Expanding the Urban-Rural Divide:

Evidence of Employment Recovery Disparities Following the 2008 Recession

Economics 970: Natural Experiments

Corbin Lubianski

A.B. Candidate, Statistics and Economics

ABSTRACT

This paper argues that metropolitan and nonmetropolitan counties have a different composition of unemployment types. This was concluded by analyzing unemployment data on metropolitan and nonmetropolitan counties between 2006 and 2016 using impulse response functions estimated by local projections with the 2008 Great Recession as the impulse shock. The results suggest that metropolitan counties have a larger composition of cyclical unemployment rate in comparison to nonmetropolitan counties. Additionally, the results suggest that nonmetropolitan counties have a larger composition of structural unemployment rate in comparison to metropolitan counties. A better understanding of the unemployment composition by metropolitan area might allocate labor force improvement programs more efficiently.

I would like to thank Ross Mattheis and my classmate colleagues in Econ 970: Natural Experiments for research advice and feedback throughout the writing process. Additionally, I would like to thank Professor Hanspeter Pfister and Liberty Vittert from Statistics 10: Introduction to Data Science for providing the initial motivation to pursue this topic and guidance to conduct exploratory data analysis.

I. INTRODUCTION

The divide between rural and urban strata appears within many aspects of society, culture, politics, industry, and education.[1] However, an area that lacks the perspective of urbanization is macroeconomics, specifically literature about the differences between urban (metropolitan or metro) and rural (nonmetropolitan or nonmetro) unemployment is strikingly sparse. More research on the disparities between metropolitan and nonmetropolitan unemployment can aid public policy in targeting urban and rural communities separately to optimize public resources and improve labor market conditions.

Unemployment, a form of labor market distress, serves as a good macroeconomics indicator of an economy's health. A low unemployment rate generally suggests a well-functioning economy, while a high unemployment rate reflects severe economic distress. The definition, the unemployment rate is the percentage of the labor force that is unemployed; yet the background surrounding unemployment rates are complex and require thoughtful analysis to infer the health of an economy. Potential workers are considered unemployed if they do not occupy a paid job, have actively looked for work recently, and are available for work.

The Great Recession (2007-2009) and the subsequent recovery period (2010-2019) reflect how uneven impacts of economic recession and recovery evolved into significantly different employment trends in rural and urban areas. Nonmetro counties lost 1.4 million jobs between 2007 and 2010 in addition to slow employment growth between 2010 and 2019. This coincided with the first recorded instance of a decline of nonmetropolitan county population between 2010 and 2016 which is associated with low demand for local goods and services (Pender et al., 2019). From 2007 to 2019, rural counties observed a decrease in labor force participation rates by 2.6 percent while urban counties observed a 0.7 percent increase which suggests that many rural workers chose between moving to urban areas for work and remaining in rural areas with fewer job prospects (Sanders, 2022). The demographic and industry shift exponentiated by the Great Recession is also critical to placing the difference of unemployment rate

[1] The terms "rural", "nonmetropolitan", and "nonmetro" are used interchangeably in addition to "urban", "metropolitan" and "metro" are also used interchangeably.

disparities into context as showcased below in Figure 1. In addition to the employment shift, the aggregated unemployment rate between 2000 and 2021 is shown below in Figure 2.

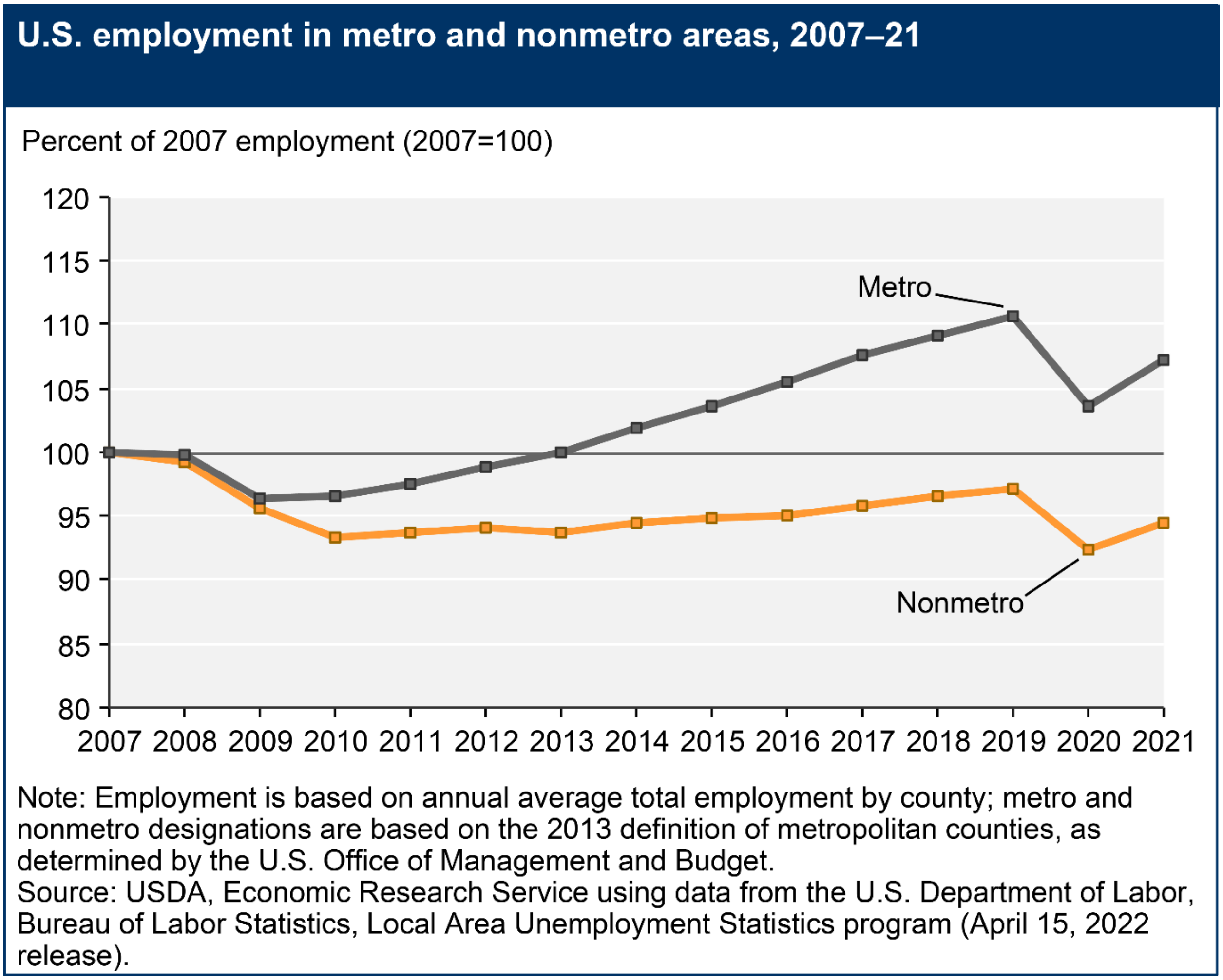

Figure 1: Data from U.S. Department of Labor showcasing the shift of employment by metro-nonmetro classification. (Sanders, 2022)

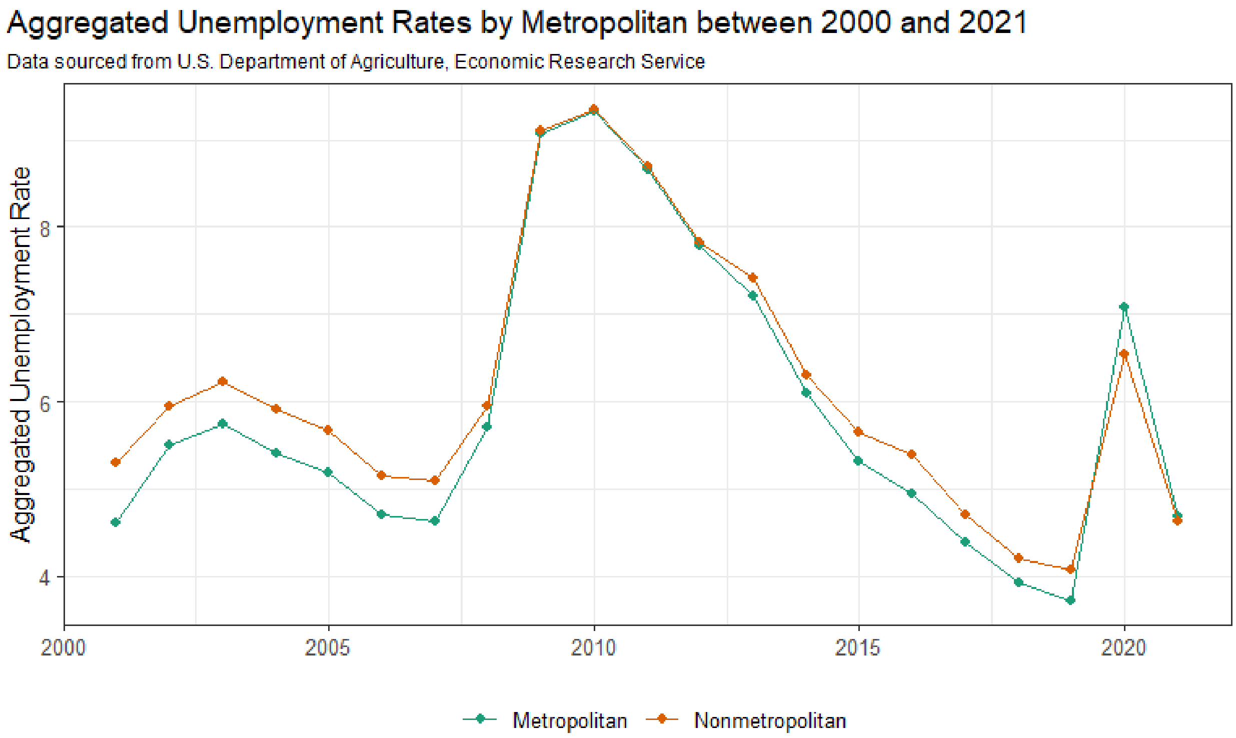

Figure 2: Data from U.S. Department of Labor showcasing corresponding changes in aggregated unemployment rate between metropolitan and nonmetropolitan counties.

Prior to the start of the Great Recession in 2007, metropolitan counties had a consistently lower unemployment rate than nonmetropolitan counties. However, the period between 2007 and 2012 saw nearly identical aggregate unemployment rates between metropolitan and nonmetropolitan counties, which after resumes the pre-2007 trends of consistently lower unemployment rate in urban areas. The shifts between the pre-recessionary (2000-2007), Great Recession (2007-2010), and recovery periods (2010-2019) suggest an underlying difference of the economies between urban and rural counties.

II. DATA and METHODOLOGY

Data collected from the Economic Research Service (ERS) at the United States Department of Agriculture provides unemployment and median household income data at the U.S. county level from 2000 to 2021 (Sanders, 2023). The data attributes include county name, FIPS code, state, metropolitan classification as determined by the ERS, civilian labor force, employed persons, unemployed persons, and the unemployment rate. The data is conveniently situated as a panel data set because there are 3129 U.S. counties that consistently have unemployment data for 22 years from 2000 to 2021.

To determine whether the behaviors of urban and rural county unemployment rates behave differently beyond the observed rates collected by the ERS, it is crucial to construct a model using time series with temporal dependence that utilizes local projections to estimate impulse response functions (Sims, 1980). A local projection model is a set of linear regressions that account for the associations between exogenous variables and an endogenous variable observed at different time points. Local projections are ideal in this context since they are more robust to misspecification, incorporate non-linearity if needed, and analytic inference is simple (Jordá, 2005). As a result, local projections are a common method to quantitatively estimate impulse response functions. In comparison, vector autoregressive (VAR) models have been shown to be mathematically equivalent to local projections such that they estimate the same impulse response. However, this only applies to asymptotic sample sizes and when the lag structure is unrestricted (Plagborg-Møller, 2021). Therefore, local projections are an ideal method to incorporate temporal dependence of unemployment rates from 2000 to 2021.

Impulse response functions require a shock variable to study its effects on the endogenous variable on other variables. The shock variable is implemented through the residual error term to determine the predictive effect of the shock variable to the endogenous variable. The local projection approach with the shock variable models the relation between the response variable h-steps ahead to the current shock locally for each horizon h. Horizons are the select years that the local projections method will estimate. Since unemployment rates between years are highly covaried with other years' unemployment rates, unemployment rate as a shock variable will easily identify how the exogenous variables have an impact.

The local projection method should consider the county gross domestic product (GDP) and civilian labor force as exogenous variables. The county-wide GDP data is available at the Bureau of Economic Analysis by selecting “CAGDP2: GDP in Current Dollars by County and MSA” in the “Gross Domestic Product (GDP)” dropdown. With a comparatively low population in rural counties in comparison to urban counties, it is possible that a low civilian labor force could more easily cause larger shifts in magnitude of the unemployment rates which might be inferred from the model that rural counties have a larger fluctuating unemployment rate yet is relatively stable. Between 2000 and 2021, nonmetropolitan counties had an average labor force size of 6932.52 while metropolitan counties had an average labor force size of 42822.46. Therefore, it is a reasonable variable to consider conditioning. Civilian labor force and GDP were transformed by a natural logarithm transformation as is customary for right skewed empirical distribution data.

Local projections aim to capture the impact of a certain event on endogenous variables. While the available years from the data span from 2000 to 2021, the relevant years to subset the panel data were determined from 2006 to 2016 to adequately estimate the 2008 Great Recession impact on unemployment rate given metropolitan classification, civilian labor force, and GDP.

Consider the following regression model with panel data:

$$y_{i,t+h} = \alpha_{i,h} + shock_{i,t}\beta_h + x_{i,t}\gamma_h + \epsilon_{i,t+h}, \ \ with\ h\ =\ 0, 1, \cdots, H\ -\ 1$$

where $\alpha_{i,h}$ denotes the county cross-section fixed effect, $x_{i,t}$ is a vector of the control variables, and $shock_{i,t}$ is the shock variable. H denotes the horizon which the model incorporates lagged parameters such that the model is incorporating previous estimates to make a better predictive model. Here, $y_{i,t+h}$ represents the unemployment rate of the county i at time $t + h$. The R package **lpirfs** is suitable for estimating local projections, especially with panel data through linear and non-linear models (Adämmer, 2019).

To extrapolate causal relationships from impulse response function, several identifying assumptions require discussion. The selection bias of the metropolitan classification must be zero for the impulse response function to equal the average treatment effect (Rambachan, 2021). The selection bias is dependent on how the assignment of whether a cross-effect is related to its past assignments, contemporaneous assignments, future assignments, and the potential outcome projection. However, beyond the assumptions of no selection bias, empirically estimating impulse response functions is challenging which local projections offer nonparametric and parametric solutions by regressing on the h-step on the cross-effect constant and assignment (Rambachan, 2021). Let Y_{t+h} be the potential outcome at lag $t + h$ and $W_{k,t}$ be the assignment of cross-effect k at time t, then the local projection estimand is defined as:

$$LP_{k,t,h} = \frac{Cov(Y_{t+h}, W_{k-t})}{Var\left(W_{k,t}\right)}.$$

To properly estimate the local projection, the potential outcome Y_{t+h} must be differentiable over assignment W_k and independent with respect to the assignment W_k.

To depart from nonparametric impulse response functions, local projections can achieve closer identification assumptions similar to ordinary least squares (OLS) regression to allow the results to be causal and interpretive. Local projections are based on the assumption that the relationship between variables is locally linear and constant over time. With uncertainty around the data-generating process, local projections do not assume any relationship about the data-generating distribution but has a trade-off of assuming no omitted variables that are correlated with the dependent and explanatory variables (Gonçalves, 2022). Lastly, local projections allow nonlinear methods and flexible specifications like the underlying covariance assumptions that may be impractical to compute in a multivariate model.

In the context of unemployment data at the county level from 2000 to 2021, local projections accommodate the county-wide level of the data since it might be unreasonable to place distributional assumptions beyond the Central Limit Theorem. By accounting for state-dependent responses more easily, the effects of fiscal policies and responses are more prevalent to showcase the impact of economic booms and slumps, especially useful in showing how the 2008 Great Recession impacted macroeconomics (Jordá, 2005).

A significant concern behind local projections' OLS characteristics is the possibility of violating the exogeneity assumption needed for a causal estimate, especially in the field of macroeconomics when many components of the economy are likely dependent on other components. To remedy this concern, panel data can correct the potential estimator inconsistency and bias of the estimates by using fixed effects to eliminate the endogeneity of the regressors. It is important to note that the fixed effects local projection estimator is consistent under weak exogeneity instead of strict exogeneity. A sufficiently large panel data of 3129 counties in 22 years is a reasonably size to assume weak exogeneity.

III. RESULTS

The computed impulse response function from a nonlinear local projections approach showcases that there is significant separation between the estimated coefficient of shock coefficient for the unemployment rate that suggests some underlying distinct characteristics that metropolitan and nonmetropolitan counties observe. Figure 3 presents the evolution of the shock coefficient throughout the horizon from 2006 to 2016 that suggests that the unemployment rate becomes less associated with lagged labor force and GDP as the economy recovers from the 2008 Great Recession. The plot shows nonlinear impulse responses for Regimes 1 (nonmetro) and Regimes 2 (metro). Figure 3 looks similar to Figure 2 as it is plotting the coefficient estimate of unemployment rate which should approximately follow the historical changes given that it is the endogenous and shock variable. During 2008, the exogenous variables civilian labor force and GDP were highly predictive of the unemployment rate which is to be expected with a major economic recession. The magnitude of the shock variable following 2008 can be interpreted as the impact of the recession's changes to the labor force and GDP diminishing the impact the variables had on the change of the unemployment rate.

While Figure 3 allows for visual evidence that metropolitan and nonmetropolitan unemployment rates recovered statistically differently after comparing their 95% confidence intervals, the OLS estimates from the nonlinear local projections approach in Table 1 allows for more nuanced discussion about the unemployment disparities. Note that while the standard deviations of the metropolitan appear to be smaller in the plot and the regression table, this represents the fact that metropolitan counties generally had a smaller unemployment rate standard deviation of 2.411% in comparison to nonmetropolitan unemployment rate standard deviation of 2.823%. This is in stark contrast to metropolitan's higher standard deviation in log GDP at 1.684 and log civilian labor force at 1.415 while nonmetropolitan has a lower standard deviation in log GDP at 1.097 and log civilian labor force at 1.031.

For the regression coefficients of the non-lagged variables in the local projections OLS estimation, the magnitudes of the log civilian labor force and the frequency of the statistically significant log GDP estimates are notably important to discuss. For the nonmetropolitan regime, log of GDP was inconsistent in testing statistically significant such that the test shown insufficient evidence to reject the null hypothesis that log GDP is statistically different from 0 in its predictive power. In 2015 and 2016 where it reached statistical significance at the $\alpha = 0.05$ level, it is difficult to reasonably contribute the significance to the 2008 Great Recession as years prior were not statistically significant.

In addition to log GDP's statistical significance from *t*-tests of the regression coefficients, the difference in magnitudes between metropolitan and nonmetropolitan coefficients is significant when glancing at the confidence intervals such that nonmetropolitan unemployment rates have a lower magnitude of its coefficient estimate of log GDP than metropolitan. Since metropolitan county unemployment rates have a much stronger association with log GDP than nonmetropolitan counties, then this might indicate that metropolitan county unemployment encompasses more cyclical unemployment.

The positive association between the log civilian labor force and unemployment rate is not surprising, yet the difference of the magnitude of the associations between metropolitan and nonmetropolitan county unemployment rates is also significant. While both regimes unsurprisingly have statistically significant labor force estimates for their respective unemployment rates, the larger magnitude of the nonmetropolitan civilian labor force coefficient estimates might indicate that nonmetropolitan counties have higher structural unemployment. Larger log civilian labor force coefficient estimates can be interpreted that as the total number of workers enter the labor force or job market, the unemployment rates are more sensitive to those increases. Larger magnitudes coefficient estimates might indicate that nonmetropolitan counties cannot easily reallocate their workforce such that they have a larger structural unemployment rate.

In Figure 2, the nonmetropolitan unemployment rate recovered slower than the metropolitan unemployment rate in addition to metropolitan unemployment rate shrinking below the nonmetropolitan unemployment rate. This relationship is not present in Figure 3 which suggests that conditional on civilian labor force and the county GDP, the county unemployment rates converged.

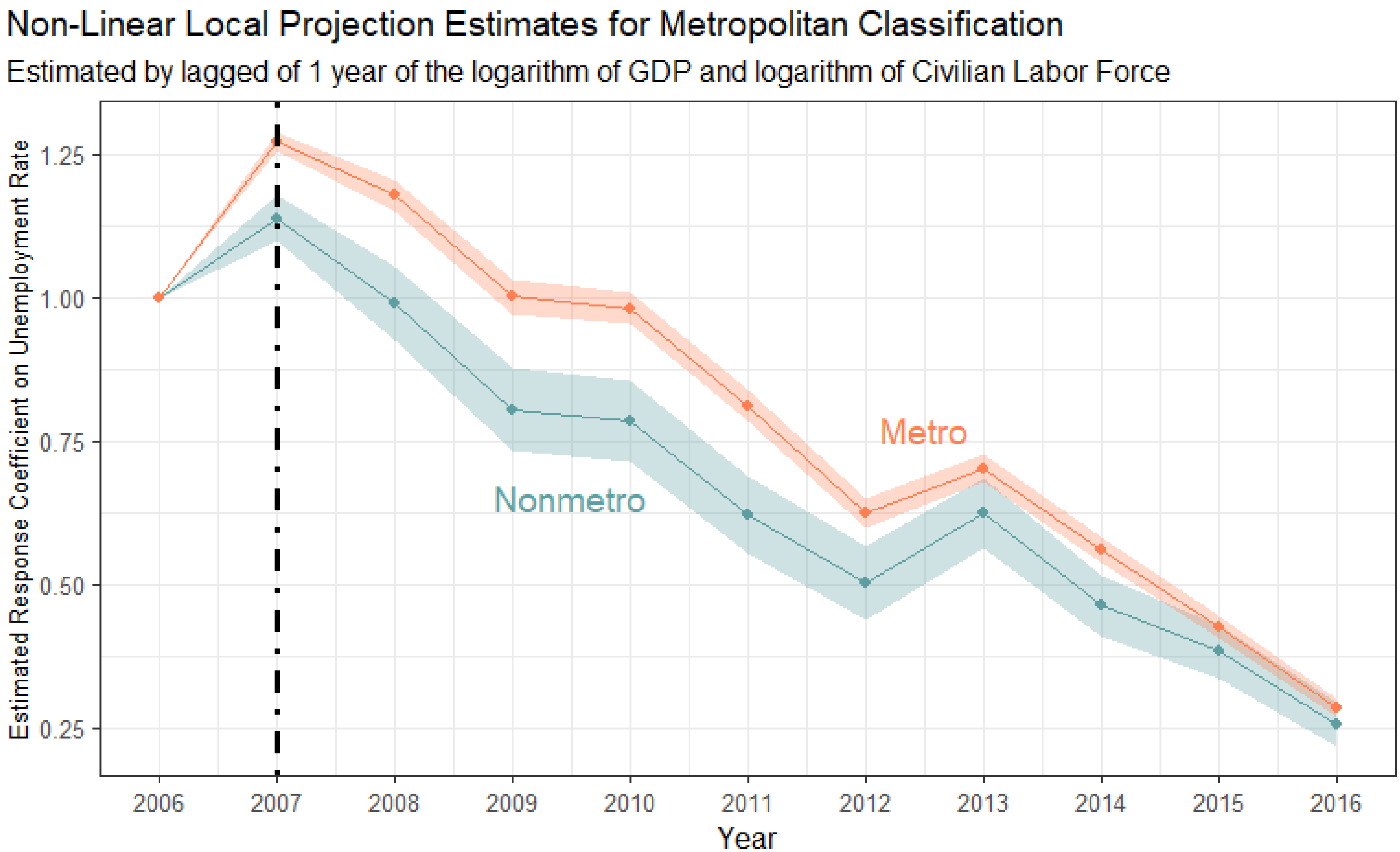

Figure 3: Estimated impulse response function of the Unemployment Rate to a GDP and Labor Force shock.

Table 1: Non-Linear Local Projection OLS Estimates

	Nonmetropolitan ($N = 1970$)			Metropolitan ($N = 1169$)		
Horizon	Shock	log(GDP)	log(Labor Force)	Shock	log(GDP)	log(Labor Force)
2006	1.000			1.000		
2007	1.138***	−0.028	8.298***	1.271***	−1.354***	5.418***
	(0.020)	(0.204)	(0.602)	(0.008)	(0.069)	(0.233)
2008	0.988***	−0.144	18.742***	1.177***	−3.518***	11.277***
	(0.032)	(0.491)	(1.351)	(0.013)	(0.165)	(0.501)
2009	0.805***	−0.377	25.918***	1.001***	−5.567***	14.948***
	(0.037)	(0.750)	(1.897)	(0.015)	(0.253)	(0.679)
2010	0.783***	−1.601*	23.777***	0.982***	−4.831***	13.673***
	(0.036)	(0.633)	(1.765)	(0.014)	(0.231)	(0.627)
2011	0.622***	−1.401*	22.881***	0.811***	−4.236***	13.100***
	(0.034)	(0.596)	(1.751)	(0.014)	(0.218)	(0.619)
2012	0.504***	−0.951	22.992***	0.623***	−4.158***	12.673***
	(0.033)	(0.659)	(1.845)	(0.013)	(0.228)	(0.643)
2013	0.625***	−0.729	21.995***	0.702***	−4.251***	11.835***
	(0.031)	(0.701)	(1.856)	(0.012)	(0.238)	(0.642)
2014	0.463***	−1.185	20.097***	0.560***	−4.896***	11.225***
	(0.027)	(0.723)	(1.917)	(0.011)	(0.240)	(0.681)
2015	0.383***	−2.454***	20.978***	0.424***	−5.529***	11.839***
	(0.023)	(0.756)	(2.104)	(0.010)	(0.257)	(0.756)
2016	0.257***	−2.908***	19.189***	0.285***	−5.377***	7.277***
	(0.019)	(0.848)	(2.818)	(0.008)	(0.289)	(1.107)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

p-values correspond to the t-test of coefficients. Lagged covariates of the predictors were excluded due to strong dependency and nearly identical results from the statistical tests.

IV. CONCLUSION

The impulse response estimates from the local projections suggest that the unemployment type composition within metropolitan and nonmetropolitan counties are statistically different. The estimates indicate that nonmetropolitan county unemployment is more sensitive to log civilian labor force than log GDP in comparison to metropolitan county unemployment having both indicators as statistically significant. Additionally, the difference of magnitudes of the log civilian labor force predictor is indicative that nonmetropolitan unemployment rate is more sensitive to changes in the log civilian labor force than changes in log GDP. As discussed previously, this is hypothesized as nonmetropolitan counties having a larger share of structural unemployment which means that whenever their labor force increases, the unemployment rate increases significantly as the local economy cannot accommodate the new workers.

On the other end, metropolitan county unemployment has a much larger significance regarding the log GDP predictor than nonmetropolitan county unemployment. The larger negative magnitude indicates that the unemployment rate in metropolitan counties is highly associated with changes in log GDP in comparison to the statistically insignificant coefficient estimates in nonmetropolitan counties. As discussed previously, this is hypothesized as metropolitan counties having a larger share of cyclical unemployment which means that shifts in the county GDP have a large predictive change in the county’s unemployment rate.

Figure 3 also supports the hypothesis of the composition of unemployment types in counties as metropolitan shock coefficients increase substantially higher from the nonmetropolitan shock coefficients yet have a larger recovery rate of a downward slope which indicates that their higher share of cyclical unemployment is associated with a faster recovery rate in comparison to nonmetropolitan counties. As the shock of the 2008 Great Recession phases out as argued by the implementation of an impulse response function, metropolitan and nonmetropolitan unemployment shock coefficients return to similar levels by 2016.

As a common tool for macroeconomic analysis, impulse response functions computed by local projections has indicated that metropolitan and nonmetropolitan county unemployment rates are statistically different in terms of the magnitude of the predictor coefficient estimates, the significance level of the coefficient estimates, and the recovery rates of the unemployment rates. The results suggest that metropolitan counties have a larger composition of cyclical unemployment rate in comparison to nonmetropolitan counties. Additionally, the results suggest that nonmetropolitan counties have a larger composition of structural unemployment rate in comparison to metropolitan counties. These results apply between the periods of 2006 to 2016 such that outside of the window indicated in the methodology, the unemployment type composition by metropolitan classification is unknown without further identifying assumptions and analysis.

These results may be crucial to more efficiently allocating resources to improve regional labor force markets. Historically, labor force programs to decrease unemployment have been primarily targeted at metropolitan areas to service the most people and have been conducted during periods of economic distress. However, with nonmetropolitan counties allegedly having a higher structural unemployment, this suggests that labor force programs should be constructed differently from metropolitan programs in order to improve the labor force efficiently.

REFERENCES

Adämmer, Philipp. (2019, December). “lprifs: An R Package to Estimate Impulse Response Functions by Local Projections.” *The R Journal* 11, no. 2 (2019, December): 421-438. https://journal.r-project.org/archive/2019/RJ-2019-052/RJ-2019-052.pdf.

Bureau of Economic Analysis. *Regional Economic Accounts: Download*. U.S. Department of Commerce, Bureau of Economic Analysis. Accessed April 27, 2023. DOI: 10.1257/0002828053828518.

Gonçalves, S., Herrera, A. M., Kilian, L., and Pesavento, E. (2022). “When Do State-Dependent Local Projections Work?” Federal Reserve Bank of Dallas, Working Papers, 2022 (2205). https://doi.org/10.24149/wp2205.

Jordá, Óscar. (2005). “Estimation and inference of impulse responses by local projections.” American Economic Review 95, 161–182. http://paulshea.com/wp-content/uploads/2013/06/Jorda.pdf.

Pender, John, Thomas Hertz, John Cromartie, and Tracey Farrigan. (2019, November 13) *Rural America at a Glance* (Economic Information Bulletin No. EIB-212). U.S. Department of Agriculture, Economic Research Service. https://www.ers.usda.gov/publications/pub-details/?pubid=95340.

Plagborg-Møller, Mikkel and Christian K. Wolf. (2021, March). “Local Projections and VARS Estimate the Same Impulse Responses.” *Econometrica* 89, no. 2 (March 2021): 955-980. https://doi.org/10.3982/ECTA17813.

Rambachan, A. and Neil Shephard. (2021, October 29). “When do common time series estimands have nonparametric causal meaning?” https://asheshrambachan.github.io/assets/files/arns_commontimeseries_causal.pdf.

Sanders, Austin. (2022, May 10). *Rural Employment and Unemployment*. U.S. Department of Agriculture, Economic Research Service. https://www.ers.usda.gov/topics/rural-economy-population/employment-education/rural-employment-and-unemployment.

Sanders, Austin. (2023, February 24) *County-level Data Sets: Download Data*. U.S. Department of Agriculture, Economic Research Service. https://www.ers.usda.gov/data-products/county-level-data-sets/county-level-data-sets-download-data/.

Sims, C. A. (1980). “Macroeconomics and Reality.” Econometrica, 48(1), 1–48. https://doi.org/10.2307/1912017.