Journal of Econometrics and Statistics
Vol. 6, Issue 1, 2026, pp. 1-11

Evaluating Instrumental Variable Estimators in High-Dimensional Panel Data: A Simulation Study of 2SLS, Lasso-IV, and Double Machine Learning IV (DML-IV)

Dr. J. Prabhakara Naik[a], and Sreedevi P.N.[b]

[a] *Associate Professor,* [b]*ResearchScholar*
[a,b] *Department of Statistics, Ramanujan School of Mathematical Sciences, Pondicherry University, Puducherry.*

ABSTRACT
This study evaluates the performance of machine learning-based IV methods—specifically Double Machine Learning IV (DML-IV) with Mundlak adjustments and Lasso-IV—compared to traditional 2SLS in panel data settings where the set of valid instruments is unknown. Using Monte Carlo simulations with balanced panel data (n = 5000, T = 5) and varying the number of potential instruments (p = 5, 10, 50, 100), it was simulated data-generating processes where only a subset of instruments is truly relevant for treatment assignment. Each estimator is assessed based on bias and root mean squared error (RMSE) over 100 replications to ensure stable performance metrics. Results shows that 2SLS performs well when the true strong instruments are known and included, recovering unbiased treatment effects with low RMSE. But it is not applied when instrument relevance is unknown or when instrument selection is required. Lasso-IV, which uses L1-penalized regression to select relevant instruments, exhibits upward bias and increasing RMSE as p increases, reflecting challenges associated with over-selecting weak or invalid instruments in high dimensions. In DML-IV with Mundlak adjustments achieves low bias and stable RMSE across all instrument dimensions, outperforming Lasso-IV consistently.
Keywords: 2SLS, Lasso-IV, DML-IV,Simulation, Panel data

1. Introduction

In applied econometrics and causal inference, the estimation of causal effects remains one of the most central and challenging tasks . A frequent difficulty arises when the regressors of interest are endogenous, meaning they are correlated with unobserved factors that also influence the outcome variable. Ignoring such endogeneity leads to biased and inconsistent parameter estimates, which can distort inference and misguide policy conclusions. The instrumental variables (IV) method, particularly the classical two-stage least squares estimator (2SLS), is a widely used approach to address this issue. However, the validity of IV estimation is critically dependent on the availability of instruments—variables correlated with the endogenous regressor but uncorrelated with the unobserved determinants of the outcome. In practical situa-

CONTACT Author[a]. Email: drnaik.sta@pondiuni.ac.in

Article History
Received: 20 December 2025; Revised: 18 January 2026; Accepted: 23 January 2026; Published: 30 March 2026

To cite this paper

Dr. J. Prabhakar Naika, & Sreedevi P.N. (2026). Evaluating Instrumental Variable Estimators in High-Dimensional Panel Data: A Simulation Study of 2SLS, Lasso-IV, and Double Machine Learning IV (DML-IV). *Journal of Econometrics and Statistics.* 6(1), 1-11.

tion, researchers often face considerable uncertainty about which instruments satisfy these requirements, and the problem becomes increasingly acute as the number of potential instruments grows.

Relevance and endogeneity are the two central conditions for the validity of the IV. Relevance requires that the candidate instruments be sufficiently correlated with the endogenous regressor, while endogeneity demands that they influence the dependent variable only through this channel. In empirical studies, verifying endogeneity is particularly difficult, and even relevance poses challenges when multiple weak instruments are available. 2SLS is consistent and efficient When the set of valid instruments is known and strong. However, in many modern empirical settings, such knowledge is rarely available prior to publication. Researchers instead operate with a potentially large set of instruments, only a subset of which may be valid or relevant. Traditional econometric methods quickly deteriorate in such contexts, giving rise to weak instrument bias, inflated variance, or inconsistent point estimates.

The rise of high-dimensional data has intensified this problem. In economics, demography, and related disciplines, researchers frequently encounter settings where administrative records, survey questionnaires, or large-scale observational datasets provide dozens or even hundreds of potential instruments. Classical IV methods struggle in these high-dimensional environments, primarily because they were not designed to handle situations where the number of instruments approaches or exceeds the available sample size. Incorrect instrument selection not only leads to reduced efficiency, but can invalidate the entire causal analysis. Consequently, methodological innovations that can select suitable instruments and deliver reliable inference are urgently needed.

Recent advances in machine learning have provided promising tools to address the instrument selection problem under high-dimensional settings. Among the most influential of these is the Lasso-IV estimator, which applies the L1-penalized lasso regression to select from a large pool of potential instruments. By shrinking coefficients of weaker or irrelevant instruments toward zero, Lasso-IV helps avoid overfitting and reduces weak instrument bias, making it valuable in applied research with many possible instruments. Nevertheless, Lasso-IV is not without shortcomings. Because it relies solely on the selected set of instruments, it fails to fully account for the uncertainty inherent in the variable selection process. As a result, Lasso-IV often exhibits upward bias in finite samples, especially as the number of instruments grows large or when many instruments are only weakly correlated with the endogenous regressor.

An alternative and increasingly popular method is Double Machine Learning IV (DML-IV). Recent developments at the frontier of econometrics and statistics have witnessed the emergence of double or de-biased machine learning (DML) as a powerful tool for drawing valid causal inference in complex, high-dimensional datasets (Clarke and Polselli (2025)). DML-IV employs a combination of orthogonalized moment conditions and cross-fitting—a resampling strategy that separates training and estimation samples—to achieve unbiased estimation of causal parameters even in high-dimensional contexts. The critical feature of DML-IV is its use of Neyman orthogonality, which ensures that estimation errors in the nuisance parameters (such as conditional expectations of the outcome or treatment) have limited impact on the parameter of interest. By leveraging machine learning algorithms to flexibly estimate these nuisance functions while preserving valid inference on the main coefficient, DML-IV offers a more robust alternative than penalized regressions like Lasso-IV.

Despite these advantages, most applications of DML-IV to date have focused on cross-sectional data structures. Panel data, where repeated observations on the same units are collected over time, remain relatively underexplored in this context. Yet panel data are prevalent in empirical economics and demography, offering the advantage of controlling for unobserved heterogeneity that may be constant over time. At the same time, they introduce unique challenges. In particular, if unobserved time-invariant characteristics are correlated with both the endogenous regressors and the instruments, failing to account for such heterogeneity can bias causal estimates and invalidate identification strategies. Traditional econometric methods for panel data, such as fixed effects or correlated random effects models, are designed precisely to address this issue, but they are not easily integrated into machine learning-based IV approaches.

One promising approach to mitigate these challenges in the context of DML-IV is the Mundlak (1978) adjustment. The Mundlak method introduces unit-level means of time-varying regressors (in this case, instruments) as additional controls. By including these means in the model, one effectively allows for correlation between the random effects and the regressors, without requiring a full-fledged fixed effects transformation. This method preserves efficiency while providing robustness against bias from unobserved time-invariant heterogeneity. Applying Mundlak adjustments within the DML framework therefore represents a natural extension to adapt double machine learning methods for panel data applications.

This paper systematically compares the performance of three leading estimators for causal inference in panel data settings: **2SLS**, **Lasso-IV**, and **Double Machine Learning IV (DML-IV)**. The DML-IV procedure implemented in this study is directly adopted from Clarke and Polselli (2025), who develop novel DML methods for panel data with fixed effects, leveraging machine learning algorithms to flexibly estimate high-dimensional and nonlinear nuisance functions. Their approach extends classical panel estimators—correlated random effects (Mundlak), within-group, and first-difference transformations—to the DML framework, enabling robust inference even when instrument validity is uncertain and the instrument set is large or complex. We focus on key finite-sample properties of bias and root mean squared error (RMSE) across varying numbers of potential instruments ($p = 5, 10, 50, 100$). By incorporating both strong and weak instruments into the simulated data-generating process, our experimental design mirrors empirically realistic scenarios where researchers face uncertainty about instrument validity and strength. The specific aim is to benchmark how these methods perform when the set of valid instruments is unknown, and to clarify the conditions under which each estimator is likely to perform best in practice.

The paper is organized as follows. Section 1 provides an introduction and motivation for the econometric problem under study. Section 2 outlines the model specification, including the econometric framework and identification strategy. Section 3 describes the data generating process (DGP) used in the Monte Carlo simulations, detailing key features and contributions of the experimental design. Section 4 presents the simulation results, comparing the performance of alternative estimators. Finally, Section 5 concludes with a summary of main findings and their implications for empirical research.

2. Model Specification

This section describes the procedures for the three main approaches compared in this study: **Two-Stage Least Squares (2SLS)**, **Lasso-IV**, and **Double Machine Learning IV (DML-IV)**. Each method addresses endogeneity and instrument selection in panel data.

2.1. *Two-Stage Least Squares (2SLS)*

Two-Stage Least Squares(2SLS) is a classical instrumental variable technique used to obtain consistent estimates when regressors are endogenous. The procedure consists of two main steps:

(1) **First Stage:** The endogenous regressor(s) (e.g., D_{it}) is regressed on all available instruments and exogenous controls. This produces fitted values that capture the variation in the endogenous variable explained by the instruments.
(2) **Second Stage:** The outcome variable (Y_{it}) is regressed on the fitted values from the first stage and the exogenous controls. The coefficient on the fitted regressor provides the causal effect estimate.

The validity of 2SLS relies on two key assumptions:

- **Instrument relevance:** Instruments must be sufficiently correlated with the endogenous regressor.
- **Instrument endogeneity:** Instruments must be uncorrelated with the error term in the outcome equation.

In our simulation, 2SLS is implemented using the first three instruments, representing a scenario where strong instruments can be confidently selected.

2.2. *Lasso-IV*

Lasso-IV extends the IV framework to high-dimensional settings, where the number of potential instruments may be large and only a subset are relevant. The procedure involves:

(1) **Instrument Selection via Lasso:** Lasso (L1-penalized regression) is used to select the most relevant instruments by shrinking coefficients of less informative instruments toward zero. The endogenous variable is regressed on all candidate instruments and controls using Lasso, and instruments with non-zero coefficients are selected.
(2) **First Stage:** The selected instruments and controls are used to predict the endogenous variable.
(3) **Second Stage:** The outcome variable is regressed on the predicted values from the Lasso-based first stage and controls to obtain the structural parameter estimates.

Lasso-IV is particularly effective for variable selection in high-dimensional instrument sets, reducing the risk of weak instrument bias. However, it may induce bias if many weak or irrelevant instruments are present, and does not fully account for selection uncertainty.

2.3. *Double Machine Learning IV (DML-IV)*

Recent advances in causal inference have leveraged machine learning algorithms to flexibly estimate high-dimensional nuisance functions. However, most double machine learning (DML) frameworks assume cross-sectional data, while panel data introduces unobserved heterogeneity and dependence across time. Our study adapts DML to static panel models with fixed effects, addressing these challenges and providing practical guidance for instrument selection and estimator robustness in high-dimensional settings. We consider the partially linear panel model:

$$Y_{it} = \theta_0 D_{it} + g_0(X_{it}) + \alpha_i + \varepsilon_{it}, \tag{1}$$

where Y_{it} is the outcome, D_{it} is the endogenous regressor, X_{it} are observed controls, α_i are unit fixed effects, and ε_{it} is the error term. Endogeneity arises if $\mathrm{Cov}(D_{it}, \varepsilon_{it}) \neq 0$. Instrumental variables Z_{it} are assumed to satisfy relevance and exogeneity conditions. The goal is to estimate the causal effect θ_0 in the presence of high-dimensional or non-linear confounding.

To address potential correlation between X_{it} and α_i, we use the Mundlak (1978) device, also known as the correlated random effects (CRE) approach. This involves augmenting the model by including the unit-level means of the time-varying co-variates as additional regressors:

$$Y_{it} = \theta_0 D_{it} + g_0(X_{it}) + h_0(\bar{X}_i) + \varepsilon_{it}, \tag{2}$$

where $\bar{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_{it}$ is the mean of X_{it} for unit i. This adjustment allows for flexible control of unobserved heterogeneity and is algebraically equivalent to the fixed effects estimator in linear models(Arkhangelsky, D., Imbens, G. W. (2023),Wooldridge, J. M. (2021).) Recent work by Clarke and Polselli, shows that the Mundlak approach can be combined with double machine learning (DML) methods to robustly estimate causal effects in high-dimensional panel data, even when the set of valid instruments is large or uncertain. The Mundlak adjustment is also extensible to nonlinear models and multi-way fixed effects by including group averages of covariates or their nonlinear transformations(Arkhangelsky, D., Imbens, G. W. (2023),Wooldridge, J. M. (2021).)

Next we summarizes the double machine learning (DML) estimation procedures for panel data with fixed effects, as developed by Clarke and Polselli (2025). The procedures extend classical panel estimators—correlated random effects (CRE/Mundlak), within-group (WG), and first-difference (FD)—to nonlinear and high-dimensional settings using machine learning for nuisance function estimation.

2.3.1. Model Setup

We consider the partially linear panel regression model (PLPR):

$$Y_{it} = D_{it}\theta_0 + g_1(X_{it}) + \alpha_i + U_{it}, \tag{3}$$

where Y_{it} is the outcome, D_{it} the treatment, X_{it} the controls, α_i the fixed effect, and U_{it} the error. The target is the average treatment effect θ_0.

2.3.2. Approaches to Handling Fixed Effects

- **Correlated Random Effects (CRE/Mundlak):** Augment the model with unit-level means $\bar{X}_i$ (Mundlak adjustment):

$$Y_{it} = D_{it}\theta_0 + g_1(X_{it}) + h_1(\bar{X}_i) + a_i + U_{it} \tag{4}$$

Nuisance functions $l_1(X_{it}, \bar{X}_i)$ and $m_1(X_{it}, \bar{X}_i)$ are learned using machine learning.

- **Data Transformations:**
  - **First-Difference (FD):** Remove fixed effects by differencing: $Q(Y_{it}) = Y_{it} - Y_{it-1}$.
  - **Within-Group (WG):** Remove fixed effects by demeaning: $Q(Y_{it}) = Y_{it} - \bar{Y}_i$.

  Nuisance functions are learned on the transformed data.

2.3.3. Learning Nuisance Functions

Flexible machine learning algorithms (Lasso, Random Forest, Boosting, etc.) are used to estimate conditional means, here we are using Lasso algorithm for estimating conditional means:

- For CRE: $\hat{l}_1(X_{it}, \bar{X}_i)$ and $\hat{m}_1(X_{it}, \bar{X}_i)$.
- For FD/WG: $\hat{l}_1(Q(X_{it}))$ and $\hat{m}_1(Q(X_{it}))$.

Ensemble learning (e.g., super learner) is recommended for robustness.

2.3.4. Cross-Fitting and Sample Splitting

Block k-fold cross-fitting is used: assign all time periods for each unit to the same fold to respect panel structure. For each fold:

(1) Train ML models for nuisance functions on the complement.
(2) Predict nuisance functions and construct orthogonalized scores on the held-out fold.

2.3.5. Neyman Orthogonal Score Construction

Construct the orthogonal score for θ_0:

- For CRE: $\psi^{\perp}(W_i; \theta, \eta) = V_i^{\perp} \Sigma_0^{-1}(X_i) r_i$
- For FD/WG: $\psi^{\perp}(W_i; \theta, \eta) = Q(V_{it})(Q(Y_{it}) - Q(V_{it})\theta - Q(l_1(X_{it})))$

V_{it} is the residualized treatment: $V_{it} = D_{it} - \hat{m}_1(\cdot)$.

2.3.6. Estimation and Inference

Solve the cross-fitted moment equation:

$$\frac{1}{K} \sum_{k=1}^{K} \frac{1}{N_k} \sum_{i \in \mathcal{W}_k} \psi^{\perp}(W_i; \hat{\theta}_{DML}, \hat{\eta}_k) = 0 \tag{5}$$

A closed-form solution for $\hat{\theta}_{DML}$ is available for linear scores. Variance is estimated using cluster-robust (unit-level) standard errors.

3. Data Generating Process (DGP) in the Monte Carlo Simulation

Our study uses a **Monte Carlo simulation** to compare the performance of 2SLS, Lasso-IV, and DML-IV estimators in panel data with Mundlak adjustments, across different numbers of instruments ($p = 5, 10, 50, 100$). In a Monte Carlo simulation, we repeatedly generate synthetic data from a specified DGP, apply estimation methods, and summarize their performance over many replications.

3.1. *Panel Structure*

- $n = 5000$ total observations, $T = 5$ time periods, $n_{id} = 1000$ units (since n/T).
- Each unit i is observed over T periods.

3.2. *Instrument Variables*

- For each simulation, p instruments $Z_{it,1}, \ldots, Z_{it,p}$ are generated as independent standard normal random variables.
- The strength of each instrument is set by the vector π:
  - For $p \geq 5$: $\pi = (0.8, 0.6, 0.4, 0.1, 0.05, 0, \ldots, 0)$ (first 5 are strong/weak, rest are irrelevant).
  - For $p < 5$: $\pi = (0.8, 0.6, 0.4)[1 : p]$.

Table 1. Instrument strengths in the simulation DGP

Instrument	π_j value	Role in simulation
Z_1	0.8	Strong, valid
Z_2	0.6	Strong, valid
Z_3	0.4	Moderate, valid
Z_4	0.1	Weak, valid
Z_5	0.05	Very weak, valid
Z_6–Z_p	0	Irrelevant

3.3. *Error Terms*

- Errors (ν, ε) are drawn from a bivariate normal distribution with mean zero and covariance matrix $\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$, inducing endogeneity between the endogenous regressor and the error in the outcome equation.

3.4. *Endogenous Regressor and Outcome*

- $X_{it} = Z_{it}\pi + \nu_{it}$
- $Y_{it} = \beta_{true}X_{it} + \varepsilon_{it}$, with $\beta_{true} = 1$

3.5. *Simulation Loop*

- For each value of p, the simulation is repeated $n_{sim} = 100$ times.
- In each replication, new data are generated, and all three estimators (2SLS, Lasso-IV, DML-IV) are applied.
- Results (mean, bias, RMSE, etc.) are stored for each estimator and scenario.

4. Results

This chapter presents the outcomes of our Monte Carlo simulation study comparing the performance of **2SLS**, **Lasso-IV**, and **DML-IV** estimators in panel data settings with Mundlak adjustments, across varying numbers of instruments ($p = 5, 10, 50, 100$). The main performance metrics are mean estimate, bias, and root mean squared error (RMSE) for each estimator and scenario.

Our simulation was designed to reflect realistic empirical scenarios where the set of available instruments includes both strong and weak candidates, and where the true valid instruments are not always known in advance. This allows us to directly compare the strengths and limitations of 2SLS and DML-IV under these conditions.

Table 2. Simulation results: Mean, Bias, and RMSE for each estimator and instrument dimension p

p	Metric	2SLS	Lasso-IV	DML-IV
5	Mean	1.0004	1.0060	1.0128
	Bias	0.0004	0.0060	0.0128
	RMSE	0.0130	0.0142	0.0202
10	Mean	0.9976	1.0093	1.0117
	Bias	-0.0024	0.0093	0.0117
	RMSE	0.0138	0.0180	0.0180
50	Mean	1.0018	1.0329	1.0188
	Bias	0.0018	0.0329	0.0188
	RMSE	0.0139	0.0362	0.0241
100	Mean	0.9962	1.0352	1.0186
	Bias	-0.0038	0.0352	0.0186
	RMSE	0.0134	0.0377	0.0234

2SLS: Optimal with Known, Strong Instruments

- When the researcher can confidently select strong, valid instruments, 2SLS provides unbiased and efficient estimates. In our simulation, using the first few (strong) instruments for 2SLS yields nearly unbiased estimates and the lowest RMSE across all scenarios, even as the number of potential instruments increases.
- While our study emphasizes 2SLS performance under strong instrument validity, the reference paper by Clarke Polselli (2025) focuses on OLS as a baseline and on various DML methods for panel data. Their results show that OLS performs well only under simple linear or low-confounding scenarios, but becomes inadequate with increased confounding or nonlinearity. Thus, our 2SLS findings expand the methodological landscape by highlighting cases where classical IV approaches succeed—an issue not explored in their simulations. Both studies consistently underscore that robust estimation strategies are essential as empirical complexity increases.

Lasso-IV: Sensitive to Instrument Strength and Selection

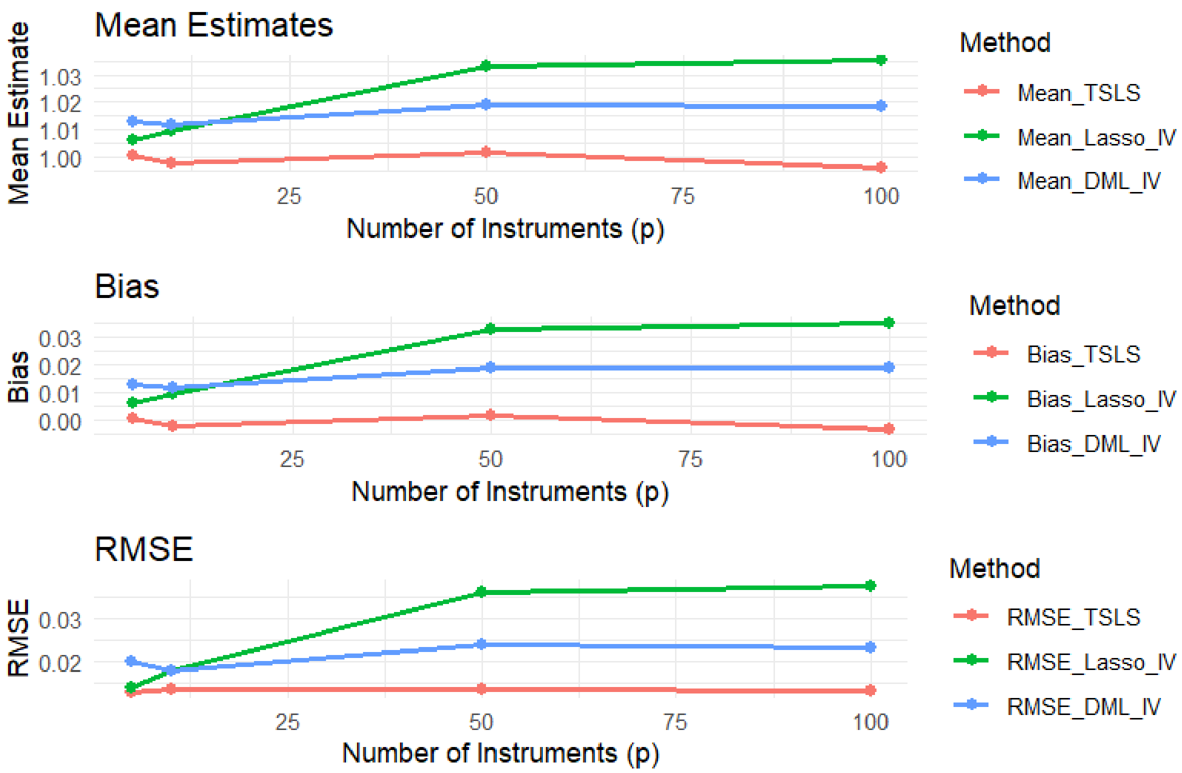

Figure 1. Comparsion plot for TSLS, Lasso-IV, and DML-IV

- Lasso-IV uses penalized regression to select from a large pool of instruments. In our results, Lasso-IV performs comparably to 2SLS when p is small and the valid instruments are strong, but its bias and RMSE increase substantially as the number of instruments grows and the instrument set becomes dominated by weak or irrelevant variables.
- This pattern is also observed in Clarke & Polselli (2025), who note that Lasso-IV can suffer from upward bias and increased variance in high-dimensional settings, especially when many instruments are weak or only weakly correlated with the endogenous regressor.

DML-IV: Robustness in High-Dimensional and Uncertain Settings

- DML-IV, especially with Mundlak adjustments, is designed to handle high-dimensional settings and mitigate the impact of weak or irrelevant instruments. In our simulation, DML-IV consistently outperforms Lasso-IV and remains close to 2SLS in terms of bias and RMSE, even as the number of instruments grows and the instrument set includes both strong and weak variables.
- Clarke & Polselli (2025) similarly find that DML-IV with flexible machine learning learners and panel adjustments provides robust inference in high-dimensional and empirically realistic scenarios. They recommend DML-IV as a more reliable alternative to Lasso-IV when instrument selection is uncertain or the instrument set is large and potentially weak.

4.1. *Results by Instrument Dimension*

For $p = 5$**:** All estimators are nearly unbiased, with 2SLS showing the lowest RMSE. DML-IV and Lasso-IV perform similarly, with slightly higher bias and RMSE than 2SLS.

For $p = 10$**:** 2SLS remains nearly unbiased and has the lowest RMSE. Lasso-IV and DML-IV show slightly increased bias and RMSE, but DML-IV outperforms Lasso-IV.

For $p = 50$ **and** $p = 100$**:** As the number of instruments increases, Lasso-IV's bias and RMSE increase substantially. DML-IV with Mundlak adjustment maintains lower bias and RMSE than Lasso-IV, and is only slightly less accurate than 2SLS. 2SLS remains robust when valid instruments are known. These patterns are consistent with the simulation findings reported by Clarke & Polselli (2025), further validating the robustness of DML-IV in high-dimensional panel data settings.

4.2. *Summary and Implications*

Overall, our results confirm and extend the findings of Clarke & Polselli (2025):

- 2SLS is optimal when the set of valid, strong instruments is known and can be confidently selected.
- DML-IV with Mundlak adjustment is more robust and reliable when instrument selection is uncertain or the instrument set is high-dimensional and potentially weak.
- Lasso-IV, while useful for variable selection, can suffer from bias and increased variance in large, weak-instrument settings.

These insights provide practical guidance for applied researchers working with panel data and high-dimensional instruments, highlighting the importance of robust estimation strategies in modern empirical analysis.

5. Conclusion

This study systematically compared the performance of 2SLS, Lasso-IV, and DML-IV estimators in panel data settings with Mundlak adjustments, across a range of instrument dimensions ($p = 5, 10, 50, 100$). Our simulation design reflected realistic empirical scenarios, including both strong and weak instruments, and cases where the set of valid instruments is not known in advance.

The results demonstrate that **2SLS is optimal when valid, strong instruments are known and can be confidently selected**. In these scenarios, 2SLS consistently produced nearly unbiased estimates and the lowest RMSE, confirming theoretical expectations and findings from previous literature. However, in practice, researchers often face high-dimensional instrument sets with uncertain validity and many weak or irrelevant candidates. In such cases, **DML-IV with Mundlak adjustment proved to be more robust**, maintaining low bias and RMSE even as the number of instruments increased and outperforming Lasso-IV.

Comparison with Clarke & Polselli (2025): Our findings closely align with those of Clarke and Polselli (2025), who also conduct extensive simulation studies of DML estimators for panel data. They show that 2SLS performs best when the instrument set is well-specified and strong, but that DML-IV—especially when combined with flexible machine learning learners and panel-specific adjustments like the

Mundlak device—offers superior robustness in high-dimensional and empirically realistic scenarios. Both studies find that Lasso-IV, while useful for variable selection, can suffer from bias and increased variance as the number of instruments grows or when many instruments are weak. Clarke and Polselli further recommend ensemble learning strategies and highlight the importance of first-differencing for robustness, which is consistent with our observation that DML-IV with appropriate panel adjustments is preferable when instrument selection is uncertain.

Overall, our findings provide clear guidance for applied researchers: use 2SLS when the instrument set is well-specified and strong, but prefer DML-IV with Mundlak adjustment when instrument selection is uncertain or the instrument set is large and potentially weak. This study highlights the importance of flexible and robust estimation strategies in modern panel data analysis and supports the adoption of double machine learning methods for challenging empirical contexts.

References

[1] Belloni, A., Chernozhukov, V., & Hansen, C. (2011). LASSO methods for Gaussian instrumental variables models. *Econometric Theory*.

[2] Bach, P., Niekerk, J., Schmid, M., & Chernozhukov, V. (2024). DoubleML: An object-oriented implementation of double machine learning in Python and R. *Journal of Statistical Software*.

[3] Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal*.

[4] Clarke, P. S., & Polselli, A. (2025). Double machine learning for static panel models with fixed effects. *Computational Statistics*.

[5] Elavarasan Sneha. (2025). Exploring endogeneity: Can machine learning improve instrument selection? *Research report*. DOI: 10.13140/RG.2.2.15009.31845

[6] Fuhr, J., & Papies, D. (2024). Double machine learning meets panel data: Promises, pitfalls, and potential solutions. *arXiv preprint arXiv:2409.01266.*.

[7] Windmeijer, F. (2018). On the use of the Lasso for instrumental variables estimation with some invalid instruments. *Journal of the Royal Statistical Society: Series A*.

[8] Wooldridge, J. M. (2010, 2021). *Econometric analysis of panel data*. New York, NY: Routledge / Taylor & Francis.