(50,50)	(100,)	(100,)	(100,)
	learning_rate	constant	constant	constant	constant	constant	constant	constant	constant	constant
SVR	kernel	linear	poly	linear	linear	linear	linear	linear	linear	linear
	C	0.1	10	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	epsilon	0.2	0.2	0.5	0.2	0.5	0.5	0.5	0.5	0.5
XGBOOST	learning_rate	0.2	0.01	0.1	0.2	0.2	0.2	0.2	0.2	0.2
	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000

Reference

(1) Grinsztajn, L.; Oyallon, E.; Varoquaux, G. Why do tree-based models still outperform deep learning on typical tabular data? *Advances in neural information processing systems* **2022**, *35*, 507-520.