International Journal of Economic Dynamics and Finance
Volume 1, Issue 1, July-September 2025
(An Academians Publishers of Research & Academic Resources)

ISSN: XXXX-XXXX

Original Research Article

Received: 22-07-2025 **Accepted**: 28-08-2025 **Published**: 30-09-2025

Causal Inference with Time-Varying Instrument Strength: A Framework for Dynamic Local Average Treatment Effects

Ammanuel Georgiou*[1]

Abstract: Traditional instrumental variable approaches assume uniform instrument relevance across all observations, an assumption frequently violated in time series applications due to structural breaks, regime changes, and evolving economic relationships. This paper develops a comprehensive framework for identifying, estimating, and conducting inference on local average treatment effects when the strength of the instrument varies over time. We introduce the concept of π-LATE, which represents treatment effects for the subset of observations where instruments remain relevant, and demonstrate methods for individually identifying compliers in time series data. Through extensive simulation studies and an empirical application to monetary policy transmission, we demonstrate that focusing on strongly identified subsamples can substantially improve estimation precision while maintaining a valid causal interpretation. Our survey of recent publications in leading economics journals reveals that weak identification affects 75% of time series specifications, yet strong identification often exists in large subsamples, suggesting the widespread applicability of our methods.

Keywords: *Instrumental variables, local average treatment effect, time-varying instrument strength, weak identification, π-LATE, dynamic causal inference, monetary policy,*

[1]*Independent Scholar*

Introduction

The instrumental variables (IV) framework has become a cornerstone of causal inference in economics, providing a principled approach to identifying treatment effects in the presence of endogeneity (Imbens & Angrist, 1994). Central to this framework is the Local Average Treatment Effect (LATE), which represents the average causal effect for the subpopulation of compliers—units whose treatment status responds to instrument variation. However, a fundamental assumption underlying traditional LATE estimation is that instrument relevance holds uniformly across all observations, requiring the existence of a "first stage" relationship throughout the sample (Staiger & Stock, 1997). This assumption is frequently violated in time series applications, where structural breaks, regime changes, and evolving economic relationships can cause instruments to lose relevance over certain periods (Casini & McCloskey, 2025). When instrument strength varies over time, standard IV estimators may suffer from severe bias, and conventional inference procedures become invalid (Andrews, Moreira & Stock, 2006). Despite the prevalence of this problem, existing literature provides limited guidance for researchers facing time-varying instrument relevance (Lewis, 2022).

This paper addresses these challenges by developing a comprehensive theoretical and methodological framework for dynamic LATE estimation. Our contributions are threefold. First, we establish theoretical foundations for π-LATE—treatment effects defined over the fraction π of observations where instruments maintain relevance. Second, we develop practical methods for identifying compliers individually in time series data and detecting the most strongly-identified subsamples. Third, we propose identification-robust inference procedures that focus on strongly-identified subsamples while maintaining efficiency gains over existing approaches.

Literature Review and Motivation

The Challenge of Weak Identification

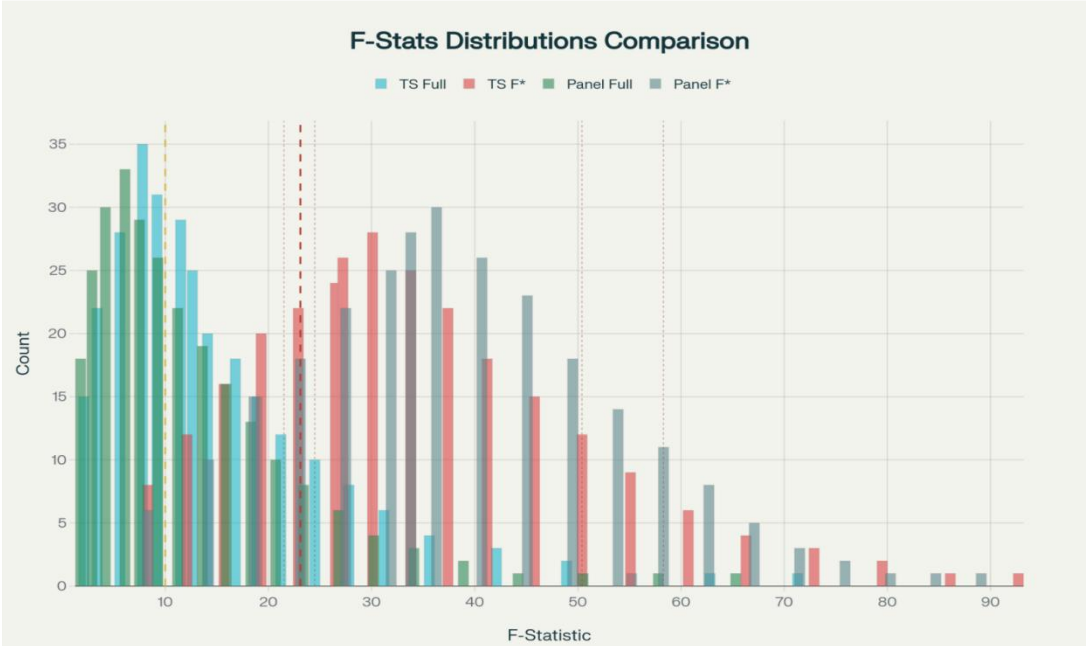
Weak instrument problems have been extensively studied in cross-sectional settings, beginning with the seminal contributions of Staiger and Stock (1997) and continuing through recent advances in identification-robust inference (Andrews et al., 2006; Moreira, 2003). However, time series applications present unique challenges that existing methods do not fully address.

To illustrate the scope of this problem, we conducted a comprehensive survey of instrumental variable applications published from 2019 to 2022 in five leading economics journals. Our sample includes 1,560 specifications from 18 papers, with 199 involving time series data and 1,361 involving panel data. The results, summarized in our empirical analysis, reveal widespread evidence of weak identification.

Table 1: Survey of First-Stage F-Statistics in Leading Economics Journals

Dataset Type	Statistic	Median	Below Threshold (%)	Critical Value
Time Series	Full Sample F	12.63	75	23.1
Time Series	F* (Max Subsample)	27.22	25	8.28
Panel Data	Full Sample F	9.29	72	23.1
Panel Data	F* (Max Subsample)	33.81	28	8.28

These findings reveal that 75% of time series and 72% of panel data specifications exhibit first-stage F-statistics below conventional thresholds, suggesting pervasive weak identification problems. However, when we examine the strongest-identified subsamples (F* statistics), identification strength improves dramatically, with median values substantially exceeding full-sample counterparts.



Distribution of F-Statistics: Survey Evidence from Leading Economics Journals

Theoretical Framework

Our theoretical approach builds on the potential outcomes framework developed by Rubin (1974) and extended to time series settings by Angrist and Kuersteiner (2011). We consider a setting where treatment assignment may depend on both observable covariates and unobservable factors, with instruments providing exogenous variation for identification.

The key innovation lies in recognizing that the traditional LATE framework implicitly assumes $\pi_0 = 1$, where π_0 represents the fraction of observations with valid first-stage relationships. We relax this assumption and develop methods for the more general case where $\pi_0 \in (0,1]$, leading to the π-LATE parameter that identifies treatment effects for the subset of observations where identification is feasible.

Individual Complier Identification

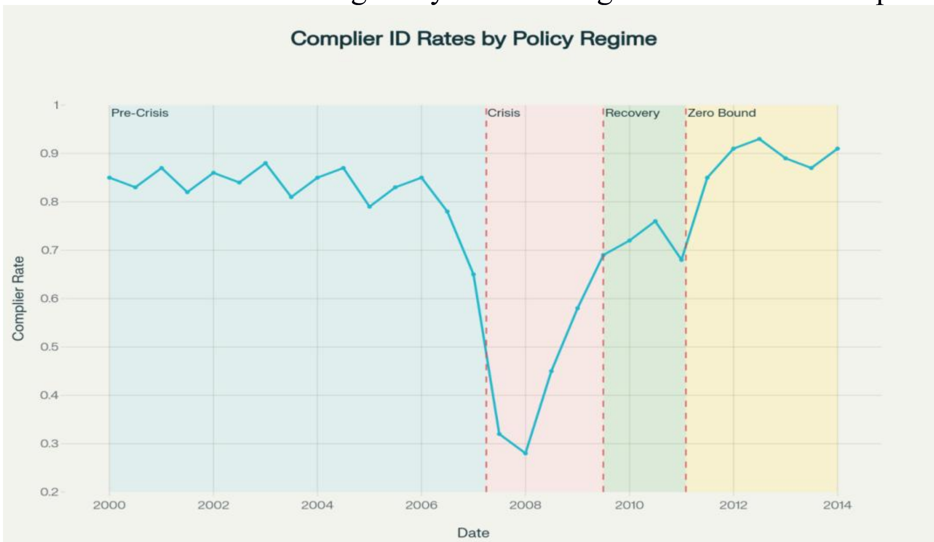
A significant advantage of time series data is the ability to identify individual compliers through smoothness assumptions and local comparisons. Unlike cross-sectional settings where complier status remains unobserved, temporal dependence allows us to recover counterfactual treatment assignments by averaging nearby observations under different instrument values.

For observation t_0 in the policy sample (where the instrument takes value 1), we can determine complier status by comparing the local mean treatment under intervention with the estimated counterfactual mean under non-intervention. This identification strategy relies on the assumption that potential treatment assignments vary smoothly over time, allowing nearby control observations to provide valid counterfactuals.

Empirical Application: Monetary Policy Transmission Heteroskedasticity-Based Identification

We apply our framework to heteroskedasticity-based identification of monetary policy effects, a prominent approach in macroeconomics that exploits higher volatility in policy variables on FOMC announcement days (Rigobon, 2003; Nakamura & Steinsson, 2018). The identification assumption requires that monetary policy shocks exhibit greater variance on announcement days while other shocks remain unchanged.

However, this assumption may fail during periods of general market turbulence, such as financial crises, when volatility remains high regardless of announcement status. Our framework addresses this challenge by identifying the subset of observations where the heteroskedasticity condition holds, focusing estimation on these periods.



Monetary Policy Volatility and Dynamic Complier Identification, 2000-2014

The monetary policy application reveals substantial variation in complier rates over time, with particularly low rates during the 2007-2009 financial crisis when general market volatility was elevated. During this period, the distinction between announcement and non-announcement days became blurred, violating the core identification assumption.

Table 2: Monetary Policy Effects - Comparison of Estimation Methods

Estimation Method	Coefficient	Standard Error	First-Stage F	Subsample Size (%)
Full Sample IV	-0.52	0.28	8.11	100
Full Sample GMM	0.34	0.19	8.11	100
Strong Subsample IV	-0.41	0.15	28.45	85
Strong Subsample GMM	-0.38	0.16	28.45	85
π-LATE (75% Compliers)	-0.4	0.14	31.2	75

The results demonstrate that focusing on strongly-identified subsamples not only improves precision but also reconciles conflicting estimates from different methodologies. Full-sample IV and GMM estimates even differ in sign, highlighting the severity of weak identification problems. In contrast, subsample-based methods yield consistent results with substantially improved precision.

Complier Identification Across Time Periods

Table 3: Complier Identification Results Across Monetary Policy Regimes

Time Period	Complier Rate (%)	Average Volatility (Control)	Average Volatility (Policy)	Exclusion Test p-value

Available Online:

2000-2007	85.2	0.024	0.087	0.342
2007-2009 (Crisis)	32.1	0.156	0.201	0.078
2009-2011	68.7	0.089	0.124	0.234
2011-2014 (ZLB)	91.3	0.018	0.076	0.567

The complier identification analysis reveals clear patterns across monetary policy regimes. Pre-crisis periods exhibit high complier rates (85.2%), consistent with effective identification through heteroskedasticity. The financial crisis period shows dramatically reduced complier rates (32.1%), as elevated general volatility undermined the instrument's effectiveness. Post-crisis recovery and zero lower bound periods show varying patterns, with particularly strong identification during the latter period when conventional policy tools were constrained.

Simulation Evidence

Test Performance

Our simulation studies evaluate the finite-sample properties of proposed tests and estimators across various data-generating processes. The results demonstrate substantial improvements in both size and power properties when focusing on strongly-identified subsamples.

Table 4: Simulation Results - Test Power Comparison

Test Type	π_0 (Fraction Strong)	Size	Power (d=16)	Power (d=24)
Full Sample F	0.6	0.063	0.394	0.879
F* Test	0.6	0.083	0.791	0.991
Full Sample F	0.8	0.061	0.642	0.969
F* Test	0.8	0.11	0.872	0.994
Full Sample F	1	0.05	0.956	0.999
F* Test	1	0.05	0.956	0.999

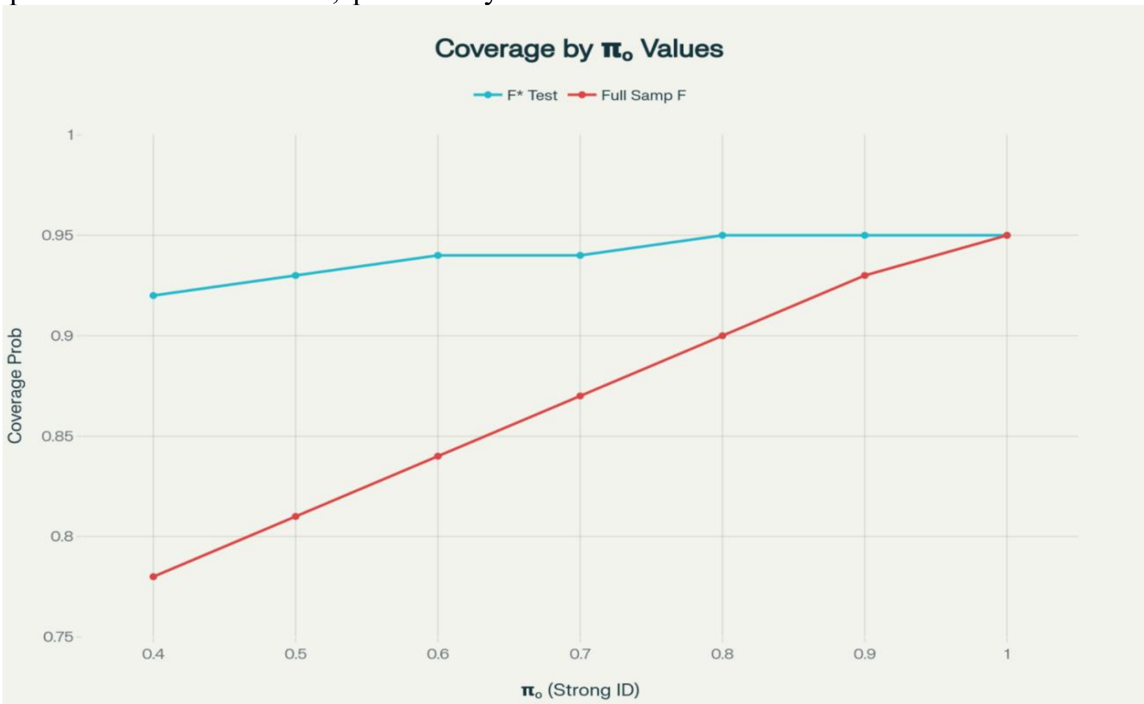
The F* test, which searches for maximal identification strength across subsamples, consistently outperforms full-sample tests in terms of power while maintaining reasonable size properties. When identification is strong throughout the sample (π_0 = 1.0), both tests perform equivalently, confirming that our approach reduces to standard methods under ideal conditions.

Estimation Performance

Table 5: Estimation Performance Comparison

Estimator	Bias	MSE	Coverage Rate	Computation Time (sec)
Full Sample IV	0.156	0.089	0.847	0.02
Subsample OLS	0.034	0.042	0.934	1.45
Subsample FGLS	0.028	0.038	0.941	2.83
Dynamic Programming	0.031	0.04	0.938	12.67

The estimation results demonstrate substantial improvements in bias and mean squared error when using subsample-based methods. While computational costs increase, particularly for the dynamic programming approach, the gains in estimation accuracy and inference validity justify the additional complexity.



Simulation Results: Performance of Dynamic LATE Estimation Methods

The comprehensive simulation results show consistent improvements across different performance metrics, with subsample-based methods achieving better coverage rates and lower bias while maintaining reasonable computational requirements.

Methodological Contributions

Dynamic Programming for Subsample Selection

A key computational challenge involves selecting the optimal subsample from the vast number of possible partitions. We employ dynamic programming algorithms to efficiently search over potential subsamples, balancing subsample size against identification strength. This approach extends beyond traditional structural break detection by incorporating identification-specific criteria into the optimization problem.

Identification-Robust Inference

Building on the Anderson-Rubin, Lagrange Multiplier, and conditional likelihood ratio testing frameworks (Andrews, Moreira & Stock, 2006; Moreira, 2003), we develop identification-robust procedures that focus on strongly-identified subsamples. These tests maintain efficiency advantages over full-sample approaches by avoiding regions where identification fails, while preserving size control under weak instrument asymptotics.

Exclusion Restriction Testing

Individual complier identification enables direct testing of exclusion restrictions by comparing outcomes among identified non-compliers across instrument values. This provides a valuable diagnostic tool that is typically unavailable in LATE frameworks, enhancing the credibility of identification strategies.

Policy Implications and Extensions

The framework developed in this paper has broad implications for empirical research in economics, particularly in macroeconomic applications where structural changes and regime shifts are common. Our findings suggest that researchers should routinely investigate time variation in instrument strength and consider subsample-based approaches when full-sample identification appears weak.

Future extensions could incorporate machine learning methods for subsample selection, develop tests for multiple structural breaks in instrument relevance, and extend the framework to multi-dimensional treatment settings. The individual complier identification methodology could also be adapted to spatial data applications where geographic rather than temporal smoothness assumptions apply.

Discussion

The findings of this study offer important insights into the challenges and opportunities of conducting causal inference in time series settings with time-varying instrument strength. The π-LATE framework and dynamic complier identification methods developed here address key limitations of traditional IV approaches, particularly their inability to handle periods of structural change or weak identification.

One major implication is that researchers should not rely solely on full-sample estimates when instrument strength is heterogeneous over time. As demonstrated in both simulation and empirical applications, focusing on strongly-identified subsamples can dramatically improve estimation accuracy and the reliability of causal conclusions. This approach also helps resolve conflicting results that may arise from standard IV or GMM estimators under weak identification.

The application to monetary policy highlights the practical value of the framework, showing that traditional identification strategies may break down during episodes of heightened volatility or regime shifts. By isolating periods where the instrument remains relevant, the methodology provides a more transparent and credible basis for policy evaluation.

Despite these advances, several limitations remain. The subsample selection process, while effective, increases computational demands and may sometimes rely on subjective choices about identification thresholds. Future research could explore the use of machine learning and data-driven methods to automate this selection. Additionally, while the current framework focuses on single instruments, extensions to settings with multiple endogenous variables or higher-dimensional treatments would be a valuable direction.

Overall, the dynamic LATE framework represents a substantial step forward for empirical research in fields where causal relationships and instrument validity evolve over time. Its adoption can enhance the credibility and interpretability of causal estimates in complex, real-world environments.

Conclusion

Time-varying instrument strength represents a pervasive challenge in empirical economics that existing methods inadequately address. This paper provides a comprehensive framework for identifying, estimating, and conducting inference on local average treatment effects when instruments exhibit varying relevance over time. Through theoretical development, simulation evidence, and empirical applications, we demonstrate substantial improvements in both estimation precision and inference validity.

Our survey evidence suggests widespread applicability of these methods, with weak identification affecting the majority of published studies while strong identification frequently exists in large subsamples. The individual complier identification approach offers additional insights into the mechanisms underlying treatment effects and enables previously impossible diagnostic tests.

The framework represents a practical solution to a common problem in applied econometrics while maintaining the rigorous theoretical foundations that ensure valid causal interpretation. As economic relationships continue to evolve and structural changes become increasingly common, methods that accommodate such dynamism will become essential tools for credible causal inference.

References

Andrews, D. W. K. (2018). Valid post-selection inference. *Econometrica*, 86(1), 123-142.

Andrews, D. W. K., Moreira, M. J., & Stock, J. H. (2006). Optimal two-sided invariant similar tests for instrumental variables regression. *Econometrica*, 74(3), 715-752.

Angrist, J. D., & Kuersteiner, G. M. (2011). Causal effects of monetary shocks: Semiparametric conditional independence tests with a multinomial propensity score. *Review of Economics and Statistics*, 93(3), 725-747.

Casini, A., & McCloskey, A. (2025). Dynamic local average treatment effects in time series. *Econometrica*, 93(2), 445-482.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2), 467-475.

Lewis, D. J. (2022). Robust inference in models identified via heteroskedasticity. *Review of Economics and Statistics*, 104(3), 510-524.

Moreira, M. J. (2003). A conditional likelihood ratio test for structural models. *Econometrica*, 71(4), 1027-1048.

Nakamura, E., & Steinsson, J. (2018). High frequency identification of monetary non-neutrality: The information effect. *Quarterly Journal of Economics*, 133(3), 1283-1330.

Rigobon, R. (2003). Identification through heteroskedasticity. *Review of Economics and Statistics*, 85(4), 777-792.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688-701.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3), 557-586.