librasophia : Pages:1–19

Original Research

Theoretical Foundations and Limitations of First-Difference and Within-Transformation Estimators in Static Panel Data Analysis

Ha Nguyen[1] and Minh Tran[2]

[1]Quy Nhon University, An Duong Vuong Street, Quy Nhon City, Vietnam.
[2]Can Tho University, 3/2 Street, Ninh Kieu District, Can Tho, Vietnam.

Abstract
Panel data analysis has become increasingly important in econometric research due to its ability to control for unobserved heterogeneity while providing greater statistical power than cross-sectional or time series data alone. This paper examines the theoretical foundations and practical limitations of first-difference and within-transformation estimators in static panel data models, with particular emphasis on their comparative performance under various data generating processes. We develop a comprehensive mathematical framework that establishes the conditions under which each estimator achieves consistency and efficiency, while also investigating their behavior in the presence of heteroskedasticity, serial correlation, and measurement error. Through rigorous theoretical analysis, we demonstrate that while both estimators eliminate time-invariant unobserved heterogeneity, they exhibit distinct properties regarding their asymptotic variance structures and finite sample performance. The first-difference estimator proves more robust to certain forms of serial correlation but suffers from amplified measurement error, whereas the within-transformation estimator maintains superior efficiency under classical assumptions but becomes inconsistent when strict exogeneity is violated. Our analysis reveals that the choice between these estimators depends critically on the underlying data generating process, the nature of the error structure, and the specific economic context. These findings have important implications for empirical researchers seeking to make informed decisions about estimation strategies in panel data applications.

1. Introduction

Panel data analysis has assisted empirical research across economics, finance, and social sciences by providing researchers with powerful tools to address fundamental identification challenges that plague cross-sectional and time series studies [1]. The ability to observe the same units over multiple time periods creates unique opportunities to control for unobserved heterogeneity, a pervasive source of bias in empirical work. However, this advantage comes with its own set of methodological challenges and trade-offs that require careful consideration.

The static panel data model forms the foundation of much empirical work in applied economics. In its most basic form, this model assumes that the dependent variable is determined by a set of explanatory variables, time-invariant unobserved heterogeneity, and a random error term. The key insight is that by exploiting the panel structure of the data, researchers can eliminate the bias arising from correlation between explanatory variables and unobserved heterogeneity, even when this heterogeneity is not directly observable. [2]

Two primary estimation strategies have emerged as the workhorses of static panel data analysis: the first-difference estimator and the within-transformation estimator. Both approaches share the common goal of eliminating time-invariant unobserved heterogeneity, but they achieve this objective through fundamentally different transformations of the data. The first-difference estimator removes unobserved heterogeneity by taking differences of consecutive observations, while the within-transformation estimator subtracts individual-specific means from each observation.

Despite their widespread use, the relative merits of these estimators remain a subject of ongoing debate. While both are consistent under standard assumptions, they differ in their efficiency properties, robustness to various forms of model misspecification, and finite sample performance. Understanding these differences is crucial for empirical researchers who must choose between these alternatives based on the specific characteristics of their data and research context. [3]

The theoretical foundations of these estimators were established in the seminal contributions of econometric theory, yet important questions remain about their comparative performance under realistic data generating processes. Recent developments in econometric theory have highlighted the importance of considering factors such as heteroskedasticity, serial correlation, and measurement error when evaluating the properties of panel data estimators. These considerations are particularly relevant given the increasing availability of large panel datasets that may exhibit complex error structures.

This paper provides a comprehensive theoretical analysis of first-difference and within-transformation estimators in static panel data models. We develop a unified mathematical framework that encompasses both estimators and allows for rigorous comparison of their properties under various assumptions about the data generating process. Our analysis extends beyond the standard textbook treatment by considering realistic departures from classical assumptions and examining their implications for estimator performance. [4]

The primary contributions of this paper are threefold. First, we provide a rigorous theoretical foundation for understanding the conditions under which each estimator achieves consistency and efficiency. Second, we develop new results regarding the comparative performance of these estimators in the presence of heteroskedasticity, serial correlation, and measurement error. Third, we offer practical guidance for researchers regarding the choice between these estimators based on observable characteristics of their data and research context.

2. Theoretical Framework and Model Specification

The foundation of our analysis rests on the standard static panel data model, which can be expressed as:

$$y_{it} = \alpha_i + x_{it}'\beta + \epsilon_{it}$$

where y_{it} represents the dependent variable for unit i at time t, x_{it} is a $K \times 1$ vector of explanatory variables, α_i captures time-invariant unobserved heterogeneity specific to unit i, β is a $K \times 1$ vector of parameters of interest, and ϵ_{it} is the idiosyncratic error term. The indices run over $i = 1, ..., N$ cross-sectional units and $t = 1, ..., T$ time periods. [5]

The fundamental challenge in estimating this model arises from the potential correlation between the explanatory variables x_{it} and the unobserved heterogeneity α_i. If $E[\alpha_i | x_{i1}, ..., x_{iT}] \neq 0$, then ordinary least squares estimation of the pooled model will yield inconsistent estimates of β due to omitted variable bias. This correlation between explanatory variables and unobserved heterogeneity is often referred to as the endogeneity problem in panel data analysis.

To address this challenge, we make several key assumptions about the data generating process. First, we assume that the idiosyncratic error term satisfies the strict exogeneity condition:

$$E[\epsilon_{it} | x_{i1}, ..., x_{iT}, \alpha_i] = 0$$

for all i and t. This assumption ensures that the explanatory variables are uncorrelated with the idiosyncratic error term, both contemporaneously and across time periods. Second, we assume that the idiosyncratic errors are independently and identically distributed across units and time periods, with constant variance σ_ϵ^2. [6]

Under these assumptions, both the first-difference and within-transformation estimators can be shown to be consistent for β. However, the specific transformation applied to the data has important implications for the efficiency properties of the resulting estimators and their robustness to departures from the classical assumptions.

The first-difference transformation eliminates the unobserved heterogeneity by taking differences of consecutive observations:

$$\Delta y_{it} = \Delta x_{it}'\beta + \Delta\epsilon_{it}$$

where $\Delta y_{it} = y_{it} - y_{i,t-1}$ and $\Delta x_{it} = x_{it} - x_{i,t-1}$. This transformation effectively removes the time-invariant component α_i from the model, leaving a relationship between the differenced variables that can be estimated using standard methods.

The within-transformation estimator, also known as the fixed effects estimator, achieves the same objective through a different approach. It subtracts the individual-specific means from each observation: [7]

$$\tilde{y}_{it} = \tilde{x}_{it}'\beta + \tilde{\epsilon}_{it}$$

where $\tilde{y}_{it} = y_{it} - \bar{y}_i$, $\tilde{x}_{it} = x_{it} - \bar{x}_i$, and $\bar{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$ represents the individual-specific mean of the dependent variable. The within-transformation removes the unobserved heterogeneity by exploiting the variation within each unit over time.

Both transformations result in consistent estimators for β under the standard assumptions, but they differ in their efficiency properties and robustness characteristics. The choice between these estimators depends on the specific features of the data generating process and the nature of potential departures from the classical assumptions.

To facilitate comparison between the estimators, we define the transformed design matrices for each approach. For the first-difference estimator, the transformed design matrix is:

$$X_\Delta = \begin{pmatrix} \Delta x_{12}' \\ \Delta x_{13}' \\ \vdots \\ \Delta x_{NT}' \end{pmatrix}$$

For the within-transformation estimator, the transformed design matrix is: [8]

$$X_W = \begin{pmatrix} \tilde{x}_{11}' \\ \tilde{x}_{12}' \\ \vdots \\ \tilde{x}_{NT}' \end{pmatrix}$$

The corresponding transformed dependent variable vectors are defined analogously. The parameter estimates for each method can then be expressed in matrix form, providing a framework for analyzing their statistical properties and comparative performance.

3. Mathematical Properties and Asymptotic Theory

The asymptotic properties of the first-difference and within-transformation estimators can be analyzed within a unified framework that highlights their fundamental similarities and differences. Both estimators belong to the class of linear estimators that eliminate the unobserved heterogeneity through data transformation, but their specific mathematical properties depend on the nature of the transformation applied.

For the first-difference estimator, the parameter estimate is given by: [9]

$$\hat{\beta}_{FD} = (X_\Delta' X_\Delta)^{-1} X_\Delta' y_\Delta$$

where y_Δ represents the vector of first-differenced dependent variables. Under the standard assumptions, this estimator is consistent for β as both N and T approach infinity, with the asymptotic distribution:

$$\sqrt{NT}(\hat{\beta}_{FD} - \beta) \to N(0, V_{FD})$$

where the asymptotic variance matrix is:

$$V_{FD} = 2\sigma_\epsilon^2 \lim_{N,T\to\infty} (NT)^{-1} (X_\Delta' X_\Delta)^{-1}$$

The factor of 2 in the variance expression reflects the fact that first-differencing doubles the variance of the idiosyncratic error term, as $\text{Var}(\Delta\epsilon_{it}) = \text{Var}(\epsilon_{it} - \epsilon_{i,t-1}) = 2\sigma_\epsilon^2$ under the assumption of no serial correlation.

For the within-transformation estimator, the parameter estimate is:

$$\hat{\beta}_W = (X_W' X_W)^{-1} X_W' y_W$$

where y_W represents the vector of within-transformed dependent variables. This estimator is also consistent for β under the standard assumptions, with asymptotic distribution: [10]

$$\sqrt{NT}(\hat{\beta}_W - \beta) \to N(0, V_W)$$

where the asymptotic variance matrix is:

$$V_W = \sigma_\epsilon^2 \left(1 - \frac{1}{T}\right) \lim_{N,T\to\infty} (NT)^{-1} (X_W' X_W)^{-1}$$

The term $(1 - T^{-1})$ reflects the degrees of freedom correction associated with the within-transformation, which subtracts the individual-specific means from each observation.

Comparing the asymptotic variance matrices reveals important differences between the estimators. When T is large, the within-transformation estimator is more efficient than the first-difference estimator, as $V_W < V_{FD}$ in the sense of the matrix ordering. However, when T is small, the efficiency ranking may be reversed, particularly if the serial correlation structure of the data favors the first-difference approach.

The efficiency comparison becomes more complex when we consider departures from the classical assumptions. In the presence of serial correlation in the idiosyncratic error term, the first-difference estimator may gain efficiency relative to the within-transformation estimator. Consider the case where the idiosyncratic errors follow an AR(1) process: [11]

$$\epsilon_{it} = \rho\epsilon_{i,t-1} + \nu_{it}$$

where $|\rho| < 1$ and ν_{it} is independent white noise with variance σ_ν^2. In this case, the variance of the first-differenced error becomes:

$$\text{Var}(\Delta\epsilon_{it}) = \text{Var}(\epsilon_{it} - \epsilon_{i,t-1}) = 2\sigma_\epsilon^2(1 - \rho)$$

where $\sigma_\epsilon^2 = \sigma_\nu^2/(1-\rho^2)$ is the unconditional variance of the error term. When ρ is positive and large, the variance of the first-differenced error is substantially smaller than $2\sigma_\epsilon^2$, potentially making the first-difference estimator more efficient than the within-transformation estimator.

The robustness properties of the estimators also differ in important ways. The first-difference estimator remains consistent under weaker assumptions about the serial correlation structure of the errors, while the within-transformation estimator may become inconsistent if the strict exogeneity assumption is violated. Specifically, if the explanatory variables are predetermined rather than strictly exogenous, meaning that $E[\epsilon_{it}|x_{i1}, ..., x_{it}] = 0$ but $E[\epsilon_{it}|x_{i,t+1}, ..., x_{iT}] \neq 0$, then the within-transformation estimator becomes inconsistent while the first-difference estimator remains consistent.

The finite sample properties of the estimators also merit consideration [12]. While both estimators are consistent in large samples, their finite sample behavior may differ substantially. The first-difference estimator uses only $T-1$ time periods for each unit, effectively reducing the sample size compared to the within-transformation estimator, which uses all T time periods. This reduction in effective sample size may lead to larger finite sample bias and variance for the first-difference estimator, particularly when T is small.

Furthermore, the presence of measurement error in the explanatory variables affects the two estimators differently. First-differencing tends to amplify measurement error, as the differenced variables contain measurement error from two time periods. If the explanatory variables are measured with classical measurement error, the first-difference estimator will suffer from greater attenuation bias than the within-transformation estimator, particularly when the measurement error is large relative to the true signal. [13]

The mathematical analysis reveals that the choice between first-difference and within-transformation estimators involves fundamental trade-offs between efficiency, robustness, and finite sample performance. These trade-offs depend critically on the specific characteristics of the data generating process and the nature of potential departures from the classical assumptions.

4. Advanced Mathematical Modeling and Estimation Theory

To develop a more sophisticated understanding of the comparative properties of first-difference and within-transformation estimators, we now present an advanced mathematical framework that incorporates complex error structures and addresses the challenges of optimal estimation in realistic panel data environments [14].

Consider the generalized static panel data model with heteroskedastic and serially correlated errors:

$$y_{it} = \alpha_i + x_{it}'\beta + \epsilon_{it}$$

where the error structure is characterized by:

$$E[\epsilon_{it}\epsilon_{js}] = \begin{cases} \sigma_{it}^2 & \text{if } i = j \text{ and } t = s \\ \sigma_{it,is} & \text{if } i = j \text{ and } t \neq s \\ 0 & \text{if } i \neq j \end{cases}$$

This specification allows for both heteroskedasticity across units and time periods, as well as serial correlation within units [15]. The covariance matrix for unit i can be expressed as:

$$\Omega_i = E[\epsilon_i\epsilon_i'] = \begin{pmatrix} \sigma_{i1}^2 & \sigma_{i1,i2} & \cdots & \sigma_{i1,iT} \\ \sigma_{i2,i1} & \sigma_{i2}^2 & \cdots & \sigma_{i2,iT} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{iT,i1} & \sigma_{iT,i2} & \cdots & \sigma_{iT}^2 \end{pmatrix}$$

Under this general error structure, the efficiency properties of the first-difference and within-transformation estimators can be analyzed using the framework of generalized least squares (GLS) estimation. The optimal linear unbiased estimator in this context is the GLS estimator, which applies the appropriate transformation to account for the error covariance structure.

For the first-difference estimator, the transformed error covariance matrix is:

$$\Omega_{\Delta,i} = D'\Omega_i D$$

where D is the $(T-1) \times T$ first-difference matrix: [16]

$$D = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

The efficient first-difference estimator is then given by:

$$\hat{\beta}_{FD,GLS} = \left(\sum_{i=1}^{N} X'_{\Delta,i} \Omega_{\Delta,i}^{-1} X_{\Delta,i} \right)^{-1} \sum_{i=1}^{N} X'_{\Delta,i} \Omega_{\Delta,i}^{-1} y_{\Delta,i}$$

For the within-transformation estimator, the transformed error covariance matrix is: [17]

$$\Omega_{W,i} = Q'\Omega_i Q$$

where Q is the $T \times T$ within-transformation matrix:

$$Q = I_T - \frac{1}{T} \iota_T \iota'_T$$

and ι_T is a $T \times 1$ vector of ones. The efficient within-transformation estimator is:

$$\hat{\beta}_{W,GLS} = \left(\sum_{i=1}^{N} X'_{W,i} \Omega_{W,i}^{-1} X_{W,i} \right)^{-1} \sum_{i=1}^{N} X'_{W,i} \Omega_{W,i}^{-1} y_{W,i}$$

The relative efficiency of these estimators depends on the specific structure of the error covariance matrix Ω_i. To illustrate this, consider the case where the errors follow a stationary AR(1) process with heteroskedastic innovations: [18]

$$\epsilon_{it} = \rho \epsilon_{i,t-1} + \nu_{it}$$

where $\nu_{it} \sim N(0, \sigma_{it}^2)$ and $|\rho| < 1$. The covariance matrix for this process is:

$$\Omega_i = \sigma_i^2 \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{pmatrix}$$

where σ_i^2 is the unconditional variance for unit i.

The efficiency comparison between the estimators requires computing the transformed covariance matrices and evaluating their relative properties. For the first-difference transformation, the covariance matrix becomes:

$$\Omega_{\Delta,i} = \sigma_i^2(1-\rho^2)\begin{pmatrix} 1+\rho^2 & -\rho & 0 & \cdots & 0 \\ [19]-\rho & 1+\rho^2 & -\rho & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1+\rho^2 \end{pmatrix}$$

This structure reveals that first-differencing introduces negative serial correlation in the transformed errors, even when the original errors are positively correlated. The magnitude of this negative correlation depends on the original correlation parameter ρ.

For the within-transformation, the covariance matrix is more complex and depends on the specific values of ρ and T [20]. The elements of $\Omega_{W,i}$ can be expressed as:

$$[\Omega_{W,i}]_{st} = \sigma_i^2 \left[\rho^{|s-t|} - \frac{1}{T^2}\sum_{k=1}^{T}\sum_{l=1}^{T}\rho^{|k-l|} - \frac{1}{T}\sum_{k=1}^{T}\rho^{|s-k|} - \frac{1}{T}\sum_{l=1}^{T}\rho^{|t-l|} \right]$$

The complexity of this expression makes it difficult to derive general analytical results about the relative efficiency of the estimators. However, we can establish several important theoretical results through asymptotic analysis.

When T is fixed and $N \to \infty$, the asymptotic relative efficiency of the first-difference estimator compared to the within-transformation estimator is:

$$\text{ARE}_{FD,W} = \frac{\text{tr}(V_W)}{\text{tr}(V_{FD})} = \frac{(1-\rho)^2}{2(1-\rho^2)}$$

where $\text{tr}(\cdot)$ denotes the trace operator. This expression shows that when $\rho > 0$, the first-difference estimator can be more efficient than the within-transformation estimator, with the efficiency gain increasing as ρ approaches unity.

The analysis becomes more complex when we consider the case where both N and T approach infinity. In this asymptotic regime, the relative efficiency depends on the rate at which T grows relative to N [21]. When T grows at a faster rate than N, the within-transformation estimator typically dominates, while when N grows much faster than T, the first-difference estimator may be preferred.

The presence of measurement error introduces additional complexity to the efficiency comparison. Suppose the observed explanatory variables are contaminated with classical measurement error:

$$x_{it}^{obs} = x_{it}^{true} + u_{it}$$

where x_{it}^{true} is the true value and u_{it} is measurement error with variance σ_u^2. The first-difference transformation amplifies this measurement error, as:

$$\text{Var}(\Delta u_{it}) = 2\sigma_u^2$$

while the within-transformation has a more complex effect on measurement error, with:

$$\text{Var}(\tilde{u}_{it}) = \sigma_u^2\left(1 - \frac{1}{T}\right)$$

The attenuation bias in the first-difference estimator is therefore more severe than in the within-transformation estimator, particularly when T is large. [22]

These advanced mathematical results provide important insights into the choice between first-difference and within-transformation estimators. The optimal choice depends on the interplay between the serial correlation structure of the errors, the degree of heteroskedasticity, the presence of measurement error, and the relative dimensions of the panel. Understanding these trade-offs is crucial for developing robust empirical strategies in panel data analysis.

5. Robustness Analysis and Diagnostic Testing

The practical implementation of first-difference and within-transformation estimators requires careful attention to the validity of the underlying assumptions and the development of appropriate diagnostic tests. This section examines the robustness properties of these estimators and presents formal testing procedures for assessing the appropriateness of each approach in specific empirical contexts.

The fundamental assumption underlying both estimators is the strict exogeneity of the explanatory variables with respect to the idiosyncratic error term [23]. However, this assumption may be violated in many empirical applications, particularly when the explanatory variables are influenced by lagged values of the dependent variable or when there are omitted variables that affect both the dependent and explanatory variables over time.

To formalize the analysis of robustness to endogeneity, consider the case where the explanatory variables are predetermined rather than strictly exogenous. In this case, we have:

$$E[\epsilon_{it} | x_{i1}, \ldots, x_{it}, \alpha_i] = 0$$

but

$$E[\epsilon_{it} | x_{i,t+1}, \ldots, x_{iT}, \alpha_i] \neq 0$$

Under this weaker assumption, the within-transformation estimator becomes inconsistent because the transformed regressors $\tilde{x}_{it} = x_{it} - \bar{x}_i$ are correlated with the transformed error term $\tilde{\epsilon}_{it} = \epsilon_{it} - \bar{\epsilon}_i$. The correlation arises because $\bar{x}_i$ includes future values of the explanatory variables that may be correlated with the current error term.

In contrast, the first-difference estimator remains consistent under the predetermined assumption, as the differenced regressors $\Delta x_{it} = x_{it} - x_{i,t-1}$ are uncorrelated with $\Delta\epsilon_{it} = \epsilon_{it} - \epsilon_{i,t-1}$ under the predetermined assumption. This robustness property makes the first-difference estimator attractive in applications where strict exogeneity is questionable.

The presence of serial correlation in the idiosyncratic error term also affects the robustness of the estimators differently. Consider the AR(1) error structure: [24]

$$\epsilon_{it} = \rho\epsilon_{i,t-1} + \nu_{it}$$

where ν_{it} is independent white noise. Under this structure, the first-difference estimator yields:

$$\Delta y_{it} = \Delta x_{it}'\beta + \Delta\epsilon_{it}$$

where $\Delta\epsilon_{it} = \epsilon_{it} - \epsilon_{i,t-1} = \nu_{it} - (1-\rho)\epsilon_{i,t-1}$. The transformed error term is serially correlated, with:

$$E[\Delta\epsilon_{it}\Delta\epsilon_{i,t-1}] = -(1-\rho)\sigma_\epsilon^2$$

This serial correlation violates the standard assumptions for ordinary least squares estimation, leading to inefficient but consistent parameter estimates.

For the within-transformation estimator, the presence of serial correlation creates a more complex error structure. The transformed error term $\tilde{\epsilon}_{it} = \epsilon_{it} - \bar{\epsilon}_i$ exhibits both serial correlation and cross-sectional correlation, as all observations for unit i share the common component $\bar{\epsilon}_i$. This structure complicates the computation of standard errors and may affect the finite sample properties of the estimator.

To address these robustness concerns, we can develop formal testing procedures for assessing the validity of the key assumptions [25]. One important test is the Sargan-Hansen test for the validity of the strict exogeneity assumption. This test compares the first-difference and within-transformation estimators under the null hypothesis that both are consistent, with differences attributed to efficiency considerations.

The test statistic is constructed as:

$$J = N(\hat{\beta}_{FD} - \hat{\beta}_W)'[V_{FD} - V_W]^{-1}(\hat{\beta}_{FD} - \hat{\beta}_W)$$

where V_{FD} and V_W are consistent estimators of the asymptotic variance matrices. Under the null hypothesis of strict exogeneity, this statistic is asymptotically distributed as $\chi^2(K)$, where K is the number of regressors. Rejection of the null hypothesis suggests that the strict exogeneity assumption is violated and that the first-difference estimator is preferred. [26]

Another important diagnostic test is the test for serial correlation in the idiosyncratic error term. For the first-difference estimator, we can test for second-order serial correlation in the differenced residuals using the Arellano-Bond test. The test statistic is:

$$AB_2 = \frac{\sum_{i=1}^{N} \sum_{t=3}^{T} \hat{\Delta\epsilon}_{it} \hat{\Delta\epsilon}_{i,t-2}}{\sqrt{\sum_{i=1}^{N} \sum_{t=3}^{T} \hat{\Delta\epsilon}_{i,t-2}^2}}$$

Under the null hypothesis of no serial correlation in the original error term, this statistic is asymptotically standard normal. The test focuses on second-order serial correlation because first-order serial correlation in the differenced residuals is expected even under the null hypothesis.

For the within-transformation estimator, we can test for serial correlation using a modified Durbin-Watson test or the Baltagi-Li test [27]. The Baltagi-Li test statistic is:

$$BL = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} (\tilde{\epsilon}_{it} - \tilde{\epsilon}_{i,t-1})^2}{\sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{\epsilon}_{it}^2}$$

This statistic has a known asymptotic distribution under the null hypothesis of no serial correlation, allowing for formal hypothesis testing.

The presence of heteroskedasticity can also be tested using appropriate diagnostic procedures. For the first-difference estimator, we can apply the Breusch-Pagan test to the differenced residuals:

$$BP_{FD} = \frac{1}{2}\left[\sum_{i=1}^{N}\sum_{t=2}^{T}\left(\frac{\hat{\Delta\epsilon}_{it}^2}{\hat{\sigma}_\Delta^2} - 1\right)\right]^2$$

where $\hat{\sigma}_\Delta^2$ is the sample variance of the differenced residuals. Under the null hypothesis of homoskedasticity, this statistic is asymptotically $\chi^2(1)$.

For the within-transformation estimator, the heteroskedasticity test requires modification to account for the within-transformation structure: [28]

$$BP_W = \frac{1}{2}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\frac{\tilde{\epsilon}_{it}^2}{\hat{\sigma}_W^2} - 1\right)\right]^2$$

where $\hat{\sigma}_W^2$ is the sample variance of the within-transformed residuals.

The results of these diagnostic tests provide important guidance for choosing between the first-difference and within-transformation estimators. When the tests indicate violations of the strict exogeneity assumption, the first-difference estimator is preferred. When serial correlation is detected, appropriate corrections to the standard errors or the use of generalized least squares methods may be necessary. The presence of heteroskedasticity suggests the need for robust standard errors or weighted least squares methods.

In addition to formal hypothesis tests, researchers should also consider the economic interpretation of the parameters and the nature of the data generating process. The first-difference estimator identifies the effect of changes in the explanatory variables on changes in the dependent variable, while the within-transformation estimator identifies the effect of deviations from individual means [29]. These different interpretations may be more or less relevant depending on the specific research question and economic context.

The robustness analysis reveals that the choice between first-difference and within-transformation estimators involves important trade-offs between consistency, efficiency, and robustness to model misspecification. Careful diagnostic testing and consideration of the economic context are essential for making informed decisions about estimation strategy in panel data applications.

6. Computational Considerations and Implementation

The practical implementation of first-difference and within-transformation estimators involves several computational considerations that can significantly affect the efficiency and accuracy of the estimation process. This section examines the numerical properties of these estimators and discusses optimal computational strategies for their implementation in large-scale panel data applications [30].

The computational complexity of both estimators is relatively modest, as they involve linear transformations of the data followed by standard least squares estimation [31]. However, the specific characteristics of each transformation create different computational challenges and opportunities for optimization. Understanding these differences is crucial for developing efficient algorithms and ensuring numerical stability in practical applications.

For the first-difference estimator, the primary computational task involves constructing the differenced variables and applying ordinary least squares to the transformed data. The differencing operation can be expressed as a matrix multiplication:

$$y_\Delta = D_N y$$

where y is the $NT \times 1$ stacked vector of dependent variables, D_N is the $N(T-1) \times NT$ first-difference matrix, and y_Δ is the resulting vector of differenced observations [32]. The matrix D_N has a block-diagonal structure:

$$D_N = \begin{pmatrix} D & 0 & \cdots & 0 \\ 0 & D & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D \end{pmatrix}$$

where each D block is the $(T-1)\times T$ first-difference matrix defined earlier [33]. This block-diagonal structure can be exploited for computational efficiency, as the transformation can be applied separately to each individual's observations without requiring the construction of the full D_N matrix.

The computational complexity of the first-difference transformation is $O(NT)$, which is linear in the total number of observations. The subsequent least squares estimation requires computing $(X_\Delta' X_\Delta)^{-1}$, which has complexity $O(K^3)$ where K is the number of regressors. When K is small relative to NT, the transformation step dominates the computational cost.

For the within-transformation estimator, the computational challenge is more complex due to the need to compute individual-specific means and subtract them from each observation. The within-transformation can be expressed as: [34]

$y_W = Q_N y$

where Q_N is the $NT \times NT$ within-transformation matrix:

$Q_N = I_{NT} - (I_N \otimes \iota_T)(I_N \otimes \iota_T')(I_N \otimes T^{-1})$

Here, I_N and I_{NT} are identity matrices of appropriate dimensions, ι_T is a $T \times 1$ vector of ones, and $\otimes$ denotes the Kronecker product. Unlike the first-difference matrix, Q_N is not block-diagonal, which complicates its efficient computation.

However, the within-transformation can be computed efficiently without explicitly constructing the transformation matrix. The algorithm involves two steps: first, compute the individual-specific means for all variables, and second, subtract these means from the corresponding observations. This approach has computational complexity $O(NT)$, the same as the first-difference transformation. [35]

The numerical stability of the estimators depends on the conditioning of the transformed design matrices $X_\Delta' X_\Delta$ and $X_W' X_W$. Poor conditioning can lead to numerical inaccuracies in the parameter estimates and their standard errors. The condition numbers of these matrices provide a measure of numerical stability:

$\kappa(X_\Delta' X_\Delta) = \frac{\lambda_{\max}(X_\Delta' X_\Delta)}{\lambda_{\min}(X_\Delta' X_\Delta)}$

$\kappa(X_W' X_W) = \frac{\lambda_{\max}(X_W' X_W)}{\lambda_{\min}(X_W' X_W)}$

where $\lambda_{\max}$ and $\lambda_{\min}$ denote the largest and smallest eigenvalues, respectively.

The first-difference transformation can worsen the conditioning of the design matrix when the explanatory variables exhibit strong persistence over time. In such cases, the differenced variables may have very small variance, leading to near-singularity of $X_\Delta' X_\Delta$. This problem is particularly acute when the explanatory variables follow near-unit-root processes. [36]

The within-transformation generally produces better-conditioned design matrices because it preserves more of the variation in the original data. However, when T is small, the within-transformation removes a substantial fraction of the total variation, which can also lead to conditioning problems.

Memory requirements constitute another important computational consideration, particularly for large panel datasets. The first-difference estimator requires storage of $(T-1)NT$ transformed observations, while the within-transformation estimator requires storage of NT transformed observations. For very large datasets, the memory savings from the first-difference approach may be significant, as it effectively reduces the sample size by approximately N observations.

The computation of standard errors requires additional consideration of the error covariance structure [37]. Under the assumption of homoskedastic and serially uncorrelated errors, the standard error calculations are straightforward for both estimators. However, when these assumptions are violated, robust standard errors must be computed using more sophisticated methods.

For heteroskedasticity-robust standard errors, the estimator of the asymptotic variance matrix is:

$\hat{V}_{robust} = (X'X)^{-1} \left(\sum_{i=1}^{N} X_i' \hat{\Omega}_i X_i \right) (X'X)^{-1}$

where X represents the appropriate transformed design matrix and $\hat{\Omega}_i$ is an estimator of the error covariance matrix for unit i. The computation of this estimator requires careful handling of the transformation-specific error structures.

For the first-difference estimator with heteroskedastic errors, the diagonal elements of $\hat{\Omega}_{\Delta,i}$ are the squared first-differenced residuals:

$[\hat{\Omega}_{\Delta,i}]_{tt} = \hat{\Delta\epsilon}_{it}^2$

For the within-transformation estimator, the computation is more complex because the within-transformed residuals are not independent within each unit. The appropriate covariance estimator must account for the correlation structure induced by the transformation. [38]

Cluster-robust standard errors provide another important option for handling correlation within units over time. The cluster-robust variance estimator is:

$\hat{V}_{cluster} = (X'X)^{-1} \left(\sum_{i=1}^{N} X_i' \hat{\epsilon}_i \hat{\epsilon}_i' X_i \right) (X'X)^{-1}$

where $\hat{\epsilon}_i$ is the vector of residuals for unit i. This estimator allows for arbitrary correlation within units while maintaining the assumption of independence across units.

The choice of standard error estimator can significantly affect inference, particularly in finite samples. Simulation studies have shown that cluster-robust standard errors tend to be more reliable than heteroskedasticity-robust standard errors in panel data applications, particularly when N is moderate and T is small.

Advanced computational techniques can further improve the efficiency of panel data estimation [39]. For very large datasets, iterative methods such as the conjugate gradient algorithm can be used to solve the normal equations without explicitly forming the cross-product matrices. These methods are particularly useful when the design matrices are sparse or have special structure.

Parallel computing techniques can also be employed to accelerate the computation of panel data estimators. The block-diagonal structure of the transformation matrices makes it natural to parallelize the computation across individuals. Each processor can handle a subset of individuals independently, with only minimal communication required for aggregating the final results.

Modern statistical software implementations of panel data estimators incorporate many of these computational optimizations [40]. However, understanding the underlying computational principles is important for researchers working with very large datasets or implementing custom estimation procedures. The choice between first-difference and within-transformation estimators may be influenced by computational considerations, particularly when dealing with datasets containing millions of observations.

The computational analysis reveals that both estimators are relatively efficient to compute, but they differ in their memory requirements, numerical stability properties, and compatibility with robust standard error calculations. These differences should be considered alongside the statistical properties when choosing between the estimators in practical applications.

7. Empirical Applications and Practical Guidelines

The theoretical analysis presented in the previous sections provides important insights into the properties of first-difference and within-transformation estimators, but translating these insights into practical guidance for empirical researchers requires careful consideration of the typical characteristics of real-world panel datasets and research questions. This section examines the empirical performance of these estimators across different application domains and provides concrete recommendations for their use. [41]

Panel data applications span a wide range of fields, from labor economics and industrial organization to international trade and development economics. Each application domain presents unique challenges and characteristics that influence the relative desirability of different estimation approaches. Understanding these domain-specific considerations is crucial for making informed methodological choices.

In labor economics applications, panel data are commonly used to study wage determination, employment dynamics, and the effects of policy interventions. A typical specification might examine the relationship between wages and various worker characteristics: [42]

$\log(wage_{it}) = \alpha_i + \beta_1 experience_{it} + \beta_2 education_{it} + \beta_3 union_{it} + \epsilon_{it}$

In this context, the unobserved heterogeneity α_i captures time-invariant worker characteristics such as ability, motivation, or other unmeasured skills that affect wages. The choice between first-difference and within-transformation estimators depends on the specific research question and the nature of the explanatory variables.

When examining the returns to experience, the first-difference estimator identifies the effect of an additional year of experience on wage growth, while the within-transformation estimator identifies the effect of deviations in experience from the individual's average experience level. The first-difference interpretation may be more natural for policy analysis, as it directly measures the wage gains from additional experience.

However, if the research interest lies in understanding how wages respond to changes in union status or other discrete variables, the within-transformation estimator may be preferable because it exploits all available variation in these variables. The first-difference estimator only uses information from periods where union status changes, potentially resulting in a substantial loss of information. [43]

The presence of measurement error in labor economics applications also favors the within-transformation estimator. Variables such as experience and education are often measured with error, and first-differencing can substantially amplify this measurement error. When measurement error is a significant concern, the within-transformation estimator typically provides more reliable results.

In industrial organization applications, panel data are frequently used to study firm behavior, market structure, and the effects of regulation. A common specification examines the determinants of firm performance:

$$performance_{it} = \alpha_i + \beta_1 size_{it} + \beta_2 market\_share_{it} + \beta_3 regulation_{it} + \epsilon_{it}$$

The unobserved heterogeneity in this context captures time-invariant firm characteristics such as management quality, technological capabilities, or organizational culture [44]. The choice between estimators depends on the persistence of the explanatory variables and the nature of the performance measure.

Firm size and market share tend to be highly persistent over time, making the first-difference estimator less efficient because the differenced variables have relatively small variance. In such cases, the within-transformation estimator is typically preferred because it makes better use of the available variation in the data.

However, when studying the effects of regulatory changes or other policy interventions, the first-difference estimator may be more appropriate because it focuses on the dynamic response to these changes. The first-difference specification directly measures how firm performance changes in response to regulatory changes, which is often the parameter of primary interest in policy analysis.

The time dimension of the panel also plays a crucial role in determining the optimal estimation strategy [45]. When T is small relative to N, both estimators face finite sample challenges, but the within-transformation estimator generally performs better because it uses all available time periods. When T is large, the efficiency advantage of the within-transformation estimator becomes more pronounced, and issues such as structural breaks or time-varying coefficients become more relevant.

The presence of gaps or missing observations in the panel creates additional complications for both estimators. The first-difference estimator requires consecutive observations to compute differences, so gaps in the data reduce the effective sample size. The within-transformation estimator can accommodate missing observations more flexibly, as it only requires the computation of individual-specific means over the available observations.

Seasonal patterns in the data present another practical consideration [46]. When the dependent variable exhibits strong seasonal variation, first-differencing may not fully eliminate these patterns, particularly if the seasonal effects vary across individuals or change over time. In such cases, the within-transformation estimator may be preferable because it removes all forms of individual-specific means, including seasonal components.

The interpretation of the results also differs between the estimators in ways that may be relevant for policy analysis. The first-difference estimator measures short-run or instantaneous effects, while the within-transformation estimator captures a form of long-run relationship by exploiting variation around individual-specific means. When the research question concerns the immediate impact of a policy change, the first-difference estimator provides more relevant evidence. When the question concerns the overall relationship between variables, the within-transformation estimator may be more informative. [47]

Diagnostic testing should be an integral part of any empirical analysis using panel data estimators. The tests discussed in the previous section provide formal procedures for assessing the validity of key assumptions, but researchers should also conduct informal diagnostic checks such as examining the time series properties of the variables and the stability of the estimates across different subsamples.

One particularly useful diagnostic approach is to compare the results from both estimators and investigate the sources of any differences. Large differences between the first-difference and within-transformation estimates may indicate violations of the strict exogeneity assumption, the presence of measurement error, or other forms of model misspecification. In such cases, additional investigation is warranted to determine the source of the discrepancy and identify the most appropriate estimation strategy.

The standard errors and confidence intervals should also be computed using methods that are robust to the most likely forms of model misspecification [48]. In most panel data applications, cluster-robust standard errors that allow for arbitrary correlation within units over time provide the most reliable basis for inference. When heteroskedasticity is suspected, additional robustness checks using weighted least squares or other methods may be warranted.

The practical guidelines emerging from this analysis can be summarized as follows. First, consider the research question and the economic interpretation of the parameters. The first-difference estimator is preferred when the focus is on dynamic effects or short-run responses, while the within-transformation estimator is preferred when the focus is on the overall relationship between variables. Second, consider the persistence of the explanatory variables [49]. Highly persistent variables favor the within-transformation estimator, while variables with substantial period-to-period variation may favor the first-difference estimator. Third, consider the potential for measurement error and endogeneity. Measurement error favors the within-transformation estimator, while endogeneity concerns may favor the first-difference estimator. Fourth, conduct comprehensive diagnostic testing to assess the validity of key assumptions and guide the choice between estimators. Finally, report results from both estimators when feasible, and investigate the sources of any substantial differences between them. [50]

These guidelines provide a framework for making informed decisions about estimation strategy in panel data applications, but they should be adapted to the specific characteristics of each research context. The choice between first-difference and within-transformation estimators ultimately depends on the interplay between the research question, the data characteristics, and the economic environment under study.

8. Conclusion

This comprehensive analysis of first-difference and within-transformation estimators in static panel data models has revealed important theoretical insights and practical implications that extend our understanding of these fundamental econometric tools. Through rigorous mathematical analysis and careful consideration of empirical applications, we have established a framework for understanding when each estimator is most appropriate and how their properties depend on the underlying data generating process.

The theoretical foundations developed in this paper demonstrate that both estimators achieve the primary objective of eliminating time-invariant unobserved heterogeneity, but they do so through fundamentally different mechanisms that create distinct implications for efficiency, robustness, and interpretation. The first-difference estimator removes unobserved heterogeneity by exploiting temporal variation within each unit, while the within-transformation estimator removes it by exploiting deviations from individual-specific means [51]. These different approaches lead to estimators with complementary strengths and weaknesses.

Our analysis of the asymptotic properties reveals that the relative efficiency of the estimators depends critically on the time dimension of the panel and the serial correlation structure of the errors. When the time dimension is large and errors are serially uncorrelated, the within-transformation estimator dominates in terms of efficiency. However, when errors exhibit positive serial correlation, the first-difference estimator can achieve superior efficiency by exploiting the reduced variance of the differenced error terms. This insight provides important guidance for researchers working with datasets that exhibit different error structures.

The advanced mathematical modeling presented in this paper extends beyond standard textbook treatments by incorporating realistic departures from classical assumptions [52]. Our analysis of heteroskedastic and serially correlated error structures demonstrates that the optimal choice between estimators depends on complex interactions between the error covariance structure, the persistence of the explanatory variables, and the dimensions of the panel. The generalized least squares framework developed here provides a unified approach for understanding these interactions and deriving optimal estimation strategies.

The robustness analysis reveals fundamental differences between the estimators that have important implications for empirical practice. The first-difference estimator maintains consistency under weaker assumptions about the exogeneity of the explanatory variables, making it more robust to certain forms of endogeneity. However, it is more susceptible to measurement error and may suffer from efficiency losses when explanatory variables are highly persistent. The within-transformation estimator achieves superior efficiency under classical assumptions but becomes inconsistent when strict exogeneity is violated and may perform poorly in the presence of strong serial correlation. [53]

Our investigation of computational considerations demonstrates that both estimators can be implemented efficiently using modern algorithms and computing resources. The first-difference estimator offers advantages in terms of memory requirements and numerical stability when explanatory variables are highly persistent, while the within-transformation estimator provides computational advantages when the time dimension is small or when missing observations create gaps in the panel structure.

The empirical applications examined in this paper illustrate how the choice between estimators should be guided by the specific research question, the characteristics of the data, and the economic context under study. In labor economics applications, where measurement error is often a concern and the focus is on long-run relationships, the within-transformation estimator is frequently preferred. In industrial organization applications, where the emphasis is often on dynamic responses to policy changes, the first-difference estimator may be more appropriate. These domain-specific considerations highlight the importance of tailoring the estimation strategy to the particular research context. [54]

The diagnostic testing procedures developed in this paper provide formal methods for assessing the validity of key assumptions and guiding the choice between estimators. The Sargan-Hansen test for strict exogeneity, the Arellano-Bond test for serial correlation, and various tests for heteroskedasticity offer researchers concrete tools for evaluating their modeling assumptions and selecting appropriate estimation strategies. These tests should be routinely applied in empirical work to ensure the reliability of the results.

Several important implications emerge from this analysis for empirical researchers. First, the choice between first-difference and within-transformation estimators should be based on careful consideration of the research question, data characteristics, and economic context, rather than arbitrary preference or convention. Second, diagnostic testing should be an integral part of the analysis to assess the validity of key assumptions and guide methodological choices [55]. Third, researchers should consider reporting results from both estimators when feasible, as differences between them can provide valuable insights into the nature of the data generating process and the robustness of the findings.

The analysis also reveals several areas where further research would be valuable. The behavior of these estimators in the presence of structural breaks, time-varying coefficients, and cross-sectional dependence deserves additional investigation. The development of more sophisticated diagnostic tests that can distinguish between different sources of model misspecification would also be beneficial. Furthermore, the extension of this analysis to dynamic panel data models and models with endogenous explanatory variables represents an important frontier for future research.

The practical guidelines developed in this paper provide a framework for making informed decisions about estimation strategy in panel data applications [56]. However, these guidelines should be viewed as starting points for analysis rather than rigid rules, as the optimal choice ultimately depends on the specific characteristics of each research context. The key insight is that understanding the theoretical properties of these estimators and their sensitivity to various assumptions is essential for conducting reliable empirical research.

In conclusion, this comprehensive analysis demonstrates that first-difference and within-transformation estimators represent powerful but distinct approaches to addressing unobserved heterogeneity in panel data models. Each estimator has comparative advantages under different conditions, and the choice between them requires careful consideration of theoretical properties, empirical characteristics, and research objectives. By providing a unified framework for understanding these estimators and practical guidance for their application, this paper contributes to the development of more rigorous and reliable empirical research in economics and related fields. The insights developed here will be valuable for researchers seeking to make informed methodological choices and for educators seeking to convey the nuances of panel data analysis to students and practitioners. [57]

References

[1] K. Walayat, D. M. K. Qamar, and D. Z. I. Iqbal, "Effectiveness of globalization and human capital on market & net income inequality in brics countries: A panel data analysis," *Journal of Arts & Social Sciences (JASS)*, vol. 9, pp. 98–108, 6 2022.

[2] Z. Xue, N. Li, H. Mu, M. Zhang, and J. Pang, "Convergence analysis of regional marginal abatement cost of carbon dioxide in china based on spatial panel data models," *Environmental science and pollution research international*, vol. 28, pp. 38929–38946, 3 2021.

[3] L. Caifei, "Empirical analysis on the effect of agricultural insurance on production——based on panel data of 31 provinces and cities in china from 2008 to 2018," *E3S Web of Conferences*, vol. 214, pp. 01013–, 12 2020.

[4] F. Emirmahmutoglu, R. Gupta, S. M. Miller, and T. Omay, "Is real per capita state personal income stationary? new nonlinear, asymmetric panel-data evidence," *Bulletin of Economic Research*, vol. 72, pp. 50–62, 6 2019.

[5] L. Keele and D. S. Small, "Instrumental variables: Don't throw the baby out with the bathwater.," *Health services research*, vol. 54, pp. 543–546, 3 2019.

[6] P. E. Ioannatos, "Brexit or euro for the uk? evidence from panel data," *Comparative Economic Studies*, vol. 63, pp. 117–138, 9 2020.

[7] G. Xu, H. Chang, H. Yang, and P. Schwarz, "The influence of finance on china's green development: an empirical study based on quantile regression with province-level panel data.," *Environmental science and pollution research international*, vol. 29, pp. 71033–71046, 5 2022.

[8] D. Shen, Y. Yang, and J. Lu, "Exploring influencing factors of seafood e-commerce efficiency: Based on panel data from 2011 to 2018," *Journal of Coastal Research*, vol. 108, pp. 322–326, 9 2020.

[9] C. Liu, Z. min Liu, S. Nicholas, and J. Wang, "Trends and determinants of catastrophic health expenditure in china 2010–2018: a national panel data analysis," *BMC health services research*, vol. 21, pp. 526–526, 5 2021.

[10] K. L. Sainani, "Instrumental variables: Uses and limitations.," *PM & R : the journal of injury, function, and rehabilitation*, vol. 10, pp. 303–308, 3 2018.

[11] L. Gilstrap, A. M. Austin, A. J. O'Malley, B. Gladders, A. E. Barnato, A. N. A. Tosteson, and J. Skinner, "Association between beta-blockers and mortality and readmission in older patients with heart failure: an instrumental variable analysis," *Journal of general internal medicine*, vol. 36, pp. 2361–2369, 6 2021.

[12] J. Bai, S. H. Choi, and Y. Liao, "Feasible generalized least squares for panel data with cross-sectional and serial correlations," *Empirical Economics*, vol. 60, pp. 309–326, 11 2020.

[13] S. Moler-Zapata, R. Grieve, A. Basu, and S. O'Neill, "How does a local instrumental variable method perform across settings with instruments of differing strengths? a simulation study and an evaluation of emergency surgery.," *Health economics*, vol. 32, pp. 2113–2126, 6 2023.

[14] Z. F. Fisher and K. A. Bollen, "An instrumental variable estimator for mixed indicators: Analytic derivatives and alternative parameterizations.," *Psychometrika*, vol. 85, pp. 660–683, 8 2020.

[15] X. Shi, B. Zhao, Y. Yao, and F. Wang, "Prediction methods for routine maintenance costs of a reinforced concrete beam bridge based on panel data," *Advances in Civil Engineering*, vol. 2019, pp. 1–12, 6 2019.

[16] X. Gan, W. Xu, and K. Yu, "Economic growth and weight of children and adolescents in urban areas: A panel data analysis on twenty-seven provinces in china, 1985–2014," *Childhood obesity (Print)*, vol. 16, pp. 86–93, 10 2019.

[17] S. Chen and B. Zhang, "Estimating and improving dynamic treatment regimes with a time-varying instrumental variable," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 85, pp. 427–453, 3 2023.

[18] D. Wang, Y. Huo, and J. Wang, "Impact of climate change on china's rice production — an empirical estimation based on panel data (1979–2011) from china's main rice-producing areas," *The Singapore Economic Review*, vol. 63, pp. 1740024–553, 6 2018.

[19] Y. Chen, Q. Zhou, X. Yang, P. Shi, Q. Shen, Z. Zhang, Z. Chen, C. Pu, L. Xu, Z. Hu, A. Ma, Z. Gong, T. Xu, P. Wang, H. Wang, C. Hao, C. Li, and M. Hao, "Influence of public health services on the goal of ending tuberculosis: Evidence from panel data in china.," *Frontiers in public health*, vol. 10, pp. 826800–, 3 2022.

[20] I. Andrews, J. H. Stock, and L. Sun, "Weak instruments in instrumental variables regression: Theory and practice," *Annual Review of Economics*, vol. 11, pp. 727–753, 8 2019.

[21] R. Y. Kim, "What makes people happy? an empirical investigation of panel data," *Applied Economics Letters*, vol. 28, pp. 91–94, 2 2020.

[22] D. Cai, J. Hu, H. Jiang, F. Ai, and T. Bai, "Research on the relationship between defense technology innovation and high-quality economic development: Gray correlation analysis based on panel data," *Managerial and Decision Economics*, vol. 44, pp. 3867–3877, 6 2023.

[23] S. Gupta, P. Seth, M. Abraham, and P. Pingali, "Covid-19 and women's nutrition security: panel data evidence from rural india," *Economia politica (Bologna, Italy)*, vol. 39, pp. 1–28, 7 2021.

[24] B. Wan and W. Min, "The role of education in urbanization: An empirical study based on china's provincial-level panel data from 2005 to 2020," *Best Evidence in Chinese Education*, vol. 15, pp. 1801–1805, 11 2023.

[25] J. Li and W.-M. Chen, "An improved grey clustering model with multiattribute spatial-temporal feature for panel data and its application," *Mathematical Problems in Engineering*, vol. 2020, pp. 1–9, 1 2020.

[26] J.-F. Hou, L. Tian, Y. Zhang, Y.-Z. Liu, J. Li, and Y. Wang, "Study of influential factors of provincial health expenditure -analysis of panel data after the 2009 healthcare reform in china.," *BMC health services research*, vol. 20, pp. 606–, 7 2020.

[27] L. Song, "Examining the relationship between intergenerational upward mobility and inequality: Evidence from panel data," *Social Indicators Research*, vol. 163, pp. 1–27, 2 2022.

[28] L. Shang and P. Xu, "Can carbon emission regulation achieve a dual target of low carbon and employment? an empirical analysis based on china's provincial panel data," *Frontiers in Energy Research*, vol. 10, 7 2022.

[29] D. Diansheng, S. Hayden, and M. Patrick, "A new approach to estimate household food demand with panel data," *Journal of Applied Economics*, vol. 21, pp. 122–136, 11 2018.

[30] W. Li, L. Xue, and P. Zhao, "Empirical likelihood based inference for varying coefficient panel data models with fixed effect," *Communications in Statistics - Theory and Methods*, vol. 51, pp. 4973–4990, 10 2020.

[31] W. Zhang, S. H. Lence, and T. Kuethe, "Are expert opinions accurate? panel data evidence from the iowa land value survey," *Land Economics*, vol. 97, pp. 875–892, 10 2021.

[32] X. Shen, H. Zhao, J. Yu, Z. Wan, T. He, and J. Liu, "Digital economy and ecological performance: Evidence from a spatial panel data in china," *Frontiers in Environmental Science*, vol. 10, 9 2022.

[33] J. D. Angrist and P. Hull, "Instrumental variables methods reconcile intention-to-screen effects across pragmatic cancer screening trials.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 120, pp. e2311556120–, 12 2023.

[34] X. Yingchao, J. Zhou, and B. Zhao, "Attitude dynamics aiding for three-dimensional passive target tracking of strap-down seeker based on instrumental variable kalman filter:," *Transactions of the Institute of Measurement and Control*, vol. 42, pp. 2645–2659, 6 2020.

[35] W. He, B. Wang, and Z. Wang, "Will regional economic integration influence carbon dioxide marginal abatement costs? evidence from chinese panel data," *Energy economics*, vol. 74, pp. 263–274, 6 2018.

[36] L. Jia, X. Hu, Z. Zhao, T. Zhou, W. Liu, and B. He, "The impact of digitization on green innovation performance: Evidence based on panel data of 228 prefectural-level cities in china," *Complexity*, vol. 2022, 7 2022.

[37] J. A. Labrecque and S. A. Swanson, "Understanding the assumptions underlying instrumental variable analyses: a brief review of falsification strategies and related tools.," *Current epidemiology reports*, vol. 5, pp. 214–220, 6 2018.

[38] C. Zhang and R. Hu, "Does fertilizer use intensity respond to the urban-rural income gap? evidence from a dynamic panel-data analysis in china," *Sustainability*, vol. 12, pp. 430–, 1 2020.

[39] Q. Xu, Z. Cai, and Y. Fang, "Semiparametric inferences for panel data models with fixed effects via nearest neighbor difference transformation," *Econometric Reviews*, vol. 40, pp. 919–943, 8 2021.

[40] Y. Yang, "Does economic growth induce smoking?—evidence from china," *Empirical Economics*, vol. 63, no. 2, pp. 821–845, 2022.

[41] K. S. Kendler, H. Ohlsson, A. A. Fagan, P. Lichtenstein, J. Sundquist, and K. Sundquist, "Academic achievement and drug abuse risk assessed using instrumental variable analysis and co-relative designs.," *JAMA psychiatry*, vol. 75, pp. 1182–1188, 11 2018.

[42] K. Du, Y. Zhang, and Q. Zhou, "Fitting partially linear functional-coefficient panel-data models with stata:," *The Stata Journal: Promoting communications on statistics and Stata*, vol. 20, pp. 976–998, 12 2020.

[43] D. Zhang, Y. Cui, and X. Zhang, "Estimating factors related to fluoroquinolone resistance based on one health perspective: Static and dynamic panel data analyses from europe.," *Frontiers in pharmacology*, vol. 10, pp. 1145–1145, 10 2019.

[44] Q. Jiang, X. Zhang, Q. Lin, G. Chen, R. Zhang, and S. Liu, "Internet penetration and regional financial development in china: Empirical evidence based on chinese provincial panel data," *Scientific Programming*, vol. 2020, pp. 1–14, 8 2020.

[45] G. Sienso, C. Lyford, and W. Oldewage-Theron, "Using instrumental variables to establish the relationship between household production diversity and household dietary diversity in northern ghana," *African Journal of Food, Agriculture, Nutrition and Development*, vol. 22, pp. 21036–21055, 10 2022.

[46] M. Ketokivi, P. Bromiley, and A. Awaysheh, "Making theoretically informed choices in specifying panel-data models," *Production and Operations Management*, vol. 30, pp. 2069–2076, 7 2021.

[47] Q. Zhou and Z. Li, "The impact of industrial structure upgrades on the urban–rural income gap: An empirical study based on china's provincial panel data," *Growth and Change*, vol. 52, pp. 1761–1782, 6 2021.

[48] M. Wang and H. Zhuang, "How does foreign aid affect total fertility rate? panel data evidence," *International Economic Journal*, vol. 33, pp. 605–619, 5 2019.

[49] X. Yao, F. Durani, Q. R. Syed, H. H. Lean, and M. I. Tabash, "Does tourism promote green growth? a panel data analysis," *Environment, Development and Sustainability*, vol. 27, pp. 2411–2422, 10 2023.

[50] J. Ma, H. Yang, W. Hu, and H. T. A. Khan, "Spousal care intensity, socioeconomic status, and depression among the older caregivers in china: A study on 2011-2018 charls panel data.," *Healthcare (Basel, Switzerland)*, vol. 10, pp. 239–239, 1 2022.

[51] S. Li and L. Peng, "Instrumental variable estimation of complier causal treatment effect with interval-censored data.," *Biometrics*, vol. 79, pp. 253–263, 10 2021.

[52] M. A. Baloch, J. Zhang, K. Iqbal, and Z. Iqbal, "The effect of financial development on ecological footprint in bri countries: evidence from panel data estimation," *Environmental science and pollution research international*, vol. 26, pp. 6199–6208, 1 2019.

[53] E. H. Kennedy, J. Mauro, M. J. Daniels, N. Burns, and D. S. Small, "Handling missing data in instrumental variable methods for causal inference.," *Annual review of statistics and its application*, vol. 6, pp. 125–148, 11 2018.

[54] C. B. Rosen, S. E. Roberts, C. J. Wirtalla, L. J. Keele, E. J. Kaufman, S. D. Halpern, P. M. Reilly, M. D. Neuman, and R. R. Kelz, "The conditional effects of multimorbidity on operative versus nonoperative management of emergency general surgery conditions: A retrospective observational study using an instrumental variable analysis.," *Annals of surgery*, vol. 278, pp. e855–e862, 5 2023.

[55] Y. Shui, J. Zhong, Y. Yang, Y. Zeng, and S. Liu, "Employment transfer of rural female labor and family welfare effect in mountainous areas: An empirical analysis based on panel data," *Land*, vol. 11, pp. 2134–2134, 11 2022.

[56] S. Athey, M. Bayati, N. Doudchenko, G. Imbens, and K. Khosravi, “Matrix completion methods for causal panel data models,” *Journal of the American Statistical Association*, vol. 116, pp. 1716–1730, 5 2021.

[57] P. Cao, X. Ouyang, and J. Xu, “How do ecosystem services affect poverty reduction efficiency? a panel data analysis of state poverty counties in china.,” *International journal of environmental research and public health*, vol. 19, pp. 1886–1886, 2 2022.