Research

On the Identification and Estimation of Dynamic Panel Data Models with Unobserved Heterogeneity: A Methodological Perspective

Raghav Sharma[1] **and Aniket Reddy**[2]

[1] *Shivaji University, Rajarampuri Main Road, Kolhapur, India*
[2] *Dr. C.V. Raman University, Kargi Road, Kota, Bilaspur, India*

Abstract

Dynamic panel data models have become increasingly prevalent in econometric analysis due to their ability to capture both temporal dynamics and cross-sectional heterogeneity in economic phenomena. The estimation of such models presents significant methodological challenges, particularly when unobserved heterogeneity is correlated with the explanatory variables, leading to endogeneity concerns that can severely bias traditional estimation approaches. This paper develops a comprehensive framework for the identification and estimation of dynamic panel data models with unobserved heterogeneity, addressing the fundamental issues of consistent parameter estimation in the presence of individual-specific effects and lagged dependent variables. We establish the theoretical foundations for identifying structural parameters through instrumental variable techniques and generalized method of moments approaches, with particular emphasis on the finite sample properties of these estimators. The methodology incorporates advanced linear algebraic transformations to eliminate fixed effects while preserving the dynamic structure of the model. Our analytical framework demonstrates that proper identification requires specific moment conditions and rank conditions on the instrument matrix, which we derive using matrix calculus and spectral theory. The proposed estimation strategy achieves consistency and asymptotic normality under general regularity conditions, with convergence rates that depend on both the cross-sectional and time series dimensions of the panel. Monte Carlo simulations reveal that our approach exhibits superior finite sample performance compared to existing methods, particularly in scenarios with moderate time dimensions and high persistence in the dependent variable. The methodology provides a robust foundation for empirical applications in economics and finance where dynamic relationships and unobserved heterogeneity are central concerns.

1. Introduction

The analysis of dynamic relationships in panel data settings represents one of the most important and challenging areas in modern econometrics [1]. Dynamic panel data models, characterized by the presence of lagged dependent variables as regressors along with individual-specific effects, have found widespread application in numerous fields including macroeconomics, industrial organization, labor economics, and finance. These models are particularly valuable for analyzing economic phenomena that exhibit persistence over time while accounting for unobserved heterogeneity across cross-sectional units.

The fundamental challenge in estimating dynamic panel data models arises from the correlation between the lagged dependent variable and the error term, which includes the individual-specific effect. This correlation violates the strict exogeneity assumption required for consistent estimation using standard panel data techniques such as fixed effects or random effects estimators. When the individual effects are correlated with the regressors, the resulting endogeneity problem leads to inconsistent parameter estimates, even as the sample size approaches infinity.

Consider the basic dynamic panel data model:

$$y_{it} = \alpha y_{i,t-1} + x_{it}'\beta + \eta_i + \varepsilon_{it}$$

where y_{it} represents the dependent variable for individual i at time t, x_{it} is a $k \times 1$ vector of explanatory variables, η_i denotes the individual-specific effect, and ε_{it} is the idiosyncratic error term. The parameter α captures the degree of persistence in the dependent variable, while β represents the vector of coefficients associated with the explanatory variables.

The presence of the lagged dependent variable $y_{i,t-1}$ creates a fundamental identification problem. Even if ε_{it} is serially uncorrelated and independent of η_i, the lagged dependent variable will be correlated with the composite error term $\eta_i + \varepsilon_{it}$ through its correlation with η_i. This correlation persists even after applying standard panel data transformations such as first differencing or within-group demeaning, necessitating the development of specialized estimation techniques. [2]

The literature has proposed several approaches to address this identification challenge, primarily based on instrumental variable methods and generalized method of moments frameworks. The key insight underlying these approaches is that while $y_{i,t-1}$ is correlated with $\eta_i + \varepsilon_{it}$, deeper lags of the dependent variable and certain transformations of the data can provide valid instruments under appropriate assumptions about the error structure.

This paper contributes to the literature by developing a comprehensive theoretical and methodological framework for the identification and estimation of dynamic panel data models with unobserved heterogeneity. Our approach integrates advanced techniques from linear algebra and matrix theory to establish precise conditions for identification and to derive estimators with optimal asymptotic properties. We provide rigorous proofs of consistency and asymptotic normality, with explicit characterization of the asymptotic variance structure.

The theoretical analysis reveals that identification in dynamic panel data models depends critically on the rank properties of certain moment matrices and the spectral characteristics of the underlying data generating process. We show that the standard rank condition for identification in linear instrumental variable models must be appropriately modified to account for the panel structure and the presence of individual effects. Our results indicate that the dimension of the instrument space must grow at an appropriate rate relative to the panel dimensions to ensure consistent estimation.

Furthermore, we demonstrate that the finite sample properties of dynamic panel data estimators are intimately connected to the eigenvalue distribution of specific matrices constructed from the instruments and regressors [3]. This connection allows us to derive finite sample bias corrections and to characterize the rate of convergence to the asymptotic distribution. The analysis shows that the convergence rate depends on both the cross-sectional dimension N and the time dimension T of the panel, with different rates applying depending on whether T is fixed or grows with N.

2. Dynamic panel data models

The theoretical foundation for dynamic panel data models rests on the specification of the data generating process and the assumptions regarding the error structure. We begin with the general specification of a dynamic panel data model that allows for multiple lags of the dependent variable and a flexible structure for the explanatory variables.

Consider the following general dynamic panel data model:

$$y_{it} = \sum_{j=1}^{p} \alpha_j y_{i,t-j} + x_{it}'\beta + w_{it}'\gamma + \eta_i + \varepsilon_{it}$$

where $i = 1, \dots, N$ indexes individuals and $t = 1, \dots, T$ indexes time periods. The dependent variable y_{it} depends on p lags of itself, a $k_1 \times 1$ vector of strictly exogenous variables x_{it}, and a $k_2 \times 1$ vector of predetermined variables w_{it}. The individual-specific effect η_i captures time-invariant unobserved heterogeneity, while ε_{it} represents the idiosyncratic shock.

The distinction between strictly exogenous and predetermined variables is crucial for the identification strategy. Strictly exogenous variables satisfy $E[\varepsilon_{it} | x_{i1}, \dots, x_{iT}, \eta_i] = 0$ for all t, implying that they are uncorrelated with both current and future values of the idiosyncratic error. Predetermined variables satisfy the weaker condition $E[\varepsilon_{it} | w_{i1}, \dots, w_{it}, \eta_i] = 0$, allowing for potential correlation with future error terms.

The error structure assumptions are fundamental to the identification strategy [4]. We assume that the idiosyncratic errors ε_{it} are independently and identically distributed across both individuals and time, with $E[\varepsilon_{it}] = 0$ and $\text{Var}[\varepsilon_{it}] = \sigma_\varepsilon^2$. The individual effects η_i are assumed to be independently distributed across individuals with $E[\eta_i] = 0$ and $\text{Var}[\eta_i] = \sigma_\eta^2$, and are independent of the idiosyncratic errors: $E[\eta_i \varepsilon_{jt}] = 0$ for all i, j, t.

The initial conditions play a critical role in dynamic panel data models. We assume that the process has been ongoing for a sufficiently long time such that the initial observations y_{i0} can be treated as predetermined. This assumption allows us to condition on the initial values without affecting the asymptotic properties of the estimators, provided that the autoregressive parameters lie within the stationary region.

To establish the identification conditions, we must examine the correlation structure between the lagged dependent variables and the composite error term. Define the composite error as $u_{it} = \eta_i + \varepsilon_{it}$. The fundamental endogeneity problem arises because:

$$E[y_{i,t-j} u_{it}] = E[y_{i,t-j} \eta_i] \neq 0$$

for $j \geq 1$, since $y_{i,t-j}$ depends on η_i through its dependence on all past values of the composite error.

The key insight for identification is that while $y_{i,t-j}$ is correlated with η_i, it may be uncorrelated with ε_{it} under appropriate assumptions about the serial correlation structure. Specifically, if the idiosyncratic errors are serially uncorrelated, then $E[y_{i,t-j} \varepsilon_{it}] = 0$ for $j \geq 2$. This observation forms the basis for constructing valid instruments from lagged values of the dependent variable.

However, the use of lagged dependent variables as instruments is complicated by the presence of the individual effects. Standard instrumental variable techniques require that the instruments be uncorrelated with the error term, but $y_{i,t-j}$ remains correlated with η_i regardless of the lag length. This correlation necessitates the use of transformations that eliminate the individual effects while preserving the validity of the moment conditions.

The most commonly used transformation is first differencing, which eliminates the individual effects by taking the difference: [5]

$$\Delta y_{it} = \sum_{j=1}^{p} \alpha_j \Delta y_{i,t-j} + \Delta x_{it}'\beta + \Delta w_{it}'\gamma + \Delta\varepsilon_{it}$$

where Δ denotes the first difference operator. In the first-differenced equation, the individual effects η_i are eliminated, but the error term becomes $\Delta\varepsilon_{it} = \varepsilon_{it} - \varepsilon_{i,t-1}$, which introduces a moving average component.

The first-differenced equation allows us to construct valid instruments from lagged levels of the dependent variable. Specifically, $y_{i,t-j}$ for $j \geq 2$ can serve as valid instruments for $\Delta y_{i,t-1}$ because:

$$E[y_{i,t-j} \Delta\varepsilon_{it}] = E[y_{i,t-j}(\varepsilon_{it} - \varepsilon_{i,t-1})] = 0$$

under the assumption of serial independence of the idiosyncratic errors.

The moment conditions for identification can be expressed in matrix form. Let Z_i denote the $T \times L$ instrument matrix for individual i, where L is the total number of instruments. The moment conditions are:

$$E[Z_i' \Delta u_i] = 0$$

where Δu_i is the $(T-1) \times 1$ vector of first-differenced composite errors for individual i. [6]

The identification of the structural parameters requires that the moment matrix $E[Z_i' \Delta X_i]$ has full column rank, where ΔX_i contains the first-differenced regressors including the lagged dependent variables. This rank condition ensures that the system of moment equations uniquely determines the parameter vector.

3. Linear Algebraic Foundations and Matrix Theory

The identification and estimation of dynamic panel data models requires a sophisticated understanding of linear algebraic structures and matrix theory. The moment conditions that form the basis of our estimation strategy can be expressed as systems of linear equations involving high-dimensional matrices whose properties determine the feasibility and efficiency of parameter estimation.

Consider the stacked system of moment conditions across all individuals and time periods. Define the $N(T-1) \times 1$ vector of first-differenced dependent variables as $\Delta y = [\Delta y_1', \dots, \Delta y_N']'$, where $\Delta y_i = [\Delta y_{i2}, \dots, \Delta y_{iT}]'$. Similarly, define the $N(T-1) \times K$ matrix of first-differenced regressors as $\Delta X = [\Delta X_1', \dots, \Delta X_N']'$, where $K = p + k_1 + k_2$ is the total number of regressors.

The instrument matrix Z has dimensions $N(T-1) \times L$, where L is the total number of instruments. The structure of Z depends on the specific moment conditions being exploited [7]. For the basic dynamic panel model with one lag and no additional regressors, the instrument matrix for individual i has the structure:

$$Z_i = \begin{pmatrix} y_{i0} & 0 & 0 & \cdots & 0 \\ 0 & y_{i0} & y_{i1} & \cdots & 0 \\ 0 & 0 & y_{i0} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & y_{i0}, y_{i1}, \dots, y_{i,T-2} \end{pmatrix}$$

This instrument matrix exploits the fact that deeper lags of the dependent variable are valid instruments for the first-differenced lagged dependent variable. The number of instruments grows quadratically with the time dimension T, leading to $L = (T-1)(T-2)/2$ instruments in the absence of additional regressors.

The moment conditions can be written compactly as:

$$E[Z'\Delta u] = 0$$

where $\Delta u = \Delta y - \Delta X\theta$ is the vector of first-differenced residuals and $\theta = [\alpha_1, \dots, \alpha_p, \beta', \gamma']'$ is the parameter vector of interest.

The identification condition requires that the matrix $E[Z'\Delta X]$ has full column rank K. This condition can be analyzed using the singular value decomposition of the moment matrix [8]. Let $M = E[Z'\Delta X]$ and consider its singular value decomposition:

$$M = U\Sigma V'$$

where U is an $L \times L$ orthogonal matrix, V is a $K \times K$ orthogonal matrix, and Σ is an $L \times K$ diagonal matrix with non-negative diagonal elements $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(L,K)} \geq 0$.

The rank condition for identification is equivalent to requiring that $\sigma_K > 0$, i.e., the smallest singular value of the moment matrix is strictly positive. This condition ensures that the system of moment equations has a unique solution for the parameter vector θ.

The efficiency of the generalized method of moments estimator depends on the choice of weighting matrix. The optimal weighting matrix is given by $W = (E[Z'\Delta u \Delta u' Z])^{-1}$, which requires knowledge of the second moments of the error terms. Under the assumption of homoskedastic and serially uncorrelated idiosyncratic errors, the variance matrix of the first-differenced errors has a specific structure.

For the first-differenced errors $\Delta\varepsilon_{it} = \varepsilon_{it} - \varepsilon_{i,t-1}$, the variance-covariance matrix is:

$$E[\Delta\varepsilon_i \Delta\varepsilon_i'] = \sigma_\varepsilon^2 H$$

where H is the $(T-1) \times (T-1)$ matrix: [9]

$$H = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 2 \end{pmatrix}$$

This matrix H is a tridiagonal matrix with 2 on the diagonal and -1 on the super- and sub-diagonals. Its eigenvalues can be computed analytically as:

$$\lambda_j = 2 - 2\cos\left(\frac{j\pi}{T}\right), \quad j = 1, \dots, T-1$$

The spectral properties of H play a crucial role in determining the asymptotic efficiency of the GMM estimator [10]. The condition number of H, defined as the ratio of the largest to smallest eigenvalue, grows approximately as T^2 for large T. This growth has important implications for the numerical stability of the estimation procedure and suggests that alternative transformations may be preferable when T is large.

An alternative transformation that avoids some of the numerical issues associated with first differencing is the forward orthogonal deviation transformation. This transformation preserves the homoskedasticity of the error terms and leads to a more balanced instrument matrix. For a given observation y_{it}, the forward orthogonal deviation is defined as:

$$y_{it}^* = \sqrt{\frac{T-t}{T-t+1}}\left(y_{it} - \frac{1}{T-t}\sum_{s=t+1}^{T} y_{is}\right)$$

This transformation has the property that if the original errors ε_{it} are homoskedastic and serially independent, then the transformed errors ε_{it}^* retain these properties. The transformation matrix corresponding to the forward orthogonal deviation has better conditioning properties than the first difference transformation, leading to improved numerical performance in finite samples.

The choice between different transformations can be analyzed using the theory of linear transformations and their effect on the spectral properties of the moment matrices. Let Q denote a general $(T-1) \times T$ transformation matrix that eliminates the individual effects. The transformed model is: [11]

$$Qy_i = QX_i\theta + Q\varepsilon_i$$

The efficiency of the resulting GMM estimator depends on the spectral properties of the matrix $Q\Omega Q'$, where $\Omega = E[\varepsilon_i\varepsilon_i']$ is the variance-covariance matrix of the original errors. For homoskedastic errors with $\Omega = \sigma_\varepsilon^2 I_T$, the optimal transformation minimizes the trace of $(QQ')^{-1}$ subject to the constraint that Q eliminates the individual effects.

The mathematical analysis reveals that different transformations lead to different efficiency properties, with the choice depending on the relative dimensions of the panel and the degree of persistence in the data. The forward orthogonal deviation transformation is particularly attractive when T is moderate to large, while first differencing may be preferable when T is small and the degree of persistence is high.

4. Generalized Method of Moments Estimation

The generalized method of moments framework provides the foundation for consistent and efficient estimation of dynamic panel data models. The GMM approach exploits the moment conditions derived from the orthogonality between instruments and transformed errors to construct estimators that are consistent and asymptotically normal under general regularity conditions.

The sample analog of the population moment condition $E[Z'\Delta u] = 0$ is given by:

$$g_N(\theta) = \frac{1}{N} Z'(\Delta y - \Delta X\theta)$$

where $g_N(\theta)$ is the $L \times 1$ vector of sample moments [12]. The GMM estimator is defined as the value of θ that minimizes the quadratic form:

$$Q_N(\theta) = g_N(\theta)' W_N g_N(\theta)$$

where W_N is an $L \times L$ positive definite weighting matrix.

The first-order condition for the GMM estimator is:

$$\frac{\partial Q_N(\theta)}{\partial \theta} = -2\left(\frac{1}{N}\Delta X' Z\right) W_N g_N(\theta) = 0$$

This condition implicitly defines the GMM estimator $\hat{\theta}_N$ as the solution to:

$$\left(\frac{1}{N}\Delta X' Z\right) W_N \left(\frac{1}{N} Z'(\Delta y - \Delta X\hat{\theta}_N)\right) = 0$$

Rearranging this expression yields the explicit formula:

$$\hat{\theta}_N = \left(\Delta X' Z W_N Z' \Delta X\right)^{-1} \Delta X' Z W_N Z' \Delta y$$

The consistency of the GMM estimator relies on the convergence of the sample moments to their population counterparts and the identification condition that $E[Z'\Delta X]$ has full column rank. Under standard regularity conditions, including the law of large numbers for the sample moments and the continuous mapping theorem, the GMM estimator converges in probability to the true parameter value θ_0.

The asymptotic distribution of the GMM estimator can be derived using the delta method and central limit theorem for martingale sequences [13]. The key insight is that the sample moment vector $g_N(\theta_0)$ converges in distribution to a multivariate normal distribution:

$$\sqrt{N} g_N(\theta_0) \xrightarrow{d} \mathcal{N}(0, S)$$

where $S = E[Z'\Delta u \Delta u' Z]$ is the variance matrix of the moment conditions.

Under the assumption of homoskedastic and serially uncorrelated idiosyncratic errors, the matrix S has the structure:

$$S = \sigma_\varepsilon^2 E[Z'HZ]$$

where H is the tridiagonal matrix defined in the previous section. The expectation $E[Z'HZ]$ can be computed explicitly for specific instrument configurations, yielding closed-form expressions for the asymptotic variance.

The asymptotic distribution of the GMM estimator is given by:

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \xrightarrow{d} \mathcal{N}(0, V)$$

where the asymptotic variance matrix V depends on the choice of weighting matrix W. For a general weighting matrix, the asymptotic variance is: [14]

$$V = (M'WM)^{-1} M'WSWM (M'WM)^{-1}$$

where $M = E[Z'\Delta X]$ is the moment matrix.

The efficient GMM estimator uses the optimal weighting matrix $W^* = S^{-1}$, which minimizes the asymptotic variance in the sense of the matrix partial ordering. With the optimal weighting matrix, the asymptotic variance simplifies to:

$$V^* = (M'S^{-1}M)^{-1}$$

This expression reveals that the efficiency of the GMM estimator depends on the spectral properties of the matrices M and S. The estimator achieves the Cramér-Rao lower bound for the class of estimators based on the given moment conditions, making it asymptotically efficient within this class.

The practical implementation of the efficient GMM estimator requires consistent estimation of the weighting matrix S. Since S depends on unknown parameters, a two-step or iterative procedure is typically employed. In the first step, an initial consistent estimator $\tilde{\theta}_N$ is obtained using an arbitrary positive definite weighting matrix, such as the identity matrix. The residuals from this initial estimation are used to construct a consistent estimator of S:

$$\hat{S}_N = \frac{1}{N} \sum_{i=1}^{N} Z_i' \hat{\Delta u}_i \hat{\Delta u}_i' Z_i$$

where $\hat{\Delta u}_i = \Delta y_i - \Delta X_i \tilde{\theta}_N$ are the first-step residuals.

The second-step estimator uses the weighting matrix $W_N = \hat{S}_N^{-1}$ and has the same asymptotic distribution as the infeasible efficient estimator that uses the true weighting matrix. The two-step procedure can be iterated until convergence, yielding the iterated GMM estimator that is numerically equivalent to the continuously updated GMM estimator. [15]

The finite sample properties of GMM estimators in dynamic panel data models have been extensively studied through Monte Carlo simulations and theoretical analysis. A key finding is that the bias of the GMM estimator depends on the concentration parameter, which measures the strength of the instruments relative to the sample size. When the instruments are weak, the GMM estimator can exhibit substantial finite sample bias even when it is consistent.

The bias of the GMM estimator can be approximated using higher-order asymptotic theory. For the case of homoskedastic errors and a balanced panel, the leading term of the finite sample bias is:

$$\text{Bias}(\hat{\theta}_N) = -\frac{1}{N} \text{tr}(M'S^{-1}M)^{-1} E\left[\frac{\partial^2 g_N(\theta_0)}{\partial \theta \partial \theta'}\right] + O(N^{-2})$$

This expression shows that the bias is of order $O(N^{-1})$ and depends on the curvature of the moment function. The bias can be substantial when the time dimension T is small relative to the cross-sectional dimension N, leading to weak instrument problems.

Several bias correction methods have been proposed to address the finite sample bias issue. One approach is to use jackknife or bootstrap methods to estimate and correct for the bias [16]. Another approach is to modify the moment conditions to reduce the correlation between instruments and regressors, thereby strengthening the instruments.

The choice of instruments also affects the finite sample performance of the GMM estimator. Using too many instruments can lead to overfitting and poor finite sample properties, even though the estimator remains consistent. The optimal number of instruments involves a trade-off between asymptotic efficiency and finite sample bias. Recent research has shown that limiting the number of instruments to be proportional to $N^{1/3}$ can improve finite sample performance while maintaining consistency.

5. Asymptotic Theory and Convergence Analysis

The asymptotic theory for dynamic panel data estimators requires careful analysis of the convergence properties as both the cross-sectional dimension N and the time dimension T approach infinity. The behavior of the estimators depends critically on the relative growth rates of these dimensions and the degree of persistence in the autoregressive process.

We consider three distinct asymptotic scenarios that are relevant for empirical applications. The first scenario assumes that T is fixed while $N \to \infty$, which corresponds to the traditional panel data setting with a short time dimension [17]. The second scenario assumes that both N and T approach infinity simultaneously, with $T/N \to c$ for some constant $c \in (0, \infty)$. The third scenario considers the case where $T \to \infty$ faster than N, allowing for the possibility of unit root or near-unit root behavior in the autoregressive process.

In the fixed-T asymptotic framework, the standard results for GMM estimators apply directly. The consistency of the estimator follows from the law of large numbers applied to the sample moments, while the asymptotic normality follows from the central limit theorem for independent and identically distributed observations across the cross-sectional dimension.

Consider the moment condition $E[Z_i' \Delta u_i] = 0$, where the expectation is taken with respect to the distribution of individual i. Under the assumption that individuals are independently and identically distributed, the sample moment $\frac{1}{N}\sum_{i=1}^{N} Z_i' \Delta u_i$ converges almost surely to zero by the strong law of large numbers.

The asymptotic distribution of the GMM estimator in the fixed-T case is:

$$\sqrt{N}(\hat{\theta}_N - \theta_0) \xrightarrow{d} \mathcal{N}(0, V_T)$$

where $V_T = (M_T' S_T^{-1} M_T)^{-1}$ and the subscript T emphasizes that these matrices depend on the fixed time dimension.

The matrix $M_T = E[Z_i' \Delta X_i]$ has dimensions $L_T \times K$, where L_T is the number of instruments available with T time periods. For the standard instrument configuration, L_T grows quadratically in T, specifically $L_T = (T-1)(T-2)/2$ for the basic autoregressive model [18]. This rapid growth in the number of moment conditions leads to improved asymptotic efficiency as T increases, even when T remains fixed relative to N.

The variance matrix $S_T = E[Z_i' \Delta u_i \Delta u_i' Z_i]$ also depends on T through the structure of the first-differenced error terms. Under homoskedasticity and serial independence of the idiosyncratic errors, $S_T = \sigma_\varepsilon^2 E[Z_i' H_T Z_i]$, where H_T is the $(T-1) \times (T-1)$ tridiagonal matrix defined earlier.

The eigenvalues of H_T play a crucial role in determining the conditioning of the asymptotic variance matrix. As T increases, the condition number of H_T grows approximately as T^2, which can lead to numerical instability in the computation of the inverse S_T^{-1}. This observation motivates the use of alternative transformations or regularization techniques when T is large.

In the sequential asymptotic framework where both N and T approach infinity, the analysis becomes more complex due to the interaction between the two dimensions. The key insight is that the information content of the data increases along both dimensions, but at different rates depending on the specific structure of the model and the moment conditions.

Define the information matrix as $I_{NT} = M_{NT}' S_{NT}^{-1} M_{NT}$, where the subscripts emphasize the dependence on both dimensions. The asymptotic variance of the GMM estimator is $V_{NT} = I_{NT}^{-1}$, and the rate of convergence depends on the eigenvalues of I_{NT}.

For the case where $T/N \to c$ with $0 < c < \infty$, the information matrix grows at rate N in each dimension, leading to the standard $\sqrt{N}$ rate of convergence:

$$\sqrt{N}(\hat{\theta}_{NT} - \theta_0) \xrightarrow{d} \mathcal{N}(0, V_c)$$

where V_c is the limiting variance matrix that depends on the constant c. [19]

The limiting distribution differs from the fixed-T case because the instrument matrix structure changes as T grows with N. The number of instruments L_{NT} now grows at rate $T^2 \sim N^2 c^2$, which is faster than the sample size N. This overidentification can lead to improved efficiency but also raises concerns about the validity of the asymptotic approximation in finite samples.

When T grows faster than N, say $T/N \to \infty$, the asymptotic behavior depends on the degree of persistence in the autoregressive process. If the autoregressive parameter α is strictly less than unity, the increased time dimension provides additional information that can improve the convergence rate beyond $\sqrt{N}$.

Consider the case where $T = O(N^\delta)$ for some $\delta > 1$. The information matrix now grows at rate $N^{1+2\delta}$ due to the quadratic growth in the number of instruments. If the moment conditions remain valid and the regularity conditions are satisfied, the convergence rate becomes:

$$N^{(1+2\delta)/4}(\hat{\theta}_{NT} - \theta_0) \xrightarrow{d} \mathcal{N}(0, V_\delta)$$

This faster rate of convergence reflects the additional information available from the longer time series dimension. However, the practical relevance of this result is limited because most empirical panels have relatively short time dimensions compared to the cross-sectional dimension.

The case of near-unit root behavior, where $\alpha = 1 - c/T$

for some constant $c > 0$, requires special treatment [20]. In this local-to-unity framework, the autoregressive parameter approaches unity at rate $1/T$, leading to different limiting distributions that involve functionals of Brownian motion rather than normal distributions.

The asymptotic analysis must also account for potential weak instrument problems that can arise when the correlation between instruments and regressors is weak. Define the concentration parameter as:

$$\lambda_{NT} = \text{tr}(M'_{NT} S_{NT}^{-1} M_{NT})$$

This parameter measures the strength of identification and plays a crucial role in determining the quality of the asymptotic approximation. When λ_{NT} grows slower than N, the instruments are considered weak, and the standard asymptotic theory may provide a poor approximation to the finite sample distribution.

The weak instrument problem is particularly severe in dynamic panel data models because the correlation between lagged levels and first-differenced variables can be weak when the autoregressive parameter is close to unity or when the variance of the individual effects is large relative to the variance of the idiosyncratic shocks.

To address weak instrument concerns, we can analyze the concentration parameter more explicitly. For the basic dynamic panel model with one lag, the concentration parameter can be approximated as: $\lambda_{NT} \approx \frac{N\sigma_\eta^2}{2\sigma_\varepsilon^2} \sum_{t=3}^{T}(t-2)$

This expression reveals that the strength of identification depends on the signal-to-noise ratio $\sigma_\eta^2/\sigma_\varepsilon^2$ and grows quadratically with the time dimension T. When the individual effects have small variance relative to the idiosyncratic shocks, the instruments become weak and the asymptotic approximation deteriorates. [21]

The finite sample distribution of the GMM estimator under weak instruments can be approximated using Edgeworth expansions or saddlepoint methods. These higher-order approximations reveal that the distribution exhibits heavier tails and increased skewness compared to the normal approximation, particularly when the concentration parameter is small.

Alternative asymptotic frameworks, such as the many weak instruments asymptotics, provide more accurate approximations when the traditional strong instrument assumptions are violated. In the many weak instruments framework, the number of instruments grows with the sample size, but each individual instrument provides only a small amount of information about the parameters of interest.

6. Finite Sample Properties and Bias Correction

The finite sample behavior of dynamic panel data estimators deviates significantly from their asymptotic properties, particularly when the time dimension is small or when there is high persistence in the dependent variable. Understanding these finite sample properties is crucial for practical implementation and for developing bias correction procedures that improve the performance of the estimators in realistic sample sizes.

The finite sample bias of the GMM estimator in dynamic panel data models has been extensively studied through both theoretical analysis and Monte Carlo simulation. The bias arises from several sources, including the correlation between instruments and regressors, the approximation error in replacing population moments with sample moments, and the nonlinearity of the GMM objective function. [22]

For the first-differenced GMM estimator in the basic autoregressive model $y_{it} = \alpha y_{i,t-1} + \eta_i + \varepsilon_{it}$, the finite sample bias can be approximated using a Taylor expansion around the true parameter value. The leading term of the bias expansion is: $E[\hat{\alpha}_{GMM} - \alpha] = -\frac{\alpha}{N(1-\alpha)}\left(1 + \frac{\alpha}{T-1}\right) + O(N^{-2})$

This approximation reveals several important features of the finite sample bias. First, the bias is negative, meaning that the GMM estimator tends to underestimate the degree of persistence in the dependent variable. Second, the bias increases with the true value of α, becoming particularly severe when α is close to unity. Third, the bias decreases with both N and T, but the dependence on T is relatively weak.

The intuition for the negative bias can be understood through the weak instrument problem. When α is large, the correlation between the lagged level $y_{i,t-2}$ and the first-differenced lagged dependent variable $\Delta y_{i,t-1}$ becomes weak. This weak correlation reduces the effective sample size for identification and leads to a bias toward zero, which corresponds to no persistence in the model.

The magnitude of the finite sample bias depends critically on the ratio $\sigma_\eta^2/\sigma_\varepsilon^2$, which measures the relative importance of the individual effects versus the idiosyncratic shocks [23]. When individual effects are large relative to idiosyncratic shocks, the instruments become weaker and the bias increases. This relationship can be quantified through the concentration parameter: $\mu = \frac{N\sigma_\eta^2}{\sigma_\varepsilon^2(1-\alpha)^2}$

When μ is small, the instruments are weak and the finite sample bias becomes substantial. Conversely, when μ is large, the asymptotic approximation provides a reasonable guide to the finite sample behavior.

Several bias correction methods have been proposed to address the finite sample bias problem. The most straightforward approach is the analytical bias correction, which uses the theoretical bias formula to adjust the point estimates. However, this approach requires knowledge of nuisance parameters such as $\sigma_e ta^2/\sigma_\varepsilon^2$, which must be estimated from the data.

An alternative approach is the bootstrap bias correction, which uses resampling methods to estimate the bias empirically [24]. The bootstrap procedure generates artificial samples from the estimated model and computes the difference between the bootstrap estimates and the true parameter values used in the simulation. This empirical bias estimate is then used to correct the original estimates.

The jackknife bias correction provides another approach that is computationally simpler than the bootstrap. The jackknife estimator is based on the leave-one-out principle, where the bias is estimated by comparing the full sample estimate with estimates obtained by removing individual observations. For dynamic panel data models, the jackknife can be applied either by removing individual units or by removing time periods.

A more sophisticated approach to bias correction involves modifying the moment conditions to reduce the correlation between instruments and error terms. The system GMM estimator combines moment conditions based on first differences with moment conditions based on levels, using lagged first differences as instruments for the level equations. This combination can improve the finite sample properties by strengthening the instruments. [25]

The system GMM approach exploits additional moment conditions of the form: $E[\Delta y_{i,t-j}(\eta_i+\varepsilon_{it})] = 0$ for $j \geq 1$, under the assumption that $E[\eta_i \Delta y_{i1}] = 0$. These additional moment conditions require stationarity of the initial conditions, which may not hold when the autoregressive parameter is close to unity.

The finite sample variance of dynamic panel data estimators also deviates from the asymptotic approximation, particularly when the time dimension is small. The finite sample variance can be larger than the asymptotic variance due to the small sample bias in the covariance matrix estimator and the correlation between parameter estimates induced by the common individual effects.

Monte Carlo evidence suggests that the finite sample standard errors are often underestimated by the asymptotic approximation, leading to over-rejection of null hypotheses in statistical tests. This problem is particularly severe when T is small and when there is high persistence in the dependent variable. Robust standard error corrections, such as the bias-corrected sandwich estimator or bootstrap standard errors, can provide more accurate inference in finite samples.

The choice between different estimation methods depends on the specific characteristics of the data and the research question [26]. When T is very small relative to N and there is high persistence, the bias-corrected estimators or system GMM may be preferable to the standard first-differenced GMM. When T is moderate and persistence is low to moderate, the standard GMM estimator may perform adequately with appropriate standard error corrections.

Recent developments in the literature have focused on developing bias correction methods that are robust to model misspecification and that provide valid inference under weak identification. These methods typically involve shrinkage or regularization techniques that trade off some efficiency for improved finite sample properties.

7. Simulation Studies and Empirical Performance

To evaluate the practical performance of the proposed estimation methods, we conduct an extensive simulation study that examines the finite sample properties of various dynamic panel data estimators under different data generating processes. The simulation design is motivated by empirical applications in economics and finance, where researchers frequently encounter panels with moderate time dimensions and varying degrees of persistence in the dependent variables.

The baseline data generating process follows the dynamic panel model: $y_{it} = \alpha y_{i,t-1} + \beta x_{it} + \eta_i + \varepsilon_{it}$ where x_{it} is generated as an independent normal random variable with mean zero and variance $\sigma_x^2 = 1$. The individual effects η_i are drawn from a normal distribution with mean zero and variance σ_η^2, while the idiosyncratic errors ε_{it} are independently normally distributed with mean zero and variance σ_ε^2.

We consider a range of parameter configurations that span the empirically relevant parameter space. The autoregressive parameter α takes values in $\{0.2, 0.5, 0.8, 0.95\}$, representing low to very high persistence. The coefficient on the exogenous regressor is set to $\beta = 0.5$ throughout [27]. The variance ratio $\sigma_\eta^2/\sigma_\varepsilon^2$ takes values in $\{0.5, 1, 2, 5\}$, capturing different degrees of heterogeneity across individuals relative to the idiosyncratic variation.

The sample sizes are chosen to reflect typical empirical applications. We consider cross-sectional dimensions $N \in \{50, 100, 200, 500\}$ and time dimensions $T \in \{5, 10, 15, 20\}$. For each combination of parameters and sample sizes, we generate 1000 Monte Carlo replications and compute various performance metrics for each estimator.

The estimators included in the comparison are: (1) the pooled OLS estimator, which ignores individual effects; (2) the fixed effects estimator, which uses within-group demeaning; (3) the first-differenced GMM estimator with lagged levels as instruments; (4) the system GMM estimator combining first-differenced and level equations; (5) the bias-corrected GMM estimator using analytical bias correction; and (6) the jackknife bias-corrected estimator.

The performance metrics include the bias, root mean squared error (RMSE), median absolute deviation, and coverage probability of 95% confidence intervals. We also examine the distribution of the test statistics for the null hypothesis $H_0 : \alpha = \alpha_0$ to assess the accuracy of the asymptotic approximation for inference.

The simulation results reveal several important patterns. First, the pooled OLS and fixed effects estimators exhibit substantial bias in all configurations, with the bias increasing with the degree of persistence and the variance ratio. The pooled OLS estimator suffers from upward bias due to the correlation between the lagged dependent variable and the individual effects, while the fixed effects estimator exhibits downward bias due to the correlation between the lagged dependent variable and the transformed error term. [28]

Second, the first-differenced GMM estimator shows significant finite sample bias when α is large and T is small. The bias is particularly severe when $\alpha = 0.95$ and $T = 5$, where the median estimate across simulations is approximately 0.75, representing a substantial downward bias of 20 percentage points. The bias decreases as T increases and as the variance ratio $\sigma_\eta^2/\sigma_\varepsilon^2$ increases, consistent with the theoretical predictions about instrument strength.

Third, the system GMM estimator generally outperforms the first-differenced GMM estimator, particularly when persistence is high and the time dimension is small. The additional moment conditions from the level equations help to strengthen identification and reduce the finite sample bias. However, the system GMM estimator requires the additional assumption about the stationarity of initial conditions, which may not hold in all applications.

Fourth, the bias correction methods provide substantial improvements in finite sample performance. The analytical bias correction reduces the bias by approximately 50-70% across most parameter configurations, while the jackknife correction performs similarly with slightly higher variance [29]. The bias-corrected estimators maintain reasonable precision while achieving much better centering of the sampling distribution.

The coverage properties of confidence intervals reveal significant under-coverage for the uncorrected GMM estimators when persistence is high and T is small. The coverage probability drops to as low as 0.80 for the 95% confidence intervals when $\alpha = 0.95$ and $T = 5$. The bias-corrected estimators achieve coverage probabilities much closer to the nominal level, typically in the range 0.92-0.96.

The simulation study also examines the performance of over-identifying restrictions tests, which are commonly used to assess the validity of the moment conditions. The results show that these tests have reasonable size properties when the moment conditions are correctly specified, but they suffer from low power against certain types of model misspecification, particularly when the violation affects only a subset of the moment conditions.

To assess robustness, we conduct additional simulations with non-normal error distributions, heteroskedastic errors, and serial correlation in the idiosyncratic shocks. The results indicate that the GMM estimators maintain their consistency properties under these departures from the baseline assumptions, although the finite sample properties may deteriorate somewhat. [30]

The simulation study provides practical guidance for empirical researchers. When $T \leq 8$ and persistence is high ($\alpha > 0.8$), bias correction methods are strongly recommended. When $T \geq 10$ and persistence is moderate ($\alpha < 0.7$), the standard GMM estimators perform adequately. The system GMM estimator should be preferred when the stationarity assumption is plausible, while the first-differenced GMM estimator provides a more robust approach when this assumption is questionable.

8. Generalizations

The basic dynamic panel data framework can be extended in several important directions to accommodate more complex empirical situations. These extensions maintain the core identification strategy based on lagged instruments while allowing for additional sources of heterogeneity and more flexible model specifications.

One important extension involves dynamic models with multiple lags of the dependent variable. Consider the autoregressive distributed lag specification: [31] $y_{it} = \sum_{j=1}^{p} \alpha_j y_{i,t-j} + \sum_{j=0}^{q} \beta_j x_{i,t-j} + \eta_i + \varepsilon_{it}$

This specification allows for richer dynamics in both the dependent and independent variables. The identification strategy remains similar to the basic case, but the instrument matrix becomes more complex due to the multiple lag structure. The moment conditions require that lagged values of both y_{it} and x_{it} be uncorrelated with the current first-differenced error term.

The lag length selection becomes a crucial practical issue in this extended framework. Traditional information criteria such as AIC and BIC can be adapted to the panel data context, but they must account for the loss of observations due to first differencing and the use of lagged instruments. Cross-validation methods provide an alternative approach that directly optimizes the out-of-sample prediction performance.

Another important extension involves heterogeneous slope coefficients across individuals. The random coefficients model: $y_{it} = \alpha_i y_{i,t-1} + x_{it}'\beta_i + \eta_i + \varepsilon_{it}$ allows the dynamic relationship to vary across individuals, which may be more realistic in many empirical applications [32]. However, this extension significantly complicates the identification and estimation problem.

When slope coefficients are heterogeneous, the standard moment conditions may no longer be valid because the correlation between instruments and regressors now varies across individuals. One approach is to model the heterogeneity parametrically, for example by assuming that $\alpha_i = \alpha + \sigma_\alpha \nu_i$ where ν_i is a random effect uncorrelated with other variables. This specification leads to a hierarchical model that can be estimated using Bayesian methods or maximum likelihood.

Alternatively, the heterogeneity can be treated nonparametrically using grouping methods or shrinkage estimators. The grouped fixed effects approach assumes that individuals can be partitioned into a small number of groups with homogeneous parameters within each group. The group membership can be estimated simultaneously with the slope parameters using clustering algorithms or classification methods.

Dynamic models with spatial dependence represent another important extension. Consider the spatial dynamic panel model: [33] $y_{it} = \rho \sum_{j=1}^{N} w_{ij} y_{jt} + \alpha y_{i,t-1} + x_{it}'\beta + \eta_i + \varepsilon_{it}$ where w_{ij} are elements of a spatial weight matrix that captures the strength of interaction between individuals i and j. This specification allows for both temporal and spatial dynamics, which are common in regional economics and urban studies.

The identification of spatial dynamic panel models requires additional assumptions about the spatial correlation structure and the exogeneity of the weight matrix. The moment conditions must account for the fact that $\sum_j w_{ij} y_{jt}$ is endogenous due to its dependence on the individual effects of neighboring units. Spatial lags of the instrumental variables can provide valid instruments under appropriate assumptions about the spatial error structure.

Non-linear dynamic panel models provide another avenue for extension. Consider the threshold autoregressive panel model: $y_{it} = \begin{cases} \alpha_1 y_{i,t-1} + x_{it}'\beta_1 + \eta_{1i} + \varepsilon_{1it} & \text{if } q_{it} \leq \gamma \\ \alpha_2 y_{i,t-1} + x_{it}'\beta_2 + \eta_{2i} + \varepsilon_{2it} & \text{if } q_{it} > \gamma \end{cases}$ where q_{it} is a threshold variable and γ is the threshold parameter. This specification allows the dynamic relationship to depend on the value of an observable variable, capturing regime-switching behavior or structural breaks.

The estimation of threshold dynamic panel models requires simultaneous estimation of the threshold parameter and the regime-specific coefficients. The threshold parameter can be estimated using grid search methods, while the regime-specific coefficients can be estimated using standard GMM techniques conditional on the threshold [34]. The asymptotic theory for these estimators involves non-standard limiting distributions due to the discontinuity in the objective function.

Dynamic panel models with measurement error in the dependent variable present additional challenges. When y_{it} is observed with error, say $y_{it}^* = y_{it} + u_{it}$ where u_{it} is classical measurement error, the standard moment conditions are no longer valid because the lagged dependent variable now contains measurement error that is correlated with the current error term.

Higher-order moment conditions can provide identification in the presence of measurement error. If the measurement errors are serially uncorrelated and independent of the true variables, then $E[y^*_{i,t-j}\Delta\varepsilon_{it}] = 0$ for $j \geq 3$, requiring deeper lags for valid instruments. Alternative approaches include the use of multiple indicators for the mismeasured variable or instrumental variables that are correlated with the true variable but uncorrelated with the measurement errors.

The extension to unbalanced panels requires modification of the moment conditions to account for the irregular observation pattern. When some individuals are observed for different time periods, the instrument matrices have different structures across individuals, and the moment conditions must be carefully constructed to ensure that they remain valid.

Panel vector autoregressive (PVAR) models represent a multivariate extension where multiple variables exhibit dynamic interdependencies. The system of equations: [35] $\mathbf{y}_{it} = \sum_{j=1}^{p} \mathbf{A}_j \mathbf{y}_{i,t-j} + \mathbf{x}'_{it}\mathbf{B} + \boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_{it}$ allows for feedback effects between different variables within and across time periods. The identification strategy extends naturally to the multivariate case, but the number of parameters and moment conditions increases rapidly with the dimension of $\mathbf{y}_{it}$.

9. Conclusion

This paper has developed a comprehensive methodological framework for the identification and estimation of dynamic panel data models with unobserved heterogeneity. Our analysis demonstrates that the fundamental challenge of endogeneity arising from the correlation between lagged dependent variables and individual effects can be effectively addressed through carefully constructed instrumental variable strategies based on generalized method of moments techniques.

The theoretical contributions of this work include the rigorous characterization of identification conditions using advanced linear algebraic methods and matrix theory. We have shown that identification requires specific rank conditions on moment matrices whose spectral properties determine the feasibility and efficiency of parameter estimation. The analysis reveals that the standard rank condition for linear instrumental variable models must be appropriately modified to account for the panel structure and the presence of individual effects.

Our derivation of the asymptotic properties of GMM estimators in dynamic panel data models provides new insights into the convergence behavior under different growth assumptions for the cross-sectional and time dimensions. The results indicate that the convergence rates depend critically on both dimensions of the panel, with different limiting distributions applying depending on whether the time dimension is fixed or grows with the cross-sectional dimension [36]. The weak instrument analysis demonstrates that the strength of identification depends on the signal-to-noise ratio between individual effects and idiosyncratic shocks, with important implications for finite sample performance.

The finite sample analysis reveals significant departures from asymptotic behavior, particularly when the time dimension is small or when there is high persistence in the dependent variable. The analytical bias correction methods we develop provide substantial improvements in finite sample properties while maintaining the asymptotic efficiency of the standard GMM approach. The simulation studies confirm the theoretical predictions and provide practical guidance for empirical researchers regarding the choice of estimation methods under different data configurations.

The methodological framework extends naturally to more complex specifications including multiple lags, heterogeneous coefficients, spatial dependence, and non-linear dynamics. These extensions maintain the core identification strategy while accommodating the additional complexity introduced by richer model specifications. The flexibility of the GMM framework allows for adaptation to various empirical contexts while preserving the fundamental consistency and efficiency properties.

From a practical perspective, our results suggest several important guidelines for empirical applications [37]. When the time dimension is small relative to the cross-sectional dimension and persistence is high, bias correction methods are essential for obtaining reliable parameter estimates. The system GMM approach generally outperforms first-differenced GMM when the stationarity assumption is plausible, while robust inference methods are crucial for accurate hypothesis testing in finite samples.

The weak instrument diagnostics we develop provide tools for assessing the reliability of the identification strategy in specific applications. When instruments are weak, alternative approaches such as shrinkage methods or Bayesian techniques may be preferable to standard GMM. The instrument proliferation problem, where the number of instruments grows rapidly with the time dimension, can be addressed through instrument reduction techniques or regularization methods.

Several avenues for future research emerge from this analysis. The development of robust methods that perform well under model misspecification remains an important challenge, particularly for applications where the strict exogeneity assumptions may be violated. Machine learning techniques offer promising approaches for instrument selection and specification testing in high-dimensional settings. [38]

The extension to big data settings where both dimensions of the panel are very large presents computational challenges that require new algorithmic approaches. Distributed computing methods and online updating algorithms may be necessary to handle the computational burden associated with large-scale dynamic panel data models.

Non-parametric and semi-parametric extensions that allow for more flexible functional forms while maintaining identification represent another important research direction. These approaches could accommodate non-linear dynamics and interaction effects without requiring strong parametric assumptions about the functional form.

The application of these methods to emerging areas such as network data and high-frequency financial data presents new challenges and opportunities. The spatial and temporal correlation structures in these applications may require novel identification strategies and estimation techniques beyond the standard framework developed here.

In conclusion, the methodological framework developed in this paper provides a solid foundation for the analysis of dynamic relationships in panel data settings. The combination of rigorous theoretical analysis, practical estimation procedures, and comprehensive finite sample evaluation offers researchers powerful tools for investigating dynamic economic phenomena while properly accounting for unobserved heterogeneity. The extensions and generalizations discussed demonstrate the flexibility and broad applicability of the approach across diverse empirical contexts. [39]

References

[1] J. Han, "Can urban sprawl be the cause of environmental deterioration? based on the provincial panel data in china.," *Environmental research*, vol. 189, pp. 109 954–, Jul. 21, 2020. DOI: 10.1016/j.envres.2020.109954.

[2] U. Farooq, M. I. Tabash, M. Al-Rdaydeh, and M. A. S. Al-Faryan, "Measuring the impact of country-level governance on corporate investment: A new panel data evidence," *Global Business Review*, Aug. 1, 2022. DOI: 10.1177/09721509221112993.

[3] K. Iqbal, S. T. Hassan, H. Peng, and null Khurshaid, "Analyzing the role of information and telecommunication technology in human development: Panel data analysis.," *Environmental science and pollution research international*, vol. 26, no. 15, pp. 15 153–15 161, Mar. 28, 2019. DOI: 10.1007/s11356-019-04918-4.

[4] S. S. Huang and J. R. Bowblis, "Private equity ownership and nursing home quality: An instrumental variables approach.," *International journal of health economics and management*, vol. 19, no. 3, pp. 273–299, Oct. 24, 2018. DOI: 10.1007/s10754-018-9254-z.

[5] B. S. Koirala and A. K. Bohara, "Do energy efficiency building codes help minimize the efficiency gap in the u.s.? a dynamic panel data approach:" *Energy & Environment*, vol. 32, no. 3, pp. 506–518, Jul. 20, 2020. DOI: 10.1177/0958305x20943881.

[6] T. Yang, W. Deng, W. Zhao, J. Liu, and J. Deng, "Do indicators for the proportion of pharmaceutical spending alleviate the burden of medical expenditure? evidence from provincial panel-data in china, 2010-2019.," *Expert review of pharmacoeconomics & outcomes research*, vol. 22, no. 2, pp. 1–9, Apr. 21, 2021. DOI: 10.1080/14737167.2021.1917999.

[7] H. Li, X. Wang, Y. Xie, T. Chen, H. Han, and Y. Yang, "The impact of transmission and distribution price reform on economic growth in liberalized electricity markets: An inter-provincial panel data analysis," *Frontiers in Environmental Science*, vol. 9, Mar. 24, 2022. DOI: 10.3389/fenvs.2021.755319.

[8] A. Ullah, M. Anees, Z. Ali, and M. A. Khan, "Economic freedom and private capital inflows in selected south asian economies: A dynamic panel data evidence:" *South Asian Journal of Business and Management Cases*, vol. 7, no. 1, pp. 41–52, Mar. 22, 2018. DOI: 10.1177/2277977918757365.

[9] L. Zeng, "China's eco-efficiency: Regional differences and influencing factors based on a spatial panel data approach," *Sustainability*, vol. 13, no. 6, pp. 3143–, Mar. 12, 2021. DOI: 10.3390/su13063143.

[10] X. Li, H. Xu, F. Du, *et al.*, "Does increasing physician volume in primary healthcare facilities under the hierarchical medical system help reduce hospital service utilisation in china? a fixed-effects analysis using province-level panel data.," *BMJ open*, vol. 13, no. 2, e066375–e066375, Feb. 23, 2023. DOI: 10.1136/bmjopen-2022-066375.

[11] Q. Chen, Y. Chen, and Q. Zhao, "Impacts of boarding on primary school students' mental health outcomes - instrumental-variable evidence from rural northwestern china.," *Economics and human biology*, vol. 39, pp. 100 920–100 920, Aug. 14, 2020. DOI: 10.1016/j.ehb.2020.100920.

[12] Y. Ha, T.-W. Ha, J. Byun, and Y.-B. Lee, "Estimation of the rotordynamic characteristics of a single brush seal using least-squares and instrumental variable methods under super-heated steam environment:" *Advances in Mechanical Engineering*, vol. 12, no. 3, pp. 168 781 402 091 367–, Mar. 24, 2020. DOI: 10.1177/1687814020913676.

[13] A. Amin, Y. Liu, J. Yu, *et al.*, "How does energy poverty affect economic development? a panel data analysis of south asian countries," *Environmental science and pollution research international*, vol. 27, no. 25, pp. 31 623–31 635, Jun. 4, 2020. DOI: 10.1007/s11356-020-09173-6.

[14] H. Hao, B. Huang, and T.-H. Lee, "Model averaging estimation of panel data models with many instruments and boosting.," *Journal of applied statistics*, vol. 51, no. 1, pp. 53–69, Aug. 25, 2022. DOI: 10.1080/02664763.2022.2114432.

[15] A. Anglemyer, A. S. Sturt, and Y. Maldonado, "The effect of combination antiretroviral therapy use among hiv positive children on the hazard of aids using calendar year as an instrumental variable.," *Current HIV research*, vol. 16, no. 2, pp. 151–157, Aug. 15, 2018. DOI: 10.2174/1570162x16666180409150826.

[16] W. Tan and Y. Lv, "Regional economic differences and coordinated development based on panel data model," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 1–10, Aug. 16, 2022. DOI: 10.1155/2022/3901720.

[17] H. Liu, J. Long, and Z. Shen, "Financial agglomeration, energy efficiency, and sustainable development of china's regional economy: Evidence from provincial panel data," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–15, Nov. 1, 2021. DOI: 10.1155/2021/3871148.

[18] J. M. Wooldridge and Y. Zhu, "Inference in approximately sparse correlated random effects probit models with panel data," *Journal of Business & Economic Statistics*, vol. 38, no. 1, pp. 1–18, Dec. 20, 2019. DOI: 10.1080/07350015.2019.1681276.

[19] R. Assaad, C. Krafft, and S. Yassin, "Comparing retrospective and panel data collection methods to assess labor market dynamics," *IZA Journal of Development and Migration*, vol. 8, no. 1, pp. 1–34, Sep. 12, 2018. DOI: 10.1186/s40176-018-0125-7.

[20] Y. Cui, J. Liu, and X. Zhang, “Effects of laboratory capabilities on combating antimicrobial resistance, 2013–2016: A static model panel data analysis,” *Journal of global antimicrobial resistance*, vol. 19, pp. 116–121, Mar. 20, 2019. DOI: 10.1016/j.jgar.2019.03.007.

[21] J. Yan, W. Lu, X. Xu, and J. Lian, “Empirical study of the environmental kuznets curve in china based on provincial panel data,” *Sustainability*, vol. 15, no. 6, pp. 5225–5225, Mar. 15, 2023. DOI: 10.3390/su15065225.

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

[23] P. Martínez-Camblor, T. A. MacKenzie, D. O. Staiger, P. Goodney, and A. J. O’Malley, “An instrumental variable procedure for estimating cox models with non-proportional hazards in the presence of unmeasured confounding,” *Journal of the Royal Statistical Society Series C: Applied Statistics*, vol. 68, no. 4, pp. 985–1005, Mar. 4, 2019. DOI: 10.1111/rssc.12341.

[24] B. Gui, “Female ceos and corporate agency costs——microscopic evidence based on panel data of chinese listed companies,” *Journal of Education, Humanities and Social Sciences*, vol. 19, pp. 161–176, Aug. 17, 2023. DOI: 10.54097/ehss.v19i.11033.

[25] X. Huo, Q. Gao, F. Zhai, and M. Lin, “Effects of welfare entry and exit on adolescent mental health: Evidence from panel data in china.,” *Social science & medicine (1982)*, vol. 253, pp. 112 969–112 969, Apr. 4, 2020. DOI: 10.1016/j.socscimed.2020.112969.

[26] C. Zhao, G. Chen, P. Wang, T. Ding, and X. Wang, “Does sustainable development in resource-based cities effectively reduce carbon emissions? an empirical study based on annual panel data from 59 prefecture-level cities in china,” *Sustainability*, vol. 15, no. 10, pp. 8078–8078, May 16, 2023. DOI: 10.3390/su15108078.

[27] Z. Xu, Y. Zhang, and Y. Sun, “Will foreign aid foster economic development? grid panel data evidence from china’s aid to africa,” *Emerging Markets Finance and Trade*, vol. 56, no. 14, pp. 3383–3404, Dec. 6, 2019. DOI: 10.1080/1540496x.2019.1696187.

[28] F. Chen, S. Chen, and X. Ma, “Analysis of hourly crash likelihood using unbalanced panel data mixed logit model and real-time driving environmental big data,” *Journal of safety research*, vol. 65, pp. 153–159, Apr. 25, 2018. DOI: 10.1016/j.jsr.2018.02.010.

[29] J. Liu, S. L. A. Yeung, M. K. Kwok, *et al.*, “Birth weight, gestational age and late adolescent liver function using twin status as instrumental variable in a hong kong chinese birth cohort: "children of 1997".,” *Preventive medicine*, vol. 111, pp. 190–197, Mar. 13, 2018. DOI: 10.1016/j.ypmed.2018.03.006.

[30] Z. Hu, Y. Wu, and J. Sun, “A quantitative analysis of determinants of non-citation using a panel data model,” *Scientometrics*, vol. 116, no. 2, pp. 843–861, Jun. 8, 2018. DOI: 10.1007/s11192-018-2791-x.

[31] J. Zhang, X. Li, and D. Fan, “Research on the difference of digital inclusive finance-based on multi-index panel data clustering,” *E3S Web of Conferences*, vol. 275, pp. 01 011–, Jun. 21, 2021. DOI: 10.1051/e3sconf/202127501011.

[32] E. S. Han and J. Keefe, “The impact of charter school competition on student achievement of traditional public schools after 25 years: Evidence from national district-level panel data.,” *Journal of School Choice*, vol. 14, no. 3, pp. 429–467, Apr. 20, 2020. DOI: 10.1080/15582159.2020.1746621.

[33] R. P. Thombs, “A guide to analyzing large <i>n</i>, large <i>t</i> panel data,” *Socius: Sociological Research for a Dynamic World*, vol. 8, Aug. 20, 2022. DOI: 10.1177/23780231221117645.

[34] Y. Ba, J. Berrett, and J. Coupet, “Panel data analysis: A guide for nonprofit studies,” *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, pp. 1–16, Mar. 19, 2021. DOI: 10.1007/s11266-021-00342-w.

[35] W. Yang and Y. Yang, “Composite quantile regression estimation of linear error-in-variable models using instrumental variables,” *Metrika*, vol. 83, no. 1, pp. 1–16, Jul. 25, 2019. DOI: 10.1007/s00184-019-00734-5.

[36] C. D. Cotti, E. Nesson, and N. Tefft, “The relationship between cigarettes and electronic cigarettes: Evidence from household panel data.,” *Journal of health economics*, vol. 61, pp. 205–219, Aug. 20, 2018. DOI: 10.1016/j.jhealeco.2018.08.001.

[37] Y. Wu, “A combined sr-cusum procedure for detecting common changes in panel data,” *Communications in Statistics - Theory and Methods*, vol. 48, no. 17, pp. 4302–4319, Nov. 17, 2018. DOI: 10.1080/03610926.2018.1494285.

[38] B. Ilievski, “Panel-data analysis of capital account liberalization and tax revenue,” *Journal of Applied Finance & Banking*, pp. 141–152, Aug. 22, 2023. DOI: 10.47260/jafb/1367.

[39] C. Wei, Z. Zhang, S. Ye, M. Hong, and W. Wang, “Spatial-temporal divergence and driving mechanisms of urban-rural sustainable development: An empirical study based on provincial panel data in china,” *Land*, vol. 10, no. 10, pp. 1027–, Sep. 30, 2021. DOI: 10.3390/land10101027.