# Labor Market Dynamics in West Virginia and the Appalachian Region

Josh Beverly

Dissertation submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Economics

Shamar L. Stewart, Co-chair

Clinton L. Neill, Co-chair

Klaus Moeltner

Kimberly L. Morgan

November 3, 2022

Blacksburg, Virginia

Keywords: Labor Force Participation, Unemployment, Appalachian region, West Virginia, Time Series

# Labor Market Dynamics in West Virginia and the Appalachian Region

Josh Beverly

(ABSTRACT)

This dissertation consists of three manuscripts analyzing labor market dynamics in West Virginia and the Appalachian Region. The first manuscript examines the dynamic effects of national, regional, and local labor market shocks on labor force participation rates in Appalachia. A dynamic factor model with time-varying loading parameters and stochastic volatility is used to explore the synchronicity and divergence between state labor force participation rates within and outside the Appalachian region. We find that the choice of time and state is crucial to the relative importance of the level of synchronization on observed change in LFPR variations. Our findings can help better target labor policy by taking advantage of the sensitivity exhibited by each state to various labor market conditions.

The second manuscript examines the dynamic effects of state, Metro/Non-Metro, and county labor market shocks on labor force participation rates in West Virginia. In the first stage, using a dynamic factor model, we find that non-metropolitan and county-specific components are dominant contributors to the observed variations in the change in West Virginia LFPRs. In the second stage, using a fixed effects panel model, we find county demographics, education levels, income, access to interstate highways, and industry composition are useful covariates for explaining the variance contributions of the state, metro/non-metro and county factors.

The third manuscript uses cointegration analysis in the presence of structural breaks to determine whether the Unemployment Invariance Hypothesis exists in West Virginia. Using monthly labor force data from 1976 - 2022, we find mixed support for the unemployment invariance, added worker effect, and discouraged worker effect hypotheses over multiple subsample periods. These results suggest that labor markets are temporally-dynamic, and a one-size-fits-all approach could prove disadvantageous to growth.

# Labor Market Dynamics in West Virginia and the Appalachian Region

Josh Beverly

(GENERAL AUDIENCE ABSTRACT)

This dissertation focues on labor market dynamics in West Virginia and the Appalachian Region. In the first of three manuscripts, we investigate how much U.S. state labor force participation rates move together nationally, and within the Appalachian Region. We find that how much labor force participation rates move together across the U.S. and within the Appalachian Region depends on the choice of time and state.

In the second manuscript, we examine how much West Virginia county labor force participation rates move together across the state and within the Metropolitan and Non-Metropolitan regions. We also study how county characteristics such as industry composition and education levels influence the variation in how much labor force participation rates move together. We find that Non-metropolitan county labor force participation rates exhibit similar dynamic behavior and that education, personal income, access to highways, and industry composition of the counties influences how much the rates move together at the different levels.

In the third manuscript, we investigate whether changes in the unemployment rate in West Virginia result influences that state's labor force participation rate in the long-run. We find that evidence of said long-run relationship albeit changing over time. We posit that the relationship dynamics are largely explained by the ensuing labor market and economic conditions. By extension, labor market policies and interventions should be timely and flexible.

# Dedication

*For my wife and son, Mariah and Mason, and to my dad, Steve.*

# Acknowledgments

I would like to thank my co-advisors, Shamar Stewart and Clinton Neill. Thank you both for your patience, support, and stability through the entire process and through several personal crises over the last few years. I would not have been able to see this realized if it were not for the both of you. Thank you for all of your time and effort training me to be a better researcher, presenter, and economist. I would also like to thank my committee members Klaus Moeltner and Kimberly Morgan for all of your help, guidance, and training during my time in the department. I would like to thank my father, Steve, who is no longer with us, for his encouragement, support and for instilling a passion for learning in me from the very beginning. Lastly, thank you Mariah, my wife, for sticking with me through thick and thin. I could not have done this without your help, your understanding, encouragement, support, and love every step of the way.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# The Dynamics of Labor Force Participation: Is All Quiet on the Appalachian Front?

## 1.1 Introduction

Not all shocks to labor force participation rates are created equal. Before the 2008-2009 Great recession, most studies posited that the national labor force participation rate, the percentage of the working-age population who are either employed or actively searching for work, was mildly procyclical, or relatively stable (Veracierto, 2008; Van Zandweghe, 2012). However, several studies since then have determined that the 2008-2009 recession adversely affected the labor force participation rate in the U.S. (Council of Economic Advisors, 2014; Van Zandweghe, 2012; Hotchkiss and Rios-Avila, 2013; Aaronson et al., 2014; Erceg and Levin, 2014). While a substantial national recession like the Great Recession is far-reaching, other exogenous shocks to labor force participation rates may only affect specific regions or states within the U.S. For example, regulatory changes and the OPEC oil embargo in the 1970s caused the price of coal to sharply increase (Van Zandweghe et al., 2017). As a result, employment, labor force participation, and earnings soared in the Appalachian region due to its heavy reliance on the coal industry, while other areas around the U.S. experienced declines in economic activity (Juhn, 1992; Black et al., 2002). Later, in the 1980s, these economic experiences were reversed due to a subsequent bust in the coal market. Figure 1 shows the dynamics and differing responses of the change in U.S. state LFPR to such shocks. The divergence of state LFPR and different economic shocks demonstrates the importance of understanding how national, regional, and state-level forces influence the change in labor force participation rates across time.

Determining the role and relative importance of national, regional, and state-level forces on the change in labor force participation rates has important policy implications. Empirical studies show that increases in LFP have large, positive effects on employment growth and national GDP (see Bryant et al., 2004; Juhn and Potter, 2006; Shoven, 2007; Cai and Lu, 2013; Bustelo et al., 2019, for example). At the peak of U.S. national labor force participation (LFP) from 1990 to 2000, labor contributed 1.34 percentage points to national economic growth in terms of output per hour (Bureau of Labor Statistics, 2021b). However, from 2000 to 2007, simultaneous with the decline in U.S. LFPR, labor contributed only 0.23 percentage points to output growth. As of 2019, labor contribution to output growth in the U.S. only recovered to 0.49 percentage points. Taylor (2016) suggests that with low unemployment rates, future growth must come primarily through increased labor force participation. Fostering labor force participation and subsequently experiencing economic growth requires improving our understanding of how driving forces at different geographic levels influence changes in labor force participation rates over time.

In this study, we examine the role and relative importance of geographic levels on the change

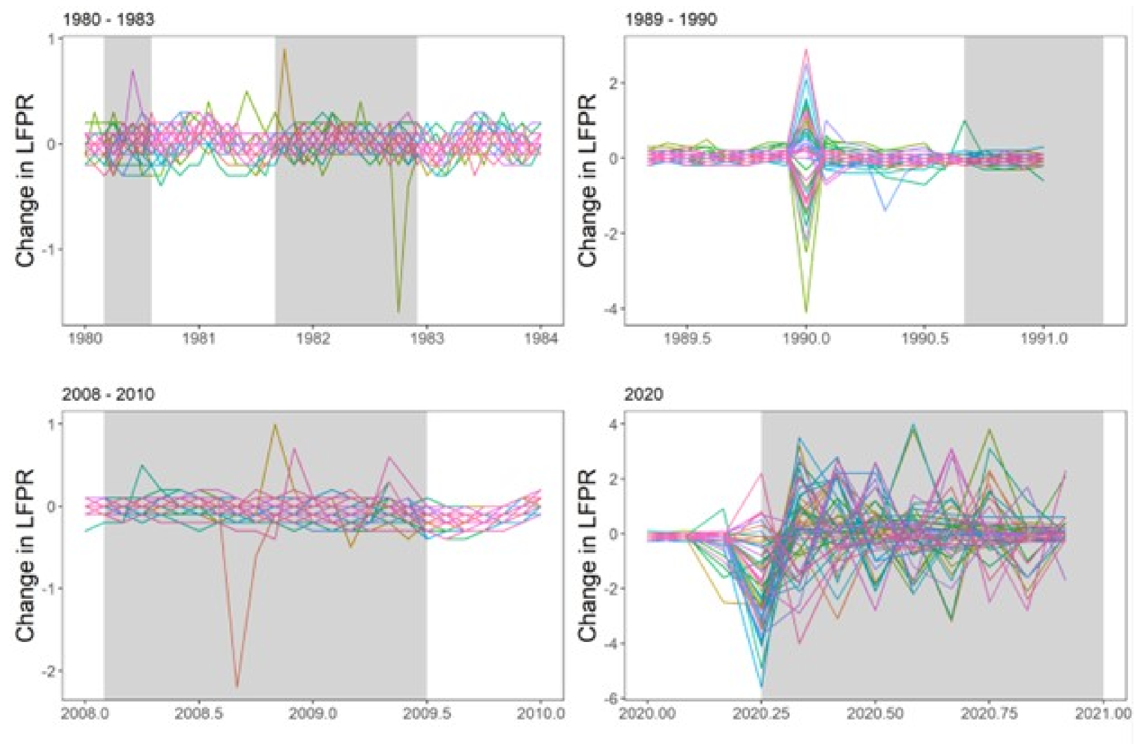


Figure 1: **Change in LFPR for Select Periods**

Note: (Top Left): Change in state LFPR over the years 1980 - 1983. These dates correspond with the Second Energy Crisis/Inflation and "Double Dip Recessions". (Top Right): Change in state LFPR over the years 1989 - 1991. These dates correspond to the S&L Crisis and Gulf War Recession. (Bottom Right):Change in state LFPR over the years 2008 - 2009. These dates correspond with the Great Recession. (Bottom Right): Change in state LFPR over the year 2020. These dates correspond with the COVID-19 pandemic and Recession. NBER-dated recessions are in gray.
Source: Bureau of Labor Statistics (BLS)

in state labor force participation rates. We investigate the extent to which state labor force participation rates move together. Specifically, we decompose state LFPR into a national (also referred to as a common component), a regional (Appalachia or Non-Appalachia), and a state-specific (idiosyncratic) factor. We, therefore, analyze the overall synchronicity and divergence of state-level LFPR data over time. We stipulate the national, regional, and state geographic levels since, arguably, influences on the change in state-level LFPR will primarily derive from an individual or joint shocks at these levels. We specifically designate the Appalachian region, which is shown in Figure 2, as our main region of focus due to the documented evidence for a strong and unique relationship between the LFPR and the geographic region itself. While the U.S. defines other geographic regions within its borders, we find none with a geographic relationship to LFP that is steeped in so much historical rhetoric and culture as we find with the Appalachian region. We discuss the empirical evidence and other support for this unique connection in Section 1.2.

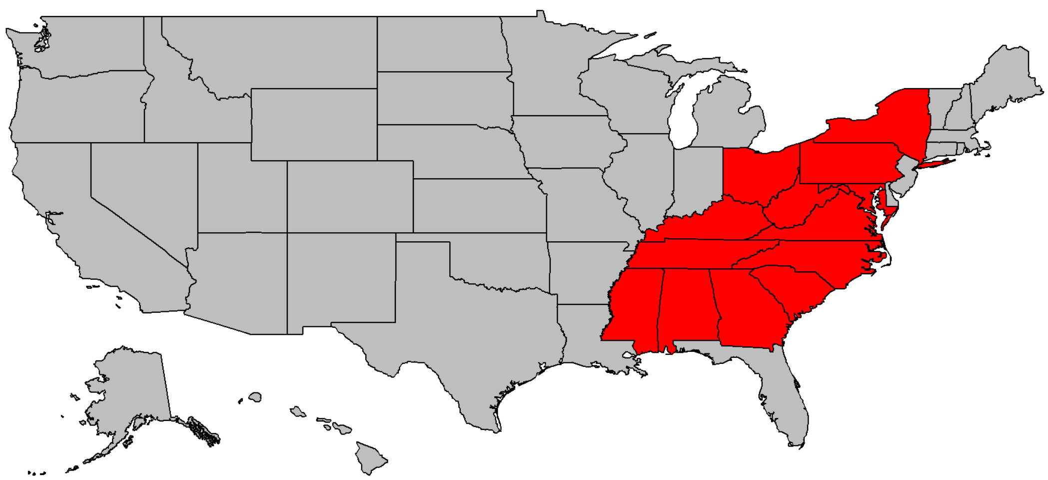

Figure 2: **The Appalachian Region**

Note: The region is officially defined by 420 counties across 13 states including Alabama, Georgia, Kentucky, Maryland, Mississippi, New York, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, Virginia and West Virginia. Since our analysis is at the state level, we include all states with at least one Appalachian county depicted above (in red) as our definition of the Appalachian region.
Source: Appalachian Regional Commission

We analyze LFPR dynamics over time and across states using a Dynamic Factor Model with time-varying and stochastic volatility. To best capture dynamics in U.S. LFPR, we solely use state, monthly, time-series labor force participation rates for all U.S. states and Washington D.C. over the 1976-2020 sample period. Including time-varying and stochastic volatility expands the standard DFM to capture the dynamics in the volatility in our national, regional, and idiosyncratic factors. Observing differential volatility across time and economic conditions is important for capturing changes in the sensitivity of our factors to labor conditions such as new or amended labor policy and major shocks like the COVID-19 pandemic, national recessions, or local natural disasters.

Our results demonstrate that the choice of year and state is crucial to the relative importance of each factor's contribution to the LFPR. This is because, in West Virginia, around 97% of variation in the change in state LFPR is explained by the Appalachian region factor in 1982, but less than 1% in 2010. In addition, our findings suggest that policy created at state or regional levels to increase labor participation and induce GDP growth may be more effective if enacted when the LFPRs are less synchronized (e.g. - where the national factor is less important). For example, across the Appalachian states, in 1999, on average only about 21% of variation in the change in state LFPR is explained by the national factor while 62% of variation is explained by state-idiosyncratic factors. We find that weaker contributions of the common factor to the observed volatilities in LFPRs and stronger synchronization with the common factor tend to coincide with business cycle expansions or economic recovery. Broader or more national policies may be appropriate during these times. We also find stronger contributions of the common factor to the observed volatilities in LFPRs and divergence from the common factor tend to coincide with periods of economic turmoil or uncertainty. More localized and time-sensitive policies may be appropriate during these times. The Appalachian region plays an important role in LFPR dynamics at specific points over the sample period, but when we exclude West Virginia, the influence of the Appalachian region gets progressively weaker.

This study bridges multiple fields and contributes to several strands of literature. From the microeconomics and regional development perspective, there is a growing interest and a burgeoning literature centered on determining how labor force participation (and general labor market dynamics) within Appalachia has changed and whether there is a structural difference in the region (see Dorsey, 1991; Isserman and Rephann, 1993; Bradley et al., 2001; Stephens and Deskins, 2018). However, these previous studies use data before the most recent economic downturns and analyze only one or two years at a time. Previous research suggests that the role of the Appalachian region in explaining LFP may change over time given local and national economic conditions (Isserman and Rephann, 1993). However, previous empirical approaches do not simultaneously capture longer horizons and the second moment of LFPR. Instead, they often assess the level and short-term trends. In addition, these studies focus on drivers of LFP through cross-sectional analysis, rather than through time-varying parameters and potential intertemporal and spatial differences. Therefore, we seek to fill the gap in the literature with the innovation that our results are not driven by the choice of year and the state of the business cycle. Given the interest in the Appalachian region, we add to this strand by assessing the synchronicity of LFPR within the region, and emphasize that this carries important implications for the region's labor and economic growth potential.

Within macroeconomics, our study adds to the body of work on labor force participation and the dynamics of national output. While we do not directly examine outcomes such as GDP or correlations with monetary policy, there is a growing consensus that labor force participation is a key component in economic recovery and growth and our study aids in that discussion with a sub-national analysis (Bryant et al., 2004; Azzoni and Silveira-Neto, 2005; Bullard et al., 2014; Erceg and Levin, 2014; Bustelo et al., 2019). Finally, we add to the literature surrounding labor markets more generally. Identification of factors, even latent ones, allows for localized labor policy to be targeted at increasing labor force participation. Understanding the geographic disparities has been of interest for decades both from a developed country (Chalmers and Greenwood, 1985; Fogli and Veldkamp, 2011) and a developing country's perspective (Kottis, 1990; Mehrotra and Parida, 2017).

The rest of this paper is organized as follows. Section 1.2, provides a discussion of the Appalachian region and the challenges related to labor force participation. Section 3.3, presents a description of the data and summary statistics. A discussion of the empirical methodology is found in Section 3.4. Discussion of our results is outlined in Section 3.5 and Section 3.6 concludes and offers potential policy recommendations.

## 1.2 Background on Appalachian Labor Force

The Appalachian region is often characterized by its economic disparity, persistent poverty, and historically low levels of skilled labor (Grossman and Levin, 1961; Rogers et al., 1997; Bollinger et al., 2011; Partridge et al., 2013; Stephens and Deskins, 2018; Appalachian Regional Commission, 2020). Labor force participation rates in Appalachia have also been consistently lower than in the rest of the U.S. over the past 45 years. However, while Appalachia performs poorly in LFP relative to the rest of the U.S., the region still accounts for approximately 31% of national GDP[1] (Bureau of Economic Analysis, 2021). This illuminates the Appalachian region's importance in terms of macroeconomic activity. As such, even small improvements in the region's LFP could have substantial impacts on national growth.

This Appalachian labor force participation gap and the potential growth for the region has received some attention from media, policymakers, government agencies, and researchers (Brainard et al., 2017; Jones, 2020). Efforts to raise levels of education and income in the area began in 1960 with a visit from then, presidential candidate John F. Kennedy (Schmitt, 2009). The Appalachian region has remained a focus of economic development since its

[1]State-level GDP is aggregated for all 13 states with at least some counties in the Appalachian region. A breakdown of the percentage of each state in the Appalachian region can be found in Table A.2 of the Appendix.

inception in 1965 with the Appalachian Region Development Act and through additional legislation aptly nicknamed the War on Poverty. However, despite efforts for improvements, Appalachia remains one of the most economically depressed regions in the U.S. (Bollinger et al., 2011) In addition to low income and education levels, geographic remoteness, reliance on the coal industry, and potentially unique culture in the region play an important role for LFPR and potential economic growth in Appalachia (Grossman and Levin, 1961; Topel, 1986; Rogers et al., 1997; Bollinger et al., 2011; Partridge et al., 2013; Stephens and Deskins, 2018; Appalachian Regional Commission, 2020).

Encompassing the central and southern portions of the Appalachian Mountain range, the Appalachia region consists predominantly of rural areas covered by forests and crops (National Land Cover Dataset, 2000). Some authors have characterized the region as remote or even geographically isolated, and studies show that rurality does influence labor force participation, other labor market outcomes, and the economy (Brainard et al., 2017; Weingarden et al., 2017; Stephens and Deskins, 2018). With an abundant natural resource endowment in terms of coal and natural gas, many parts of the region have historically been reliant on the timber, gas, and coal mining industries as the main source of employment. This strong reliance on natural resource industries creates a regional connection to the LFPR that is susceptible to regional shocks and policy changes that other regions do not experience.

There is also anecdotal evidence for a unique Appalachian culture that drives behavior and labor decisions. Fernandez and Fogli (2009) support the idea that beliefs or culture can impact present-day labor decisions. For Appalachia, this is often attributed to an assumed frontier-like attitude due to the physical isolation of the mountains and rugged terrain in the region (Billings, 1974). From medical literature, Behringer and Friedell (2006) describe cultural factors for the Appalachian region, such as religious fatalism, medical communication barriers, and deep cultural or economic connections to potentially dangerous industries that contribute to the observed lower health, education, income, and labor force participation in the region. A few empirical studies have examined whether unique Appalachian culture or behavior that drives labor force participation exists, but no consensus has been reached (Dorsey, 1991; Isserman and Rephann, 1993; Stephens and Deskins, 2018).

Dorsey (1991) suggests that an "Appalachian effect" or unique Appalachian culture does persistently decrease the LFPR for the region. West Virginia is of particular interest to Dorsey (1991) as it is the only state entirely encompassed by the Appalachian region. West Virginia stands out relative to the other Appalachian region states as it exhibits persistently lower labor force participation rates (Dorsey, 1991) and ranks higher in negative health indicators (Raghupathi and Raghupathi, 2018). Using state-level data for 1987, Dorsey finds that traditional economic and demographic variables have little explanatory power and contends that cultural differences explain most of the variation in LFP. However, Isserman and Rephann (1993) criticize Dorsey (1991) for using only one year of data as it may produce misleading conclusions.

To expand the methodological rigor, Isserman and Rephann (1993) utilize multiple model specifications (including the one used in Dorsey (1991)) separately on 1980, 1987, and 1991 county-level data. Isserman and Rephann (1993) find a small, negative Appalachian effect for male and female workers in 1980 in only one specification. Other specifications revealed small negative Appalachian effects for 1980, 1987, and 1991 for female workers but only for male workers in 1987. While the authors conclude that their results depend greatly on the year of data chosen for analysis, they also only use one year of data for each of their specifications. The authors posit that the choice of year may explain the contrasting results with Dorsey (1991) even given the different geographical scales between the two studies.

Stephens and Deskins (2018) use county-level data to investigate the drivers of LFP and the differences between rural and urban areas and between the Appalachian and Non-

Appalachian regions. The authors first find LFPR in rural counties in the Appalachian region to be about 1.5 percentage points lower than rural counties outside the region. They also find that the factors accounted for in the analysis, via an Oaxaca-Blinder decomposition, explain much of the variation between rural and urban areas. They attribute the variation unexplained by the known factors, 1.1 percentage points, to be a potential "Appalachian Effect" on LFP.

The connection to LFPR through the regional characteristics and debate described in this section motivate our use of Appalachia as our region of interest. With this strong connection between the LFPR and the regional economy, previous research suggests that the role of the Appalachian region in explaining LFP may change over time given local and national economic conditions (Isserman and Rephann, 1993). By including the Appalachian region in our analysis, we can measure the influence of the regional shocks on Appalachian state LFPR and if comovement of LFPR in these states persists over time.

We add to the debate above by contending, along with Isserman and Rephann (1993), that the year chosen for analysis may influence results regarding an "Appalachian Effect" on LFPR. We suggest that regional and national shocks may exacerbate the region's already poor performance in LFPR and other indicators disproportionally compared to other areas in the U.S. and may explain the "Appalachian Effect" for these points in time. In this paper, we demonstrate that the comovement or divergence in the change in state LFPR varies over time, geographical level, and shock. We also show that the years of data and subsequent conclusions in previous studies align with periods of stronger regional and national synchronization of change in state LFPR due to regional or national shocks.

## 1.3 Data

To investigate the synchronicity and response of the Appalachian region labor force to changing economic environments, we use monthly labor force participation rates for the 50 U.S. states and Washington D.C. over the period January 1976– December 2020.[2] We estimated our model using the first difference of the LFPRs and the differenced data can be seen in Figure 3.[3] This data is collected from the Bureau of Labor Statistics (BLS)[4]. There is a notable difference in each state's response during periods of recessions, financial crises, and the COVID-19 pandemic. For example, Maryland and Virginia exhibited relatively large negative changes around the 1990/91 and 2007/2009 recessions. However, Mississippi and Alabama show positive changes in the LFPR during the same periods. As expected, most of the states exhibit significant declines in the LFPR during the pandemic period.

While the unemployment rate is popular for empirical analysis and economic policies, we use the LFPR as it provides a truer representation of labor market conditions (Juhn and Potter, 2006). That is, the unemployment rate does not always reflect that people have dropped out of the labor force (Juhn and Potter, 2006; Hotchkiss and Rios-Avila, 2013; Stephens and Deskins, 2018). An economy might simultaneously experience a high level of discouraged workers (individuals who give up looking for a job and fall out of the labor force) and a low unemployment rate (Hotchkiss and Rios-Avila, 2013).[5] At face value, this would signal improving economic conditions and a thriving labor market. Consequently, unemployment rates in distressed areas can be comparable to the national average when

[2] All Augmented Dickey-Fuller tests supported the conclusion of a unit root process and high persistence.

[3] In Section 3.5, we present the estimation results from our DFM-TV-SV model with the 1976-2020 data. Given the visible and large decreases in the labor force participation rates (Figure 3) during the COVID-19 period, we also re-estimated the model excluding data for 2020. The results were quantitatively similar and are available upon request.

[4] Retrieved at: `https://download.bls.gov/pub/time.series/la/`

[5] In addition, unemployment does not gauge the size of the underground or "informal" economy – as evidenced by the fact that some developing countries have low official unemployment rates (Bradley et al., 2001).

labor force participation remains low. For example, since 2000, West Virginia reported an average rate of unemployment of 6.2% compared to the national average of 6%[6]. Yet, as discussed earlier, West Virginia has persistently lower LFPR as compared to the rest of the country.

The LFPR represents the percentage of the civilian and noninstitutional working-age population that is either working or actively looking for work. Table B.1 highlights that the LFPR varies within the Appalachian region and across all states. Over the sample period, West Virginia has the lowest rates in the country. Within the Non-Appalachian region, New Mexico and Oklahoma exhibit the lowest labor force participation rates, while Alaska and Minnesota have the highest. It should be noted that within the Appalachian region, South Carolina and Maryland exhibit higher labor force participation rates but also make up the smallest percentages (in terms of the number of counties) of the Appalachian region[7].

Given the conflicting results of previous literature and the different choices in the years studied, we acknowledge that the labor force dynamics of the Appalachian region may change over time with economic conditions. Moreover, (counter)cyclical factors play a large role in national and sector-specific labor markets. The global COVID-19 pandemic and Great Recession were felt worldwide. The U.S. dropped 31 places in international LFPR rankings between 2000 and 2020 (World Bank, 2021). Utilizing monthly LFPR data over 45 years allows us to account for long-term trends, major economic events, and measure the evolution and relative importance of the Appalachian region. While more granular data is arguably better, we use state-level LFPR data given the unavailability of monthly county-level data for a similar sample period. This aggregation reflects a potential drawback to our choice of analysis at the state level. However, given the large number of counties and equivalents across the U.S. (3143) and the computational burden of our model estimation, we are restrained to

[6] Visualizations of unemployment rate data are available upon request.

[7] (See Table A.2 in the Appendix)

a state-level analysis. Regardless, research at this level helps fill the gap in the analysis of statewide participation rates as many studies on individual labor force participation decisions already exist. Additionally, using aggregate state participation rates allows for a focus on regional differences and on actionable policy at the state level.

## 1.4 Structural Model

We consider a Dynamic factor model with Time-Varying Stochastic Volatility (DFM-TV-SV) in the spirit of Del Negro and Otrok (2008). In general, the dynamic factor model is a dimension reducing technique which models the co-movements of a high-dimensional vector of time series variables (the LFPR) as a function of a few latent dynamic factors (Stock and Watson, 2011).

Using a similar state space analysis, Stock and Watson (2016) posit that comovements of many macroeconomic variables can be described by a unobserved single index or dynamic factor. We build off this premise and model changes in state LFPR as functions of a national, regional, and idiosyncratic (state-specific) factors. Restricting our latent factors of LFPR to a small number in our dynamic factor analysis is consistent with standard dynamic equilibrium macroeconomic theory (Stock and Watson, 2016). To this end, we employ the MCMC estimation method to estimate this general model using a panel of state LFPR data in the U.S. for the past few decades.

To our knowledge, ours is the first study to apply the DFM framework to regional U.S. labor force participation. Other studies have used this methodology to investigate variables such as output growth (Bian et al., 2020), bond yield (Bhatt et al., 2017), changes in business cycles (Del Negro and Otrok, 2008), labor market conditions (Chung et al., 2014), inflation (Mumtaz and Surico, 2012), equity market valuations (Ma et al., 2018), commodities (West

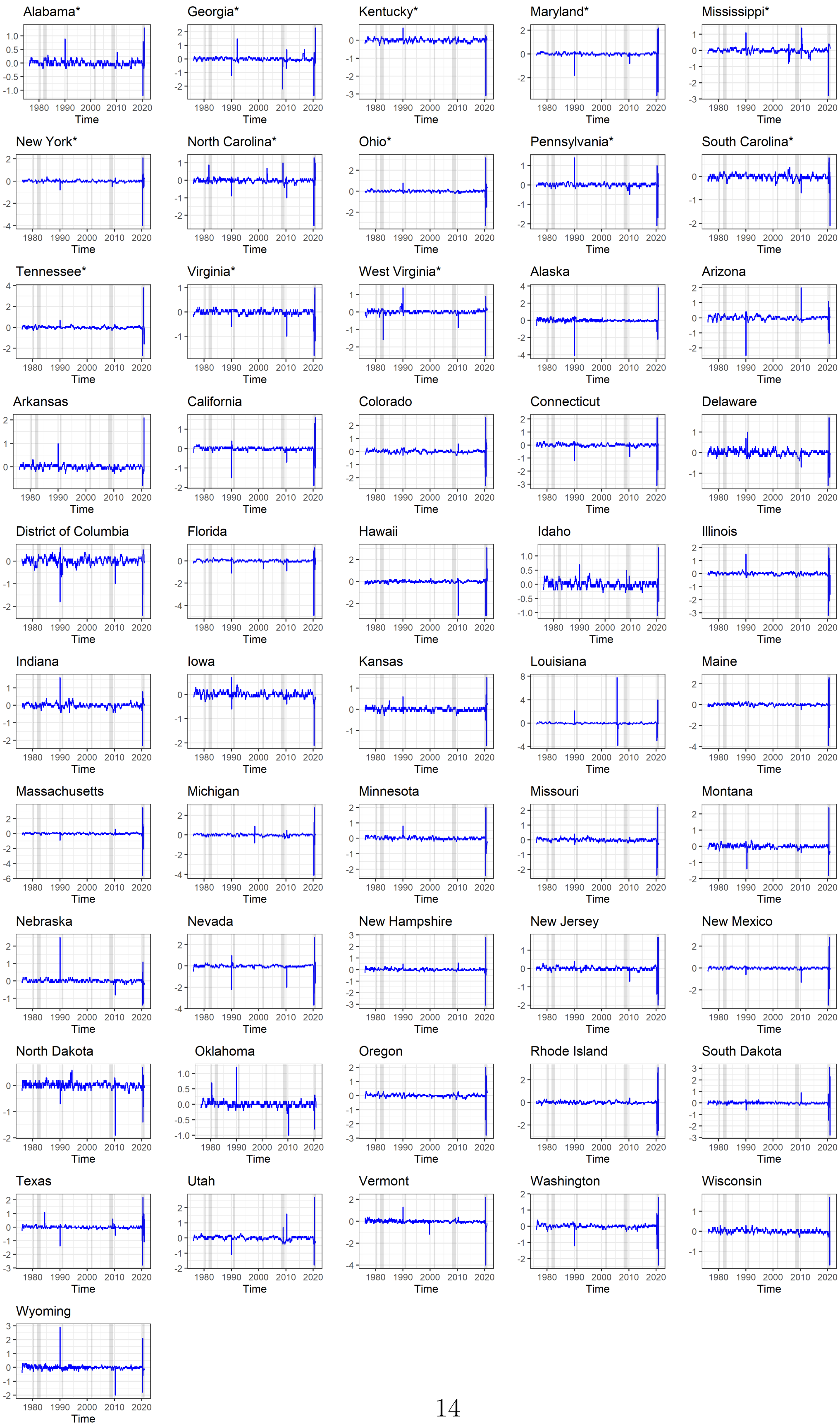




Figure 3: **Change in U.S. State Labor Force Participation Rates**

and Wong, 2014), oil (Aastveit et al., 2015), and cattle prices (Foster et al., 1995; Walburger and Foster, 1998).

### 1.4.1 Standard Dynamic Factor Model

We consider the following specification for our measurement equation:

$$y_{i,t} = \omega_{i,t}\mathscr{C}_t + \widetilde{\beta}_{i,t}\mathscr{R}_t + \xi_{i,t} \tag{1}$$

where $y_{i,t}$ is the change in labor force participation rate for state (and Washington, D.C.) i = 1,2,...51 at month $t$. $\mathscr{C}_t$ is the national or common factor that affects $y_{i,t}$. $\mathscr{R}_t$ is a vector that contains the two regional factors, $\mathscr{R}_{j,t}$, $j = 1, 2$ corresponding to the Appalachian and Non-Appalachian regions, respectively.[8] $\xi_{i,t}$ is the idiosyncratic or state-specific factors. The idiosyncratic factors account for movement by each state after the national and regional factors are removed. Since the geographical characteristics of the comovements are unobserved, we infer them from factor loadings which are the coefficients of the vectors of the lagged factors.

The national factor's loading parameter, $\omega_{i,t}$, captures the exposure to a national (common) factor. The row vector $\widetilde{\beta}_{i,t}$ has a non-zero time-varying regional loading parameter $\beta_{i,t}$ for the position corresponding to the region for state $i$ and zeros for all other elements. Accordingly, each regional factor, $\mathscr{R}_{j,t}$, captures changes in the LFPR specific to each region and is separately identified by setting Appalachian region loadings to zero for Non-Appalachian region states and Non-Appalachian region loadings to zero for Appalachian states. We capture the dynamics of each factor by including time-varying factor loading parameters.

[8] Thirteen (13) states with counties within the Appalachian region are included in the Appalachian region factor and the 38 other states (including Washington D.C.) for the Non-Appalachian region factor. For brevity, we report only the results for the Appalachian region. The full results will be made available upon request.

The transition equations for each factor evolve as stationary processes:

$$\mathscr{C}_t = \sum_{p=1}^{P} \phi_p^{\mathscr{C}} \mathscr{C}_{t-p} + e^{h_t^{\mathscr{C}}} \cdot \nu_t^{\mathscr{C}}; \quad \nu_t^{\mathscr{C}} \sim i.i.d.\, \mathscr{N}(0, \sigma_{\mathscr{C}}^2) \tag{2}$$

where $\phi_p^{\mathscr{C}}$ is the autoregressive coefficient for the national factor, $P = 2$. $e^{h_t^{\mathscr{C}}}$ represents the stochastic volatility components, and $\nu_t^{\mathscr{C}}$ the innovation to the law of motion for the national or common factor.

$$r_{j,t} = \sum_{l=1}^{L} \phi_{j,t}^{\mathscr{R}} r_{t-l} + e^{h_{j,t}^{\mathscr{R}}} \cdot \nu_{j,t}^{\mathscr{R}}; \quad \nu_{j,t}^{\mathscr{R}} \sim i.i.d.\, \mathscr{N}(0, \sigma_{j,s}^2) \tag{3}$$

where $\phi_{j,t}^{\mathscr{R}}$ is the autoregressive coefficient for each regional factor, $L = 2$, $e^{h_{j,t}^{\mathscr{R}}}$, the stochastic volatility components, and $\nu_{j,t}^{R}$ the innovation to the law of motion for the regional factor.

$$\xi_{i,t} = \sum_{q=1}^{Q} \phi_q \xi_{t-q} + e^{h_{i,t}^{\mathcal{S}}} \cdot \nu_{i,t}^{\mathcal{S}}; \quad \nu_{i,t}^{s} \sim i.i.d.\, \mathscr{N}(0, \sigma_i^2) \tag{4}$$

where $\phi_q$ is the autoregressive coefficient for the idiosyncratic shock, $Q = P = L = 2$, $e^{h_{j,t}^{r}}$, the stochastic volatility components, and $\nu_{i,t}^{s}$ the innovation to the law of motion for the idiosyncratic factor. For proper identification, we follow the literature and assume that $\nu_t^{\mathscr{C}}, \nu_{j,t}^{\mathscr{R}}$, and $\nu_{i,t}^{\mathcal{S}}$ are orthogonal to each other.

### 1.4.2 Dynamic Factor Model with Time-Varying Stochastic volatility

To capture the dynamics in the volatility over time, we expand the standard DFM to include stochastic volatility in the laws of motion of the national, regional, and idiosyncratic factors (Equations 11 – 15). This extension assumes random, rather than constant, innovations (error terms) of each factor.[9] In particular, we observe differential volatility across time and economic conditions. Importantly, this assumption and specification allows us to capture changes in the sensitivity of our factors to labor conditions over our sample. To this extent,

[9]Formally, the stochastic volatility model assumes that the variance of the error term is itself normally distributed.

we are able to capture potential volatility changes due to new or amended labor policy and major shocks to the local economies like the COVID-19 pandemic and natural disasters.

Formally, the innovations, $e^{\bullet}$, vary over time and each stochastic volatility term, $h_{\bullet}$, evolve according to a random walk process without drift such that:[10]

$$h_t^{\mathscr{C}} = h_{t-1}^{\mathscr{C}} + \sigma_{\mathscr{C}}^h \cdot \eta_t^{\mathscr{C}}; \quad \eta_t^{\mathscr{C}} \sim i.i.d.\mathscr{N}(0,1) \tag{5}$$

$$h_{j,t}^{\mathscr{R}} = h_{j,t-1}^{\mathscr{R}} + \sigma_{j,\mathscr{R}}^h \cdot \eta_{j,t}^{\mathscr{R}}; \quad \eta_{j,t}^{\mathscr{R}} \sim i.i.d.\mathscr{N}(0,1) \tag{6}$$

$$h_{i,t}^{\mathscr{S}} = h_{i,t-1}^{\mathscr{S}} + \sigma_i^h \cdot \eta_{i,t}^{\mathscr{S}}; \quad \eta_{i,t}^{\mathscr{S}} \sim i.i.d.\ \mathscr{N}(0,1) \tag{7}$$

where $\sigma_{\mathscr{C}}^h, \sigma_{j,\mathscr{R}}^h, \sigma_i^h$ are the standard deviations of the innovation to each law of motion respectively and $\eta_t^{\mathscr{C}}, \eta_{j,t}^{\mathscr{R}}$, and $\eta_{i,t}^{\mathscr{S}}$ are the volatility shocks. We also assume that, $\eta_t^{\mathscr{C}}, \eta_{j,t}^{\mathscr{R}}$, and $\eta_{i,t}^{\mathscr{S}}$ are orthogonal to each other.

Lastly, for identification, we follow the standard normalization procedures used in the macroeconomics literature (See Del Negro and Otrok, 2008; Bhatt et al., 2017, for example). First, given that the scale of the factor loadings and the standard deviations for each factor cannot be separately identified, we restrict the shocks of the national and regional factors $\sigma_{\mathscr{C}}^2 = \sigma_{1,\mathscr{R}}^2 = \sigma_{2,\mathscr{R}}^2 = 1$. Second, since the scale of stochastic volatility term $h_{\bullet}$ is determined by the initial condition, we constrain each $h$ in the stochastic volatility equations (12 – 16) to a starting value of zero. That is, $h_0^{\mathscr{C}} = h_{j,0}^{\mathscr{R}} = h_{i,0}^{\mathscr{S}} = 0$. This assumes no stochastic volatility before the sample period but allows for derivation of an ergodic distribution for the initial conditions (Del Negro and Otrok, 2008).

[10] Del Negro and Otrok (2008) opines that policy or structural changes occurring over time are permanent and not transitory. We, therefore, model the time-variation as a drift rather than a stationary process. This is a departure from previous studies on Appalachia's LFPR (Dorsey, 1991; Isserman and Rephann, 1993; Stephens and Deskins, 2018) as they often do not account for long term trends, or changes in national- and state-level labor force conditions. We contend that our approach is more flexible and better accounts for potential long term trends and structural changes in LFP conditions. This DFM-TV-SV model approach therefore fill the gaps in the previous literature.

### 1.4.3 Gibbs-Sampling Algorithm

Following Del Negro and Otrok (2008); Bhatt et al. (2017); Bian et al. (2020), we estimate our model via a Monte Carlo Markov Chain (MCMC) Bayesian estimation utilizing the Gibb-Sampling Algorithm a lá Kim et al. (1999). Below, we provide a brief description of our model estimation. For additional information about our execution of the procedure and the Gibb Sampler, the interested reader is directed to the technical appendix of Bhatt et al. (2017).

For notational ease, let $\Xi$ be the collection of time-varying coefficients and hyperparameters such that

$$\Xi = \left(\omega^{T\prime}, \beta^{T\prime}, \varphi_{\mathscr{C}}', \varphi_{\mathscr{R}}', \varphi_{\mathscr{S}}', g^2, \{h_{1,t}^{\mathscr{C}}\}_{t=1}^{T}, \{h_{1,t}^{\mathscr{R}}\}_{t=1}^{T}, \{h_{2,t}^{\mathscr{R}}\}_{t=1}^{T}, \{\{h_{1,t}^{\mathscr{S}}\}_{t=1}^{T}\}_{i=1}^{51\prime}\right),$$

where $\omega^T = \{(\omega_1, \omega_2, \dots, \omega_{51})'\}_{i=1}^{T}$ and $\beta^T = \{(\widetilde{\beta_1}, \widetilde{\beta_2}, \dots, \widetilde{\beta_{51}})'\}_{t=1}^{T}$ denote our time-varying coefficients. $\varphi_{\mathscr{C}} = (\phi_1^{\mathscr{C}}, \phi_2^{\mathscr{C}})$, $\varphi_{\mathscr{R}} = (\phi_{1,1}^{\mathscr{R}}, \phi_{1,2}^{\mathscr{R}}, \phi_{2,1}^{\mathscr{R}}, \phi_{2,2}^{\mathscr{R}})$, $\varphi_{\mathscr{S}} = (\phi_{1,1}, \phi_{1,2}, \phi_{2,1}, \phi_{2,2}, \dots, \phi_{51,1}, \phi_{51,2})$, and $g^2 = \{\sigma_i^2\}_{i=1}^{51})$ are the time invariant variances. Lastly, the $h_{\bullet}$ represent the latent stochastic volatilities.

1. Draw the common and regional factors conditioned on the time-varying factor loadings, the autoregressive coefficients of the national and idiosyncratic components, the time invariant variance, and the stochastic volatilities.

$$f\left(\{\mathscr{C}_t\}_{t=1}^{T}, \{\mathscr{R}_{1,t}\}_{t=1}^{T}, \{\mathscr{R}_{2,t}\}_{t=1}^{T} \,\middle|\, \Xi\right)$$

Given the presence of stochastic volatility, this process requires modification from the original procedure outlined in Chib and Greenberg (1994). This modification is described in detail in Del Negro and Otrok (2008).

2. Take a random draw of the AR(Q) and variance parameters for the idiosyncratic factor conditioned on the national factor, regional factors, time-varying factor loadings, and the idiosyncratic stochastic volatility.

$$f\left(\varphi_S, g^2 \middle| \{\mathscr{C}_t\}_{t=1}^T, \{\mathscr{R}_{1,t}\}_{t=1}^T, \{\mathscr{R}_{2,t}\}_{t=1}^T, \omega, \widetilde{\beta}, \{h_{i,t}\}_{t=1}^T\right)$$

3. Get a random draw of the time-varying loadings parameters, conditioned on the national factor, regional factors, the autoregressive coefficients of the national factor, the time invariant variances, and idiosyncratic stochastic volatility.

$$f\left(\omega, \widetilde{\beta} \middle| \{\mathscr{C}_t\}_{t=1}^T, \{\mathscr{R}_{1,t}\}_{t=1}^T, \{\mathscr{R}_{2,t}\}_{t=1}^T, \varphi_c, \sigma^2, \{h_{i,t}\}_{t=1}^T\right)$$

Since we assume the errors, conditional on the factors in Equation 9, and the innovations in the factor loadings are independent across $i$, we can draw the time-varying loadings one at a time. This diminishes the effect of dimensionality and aid in efficiency.

4. Take a random draw of the AR parameters of the national and regional factors, conditioned on their respective loading factor and stochastic volatilities.

$$f\left(\varphi_{\mathscr{C}} \middle| \{\mathscr{C}_t\}_{t=1}^T, \{h_{1,t}^{\mathscr{C}}\}_{t=1}^T)\right.$$
$$f(\varphi_{\mathscr{R}} \left| \{\mathscr{R}_{1,t}\}_{t=1}^T, \{\mathscr{R}_{2,t}\}_{t=1}^T, \{h_{1,t}^{\mathscr{R}}\}_{t=1}^T, \{h_{2,t}^{\mathscr{R}}\}_{t=1}^T\right)$$

5. Get a random draw of the time invariant and time-varying stochastic volatility for the national, regional and idiosyncratic components, conditioned on the factor loadings and autoregressive parameters. This step follows the algorithm from Kim et al. (1998)

$$f\left(\{h_{1,t}^{\mathscr{C}}\}_{t=1}^T, \sigma_{\mathscr{C}}^h \middle| \{\mathscr{C}_t\}_{t=1}^T, \varphi_{\mathscr{C}}\right)$$
$$f\left(\{h_{1,t}^{\mathscr{R}}\}_{t=1}^T, \{h_{2,t}^{\mathscr{R}}\}_{t=1}^T, \sigma_1^h, \sigma_2^h \middle| \{\mathscr{R}_{1,t}\}_{t=1}^T, \{\mathscr{R}_{2,t}\}_{t=1}^T, \varphi_{\mathscr{R}}\right)$$
$$f\left(\{h_{1,t}^{\mathcal{S}}\}_{t=1}^T, \sigma_i^h \middle| \{\mathscr{C}_t\}_{t=1}^T, \{\mathscr{R}_{1,t}\}_{t=1}^T, \{\mathscr{R}_{2,t}\}_{t=1}^T, \omega, \widetilde{\beta}, \varphi_{\mathcal{S}}\right)$$

6. **Repeat steps 1 - 5**: $(B + K)$ number of times where B is the number of burn-ins or draws discarded in order to reach confidence in the initial conditions imposed. K is the number of keepers or draws that are saved after the allotted burn-in values have

been reached. We use B = 10,000 and K = 40,000 draws respectively.

## 1.5 Results

Over time, we find that the economic structure of a state and its connection to the national and regional economies have different sensitivities and structural breaks. This provides a strong justification for using a model with time-varying loading parameters and stochastic volatility. In Figures 4 and 5, we plot the time-varying loading parameters (posterior medians) of an unobserved national/common factor, for the Appalachian region and selected Non-Appalachian states together with their 90% confidence intervals.[11] These national factor loadings reflect changes in the sensitivity or a measure of synchronization of each state with the national factor. The tight confidence intervals around our median estimates indicate a fairly low level of parameter uncertainty. In Figure 6, we plot the time-varying loading parameters (posterior medians) for an unobserved Appalachian regional factor for the Appalachian states, together with their 90% confidence intervals[12] The Appalachian regional factor loadings reflect changes in the sensitivity or a measure of synchronization of each Appalachian state with the regional factor.

### 1.5.1 National and Regional Factors Loadings

We observe considerable time variations in the national factor loading parameters across states. We see that the lower bound of the 90% confidence bands in the latter part of the sample period is above the upper bound of the confidence bands in the early part of the sample period. Not only does this provide further justification for using a general approach

[11] For the sake of brevity, we have suppressed the results for the rest of the non-Appalachian states. The full results are available from the authors upon request.

[12] Results for Non-Appalachian regional factor loadings are available upon request.

for modeling comovement among state LFPR, it also justifies our extension of the standard DFM with time-varying parameters.

Additionally, despite the substantial time variations, the dynamics and overall shape of the national factor loadings over time are similar across states and are mostly positive. The negative loading parameters at the beginning of the sample period indicate that each state's LFPR had a relatively sensitive and negative correlation with the national factor. Noticeably, there is a change in sign of the factor loadings near 1990. Most states exhibit near-zero correlation with the national factor in 1990 which is supported by the divergence in the change in state LFPR time series seen in Figure 1. After the 1990 recession, the loading parameters are mostly positive indicating that change in LFPR for each state is relatively sensitive and positively correlated with the national factor. Around the Great Recession (2008 − 2009), we observed a mixed response across states. In general, most states decrease in sensitivity, indicating a divergence of state LFPR from the national factor. The responses in Appalachian states such as Alabama, Georgia, and Kentucky were a bit more extreme than their counterparts. Interestingly, for Georgia, the positive correlation with the national factor becomes negative and for Pennsylvania, the sensitivity to the common factor increased significantly at the onset of the crisis and quickly decreased thereafter. The sensitivity or correlation to the common factor for other states such as California and Iowa, seen in Figure 5, increased through the recession, and for Oregon remained the same.

Turning our attention to the loading parameters of the Appalachian regional factors. Figure 6 reveals that the sensitivity of statewide LFPRs is much more heterogeneous (than their national factor counterpart). Although the median estimates are, by and large, near-zero, we observe a large degree of parameter uncertainty. This was especially true during the COVID-19 pandemic period when the range of the upper and lower bounds widened most.

West Virginia is a notable exception here. The confidence bounds are much tighter around

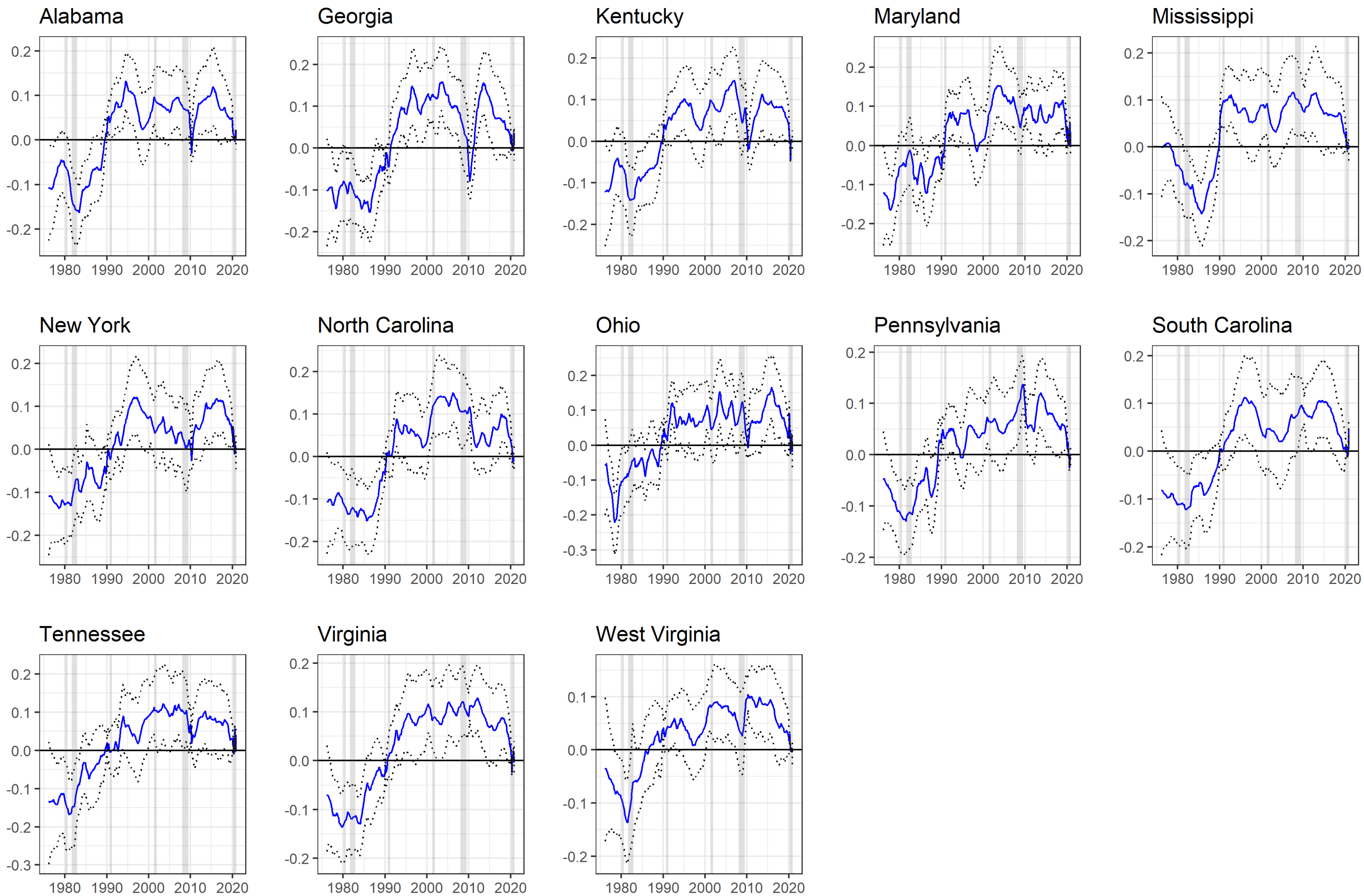


Figure 4: **National Factor Loadings by States in the Appalachian Region**

Note: Shaded regions are the NBER-dated recessions. The blue solid lines represent the median of the posterior distribution. Dashed lines represent the 5th and 95th percentiles.

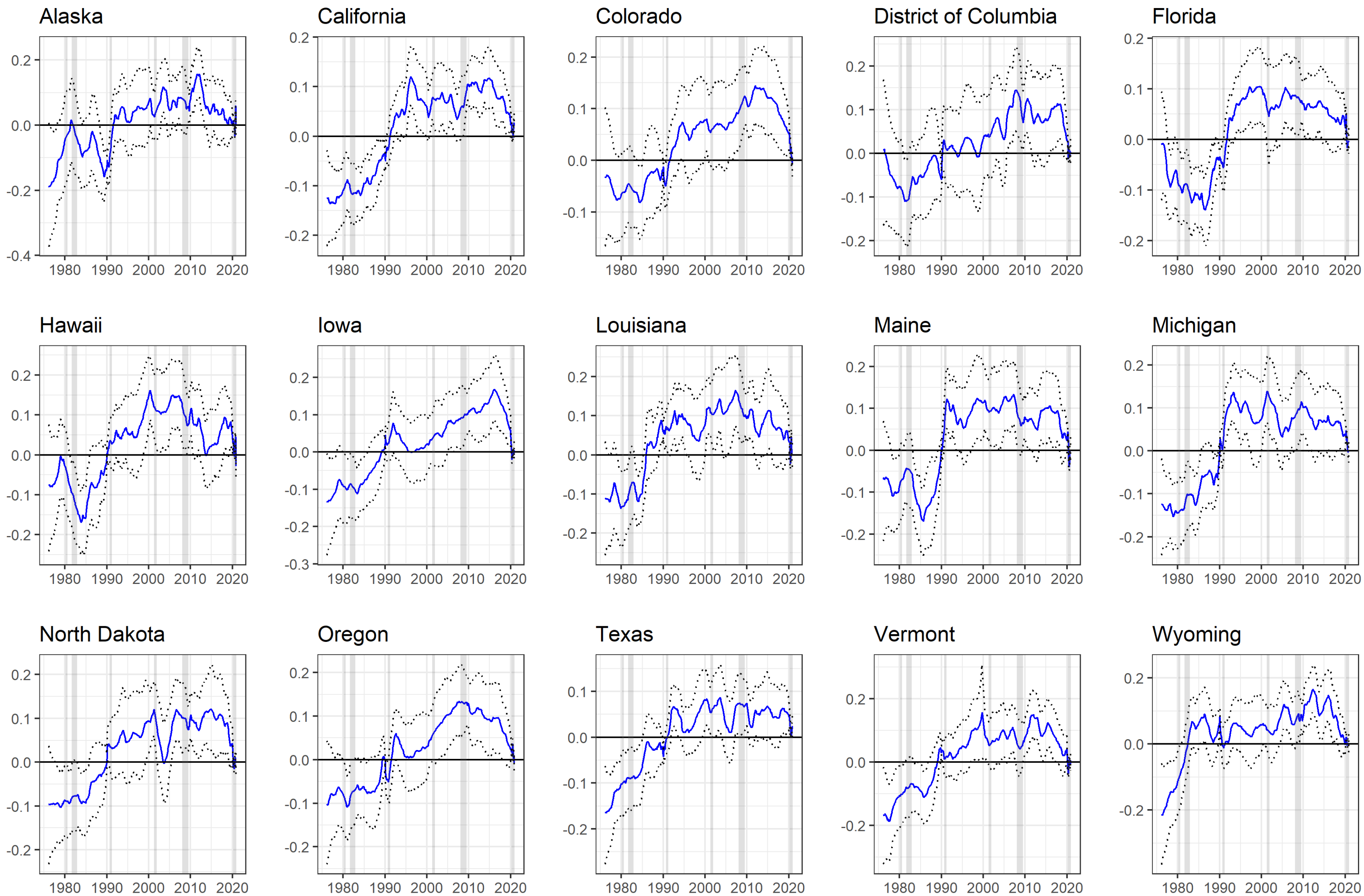


Figure 5: **National Factor Loadings by States Not in the Appalachian Region**

Note: Shaded regions are the NBER-dated recessions. The blue solid lines represent the median of the posterior distribution. Dashed lines represent the 5th and 95th percentiles.

that state's median estimates during the early and late 1980s. Moreover, during theses periods West Virginia has a much larger connection to the regional economy, compared to other states – the estimates ranged from $\pm 0.4$. In the early 1980s, West Virginia exhibits strong and negative regional factor loadings. This would indicate a strong sensitivity or synchronization and a negative relationship with the regional economy. West Virginia and the Appalachian region experienced a coal bust in this period which precipitated a high level of unemployment (Black et al., 2005). This was further exacerbated by the national recessions in the early 1980s. Our model not only distinguishes the national sensitivity from the regional, but our results also indicate that while the connection to the national economy was relatively strong during this time, the connection or influence of the regional economy was stronger.

Approaching the end of the 1980s, the regional factor loadings for West Virginia gradually increased. This strong comovement between the that state's LFPR and the Appalachian factor coincides with the labor growth and expansion in 1987-1988 (Howe and Parks, 1989) and the Pittston Coal Strike of 1989 (Birecree, 1996). Moreover, much of the regional economy shifted away from the reliance on coal and restructured the West Virginia labor market into other industries (Stevens, 1986). Together with the decline in the sensitivity to the national factor (Figure 4) for West Virginia during this time, our results point to a potentially more insulated West Virginian economy. In short, given its state-specific labor market characteristics, West Virginia displayed a strong connection to the regional economy making West Virginia susceptible to regional shocks.

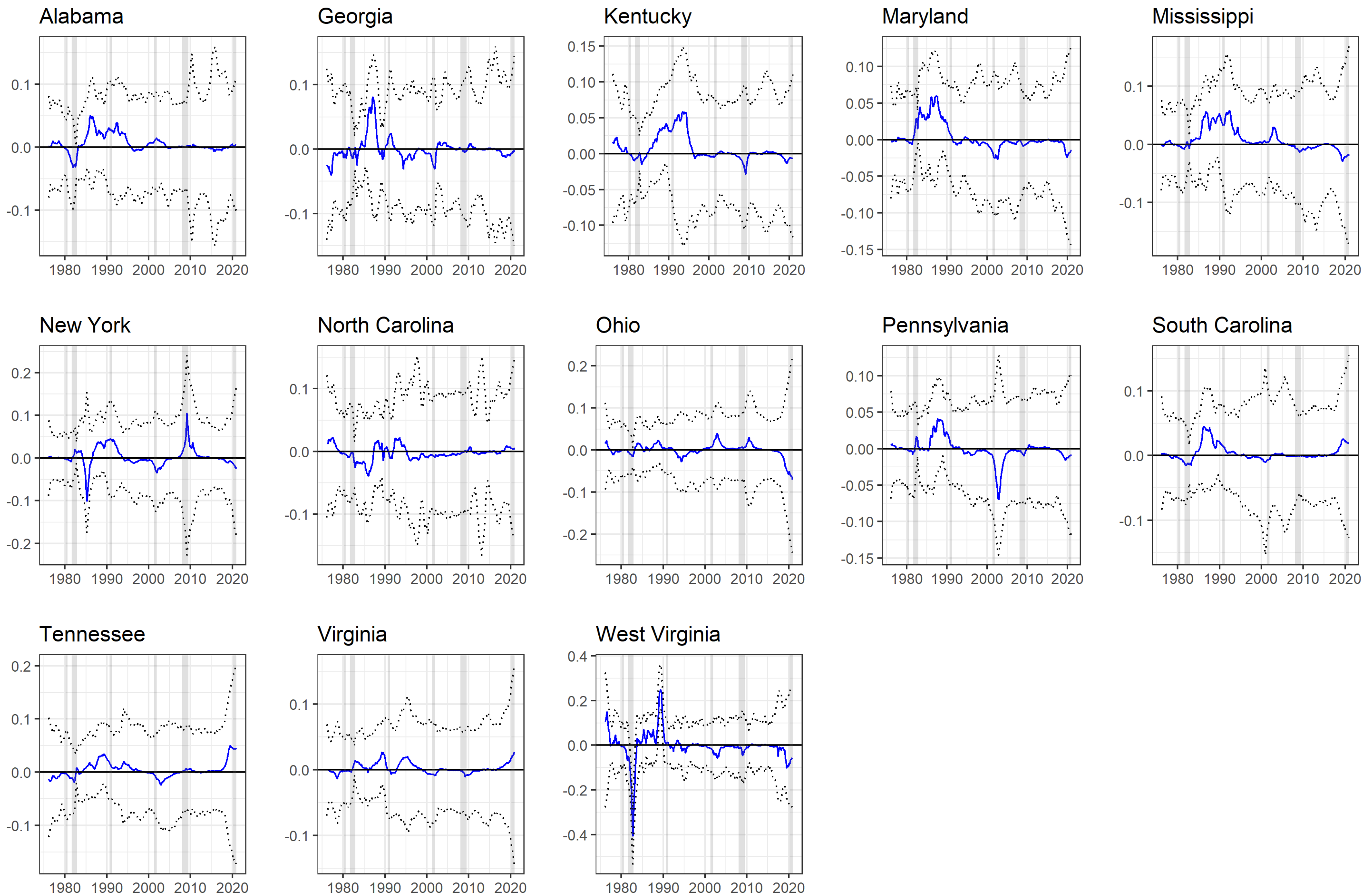


Figure 6: **Regional Factor Loadings by State in the Appalachian Region**

Note: Shaded regions are the NBER-dated recessions. The blue solid lines represent the median of the posterior distribution. Dashed lines represent the 5th and 95th percentiles.

### 1.5.2 Variance Decompositions

From Equation 9, our model implies the following variance decomposition structure:

$$\mathrm{Var}(y_{i,t}) = \omega_{i,t}^2 \ \mathrm{Var}(\mathscr{C}_t) + \widetilde{\beta}_{i,t} \ \mathrm{Var}(\mathscr{R}_t)\widetilde{\beta}'_{i,t} + \ \mathrm{Var}(\xi_{i,t}) \tag{8}$$

The fraction of volatility due to say, the national factor, $\mathscr{C}$, would be:

$$\frac{\omega_{i,t}^2 \ \mathrm{Var}(\mathscr{C}_t)}{\mathrm{Var}(y_{i,t})}$$

Below, we discuss the contribution of each of the three components (of Equation 10) to the state LFPR.

Figures 7 and 8 plot the percentage contributions of the national, regional, and state factors to the total change in LFPR variations for all states within Appalachian and selected, non-Appalachian states [13]. These plots allow us to ascertain the relative importance of each factor in explaining labor market dynamics. Despite an obvious heterogeneity across space and time, we observe that the national and idiosyncratic factors are consistently the dominant contributors. This implies that much of the variations in the state LFPRs is explained by either national labor market trends or individual state circumstances.

Interestingly, concerning time, the contribution of the national factor was most pronounced during periods of recessions, financial crises, or the COVID-19 pandemic. The national factor dominated during the early 1980s, 1990, early 2000s, 2009-2010 and 2020 for most states in and out of the Appalachian region. At times, close to 100% of the variations in the LFPRs variation was explained by the national factor which corresponds to the zero correlation of the state changes in LFPR with national factor seen in Figures 4 and 5. This indicated the national shock led to divergence in the change in LFPR across states. However, outside these periods, the idiosyncratic factor is more important. In other periods, states appear to

[13] Our estimation algorithm included all 50 states (plus Washington D.C.). The full results are available upon request.

be more insulated and more susceptible to state-specific labor and economic shocks.

Additionally, concerning location, we see the idiosyncratic factor is more important in states such as Alaska, Colorado, Iowa, Oregon, and Washington D.C in Figure 8. These states are either far away from the rest of the county or have a unique economic structure that seems to drive these results. Conceivably, geographic isolation at the outskirts or rural center of the U.S. leads to a smaller connection with national trends and more susceptibility to idiosyncratic shock. We found the case of West Virginia to be rather curious as well. Unlike its Appalachian counterparts, the state appeared to be much more insulated outside of national crises windows. In fact, over the sample period, West Virginia again exhibited the closest connection to the Appalachian factor. Figure 9 reveals significant heterogeneity across the states in Appalachian region, but the Appalachian factor explains a large portion of West Virginia LFPR dynamics compared to the other states. On several occasions, the computed contribution surpassed 75%. In the mid-1980s, the Appalachian factor explained almost 100% of the change in LFPR for West Virginia. Most other states, barring New York, experienced much smaller contributions from the Appalachian factor.[14] For the remaining states, incidents of increases in the relative importance of the Appalachian factor appear to coincide with periods of economic recovery and booms.

Additionally, Figure 9 indicates that most states in the Appalachian region exhibit a slightly decreasing trend in the overall importance of the Appalachian region factor throughout the sample period. This indicates that state labor force participation is becoming less influenced by regional shocks or trends (and more influenced by increasingly important national shocks and trends). We conclude, therefore, that changes in the LFPRs are more largely attributable to a state's connection to the national LFPR dynamics and its local labor market during

[14] In 1985 and 2009 around 60% of the variation in New York's LFPR was explained by the Appalachian Factor. These peaks appear to coincide more with recessionary periods in the early 1990s, 2000s and the Great Recession in 2008-2009.

times of economic prosperity. In periods of national turmoil, however, the national factor clearly explains much of the variation but state LFPRs respond more in connection to the idiosyncratic factors. In the discussion that follows, we attempt to place our findings in the context of the extant literature.

Existing studies pose the open-ended question about whether an "Appalachian Effect" and/or cultural factors contribute to lower LFPR in the long run for the Appalachian region. As we briefly discussed in Section 1.2, a consensus has yet to be reached and the question remains unanswered. Since our model and results do not attempt to measure a cultural effect, we focus instead on the temporal aspect of the question. Isserman and Rephann (1993) argue that a cultural effect would be persistent and not "ebb and flow" over time. The authors conclude that if empirical results depend on the choice of year, then the gap in LFPR between the Appalachian region and the rest of the U.S. cannot be caused by an Appalachian culture. The results for the portion of the variance explained by our Appalachian region factor support this conclusion. For example, the Appalachian region factor explains a relatively larger portion of the variance for most of the states in the Appalachian region in the mid-1980s, as seen in Figure 9. This coincides with the strongest Appalachian region effects found by Isserman and Rephann (1993) and Dorsey (1991) with 1987 data. We argue that the gap in LFPR captured by these previous studies in 1987 are potentially driven by a regional shock, such as the coal bust during the late 1980s. Again, this highlights the fact that the choice of year is a critical driver of the postulation in the current literature.

Comparing our findings with more recent studies, Stephens and Deskins (2018) determine that Appalachian counties are 7.1 percentage points lower in LFPR than similar non-Appalachian counties. Through a Blinder-Oaxaca decomposition, the authors determine that of the 7.1 percent difference, 1.1 percentage points remained unexplained by their model. They posit this as evidence supporting an "Appalachian Effect". However, our results show

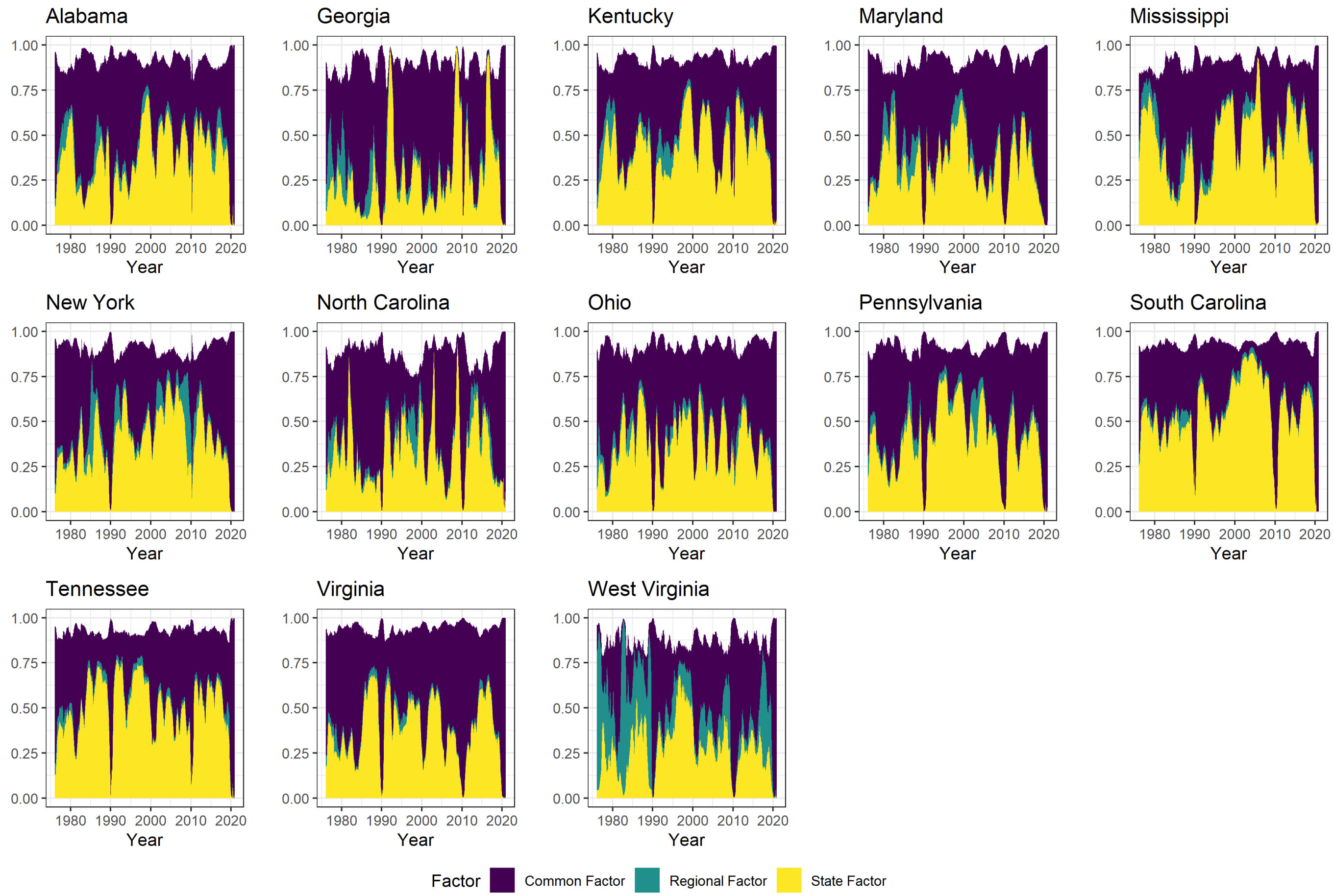


Figure 7: **Variance Contribution of Factor by State in the Appalachian Region**

Note: Colored areas represent the percent contribution of each factor to observed variation in the change in LFPR. Percent contributions are respective medians of the posterior distribution.

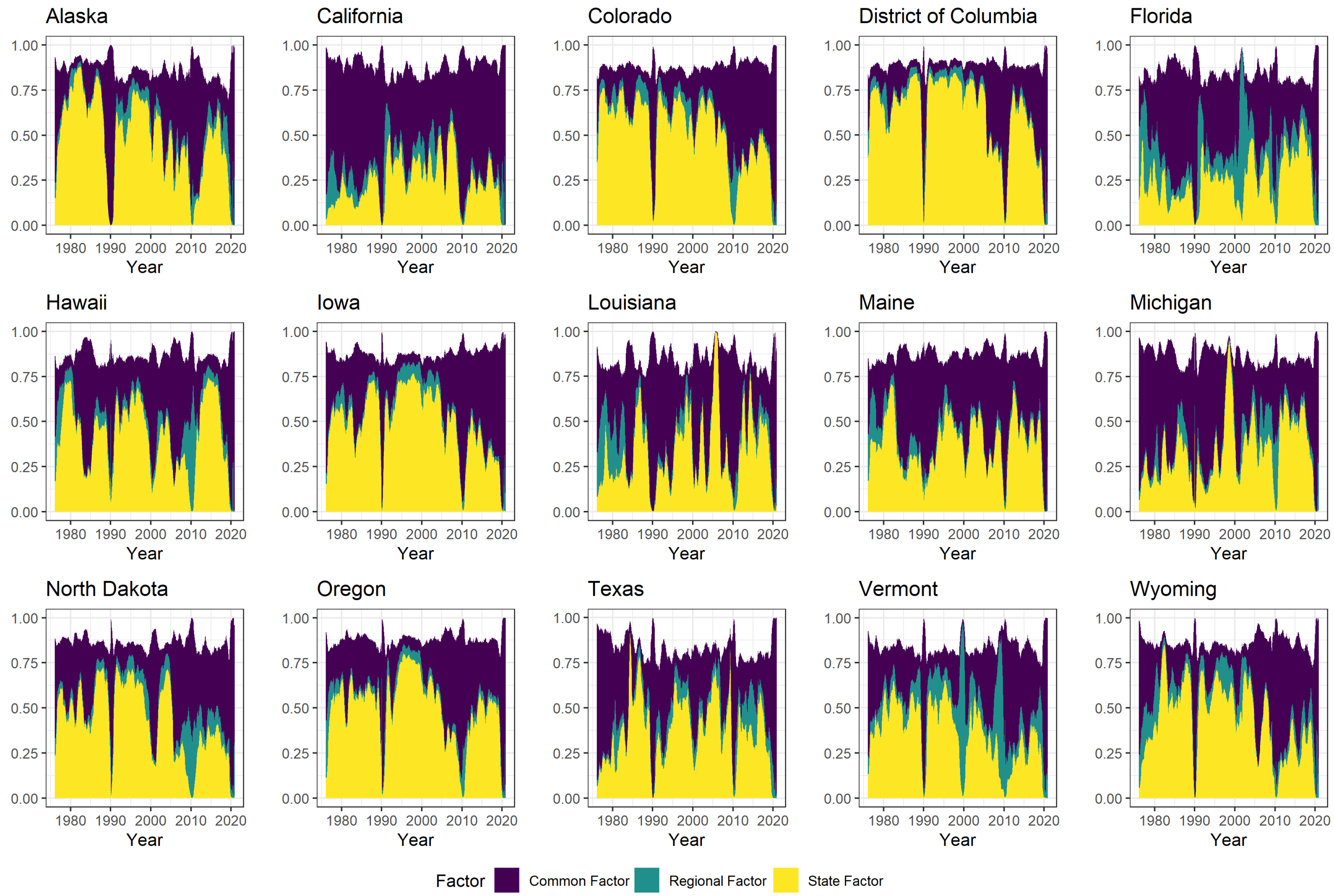


Figure 8: **Variance Contribution of Factor by State**

Note: Colored areas represent the percent contribution of each factor to observed variation in the change in LFPR. Percent contributions are respective medians of the posterior distribution.

a weak relationship between change in state LFPR and the regional economy in 2000 and 2010. Figure 6 shows that for these years the regional factor contributes little to the variance in the change of LFPR. Moreover, we find a stronger relationship with the national economy over these years. For example, on average across the Appalachian states, the national factor explains over 88% of the variation in the change in LFPR in April of 2010. This may be explained by national economic trends and events such as decreases in unemployment and layoffs throughout 2010 and the signing of the Jobs Bill (HIRE) and the Affordable Care Act (ACA) in March 2010, by President Obama. Figure 7 shows that for Maryland, North Carolina, Pennsylvania, South Carolina, Virginia, and West Virginia, the variation explained by the national factor exceed 98% for the middle part of 2010. Again, since our methodology centers around changes in LFPR, the weaker influence of the Appalachian region factor for 2000 and 2010 for most of the states in the region may simply reflect the lack of major regional economic or labor market events during these years. However, given our results, we suggest that the unexplained variation in Stephens and Deskins (2018) may be related to a disproportionate effect of national economic shocks or events during this time on an already economically distressed Appalachian region.

Figure 13 presents the average cross-state time-varying correlation. There is a general upward trend in the median correlation statistic over the sample period. In congruence with our results from Figures 4 and 5, we especially see evidence of the divergence in the change in state LFPR in 1990 and 2020. In Figure 13 there is an asymmetric response in the correlation during periods of crisis. During the recession of the late 1990s and the Great Recession (2008-2009), the average cross-state correlation increased. This indicates that during and immediately following these episodes, the LFPRs across states were becoming more synchronized. To this effect, policies aimed at dampening labor market shocks and encouraging recovery could prove more effective as a one-size-fits-all approach becomes more

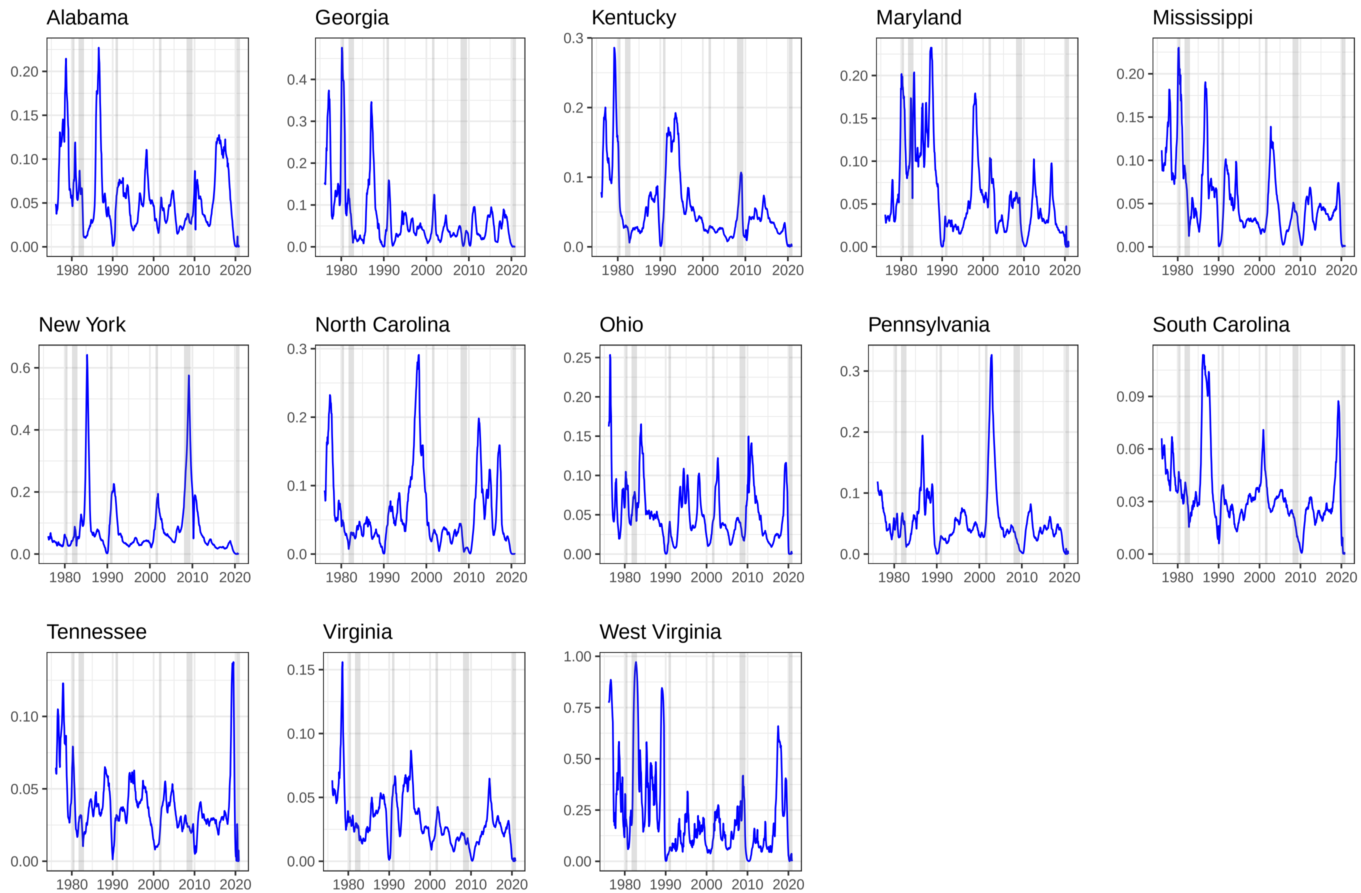


Figure 9: **Variance Contribution of the Appalachian Regional Factor by State**

Note: Shaded regions are the NBER-dated recessions. The blue solid lines represent the median of the posterior distribution.

useful. Figure 13 also show that over time cross-state correlation is increasing in general. Together with our previous results from the variance contribution and factor loadings, we can then conclude that changes in state LFPRs are becoming more connected with the national factor over time.[15]

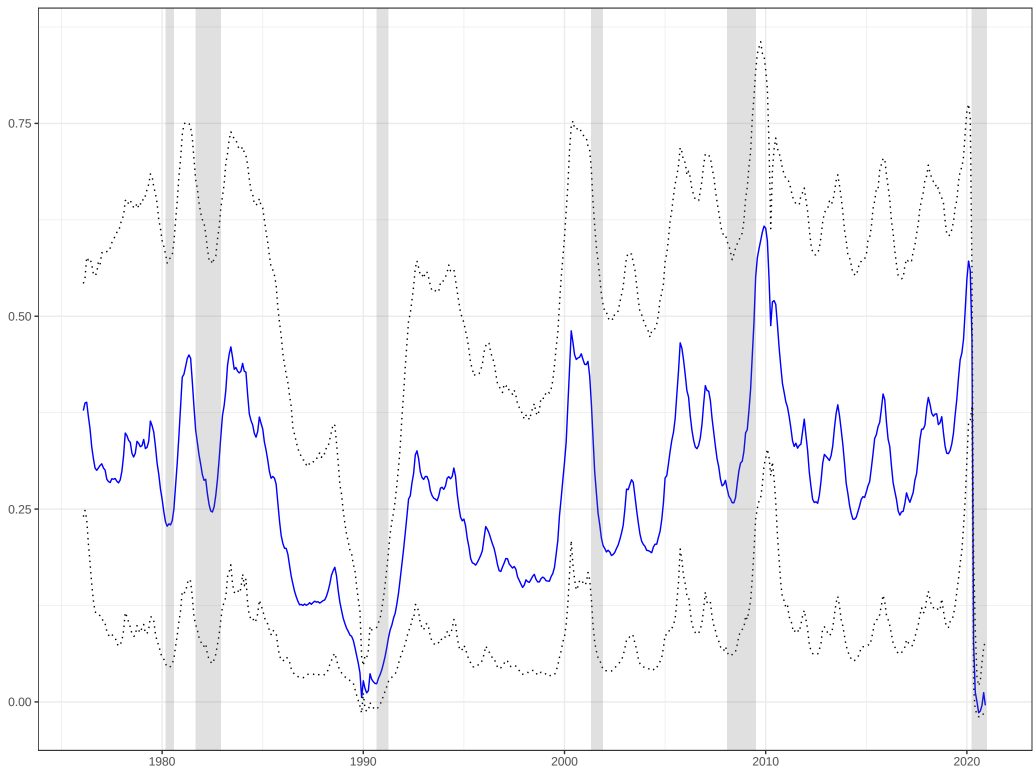


Figure 10: **Average Cross State Correlation (All States)**

Note: Shaded regions are the NBER-dated recessions. The blue solid lines represent the median of the posterior distribution. Dashed lines represent the 5th and 95th percentiles.

[15]Since LFPs were unusually low in 2020, we checked the robustness of our result re-estimated our model results to be at least partially driven by the outlying market conditions during that time. Our results excluding 2020 and the sensitivity of our model and findings to the 2020 data are similar to our primary results and so they are excluded for brevity but are available upon request.

## 1.6 Conclusion

In this paper, we demonstrate that comovement or divergence of the change in state LFPR varies over time, geographic level, and economic and labor market fluctuations. In particular, we examine the relative contribution of a latent regional factor to the labor force dynamics of states in Appalachia over time. Using a dynamic factor model with time-varying parameters and stochastic volatility, we show that the choice of year and state together with the economic conditions are crucial to the relative importance of each factor's contribution to the LFPR.

We determine that national and state-specific factors play a dominant role in explaining the change in LFPR variations for most U.S. states during periods of recessions and general economic downturns. During large national recessions, we find divergence in the change in U.S. state LFPR. However, during times of recovery and expansion, change in state LFPR is more synchronized with the national factor. We determine that the influence of the Appalachian factor increases during economic expansions or recoveries, albeit still smaller in magnitude than the common and idiosyncratic components. We conclude that these findings may influence the results of cross-sectional studies using LFPR given the years and economic and labor market conditions over which the data covers.

Consistent with other studies, we find that West Virginia displays the strongest connection to the Appalachian region factor. The regional factor persistently contributes more than 1/2 of the volatility observed in that state's change in LFPR. This is far higher than any of the other 12 states. Additionally, the change in LFPR for West Virginia synchronizes greatly with the Appalachian region factor during two periods of regional labor market shifts (Coal Boom and Bust). This curious observation offers an avenue for further investigation. Our approach and results highlight the need for policies that accommodate both the state of the economy and state-specific characteristics.

### 1.6.1 Policy Implications

Our results are important for policymakers and potential improvements in regional and national output growth. Federal labor policy is more effective when LFP is highly synchronized across the nation. However, when regions of the U.S. exhibit distinct behavior or if states themselves exhibit more individual behavior, then more localized and targeted labor policies would be more efficient.

In short, given the relationship between LFP and output, increasing regional and local participation would stimulate output growth. However, our results point to the need for disparate government responses during crises and booms. There is a need for a federal response during periods of economic growth and economic recovery as states become more homogeneous and connected to the national factor. During periods of economic crisis and pandemics, there is a greater need for state and region-specific policies. By varying the level of policy interventions during different stages of the business cycle, economically distressed areas would experience more targeted labor market policies than a one-size-fits-all prescription. It is also important to note that state LFPRs are gradually growing more synchronized and connected to the national factor. While this presents opportunities to implement more effective federal-level policies to assist depressed labor markets, it also reduces the nation's ability to absorb labor market shocks. This has important long-term implications for the Appalachian region since economically distressed areas are already more vulnerable to economic shocks.

It is also important to note that while the Appalachian region factor explains significant portions of the variance in West Virginia throughout the sample period, we find that the connection or synchronization with the Appalachian region factor is near-zero. Zero correlation with the Appalachian region factor is most likely due to the lack of regional shocks induced by a gradual shift from dependence on natural resource employment to other industries. However, given the level of contributions of the Appalachian region to the observed variance in the change in LFPR in recent years (Figure 9), we expect that West Virginia is not completely impervious to regional economic shocks.

### 1.6.2 Limitations and Avenues for Future Work

Given that this study is limited to state-level data we do not address concerns in (Isserman and Rephann, 1993) related to the idea that the geographic level of the data may significantly contribute to certain findings.[16] A future avenue for research could focus on a more disaggregated analysis to account for differences across the rural-urban spectrum and county inclusion/exclusion within Appalachia. While we specifically emphasize the macroeconomic nature of state-level LFPR, further investigation into the impact on the rural/urban divide of these results may be warranted. This is buoyed by the fact that West Virginia appears to be structurally different– as evidenced by the persistently low labor force participation and relatively large variance contributions of the Appalachian region factor. While it is "all is quiet on the Appalachian front" regarding regional results for most Appalachian states, West Virginia stands out. Since West Virginia is the only state with all counties designated in the Appalachian Region, further research may provide insight into narrowing down problematic sub-regions and why certain areas remain economically distressed. Moreover, this will allow for direct comparability with the extant literature.

Lastly, our results prompt questions about the relationship between the national, regional, and state factors and known drivers of labor force participation. Some examples pertain to when and how West Virginia adjusts to shocks in LFP, and how much of the variation and error realization of the latent factors are explained by unexpected changes in other

[16] Due to the high computational burden of our model and the curse of dimensionality, we were not able to explore this avenue. We would expect to see more diversity within and between states and counties and a potentially more pronounced Appalachian factor.

factors and included variables. More research is needed to ameliorate decades of low labor force participation and maximize the growth potential for West Virginia and the rest of the Appalachian region.

## 1.7 Appendix A

Table A.1: **State LFP Descriptive Statistics**

| States | Mean | Median | Min | Max | S.D. |
|---|---|---|---|---|---|
| Alabama | 60.50 | 60.90 | 55.90 | 64.50 | 2.29 |
| Alaska | 70.87 | 71.90 | 61.10 | 75.30 | 2.72 |
| Arizona | 63.16 | 63.50 | 59.10 | 67.10 | 2.00 |
| Arkansas | 60.98 | 61.15 | 56.20 | 64.20 | 2.02 |
| California | 65.12 | 65.70 | 59.20 | 68.00 | 1.77 |
| Colorado | 70.54 | 70.60 | 64.90 | 74.30 | 2.06 |
| Connecticut | 67.64 | 67.60 | 63.30 | 71.30 | 1.71 |
| Delaware | 65.95 | 66.60 | 60.10 | 70.90 | 3.00 |
| District of Columbia | 67.50 | 67.40 | 63.00 | 72.10 | 2.20 |
| Florida | 60.63 | 61.40 | 54.90 | 63.70 | 2.32 |
| Georgia | 65.95 | 66.30 | 59.40 | 69.30 | 2.34 |
| Hawaii | 65.58 | 66.40 | 56.20 | 69.90 | 2.54 |
| Idaho | 66.82 | 66.60 | 62.70 | 71.40 | 2.33 |
| Illinois | 66.42 | 66.20 | 60.40 | 70.00 | 1.70 |
| Indiana | 66.09 | 66.10 | 61.20 | 70.90 | 1.96 |
| Iowa | 69.72 | 70.00 | 64.10 | 73.50 | 2.53 |
| Kansas | 68.98 | 69.10 | 64.90 | 71.50 | 1.61 |
| Kentucky | 61.51 | 62.00 | 56.00 | 63.70 | 1.59 |
| Louisiana | 60.54 | 60.80 | 54.60 | 68.70 | 1.45 |
| Maine | 64.84 | 65.15 | 58.60 | 68.80 | 2.35 |
| Maryland | 68.78 | 69.00 | 63.00 | 71.50 | 1.65 |

Continued on next page

**Table A.1 – continued from previous page**

| States | Mean | Median | Min | Max | S.D. |
|---|---|---|---|---|---|
| Massachusetts | 66.93 | 67.10 | 60.40 | 69.40 | 1.25 |
| Michigan | 64.09 | 64.20 | 57.40 | 68.80 | 2.34 |
| Minnesota | 71.87 | 71.50 | 65.40 | 75.70 | 2.37 |
| Mississippi | 59.63 | 59.80 | 53.30 | 63.30 | 2.39 |
| Missouri | 66.18 | 66.00 | 59.80 | 71.00 | 2.65 |
| Montana | 65.95 | 66.60 | 61.40 | 69.00 | 1.91 |
| Nebraska | 70.59 | 71.30 | 64.80 | 74.10 | 2.53 |
| Nevada | 68.72 | 69.70 | 58.00 | 73.50 | 3.37 |
| New Hampshire | 70.35 | 70.90 | 65.10 | 73.60 | 1.84 |
| New Jersey | 65.36 | 65.90 | 61.40 | 67.60 | 1.48 |
| New Mexico | 61.56 | 62.40 | 55.00 | 63.90 | 2.07 |
| New York | 61.42 | 61.60 | 56.80 | 63.60 | 1.41 |
| North Carolina | 65.68 | 66.60 | 56.20 | 69.00 | 2.52 |
| North Dakota | 69.65 | 70.50 | 62.30 | 74.70 | 2.96 |
| Ohio | 64.82 | 64.65 | 59.80 | 67.70 | 1.78 |
| Oklahoma | 62.96 | 63.60 | 58.90 | 65.50 | 1.60 |
| Oregon | 65.71 | 66.00 | 59.20 | 68.90 | 2.31 |
| Pennsylvania | 62.56 | 63.10 | 58.30 | 65.30 | 1.88 |
| Rhode Island | 66.13 | 66.30 | 59.40 | 68.40 | 1.44 |
| South Carolina | 63.30 | 63.90 | 56.60 | 66.90 | 2.73 |
| South Dakota | 69.92 | 70.10 | 64.30 | 73.20 | 2.32 |
| Tennessee | 62.90 | 62.90 | 58.00 | 67.20 | 2.02 |
| Texas | 66.90 | 67.30 | 60.20 | 69.40 | 1.96 |

Continued on next page

**Table A.1 – continued from previous page**

| States | Mean | Median | Min | Max | S.D. |
|---|---|---|---|---|---|
| Utah | 69.22 | 69.40 | 62.50 | 73.40 | 2.77 |
| Vermont | 69.34 | 70.45 | 60.90 | 72.60 | 2.36 |
| Virginia | 67.58 | 67.80 | 63.20 | 70.90 | 1.54 |
| Washington | 66.20 | 66.30 | 60.60 | 69.90 | 2.22 |
| West Virginia | 54.09 | 54.65 | 51.00 | 56.20 | 1.55 |
| Wisconsin | 69.76 | 69.30 | 65.40 | 74.50 | 2.44 |
| Wyoming | 69.69 | 70.35 | 64.10 | 72.40 | 2.01 |

Note: Statistics reflect the state-level labor force participation rates over the sample period January 1976 – December 2020. S.D refers to the standard deviation.

Table A.2: **Composition of Appalachia by State**

| State | Counties in Appalachia (%) | Percent of Appalachia |
|---|---|---|
| Alabama | 55.22 | 8.81 |
| Georgia | 23.27 | 8.81 |
| Kentucky | 45.00 | 12.86 |
| Maryland | 12.50 | 0.71 |
| Mississippi | 29.27 | 5.71 |
| New York | 22.58 | 3.33 |
| North Carolina | 29.00 | 6.90 |
| Ohio | 36.36 | 7.62 |
| Pennsylvania | 77.61 | 12.38 |
| South Carolina | 13.04 | 1.43 |
| Tennessee | 54.74 | 12.38 |
| Virginia | 18.38 | 5.95 |
| West Virginia | 100.00 | 13.10 |

Table A.3: **Composition of Regions By State**

| States | Non-Appalachia | Appalachia | Plains | Mideast | Great Lakes |
|---|---|---|---|---|---|
| Alabama | | ✓ | | | |
| Alaska | ✓ | | | | |
| Arizona | ✓ | | | | |
| Arkansas | ✓ | | | | |
| California | ✓ | | | | |
| Colorado | ✓ | | | | |
| Connecticut | ✓ | | | | |
| Delaware | ✓ | | | ✓ | |
| District of Columbia | ✓ | | | ✓ | |
| Florida | ✓ | | | | |
| Georgia | | ✓ | | | |
| Hawaii | ✓ | | | | |
| Idaho | ✓ | | | | |
| Illinois | ✓ | | | | ✓ |
| Indiana | ✓ | | | | ✓ |
| Iowa | ✓ | | ✓ | | |
| Kansas | ✓ | | ✓ | | |
| Kentucky | | ✓ | | | |
| Louisiana | ✓ | | | | |
| Maine | ✓ | | | | |
| Maryland | | ✓ | | ✓ | |
| Massachusetts | ✓ | | | | |
| Michigan | ✓ | | | | ✓ |

Continued on next page

**Table A.3 – continued from previous page**

| States | Non-Appalachia | Appalachia | Plains | Mideast | Great Lakes |
|---|---|---|---|---|---|
| Minnesota | ✓ | | ✓ | | |
| Mississippi | | ✓ | | | |
| Missouri | ✓ | | ✓ | | |
| Montana | ✓ | | | | |
| Nebraska | ✓ | | ✓ | | |
| Nevada | ✓ | | | | |
| New Hampshire | ✓ | | | | |
| New Jersey | ✓ | | | ✓ | |
| New Mexico | ✓ | | | | |
| New York | | ✓ | | ✓ | |
| North Carolina | | ✓ | | | |
| North Dakota | ✓ | | ✓ | | |
| Ohio | | ✓ | | | ✓ |
| Oklahoma | ✓ | | | | |
| Oregon | ✓ | | | | |
| Pennsylvania | | ✓ | | ✓ | |
| Rhode Island | ✓ | | | | |
| South Carolina | | ✓ | | | |
| South Dakota | ✓ | | ✓ | | |
| Tennessee | | ✓ | | | |
| Texas | ✓ | | | | |
| Utah | ✓ | | | | |
| Vermont | ✓ | | | | |

Continued on next page

**Table A.3 – continued from previous page**

| States | Non-Appalachia | Appalachia | Plains | Mideast | Great Lakes |
|---|---|---|---|---|---|
| Virginia | | ✓ | | | |
| Washington | ✓ | | | | |
| West Virginia | | ✓ | | | |
| Wisconsin | ✓ | | | | ✓ |
| Wyoming | ✓ | | | | |

Note: We define states to be included in the Appalachian region if they have at least one county located in the region as defined by the Appalachian Region Commission (ARC). States included in Plains, Mideast and Great Lakes regions are defined by the U.S. Bureau of Economic Analysis.

### 1.7.1 Results Continued

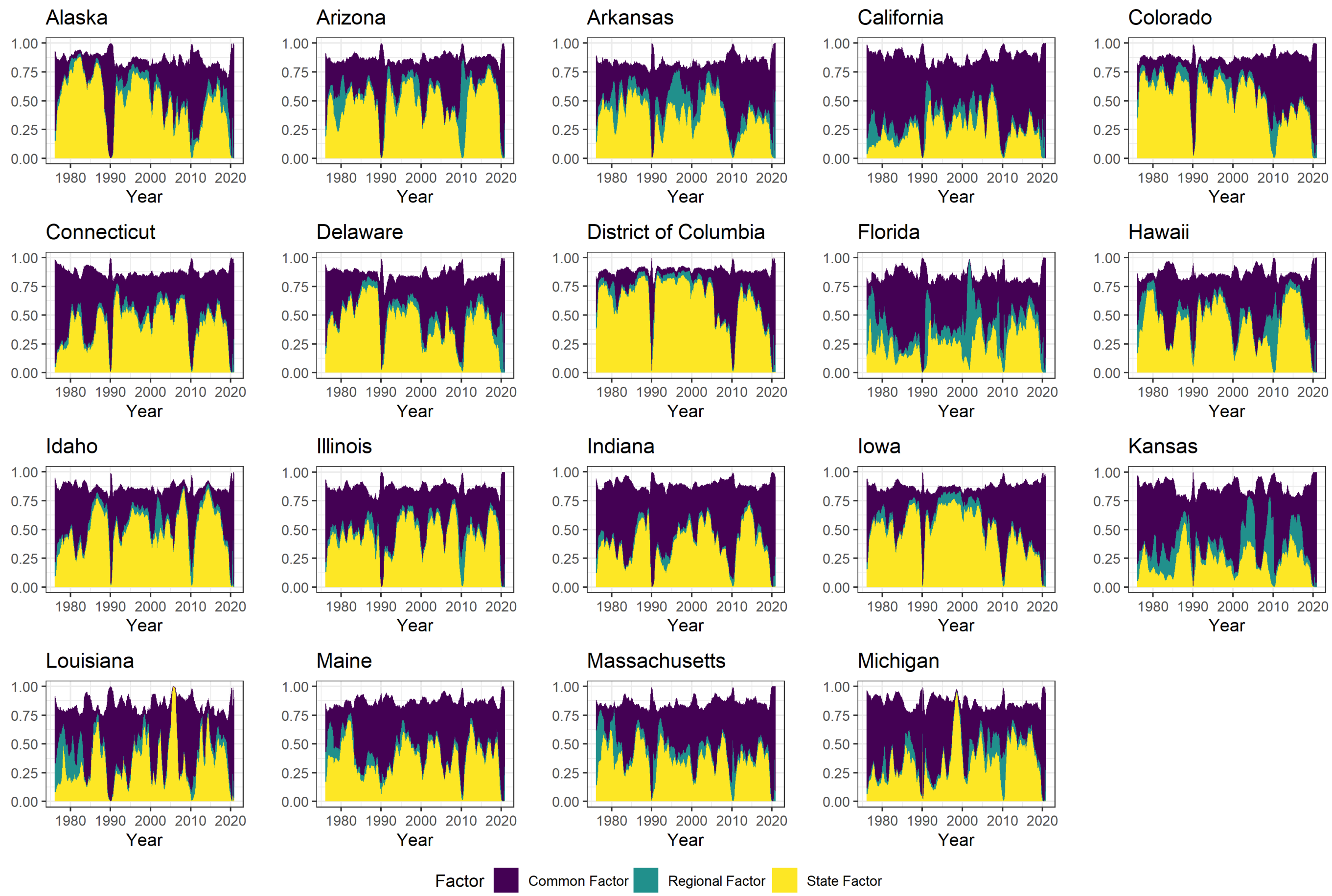


Figure A.1: **Variance Contribution of Factor by State Not in the Appalachian Region**

Note: These estimations include 2020 and the rest of the state results that are not included in the main body of the paper.

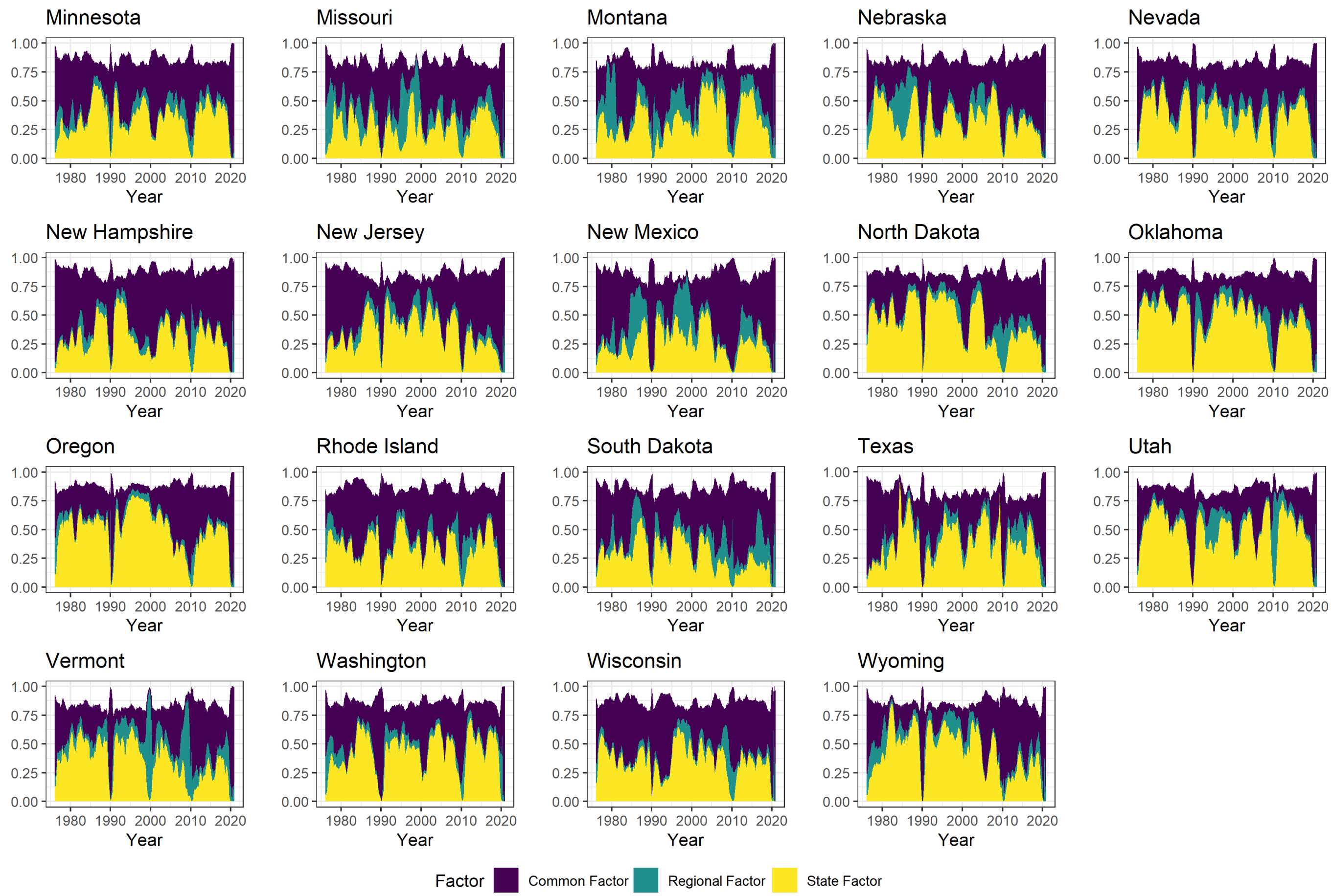


Figure A.2: **Variance Contribution of Factor by State Not in the Appalachian Region**

Note: These estimations include 2020 and the rest of the state results that are not included in the main body of the paper.

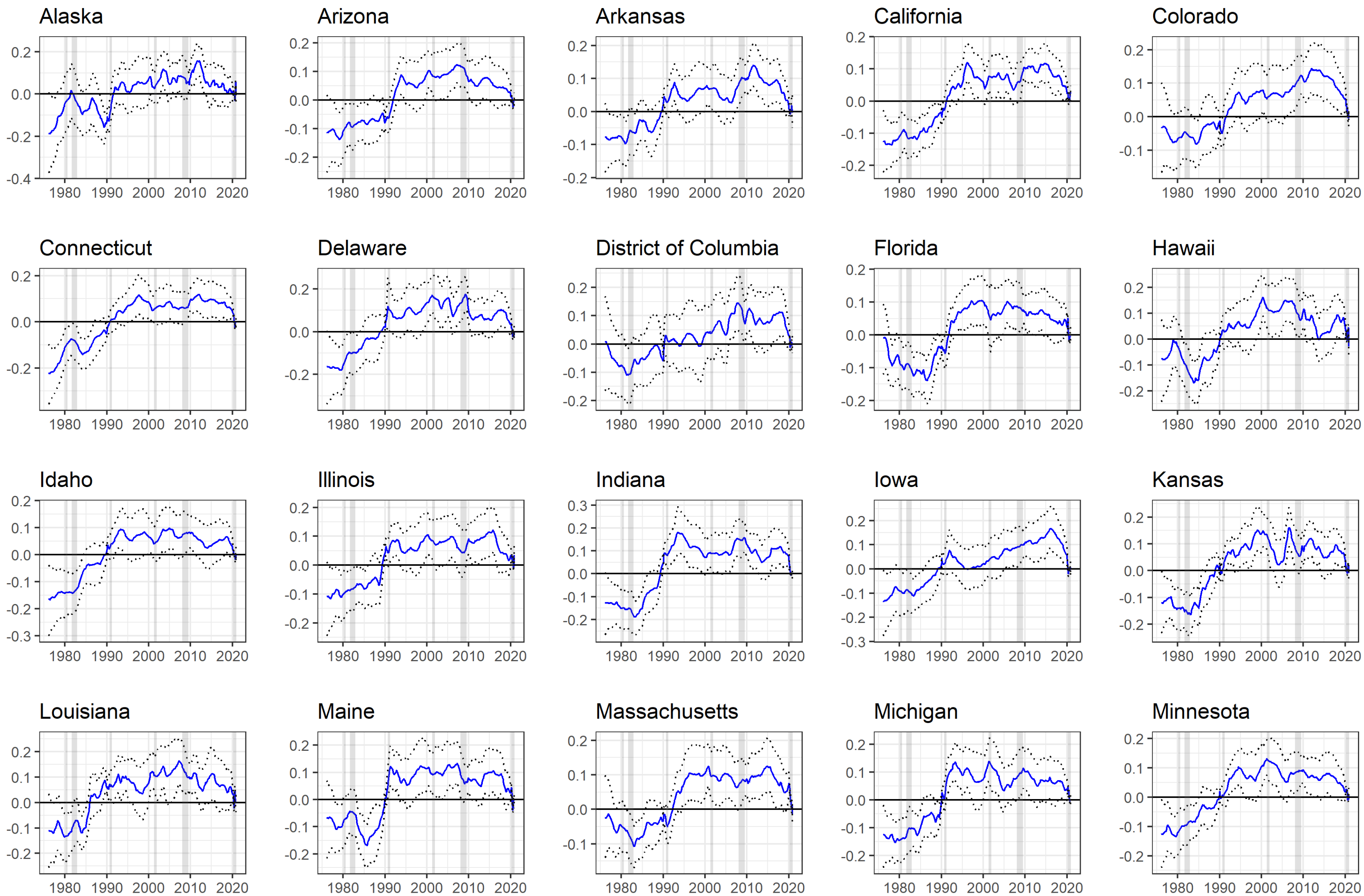


Figure A.3: **National Factor Loadings by State**

Note: Shaded regions are the NBER-dated recessions. The blue solid lines represent the median of the posterior distribution. Dashed lines represent the 5th and 95th percentiles. These include state results that are not included in the main body of the paper.

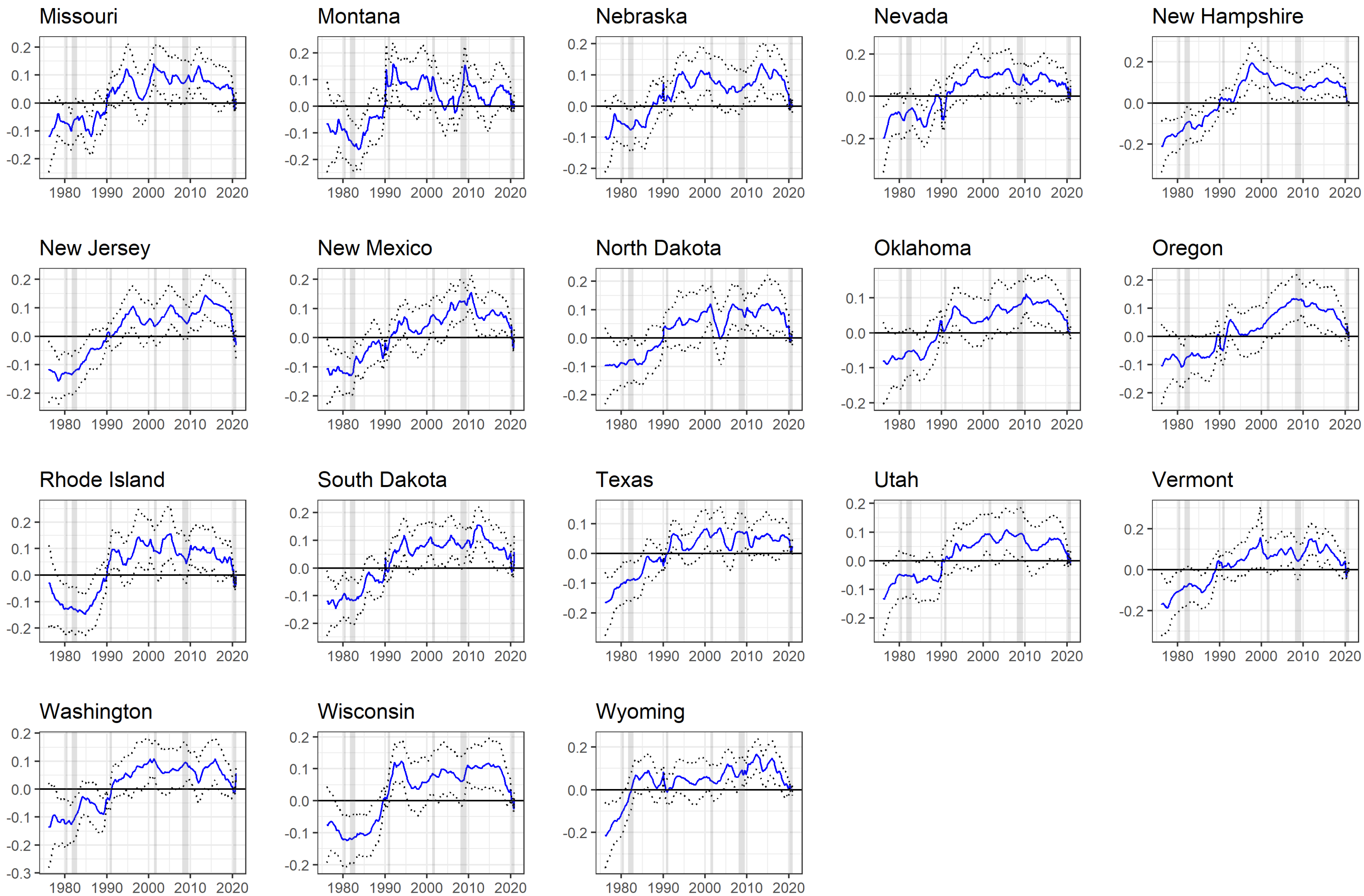


Figure A.4: **National Factor Loadings by State**

Note: Shaded regions are the NBER-dated recessions. The blue solid lines represent the median of the posterior distribution. Dashed lines represent the 5th and 95th percentiles. These include state results that are not included in the main body of the paper.

# Chapter 2

# Labor Force Participation Rate Comovements in Metro/Non-Metro West Virginia Counties

## 2.1 Introduction

The decline of the U.S. labor force participation rate (LFPR) in recent years has sparked a growing interest in various aspects of labor force participation (LFP) as an indicator has for future employment and economic growth. One such topic that has received attention from scholars regards spatial differences in LFPRs. Differences, especially across the urban/rural divide, have been found in various countries and across demographic subgroups such as women and poor households (see Kilkenny and Huffman, 2003; Maurer-Fazio et al., 2005; Fogli and Veldkamp, 2011; Mryyan, 2014; Stephens and Deskins, 2018; Sakanishi, 2020, for example). However, measuring the degree of synchronicity of LFPRs within and across urban and rural areas has received relatively little attention in the literature. Specifically, synchronized labor force participation (LFP) decisions linked within economically distressed economies, such as the state of West Virginia, may dramatically impact local labor markets and contribute to persistent labor market distress. The stronger the linkages between areas, the greater the impact labor market shocks will have on the larger economy. These compounded effects imply potential obstacles to overcome in increasing labor force participation and realizing labor and economic growth in rural and distressed areas of the U.S.

West Virginia, compared to the rest of the U.S., exhibits persistent economic disparity resulting from a number of factors. The literature has identified several contributors to this low economic performance such as low income, poor education outcomes, inadequate health-care resources, high poverty levels, geographic isolation, and low LFPRs (Billings, 1974; Stephens and Deskins, 2018; Isserman and Rephann, 1993; Behringer and Friedell, 2006; Muntaner and Barnett, 2000). Efforts to raise education and income levels in the area began in 1960 with a visit from then-presidential candidate John F. Kennedy to McDowell County, West Virginia. Since then, West Virginia has remained a focus of economic development given the development of the Appalachian Region Development Act in 1965. However, persistently poor performance in economic indicators such as GDP, unemployment, and LFP continues to impede economic growth and prosperity in the state. Consequently, West Virginia remains one of the most economically depressed regions in the U.S. Since improvements in LFPRs directly increase employment growth and GDP(see Bryant et al., 2004; Juhn and Potter, 2006; Shoven, 2007; Cai and Lu, 2013; Bustelo et al., 2019, for example), West Virginia has an opportunity to expand upon its own economic growth, development, and prosperity.

While the inequality around skills, income, and other indicators in West Virginia have not gone unnoticed by policymakers or the media, we also focus on West Virginia since LFP in the state has received relatively little attention by economists (Dorsey, 1991). The few studies that include West Virginia in their regional LFPR analyses find peculiar results for West Virginia and call for further research to investigate the state more (Stephens and Deskins, 2018; Isserman and Rephann, 1993; Dorsey, 1991; Beverly et al., 2022). The overall lack of attention on West Virginia over time, however, is unfortunate since knowledge of county LFPR comovements is essential to understanding labor market inequality. Also, how LFPR comovements are related to spatial influences is critical for improving regional economic development efforts that can be applied to other areas as well. Answering broader questions on the effects of labor market shocks in local economies may better inform labor policy and help reverse persistent under performance in distressed areas.

In the first stage of this study, we examine the role and relative importance of state, county, and metropolitan/non-metropolitan influences on the change in West Virginia county LFPRs. We decompose West Virginia county LFPRs into latent factors that measure comovement at these levels. The state, county, and metropolitan and non-metropolitan factors are estimated using a Dynamic Factor Model (DFM) with time-varying (TV) and stochastic volatility (SV) parameters. We assume that these few factors which capture state, metropolitan/non-metropolitan, and county comovements describe changes in each county's LFPR. We then determine the relative importance of each factor by calculating the percent of the variance in each county's LFPR at each time period that is explained by each factor. In the second stage of the study, we investigate these percentages further by regressing these percent contributions of the factors to LFPR variations on county characteristics such as demographics, industry composition, life expectancy, etc. This part of the analysis allows us to determine which county characteristics best explain a county's sensitivity to state and metropolitan/non-metropolitan influences on LFPRs and through which channels LFPRs can be increased.

Most other studies on labor dynamics tend to focus on dynamics of unemployment (see, Elsby et al. (2013, 2009); Petrongolo and Pissarides (2008); Fujita and Ramey (2006) for example). While the unemployment rate is often the popular choice for assessing economic health, it has become less reliable as an indicator (Juhn and Potter, 2006). For example, high unemployment may lead workers to become discouraged in their job search and drop out of the labor force. This lowers the unemployment rate and falsely indicates improvements in the economy. We study the LFPR since it better accounts for individuals who have dropped out of the labor force and therefore provides a truer representation of labor market conditions (Hotchkiss and Rios-Avila, 2013; Stephens and Deskins, 2018). One study that does focus on labor force participation (LFP) dynamics is Epstein

(2018). In this study, the author analyzes the dynamics between LFP, Gross Domestic Product (GDP), and labor force productivity. Epstein (2018) focuses on the cyclical nature of LFP does not directly study comovements in LFPRs or economically distressed regions.

As it is well established that LFPRs vary across counties and metropolitan and non-metropolitan areas, several studies highlight the importance of distinguishing between rural and urban populations and controlling for spatial differences in empirical LFP or unemployment studies (Stephens and Deskins, 2018; Weingarden et al., 2017; Phimister et al., 2002; Elhorst, 2003; Kilkenny and Huffman, 2003). A select number of studies have accounted for rural/urban differences in dynamic LFP analyses. In Voicu (2001) the authors use rural and urban status in a dynamic LFP search model to study individual LFP decisions in Romania. Hamrick (1997) investigates how rural and urban labor markets respond to business cycle movements. To our knowledge, ours is the first study to investigate county comovements in the change in LFPRs over time. Our research at the sub-state level shifts the focus of labor policy from national interventions to regional differences and actionable policy within state borders.

In our two-stage analysis, we first find that in recent years a common state-wide component has influenced change in West Virginia county LFPRs, signaling a synchronization of county labor markets. This makes West Virginia counties and the state overall more vulnerable to broad labor market shocks. We also find that non-metropolitan county LFPRs are persistently correlated with the collective movements of non-metropolitan LFPRs over time. This, again, implies a vulnerability for non-metropolitan counties to shocks, but also provides an opportunity to target many counties simultaneously with non-metropolitan specific policy, given that these areas are known to have lower LFPRs and poorer economic activity.

However, we also find evidence of idiosyncratic behavior across the state as well. About 57% and 65% of the variation in change in metropolitan and non-metropolitan county LFPRs is explained by county-specific components, respectively. This calls for targeting labor and economic issues more individually across the state rather than implementing state-wide "one-size-fits-all" solutions.

Lastly, we find that county demographics, education levels, income, access to interstate highways, and industry composition are strongly related to the importance of our estimated state, county, and metropolitan and non-metropolitan factors. For non-metropolitan counties specifically, policymakers should focus on formalizing informal/nonstandard work, and increasing wage income, employment opportunities, and work force development programs. For West Virginia in general, policymakers should work to support and diversify key industries in West Virginia through building infrastructure, creating accessibility, and developing a healthy, well educated, and well trained work force. A united front and a targeted focus on programs that build on the State's unique assets and capitalize on industries such as manufacturing and gas, where West Virginia outperforms its neighboring states, provides a unique opportunity to increase jobs, infrastructure, wages, and attract potential businesses.

The remainder of the paper follows the following organization. In Section 3.3, we present summary statistics and describe the data we use in our DFM (stage one) and regression (stage two) analyses. Section 3.4 presents a discussion of the empirical methodology for both stages. We discuss our results in Section 3.5. Section 3.6 concludes and offers potential policy recommendations.

## 2.2 Data

### 2.2.1 Labor Force Participation Rates

This study expands on the work of Beverly et al. (2022) by using county-level LFPRs in a DFM-TV-SV model. We calculate the LFPR according to the BLS definition as the fraction of the civilian and non-institutional population between the ages of 16-64 participating in the labor force[1]. We obtain labor force and population data over 1990-2020 from the BLS and the U.S. Census Bureau,

[1] The BLS also defines the labor force to be individuals who are either working or actively searching for work.

respectively. [2]. By utilizing 30 years of monthly LFPR data, we can capture long-term and major economic trends that affect the state of West Virginia and measure the influence of metropolitan and non-metropolitan regions.

To define metropolitan and non-metropolitan, we follow the USDA which sorts each county into nine subcategories defined by population and proximity to a metropolitan areas, described in Table 1. Counties are divided into three metropolitan and six non-metropolitan subcategories. We use these general metropolitan or non-metropolitan classifications to distinguish the counties in West Virginia for our analysis into metropolitan and non-metropolitan regions. That is, we classify counties with RUCC (1, 2, and 3) as metropolitan and RUCC (4, 5, 6, 7, 8, and 9) as non-metropolitan. As a result, thirty-four (34) counties are classified as non-metropolitan and twenty-one (21) are classified as metropolitan.

Table 1: **USDA Rural-Urban Continuum Code Descriptions**

| Code | Description |
|---|---|
| **1** | Counties in metro areas of 1 million population or more |
| **2** | Counties in metro areas of 250,000 to 1 million population |
| **3** | Counties in metro areas of fewer than 250,000 population |
| **4** | Urban population of 20,000 or more, adjacent to a metro area |
| **5** | Urban population of 20,000 or more, not adjacent to a metro area |
| **6** | Urban population of 2,500 to 19,999, adjacent to a metro area |
| **7** | Urban population of 2,500 to 19,999, not adjacent to a metro area |
| **8** | Completely rural or less than 2,500 urban population, adjacent to a metro area |
| **9** | Completely rural or less than 2,500 urban population, not adjacent to a metro area |

Note: For the purposes of this paper, counties in Codes 1 – 3 are classified as Metropolitan and 4 – 9 as Non-metropolitan.

To render each county LFPR data series stationary for use in our first stage model (DFM), we first difference the data. All Augmented Dickey-Fuller tests support the conclusion of a unit root process and high persistence (Dickey and Fuller, 1979). Summary statistics aggregated by RUCC status in Table 2 highlight that the LFPR varies within and across the metropolitan/non-metropolitan

[2]Monthly county labor force values are divided by annual county civilian/non-institutional populations between the ages of 16-64 to calculate pseudo-monthly LFPRs for each county.

divide. Non-metropolitan counties exhibit the lowest average LFPR and the highest. Summary statistics for individual counties are presented in Table B.1. The LFPR for non-metropolitan counties appears less stable than the metropolitan counterparts given the larger reported standard deviations in both Tables 2 and B.1. Since we use differenced data in our model, we also plot the change in LFPRs aggregated by RUCC status in Figure 11. From the figure, one can see that the change in LFPRs across these two groups varies over time as well. Figure 11 also shows that non-metropolitan LFPR exhibit larger variance than metropolitan rates. For individual county LFPR change plots, see Figure B.1 of the Appendix.

Table 2: **Summary Statistics for Labor Force Participation Rates by RUCC Status**

| | Counties | Mean | SD | Minimum | Maximum |
|---|---|---|---|---|---|
| **Metropolitan** | 21 | 66.22 | 7.32 | 45.30 | 84.73 |
| **Non-Metropolitan** | 34 | 62.82 | 8.78 | 38.82 | 97.80 |
| **Total** | 55 | 64.11 | 8.78 | 38.82 | 97.80 |

Note: We use the 2013 USDA Rural-Urban Continuum Codes (RUCC) to classify each county in the metropolitan or non-metropolitan categories. Descriptions of the codes can be seen in Table 1. SD refers to standard deviation.

### 2.2.2 County Characteristics

In the second stage of our analysis, we examine the relationships between county characteristics and the proportion of variance explained by our estimated factors presented in Section 2.4.3. Through this analysis we determine through which channels state, RUCC, and county factors affect county LFPR changes. Since most county variables are unavailable at the monthly frequency, we collect annual county data to create a panel of the census years 1990, 2000, 2010, and 2020. We consider 22 explanatory variables potentially influencing a county's sensitivity to joint LFPR comovements over time. We briefly describe these variables next but provide summary statistics and additional descriptions in Tables B.2 and B.3 of the Appendix.

First, we include measures of industry composition (i.e., the share of government, agriculture, man-

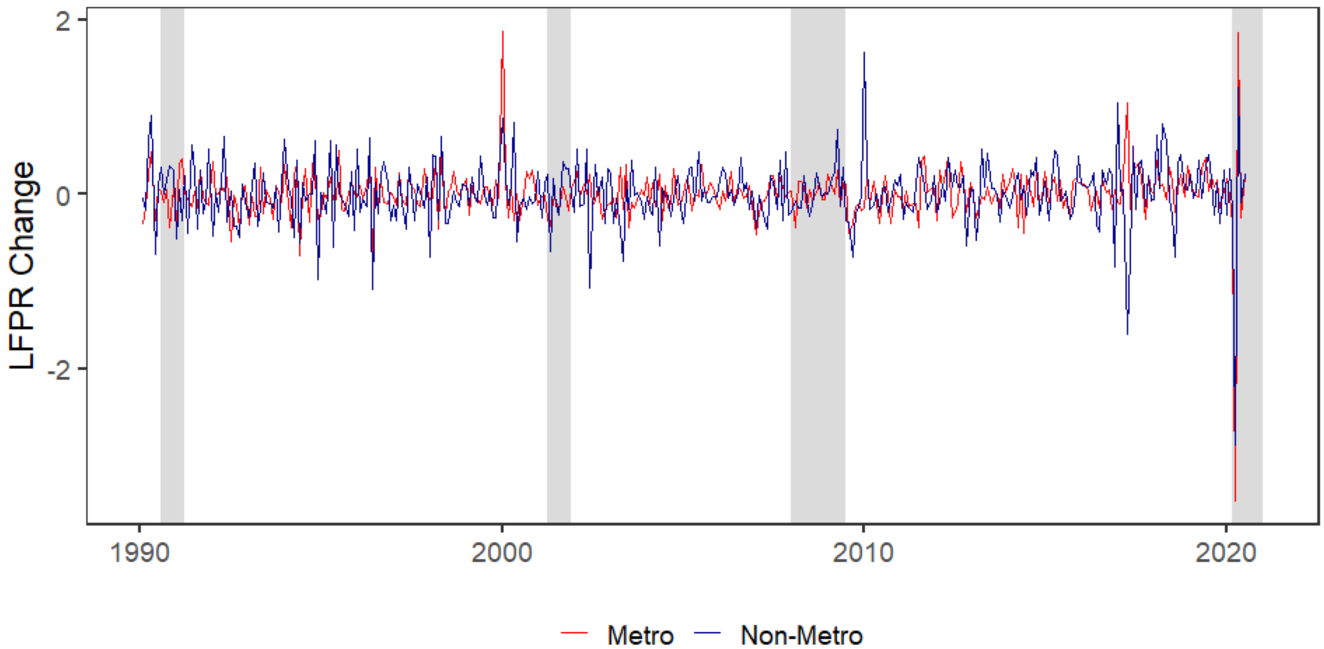


Figure 11: **Change in Labor Force Participation Rates by RUCC Status**

Note: Shaded regions are the NBER-dated recessions.

ufacturing, coal mining, and non-farm employment in a county) constructed from local area data from the Bureau of Economic Analysis (BEA). Different industry sectors may induce change in LFPRs. Isserman and Rephann (1993); Chinitz (1961) point out that shift work, such as manufacturing and mining jobs, may constrain the LFPR of partners in the household. Given the dependencies of counties, regions, or states on certain industries, the composition of industry in a county may be important in explaining the sensitivity to each factor. Beverly et al. (2022) note that shocks, specifically to major local or regional industries like the coal/gas industry in West Virginia and the Appalachian region, impact changes in LFPR.

In addition to each county's industry composition, we also use the number of natural gas wells and the volume of natural gas produced in each county in thousands of cubic feet (MCF) as controls. Since persistent unemployment and negative economic opportunity may also affect LFPRs, we include the unemployment rate for each county. Workers living in a county with low economic opportunities may become discouraged, which would induce changes in LFPRs at the county, regional, and state levels (Isserman and Rephann, 1993; Stephens and Deskins, 2018). Therefore, we assert that county labor market conditions may also explain the level of variance explained by the different factor levels we have estimated.

Studies have shown that LFPR differences persist across racial, gender, and age demographics (Stephens and Deskins, 2018; Cajner et al., 2017; Compton and Pollak, 2014). Changes in these demographics may also affect how much state, RUCC, and county factors explain the variation in the change in LFPRs through migration patterns/restrictions and concentrations of these groups throughout the state. We use data from the decennial U.S. Census to control for these differences. We calculate the shares of the population for each county that are African American and other races. We use the share of the population that is white or Caucasian as the base category. We also include the shares of each county's population that is female under 25, between 25 and 54 years of age, and between the ages of 54 - 65. We use the share of the population over 65 as the reference category.

Additionally, studies show that individuals with higher education are more likely to migrate and, therefore, more likely to participate in the labor force (Weingarden et al., 2017). Changes in a county's education levels may also explain a factor's sensitivity to influencing changes in LFPRs. To control for these differences across levels of education, we use Economic Research Service (ERS) education data from the United States Department of Agriculture (USDA) for 1990, 2000, and 2010, in conjunction with education data from U.S. Census 2020 decennial census and the American Community Survey 2020. From these sources, we calculate the fractions of the county population over 25 with a high school diploma (or equivalent), and some college experience.

We additionally control for geographical county characteristics that may be related to the variance in the change in LFPRs explained by a given factor. We include the land area in square miles for each county from the U.S. Census Bureau and the average rainfall in inches for each county for our panel years from the National Centers for Environmental Information (NCEI). We also include county highway data from the U.S. Department of Transportation's Federal Highways Administration to create an indicator variable for counties with at least one interstate highway. Geographic variables like these may positively and negatively affect the LFPR in a given county (Isserman and Rephann, 1993; Stephens and Deskins, 2018; Rickman and Wang, 2017; Marston, 1985). Infrastructure and wealth in metropolitan and non-metropolitan counties may attract workers and businesses, therefore affecting the LFPR and the levels of variation in the change in LFPRs explained by our factors.

Lastly, to capture a county's overall lifestyle and other economic outcomes, we first use the Institute for Health Metrics and Evaluation (IHME) data on life expectancy as a proxy for county health. Poor health can lead to difficulties finding and keeping jobs, leading to withdrawal from the workforce. Poor health induced by unhealthy habits increases morbidity and can also cause changes in LFPRs. Behringer and Friedell (2006) note strong health trends in Appalachian and, by extension, the state of West Virginia[3]. We suggest that these regional/cultural health trends potentially link labor markets across counties and are therefore important in determining how well our spatial factors explain the variation in county LFPRs. Additionally, we gather poverty rate data from the U.S. Census Bureau, Temporary Aid to Needy Families (TANF) data compiled by West Virginia Kids Count from the U.S. Department of Health and Human Resources, and personal income data from the BEA. We include these indicators since income from transfer payments may cause disincentives to work. Also, the number of people accepting transfer income depends on the local and state economies, therefore linking assistance spatially (Isserman and Rephann, 1993).

## 2.3 Methodology

### 2.3.1 Stage 1

In the first stage of this analysis, we use the dynamic factor model with time-varying factor loadings and stochastic volatility (DFM-TV-SV), developed by Del Negro and Otrok (2008). The DFM-TV-

[3]Every county in West Virginia is a part of the Appalachian Region.

SV model decomposes the variation in the change of each county's LFPR into three components: state, metropolitan/non-metropolitan status, and each county's idiosyncratic movements.

Formally, our measurement equation is given by

$$y_{i,t} = \lambda_{i,t}\mathscr{S}_t + \widetilde{\gamma}_{i,t}\mathscr{M}_t + \varepsilon_{i,t} \tag{9}$$

where $y_{i,t}$ is the change in labor force participation rate for West Virginia county i = 1,2,...55 at month $t$. $\mathscr{S}_t$ is the state or common factor that affects $y_{i,t}$. $\mathscr{M}_t$ is a vector that contains the two factors, $\mathscr{M}_{j,t}$, $j = 1, 2$ corresponding to the non-metropolitan and metropolitan county classifications designated by the USDA (USDA, 2013). The state loadings, $\lambda_{i,t}$ and $\widetilde{\gamma}_{i,t}$, measure the responses of each county's LFPR to changes in state and RUCC factors, respectively. Increases in the state loadings for county i, for example, means that changes in county i's LFPR respond more strongly to the state LFPR factor. Finally, $\varepsilon_{i,t}$ is the idiosyncratic or county-specific factors which captures county influences or shocks on LFPRs after the national and RUCC influences are removed.

The specification in Equation 9 extends the standard constant parameter Dynamic Factor Model (DFM) by allowing the factor loading parameters to vary over time. We also impose that the three factors are orthogonal for similar identification purposes as the (DFM). However, allowing for time-varying loading parameters enables us to capture the changes in the relative contribution of each factor to the change in LFPR variation over time. As a result, our specification in Equation 9 implies the following variance decomposition structure:

$$\mathrm{Var}(y_{i,t}) = \lambda_{i,t}^2 \ \mathrm{Var}(\mathscr{S}_t) + \widetilde{\gamma}_{i,t} \ \mathrm{Var}(\mathscr{M}_t)\widetilde{\gamma}_{i,t}' + \ \mathrm{Var}(\varepsilon_{i,t}) \tag{10}$$

To capture the dynamics in the volatility over time, we also expand the standard DFM to include stochastic volatility in the laws of motion of the national, regional, and idiosyncratic factors (Equations 11 – 15). This extension assumes random, rather than constant, innovations (error terms) of each factor.[4] In particular, we observe differential volatility across time and economic conditions. Importantly, this assumption and specification allow us to capture changes in the sensitivity of our factors to labor conditions over our sample. To this extent, we can capture potential volatility changes due to new or amended labor policy and major shocks to the local economies like the COVID-19 pandemic and natural disasters.

The transition equations for each factor evolve as stationary processes:

$$\mathcal{S}_t = \sum_{p=1}^{P} \phi_p^{\mathcal{S}} \mathcal{S}_{t-p} + e^{h_t^{\mathcal{S}}} \cdot \nu_t^{\mathcal{S}}; \quad \nu_t^{\mathcal{S}} \sim i.i.d.\,\mathcal{N}(0, \sigma_{\mathcal{S}}^2) \tag{11}$$

where $\phi_p^{\mathcal{S}}$ is the autoregressive coefficient for the state factor, $P = 2$. $e^{h_t^{\mathcal{S}}}$ represents the stochastic volatility components, and $\nu_t^{\mathcal{S}}$ the innovation to the law of motion for the state or common factor. The stochastic volatility for the state factor is assumed to follow a random walk process:

$$h_t^{\mathcal{S}} = h_{t-1}^{\mathcal{S}} + \sigma_{\mathcal{S}}^h \cdot \eta_t^{\mathcal{S}}; \quad \eta_t^{\mathcal{S}} \sim i.i.d.\mathcal{N}(0, 1) \tag{12}$$

where $\sigma_{\mathcal{S}}^h$ is the standard deviation of the innovation to the law of motion for the state factor and $\eta_t^{\mathcal{S}}$ is the volatility shock.

Likewise, the metropolitan and non-metropolitan RUCC factors are assumed to follow AR(l) processes:

$$m_{j,t} = \sum_{l=1}^{L} \phi_{j,t}^{\mathcal{M}} r_{t-l} + e^{h_{j,t}^{\mathcal{M}}} \cdot \nu_{j,t}^{\mathcal{M}}; \quad \nu_{j,t}^{\mathcal{M}} \sim i.i.d.\,\mathcal{N}(0, \sigma_{j,s}^2) \tag{13}$$

where $\phi_{j,t}^{\mathcal{M}}$ is the autoregressive coefficient for each regional factor, $L = 2$, $e^{h_{j,t}^{\mathcal{M}}}$, the stochastic volatility components, and $\nu_{j,t}^{R}$ the innovation to the law of motion for the regional factor. The stochastic volatility for the regional factors are also assumed to follow random walk processes:

$$h_{j,t}^{\mathcal{M}} = h_{j,t-1}^{\mathcal{M}} + \sigma_{j,\mathcal{M}}^h \cdot \eta_{j,t}^{\mathcal{M}}; \quad \eta_{j,t}^{\mathcal{M}} \sim i.i.d.\mathcal{N}(0, 1) \tag{14}$$

[4] Formally, the stochastic volatility model assumes that the error term's variance is normally distributed.

where $\sigma^h_{j,\mathscr{M}}$ is the standard deviation of the innovation to each law of motion of the regional factor and $\eta^{\mathscr{M}}_{j,t}$ is the volatility shock.

The idiosyncratic factor for each county similarly follows a stationary AR(q) process:

$$\varepsilon_{i,t} = \sum_{q=1}^{Q} \phi_q \varepsilon_{t-q} + e^{h^{\mathscr{S}}_{i,t}} \cdot \nu^{\mathscr{S}}_{i,t}; \quad \nu^s_{i,t} \sim i.i.d.\, \mathscr{N}(0, \sigma^2_i) \tag{15}$$

where $\phi_q$ is the autoregressive coefficient for the idiosyncratic shock, $Q = P = L = 2$, $e^{h^r_{j,t}}$, the stochastic volatility components, and $\nu^s_{i,t}$ the innovation to the law of motion for the idiosyncratic factor. For proper identification, we follow the literature and assume that $\nu^{\mathscr{S}}_t, \nu^{\mathscr{M}}_{j,t}$, and $\nu^{\mathscr{S}}_{i,t}$ are orthogonal to each other. The stochastic volatility for each county factor follow random walk processes:

$$h^{\mathscr{S}}_{i,t} = h^{\mathscr{S}}_{i,t-1} + \sigma^h_i \cdot \eta^{\mathscr{S}}_{i,t}; \quad \eta^{\mathscr{S}}_{i,t} \sim i.i.d.\ \mathscr{N}(0, 1) \tag{16}$$

where $\sigma^h_i$ is the standard deviation of the innovation to each law of motion and $\eta^{\mathscr{S}}_{i,t}$ is the volatility shock. We also assume that, $\eta^{\mathscr{S}}_t, \eta^{\mathscr{M}}_{j,t}$, and $\eta^{\mathscr{S}}_{i,t}$ are orthogonal to each other.

Lastly, we follow the macroeconomics literature for standard normalization procedures (See Del Negro and Otrok, 2008; Bhatt et al., 2017, for example). We first restrict the shocks of the national and regional factors $\sigma^2_{\mathscr{S}} = \sigma^2_{1,\mathscr{M}} = \sigma^2_{2,\mathscr{M}} = 1$. This is because the scale of the factor loadings and the standard deviations for each factor cannot otherwise be separately identified. Additionally, we constrain each $h$ in the stochastic volatility equations (12, 14, 16) to a starting value of zero since the scale of stochastic volatility term $h_\bullet$ is determined by the initial condition. (i.e $h^{\mathscr{S}}_0 = h^{\mathscr{M}}_{j,0} = h^{\mathscr{S}}_{i,0} = 0$) While this implies no stochastic volatility before the sample period, it allows us to derive equally representative distributions for the initial conditions (Del Negro and Otrok, 2008).

We estimate the above DFM-TV-SV using the Monte Carlo Markov Chain (MCMC) Bayesian estimation method. We follow Kim et al. (1999) and Kim et al. (1998) for the standard Gibbs-Sampling algorithm and procedure to draw stochastic volatility, respectively. Steps and additional information about the Gibbs-Sampling algorithm can be found in the technical appendix of Bhatt et al. (2017) as well as in Section 3 of Beverly et al. (2022)

### 2.3.2 Stage 2

After estimating Equation 10 from the first stage of our analysis, we then compute each factor's contribution to the total variability in the change in LFPR for county i. From Equation 10, the fraction of volatility due to the state factor, $\mathscr{S}$, would be:

$$\theta_{it}^{S} = \frac{\lambda_{i,t}^{2}\ \mathrm{Var}(\mathscr{S}_t)}{\mathrm{Var}(y_{i,t})} \tag{17}$$

The fractions of volatility due to the RUCC and county factors are calculated similarly and are denoted $\theta_{it}^{\mathscr{M}}$ and $\theta_{it}^{i}$, respectively.

For the second stage of our analysis, we investigate how well county characteristics explain a county's sensitivity to state, RUCC, and idiosyncratic influences. To do this, we regress the proportion of the change in LFPR variance explained by the state, RUCC, and county factors ($\theta_{it}^{S}$, $\theta_{it}^{\mathscr{M}}$, $\theta_{it}^{i}$ respectively) on the 22 explanatory variables explained above for the years 1990, 2000, 2010, and 2020[5]. From the Gibbs-Sampler in our initial estimation, we keep 40,000 draws of $\theta_{it}^{S}$, $\theta_{it}^{\mathscr{M}}$, and $\theta_{it}^{i}$. We reduce these draws to 1000 random values per month for each year of our panel. We then randomly select 5000 total for each panel year to keep for this second stage analysis[6]. We then draw one $\theta_{i}^{S}$ ($\theta_{i}^{\mathscr{M}}$, $\theta_{i}^{i}$) from each panel year and regress these four values on the 22 explanatory variables. We repeat this regression 5000 times for all the kept $\theta$'s and the same explanatory variables, which produces distributions for the $b_k$'s of our regressors[7]. A single panel regression model is given by:

[5]We recognize the philosophical inconsistency in evaluating the panel regressions in a frequentist framework while estimating the DFM-TV-SV (i.e. $\theta_{it}^{S}$, $\theta_{it}^{\mathscr{M}}$, $\theta_{it}^{i}$) using Bayesian methodology. The literature does commonly mix Bayesian and frequentist approaches (Kose et al., 2003).

[6]January is missing for 1990 due to the differenced data. 2020 only has January - July.

[7]While using frequentist methodology, we use 5000 draws/regressions as a way to keep a Bayesian theme and better integrate the two approaches philosophically. Visualization of the distribution of coefficients can be seen in Figures B.2 - B.5 of the Appendix.

$$\theta_{it}^{S}|_1^{5000} = b_{it} + \sum_{k=1}^{23} b_k X_{k,i,t} + e_{it}^{S} \tag{18}$$

where $\theta_{it}^{S}|_1^{5000}$ is one of 5000 draws of the proportion of the variance in the change in a county's LFPR explained by the state factor for county $i$ ($i = 1, ..., 55$), and for time t (1990, 2000, 2010, 2020). $X_{k,i,t}$ represents the value for characteristic $k$ ($k = 1, ..., 22$) in county $i$ at time t. Similar regressions are used to explain $\theta_{it}^{\mathscr{M}}$ and $\theta_{it}^{i}$ and are also repeated 5000 times. We estimate each iteration of Equation 18 using fixed effects with clustered standard errors to adjust for autocorrelation and heteroskedasticity. Given that $\theta_{it}^{S}$, $\theta_{it}^{\mathscr{M}}$ and $\theta_{it}^{i}$ sum to one in a single iteration by construction, the coefficients from one of the regressions will potentially be redundant. For clarity, we report the full results in Table B.4. Since the magnitudes of the coefficients are not particularly meaningful, we focus rather on any evidence for connections between the county characteristics and the dependent variables and the signs of these relationships. In Table 3, we report only the variables with the strongest relationships and present the percentage of the coefficient distribution that is above zero. In Figures 17, 18, and 19 we plot the distributions of these variables along with the means of the distributions and the Highest Posterior Density Intervals (HPDI) as credible intervals for our results.

## 2.4 Results

In this section, we plot the results from our DFM-TV-SV model described above. Figure 12, displays the time-varying loading parameters (posterior medians) of an unobserved state factor averaged across metropolitan (Panel A) and non-metropolitan (Panel B) counties together with their 90% confidence intervals. In support of our findings for the state factor loadings, we display the average cross-county correlation for all West Virginia counties (Panel A), metropolitan counties (Panel B), and non-metropolitan counties (Panel C) in Figure 13. In Figure 14, we plot the time-varying loading parameters (posterior medians) for unobserved metropolitan and non-metropolitan factors averaged across metropolitan (Panel A) and non-metropolitan (Panel B) counties together with their 90% confidence intervals. Individual county results for the state factor loadings and metropolitan/non-metropolitan factor loadings are available upon request. We present our variance decomposition results from Equation 10, averaged over the entire sample period and averaged across metropolitan/non-metropolitan counties in Figures 15 and 16, respectively. Individual time-varying results for each county are also available upon request.

### 2.4.1 State Factor Loadings

The state factor loadings reflect the sensitivity or correlation between individual West Virginia county LFPRs and the state factor, which captures a state-level unobserved effect in the movement for all West Virginia county LFPRs. We average individual county results by RUCC status, and Figure 12 shows that we find a near-zero correlation with the state factor for most periods for both metropolitan (Panel A) and non-metropolitan (Panel B) counties. The near-zero correlation over time indicates a weak relationship between changes in the state factor and changes in either metropolitan or non-metropolitan LFPRs.

However, Figure 12 also shows that after 2010, the loadings on the state factor are positive for metropolitan (Panel A) counties. These positive loadings indicate that, as of relatively recently, increases in the state factor signal increases in the change in metropolitan LFPRs. That is, changes in metropolitan LFPRs have become more influenced by a state or common component. For non-metropolitan counties, loadings on the state factor are also positive in 2019 and 2020. This increased connection to a common component renders metropolitan and non-metropolitan areas more susceptible to recessions or other labor market shocks. With larger decreases in LFPR levels, metropolitan and non-metropolitan areas would experience larger relative declines in employment and economic activity, evidenced by the steep declines in LFPR levels during the COVID-19 pandemic. The increasing trend in the average cross-county correlation, seen in Figure 13, further supports this finding and suggests that West Virginia county LFPRs are becoming synchronized.

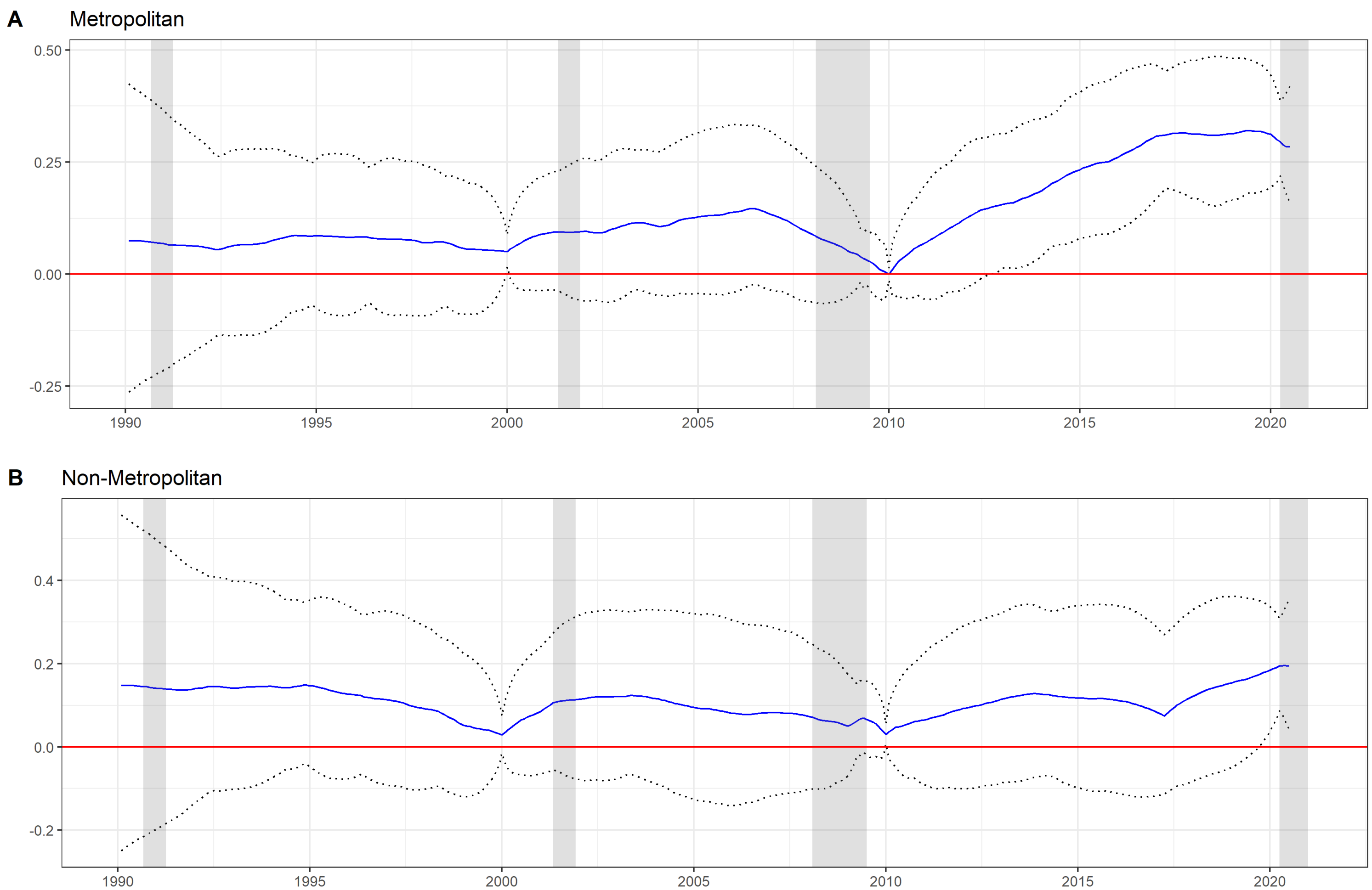


Figure 12: **Average State Factor Loadings by RUCC Status**

Note: Shaded regions are the NBER-dated recessions. The blue solid lines represent the average of the medians of the posterior distribution. Dashed lines represent the 5th and 95th percentiles. Panel (A) shows the state factor loadings averaged across the 21 metropolitan counties and Panel (B) shows the state factor loadings averaged across the 34 non-metropolitan counties.

Outside the COVID-19 pandemic window, this increasing trend coincides with increasing LFPRs in West Virginia which presents policymakers with an opportunity to take advantage of potential larger increases in LFPRs associated with the synchronized county LFPRs.

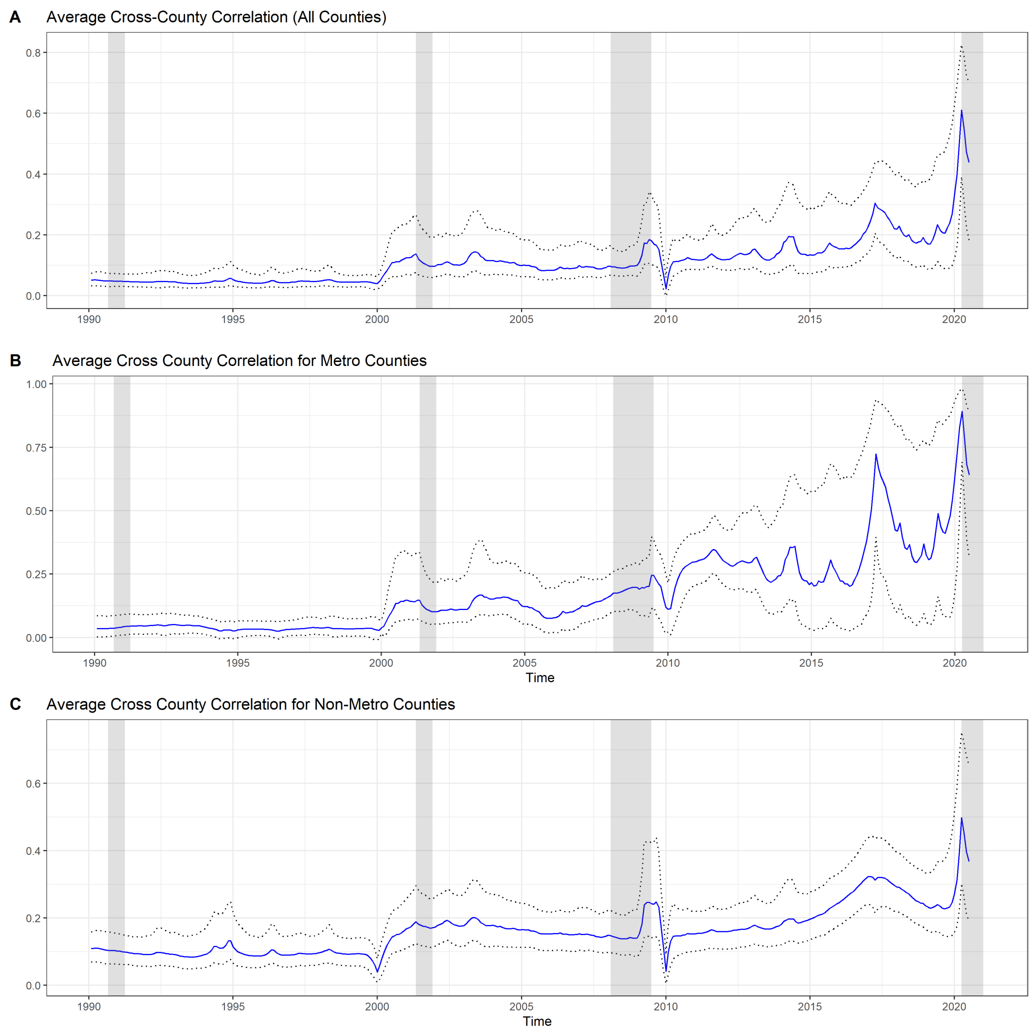


Figure 13: **Average Cross-County Correlation**

Note: Shaded regions are the NBER-dated recessions. The blue solid lines represent the medians of the posterior distribution averaged across each group: All West Virginia counties (Panel A), all non-metropolitan counties (Panel B), and all metropolitan counties (Panel C). Dashed lines represent the 5th and 95th percentiles.

### 2.4.2 Metropolitan and Non-Metropolitan Factor Loadings

Like the state factor loadings, the RUCC factor loadings reflect the sensitivity or correlation between individual West Virginia county LFPRs and their respective RUCC factor (metropolitan vs. non-metropolitan). These factors capture unobserved effects in the movement of county LFPRs by RUCC classification and apart from any state factor changes. Once again, we average these individual county sensitivities to the RUCC factors across counties in each designation. Figure 14 Panel (A) shows that metropolitan factor loadings for metropolitan counties are non-zero during and immediately following the 2008-9 recession. These positive factor loadings indicate that changes in metropolitan LFPRs were sensitive to a metropolitan component apart from any statewide component. During this period, as the metropolitan factor increases, change in metropolitan county LFPRs increase as well, which implies increased instability in metropolitan LFPRs during the 2008-09 recession. This response to the 2008-09 recession compared to the previous two recessions that occurred during the sample period most likely derives from the severity of the recession and the unusually steep decline in LFPRs in 2010 (Elsby et al., 2011; Daly et al., 2009).

In Figure 14 Panel (B), we plot the non-metropolitan factor loadings. We find that they are positive over the entire sample period. These persistent positive factor loadings indicate that increases in the non-metropolitan factor signal increases in the change in LFPRs for non-metropolitan counties over the last three decades. That is, as non-metropolitan county LFPRs become more synchronized, changes in individual non-metropolitan county LFPRs increase. This persistent correlation between non-metropolitan county LFPRs and the non-metropolitan factor demonstrates a strong relationship in the comovement of non-metropolitan counties and implies that synchronized non-metropolitan LFPRs are less stable and more susceptible to larger swings in LFPR levels. These findings align well with Hamrick (1997) who finds non-metropolitan areas to be more sensitive to business cycle, exchange rate, and unemployment rate movements compared to metropolitan areas in general.

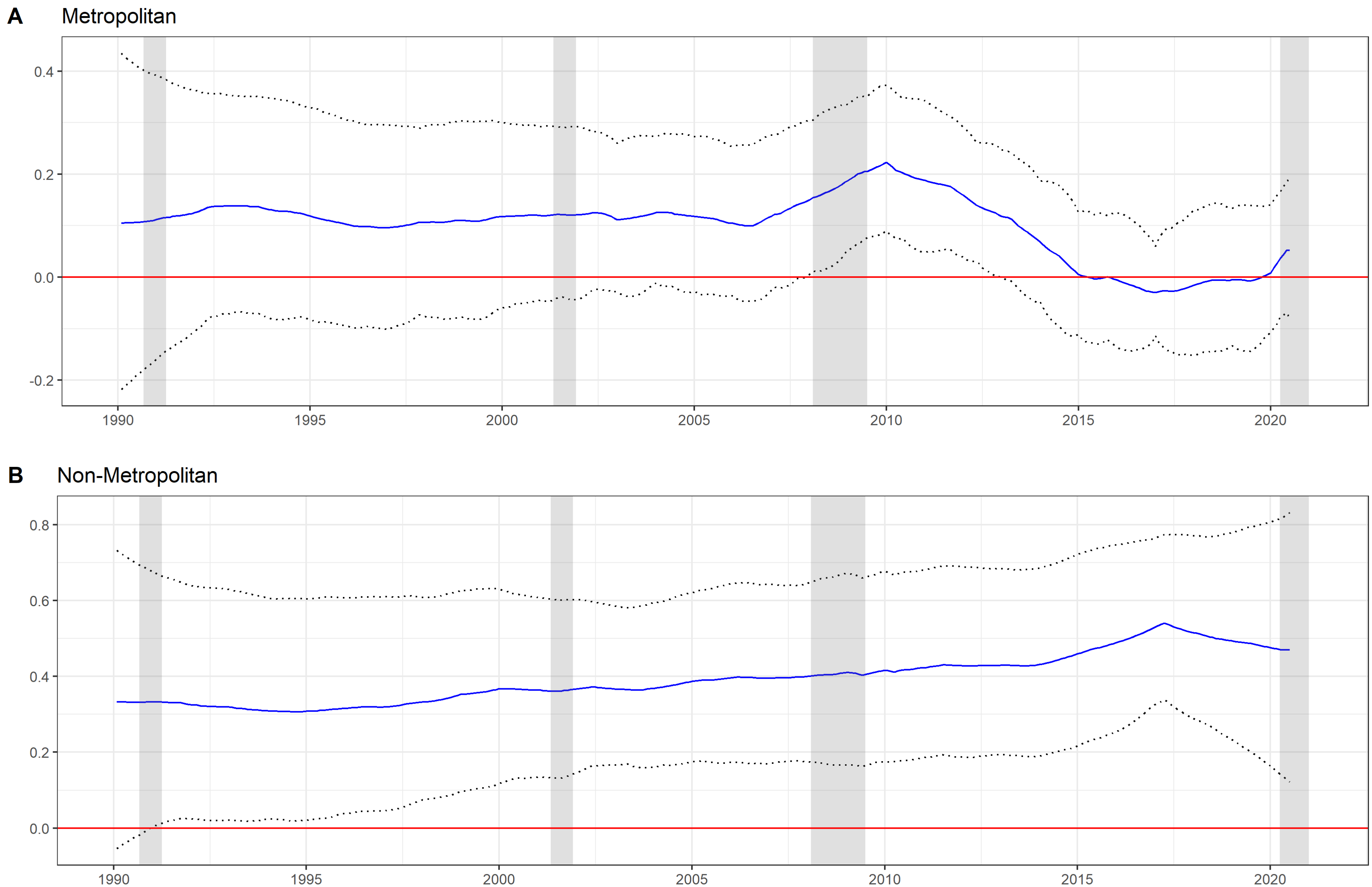


Figure 14: **Average RUCC Factor Loadings by RUCC Status**

Note: Shaded regions are the NBER-dated recessions. The blue solid lines represent the average of the medians of the posterior distribution. Dashed lines represent the 5th and 95th percentiles. Panel (A) shows the metropolitan factor loadings averaged across the 21 metropolitan counties and Panel (B) shows the non-metropolitan factor loadings averaged across the 34 non-metropolitan counties.

### 2.4.3 Variance Decompositions

Next, we turn to our variance decomposition results (from Equations 10 and 17) and assess the strength of inter-state comovements in the change of West Virginia county LFPR. These results allow us to determine the relative importance of each factor in explaining West Virginia county LFPR dynamics. Figure 15 maps the percentage contributions of the state, RUCC, and county factors to the total change in LFPR variations averaged over our sample period, (i.e. $\bar{\theta}_i^S$, $\bar{\theta}_i^{\mathscr{M}}$, $\bar{\theta}_i^i$, respectively.) Figure 15 Panel (A) shows the average percent of variance explained by the state factor for each county[8]. State influences on the change in LFPRs are relatively important for metropolitan counties and especially the counties surrounding the state capital (Kanawha County: 38%, Boone County: 37%, Putnam County: 36% and Lincoln County: 48%). While this is expected, the overall weak contribution of the common state factor (20%) demonstrates the stark heterogeneity at the county level.

The average percent of variance explained by the metropolitan and non-metropolitan factors for each county seen in Figure 15 Panel (B) supports our non-metropolitan factor loading findings. Non-metropolitan counties, especially in the eastern part of the state, show relatively large percentage contributions of the non-metropolitan factor to LFPR variations (Pendleton County: 43%, Hardy County: 33%, Webster County: 47%, and Monroe County: 42% for example). These counties are more rural and isolated, and these results demonstrate that non-metropolitan status plays an important role in explaining the change in WV county LFPR variability.

There is a stark contrast between Figure 15 Panel (A) and Panel (C), which shows the average percent of variance explained by the county idiosyncratic factor for each county. For counties outside the state capital area, the idiosyncratic state factor plays a leading role in explaining the change in WV county LFPR variability (65%). This dominant role is especially stark for northern and panhandle counties. With proximity to Pennsylvania, Maryland, and Virginia and greater distance from the state capital, we expect this result reflects extra-state influence, resulting in

[8] For reference, a West Virginia map with county names and metropolitan/non-metropolitan status can be seen in Figure B.1 of the Appendix

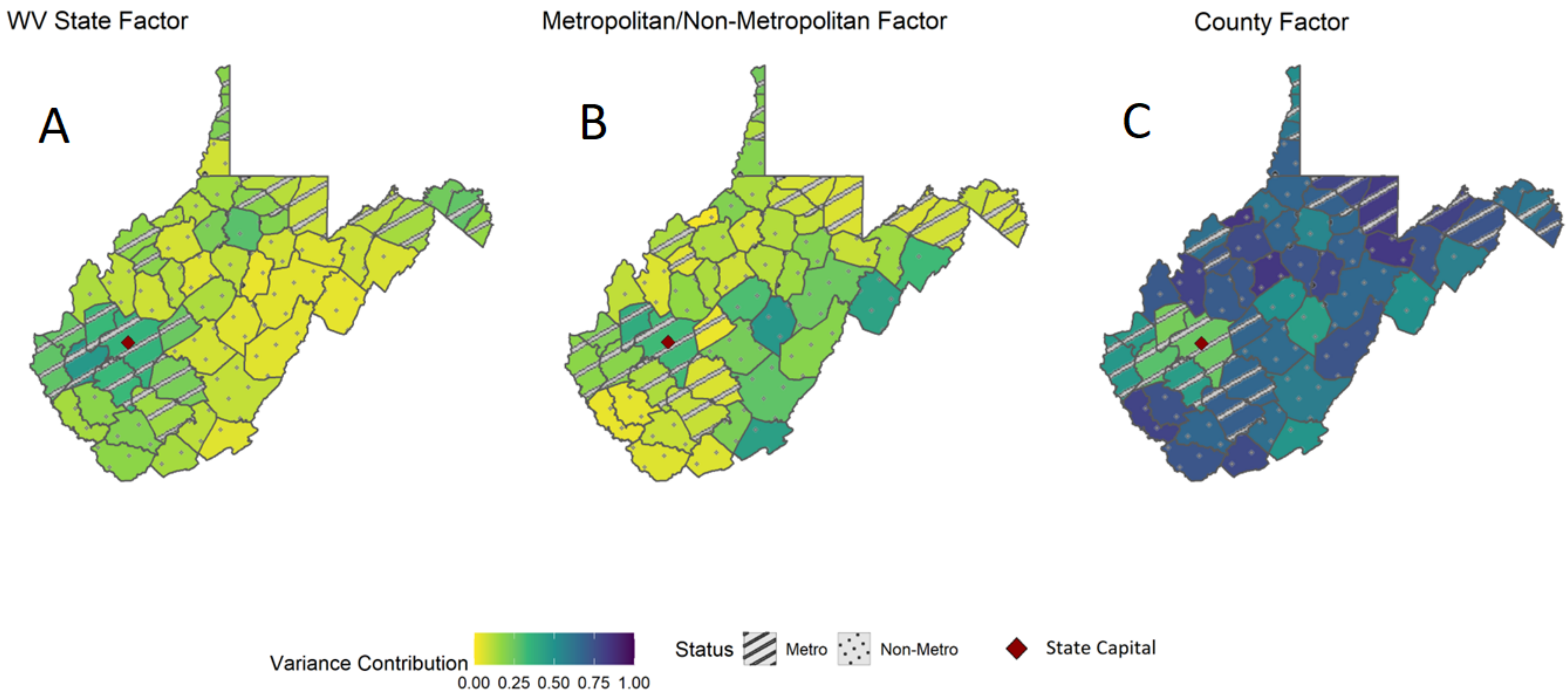


Figure 15: **Average Variance Contribution by Factor in West Virginia Over Time**

Note: Colored areas represent the percent contribution of each factor to observed variation in the change in LFPR. Percent contributions are respective medians of the posterior distribution. Panel (A) shows the percent of the variance explained by the state factor for each county averaged over the entire sample period. Panel (B) shows the percent of the variance explained by the metropolitan and non-metropolitan factors for respective counties averaged over the entire sample period. Panel (C) shows the percent of the variance explained by the county-specific factor for each county averaged over the entire sample period.

idiosyncratic behavior compared to other West Virginia counties. This, again, demonstrates the lack of an overall West Virginia state influence across counties.

We additionally present the average variance contribution of each factor to change in LFPR variations over time in Figure 16. Importantly, we want to note that these results reveal the length of shocks to the change in county LFPR variations. We find that state-wide shocks, such as national recessionary periods, at most last only a few years before other influences take over. While the impacts of these shocks may induce longer lasting effects, the shocks themselves provide a short window for actionable policy. Taylor (2016) suggests more permanent and predictable institutional reforms may increase long-run growth rates and boost growth in the short run. In light of this, we suggest targeted long term strategies may help economically distressed areas like West Virginia break the persistent cycles of economic disparity.

Figure 16 also supports our previous results regarding the contribution of the state and county factors to West Virginia LFPR variations. Figure 16 Panel (A) shows that the state factor contributes larger percentages to LFPR variations in metropolitan counties (27% compared to 16% in non-metropolitan counties) but also shows that the influence of the state factor is growing over time. In contrast, Figure 16 Panel (B) shows that the dominant influence of the county factor for non-metropolitan counties (excluding a few state-wide shocks) remains rather stable over time and explains about 65% of the variation overall. Figure 16 Panel (B) also displays that the contributions of the non-metropolitan factor rose after 2000 and have remained relatively stable in the last two decades. About 18% of the total variation in non-metropolitan LFPRs is explained by the non-metropolitan factor. Overall, these results again demonstrate that while individual county LFPRs behave mostly independently from others, in recent years the metropolitan county LFPRs are influenced more by state-wide trends and the non-metropolitan counties are influenced more by non-metropolitan trends.

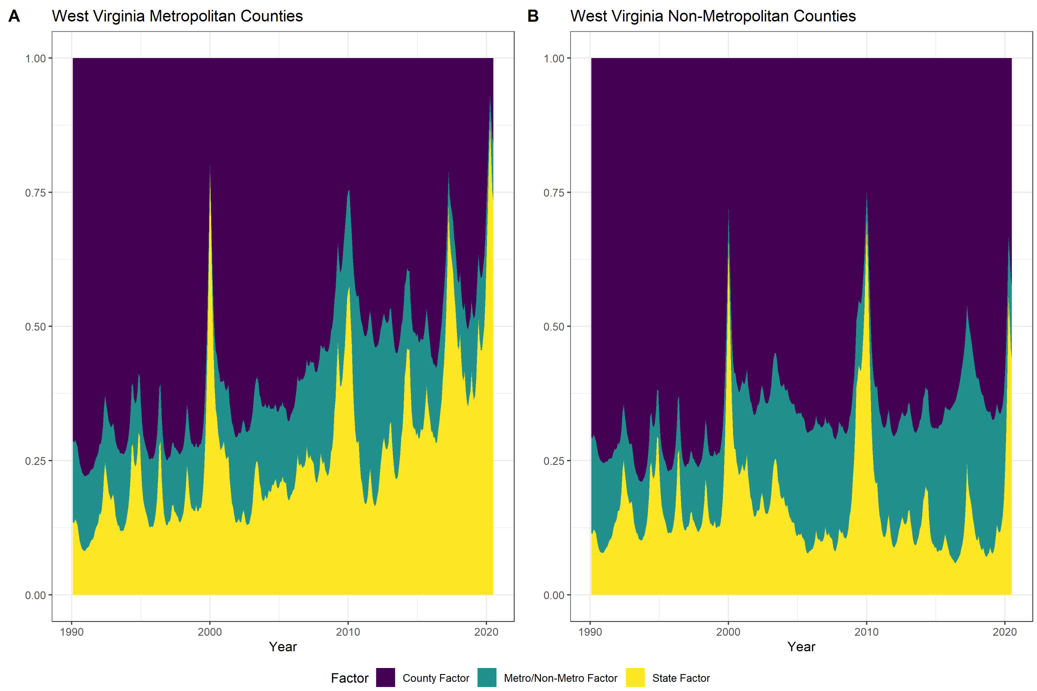


Figure 16: **Average Variance Contribution by Factor in West Virginia**

Note: Colored areas represent the percent contribution of each factor to observed variation in the change in LFPR. Percent contributions are respective means of the posterior distribution. Panel (A) shows the percent contribution of each factor overtime average across all metropolitan counties. Panel (B) shows the percent contribution of each factor overtime average across all non-metropolitan counties.

### 2.4.4 Stage 2 results: Do County Characteristics Matter?

Next, we present the regression results from the second stage of our analysis regarding how well county characteristics explain a county's sensitivity to state, RUCC, and idiosyncratic influences. After performing the panel regressions 5000 times for each factor, we save the 5000 coefficients and corresponding p-values for each regressor. We focus on the sign of the majority of the coefficients for significant variables. As such, we report the percentage of the distributions of significant variables that is above or below zero in Table 3. These correspond to the respective distributions in Figures 17, 18, and 19. In these figures, the means are plotted along with the upper and lower bounds of the 95% Highest Posterior Density Interval (HPDI). We report the results of the 13 variables where over 800 of the 5000 p-values saved are statistically significant at the 5% level[9]. Full results that include the magnitudes of the mean of each coefficient distribution can be found in Table B.4 of the Appendix.

Table 3: **Percentage of the Panel Regression Coefficient Distributions Above Zero**

| Variable | State | Metro | Non-Metro | County |
|---|---|---|---|---|
| % Government Jobs | | 4.84% | | |
| % Agriculture Jobs | | | 2.06% | |
| % Manufacturing Jobs | | | | 96.98% |
| Unemployment Rt | 7.14% | | | |
| Life Expectancy | | | | 8.68% |
| % Female < 25 | 8.88% | | | 90.16% |
| Personal Income | 98.92% | | 3.44% | |
| % with HS Diploma | | | | 65.96% |
| % with Some College | 8.66% | 98.32% | | |
| % Receiving TANF | 32.42% | | | |
| Gas Production | 89.18% | | 5.62% | 18.02% |
| Number of Gas Wells | 14.48% | | | 97.02% |
| Interstate | 27.16% | 93.26% | | 36.36% |
| Draws | 5000 | 5000 | 5000 | 5000 |
| Average $R^2$ | 0.38 | 0.37 | 0.25 | 0.4 |

Note: This table reports the percentage of the distribution of coefficients from the 5000 regressions that are above zero. Only the regressors (Column 1) with p-values that reach a significance threshold are reported. The full results and magnitudes of the coefficients can be found in Table B.4 of the Appendix. We report variables in this table if over 800 of the 5000 p-values saved are statistically significant at the %5 level.

[9]This threshold is used after adjusting for false discovery rates. Results for variables that have fewer than 800 significant p-values can still be found in Table B.4 of the Appendix.

Our results suggest a strong link between several county characteristics and the proportion of the variance in the change in LFPRs explained by the state (common) factor. First, there are positive links between the importance of the state factor with personal income and gas production. As personal income and gas production in counties increase, the state factor explains more of the variation in the change in LFPRs. These results make sense for West Virginia. First, the natural gas industry has become West Virginia's primary industry, replacing the coal industry in recent years. West Virginia ranks 5th in U.S. energy production and the West Virginian gas reserves in 2019 contributed 10% of the U.S. total. As the state's main industry, it makes sense with greater natural gas production, state labor markets are more integrated. However, this highlights a vulnerability for West Virginia. More integrated labor markets around a single or small number of industries makes West Virginia more susceptible to labor market shocks through sources such as natural resource and other environmental regulation.

Like gas production, more integrated labor markets at the state level are associated with increases in personal income as well. Figure 17 shows that zero lies outside of the 95% HPDI indicating we can be 95% re that increases in personal income positively impact the influence of the state factor in LFPR changes over time. Isserman and Rephann (1993) point out that with higher wages, LFPRs tend to increase. However, with higher non-wage income, LFPRs tend to decrease. Given the lower LFPR levels in the state, we suggest the latter case explains the association with state-wide integration of labor markets and conditions. In 2010, government transfer payments made-up nearly 25% of the West Virginia's personal income, which is more than any other state. This large contribution derives from a larger elderly and unhealthy population in the state. This link manifested especially at the state level demonstrates important implications for the future of the labor market and overall economy in West Virginia. As non-wage personal income increases with larger shares of elderly or disabled population, labor force participation rates, employment growth, and economic activity will continue to decrease in West Virginia.

Secondly, Table 3 and Figure 17 show negative links exist between the importance of the state factor in explaining change in LFPR variations and the unemployment rate, share of the population that

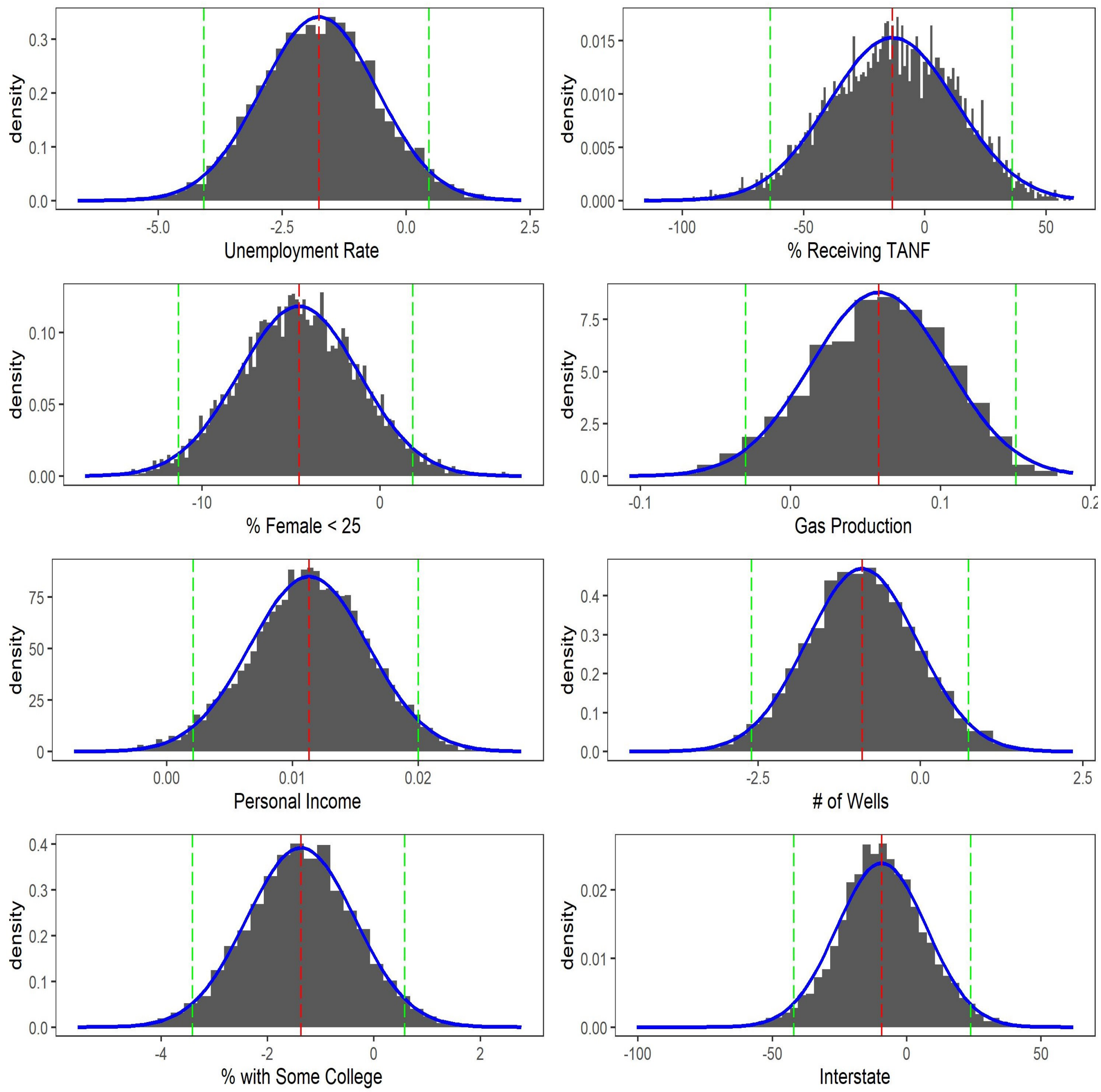


Figure 17: **Distributions of Regression Coefficients - State Factor**

Note:The variables reported display the strongest connection with the state factor as over 800 of p-values saved from the 5000 regressions are significant at the 5% level.The Red lines represent the mean of each distribution and blue lines are normal distributions with the same mean and standard deviation as the data mapped over each distribution. Green lines represent the lower and upper bounds of the 95% Highest Posterior Density Interval (HPDI) is female below the age of 25, share of population 25/older that has some college, the number of gas wells, share of families with children that are receiving TANF, and the existence of an interstate. Increases in each of the variables are associated with less importance of the state factor or less integration at the state-level. In theory, increases in the unemployment rate, female population, and transfer income are all associated with potentially varying labor force decisions (Isserman and Rephann, 1993; Phimister et al., 2002; Fernández, 2013). For example, high unemployment rates can illicit higher LFPRs given the economic turmoil but can illicit lower LFPRs as well if many workers cannot find jobs and end up dropping out the labor force. Individual responses depend on households, counties, and economic status and so divergence of these responses explain the less state-wide influence. Higher education is negatively related to the proportion of the variance explained by the state factor because higher education is usually associated with urban areas and higher LFPRs (Bowen and Finegan, 1966; Isserman and Rephann, 1993; Stephens and Deskins, 2018). We find this to be the case, as the link between higher education and the metropolitan factors is strong. Figure 18 shows that zero lies outside the 95% HPDI supporting a strong non-zero and non-negative relationship. Since rural/non-metropolitan areas typically have less education and lower LFPRs, the proportion of the variance explained by the state factor would naturally decrease as higher education increased.

In contrast to natural gas production, the number of gas wells decreases the importance of the state factor in explaining variation in the change in LFPRs. While these two variables are related, these results highlight the different relationships these variables have on the state. Production, processing, storage, and transportation of natural gas support more jobs, counties and other industries which most likely accounts for its association with the state-wide factor. Increasing the number of wells in a given county would not link LFPRs across West Virginia and may only benefit the individual counties. Accordingly, we find that the link between the number of gas wells and the county factor is stronger and positive. Figure 19 shows that zero is just about outside of the 95% HDPI, indicating strong certainty of negative relationship. For the state factor, we lastly find a strong link to the interstate system. We find that having at least one interstate decreases the importance of the state factor (and county factor) and increases the importance of the metropolitan factor. Interstates systems typically connect infrastructure or urban areas which also tend to have higher LFPRs. Since counties with interstate highways are physically linked, it not surprising that for these counties, the metropolitan factor captures more of the variation in the change in LFPRs.

Lastly, we find several other interesting links to the metropolitan, non-metropolitan and county factors that we have yet to mention. First, we find a strong link between each of these factors and a different industry. The non-metropolitan factor is negatively related to the share of industry in agriculture and natural gas production. Figure 18 shows that zero lies outside the 95% HPDI for agriculture and close to outside the bands for natural gas production. As the share of industries in these areas increase, labor markets become less integrated. In rural areas and for the agricultural sector, it is easier to engage in underground activities or informal work (Isserman and Rephann, 1993). As workers can secure food, heating materials, and income outside the formal economy the influence of a common non-metropolitan influence naturally decreases. We suggest that these many different options emerge as divergence in LFPRs across non-metropolitan counties and explain the negative relationship with proportion of LFPR variations explained by the non-metropolitan factor.

Next, we find that the metropolitan factor is negatively associated to the share of industry in a county that is government related. This implies that as the share of employment in government jobs increases in metropolitan counties, the metropolitan factor becomes less important in explaining change in LFPR variations. In congruence with Stephens and Deskins (2018) we suggest that the stability of government jobs, even through periods of economic turmoil, explain why the metropolitan factor explains less variation in the change in LFPRs. Additionally, we find a strong positive link between the county factor and the share of industry in manufacturing as seen in Table 3 and Figure 19. Manufacturing jobs are often highly unionized which may result in less job access and employment opportunity (Isserman and Rephann, 1993). Once again, decisions in response to job loss or economic turmoil vary by household, county, and economic status which most likely drives the significant idiosyncratic result. Lastly, we find links between the county factor and the life expectancy and share of the population over 25 that has a high school diploma. Like our other

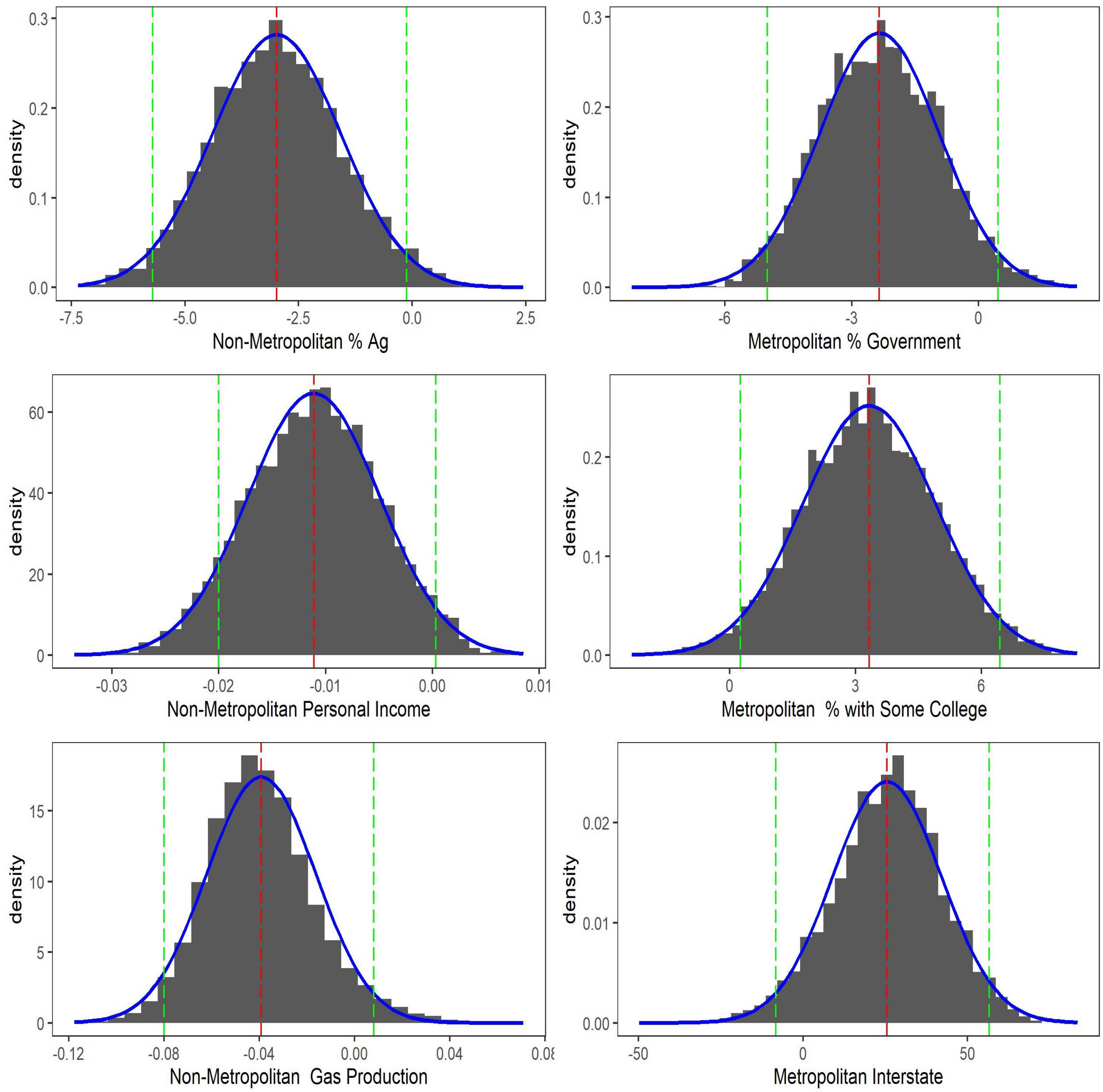


Figure 18: **Distributions of Regression Coefficients - Metro and Non-metro Factors**

Note:The variables reported display the strongest connection with the metropolitan and non-metropolitan factors as over 800 of p-values saved from the 5000 regressions are significant at the 5% level.The Red lines represent the mean of each distribution and blue lines are normal distributions with the same mean and standard deviation as the data mapped over each distribution. Green lines represent the lower and upper bounds of the 95% Highest Posterior Density Interval (HPDI)

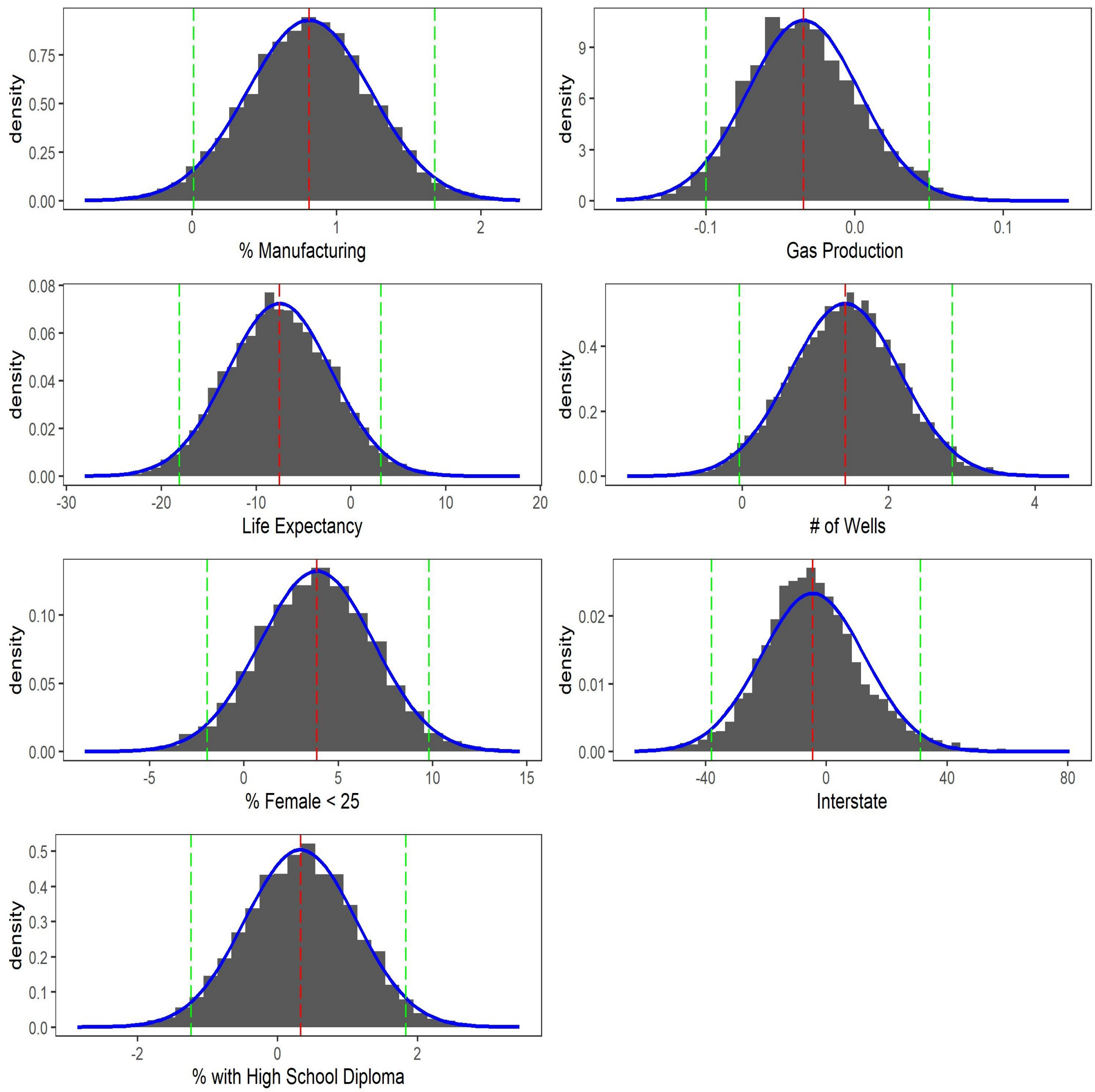


Figure 19: **Distributions of Regression Coefficients - County Factor**

Note:The variables reported display the strongest connection with the county factor as over 800 of p-values saved from the 5000 regressions are significant at the 5% level.The Red lines represent the mean of each distribution and blue lines are normal distributions with the same mean and standard deviation as the data mapped over each distribution. Green lines represent the lower and upper bounds of the 95% Highest Posterior Density Interval (HPDI) idiosyncratic results, neither of these variables illicit common behavior in LFP decisions across the state.

## 2.5 Conclusion

In this paper, we examine the role and relative importance of state, county, and metropolitan and non-metropolitan levels on the change in West Virginia county LFPRs. We focus on West Virginia LFPR dynamics given the historically poor performance in economic indicators in the state and the relatively little attention West Virginia has received from economists. We decompose West Virginia county LFPRs into state, county, and metropolitan and non-metropolitan latent factors or measures of LFPR comovements that are estimated using a Dynamic Factor Model (DFM) with time-varying (TV) and stochastic volatility (SV) parameters. We apply this model to West Virginia county data spanning Jan 1990-July 2020. Secondly, we examine the relationships between county characteristics and the proportion of variance explained by our estimated factors. Through this part of the analysis, we determine which county characteristics best explain a county's sensitivity to state and metropolitan/non-metropolitan influences on LFPRs.

Summarizing our findings, we first find the correlation between metropolitan and non-metropolitan county LFPR series and the state factor to be near-zero for most periods in the sample. Near the end of the sample, the metropolitan counties followed by the non-metropolitan counties show positive correlations with the state factor indicating recent influence on the county LFPRs by a common state trend and increased synchronization of county LFPRs across the state. We also find positive correlation between the non-metropolitan factor and non-metropolitan counties for most periods in the sample which indicates a strong and persistent non-metropolitan trend. While the correlation between metropolitan counties and the metropolitan factor are near-zero for most of the sample period, we find positive correlation following the 2008-09 recession indicating a unique metropolitan influence induced by the severity of the Great Recession.

Secondly, this study determines the relative importance of the factors by measuring the contribution of each factor to county LFPR variations. We find that the county factor dominates in contributing to county LFPR variations indicating significant idiosyncratic behavior across the state. Additionally, we find relatively large contribution of the non-metropolitan factor indicating noticeable rural or non-metropolitan behavior as well. Regressing these variance contributions on county characteristics reveals main drivers for the importance of each factor. County demographics, education levels, income, access to interstate highways, and industry composition are strongly related to the importance of our estimated state, county, and metropolitan and non-metropolitan factors. Larger positive increases in LFPRs are associated with increases in education, infrastructure, and industry.

### 2.5.1 Conclusions and Policy Implications

The results of our analysis have important implications for policy makers in West Virginia and other rural or struggling economic areas. First, considering differences in metropolitan and non-metropolitan areas should be a primary concern for West Virginia policymakers to see employment and LFP growth in the state. Given the strong idiosyncratic behavior and correlation across non-metropolitan counties that has persisted for decades, policies should be tailored for metropolitan, non-metropolitan, and even individual counties. Non-metropolitan counties, especially, are more susceptible to larger LFPR changes and as they are more sensitive to economic shocks such as recessionary periods and unemployment rate movements. For non-metropolitan counties, our results suggest that West Virginia policymakers should focus on strategies such as formalizing informal/nonstandard work, and increasing wage income, employment opportunities, and work force development programs. Ideas for bringing informal sectors into formal employment may include creating flexible work schedules, developing occupational skills training, encouraging self-employment reporting, and tax credits to encourage conversion of non-standard workers to standard workers within businesses (O'Leary and Boettner, 2015). Several work force development programs already exist in the state and include Jobs & Hope West Virginia (services to overcome substance use disorders (SUDs) and barriers to employment), Blue Ribbon Task Force (Aligns state community and technical colleges with the state workforce development system), Empower WV (connects high school seniors with workforce pathways and the state workforce development system) and Workforce Innovation and Opportunity (serves low-income and low-skilled workers). Other ideas to support workforce development in the state include prioritizing customized job training, expanding current programs, and implementing programs that have been successful in other states (O'Leary and Boettner, 2015).

Additionally, since changes in West Virginia county LFPRs have become relatively more synchronized in recent years, policymakers may be able to take advantage of this cross-correlation and more effectively apply state-wide policy. Our results indicate that increases in education, infrastructure, jobs, wage income, and industry are associated with larger positive increases in LFPRs. Our results specifically highlight the importance of diversifying key industries in West Virginia. In 2019, West Virginia Governor, Jim Justice, established the Downstream Jobs Task Force to build up the manufacturing industry in the state by developing infrastructure in preparation for "The Appalachian Energy and Petrochemical Renaissance" (or so-called revival in energy and petrochemical industries). Other policy and programs like this are needed to build on the State's unique assets. Capitalizing on industries where West Virginia outperforms its neighboring states provides a unique opportunity to increase jobs, infrastructure, wages, and attract potential businesses.

### 2.5.2 Avenues for Future Work

Finally, additional research on West Virginia labor and other distressed areas in the U.S. and around the world is needed to help reverse the historic cycles of economic despair that plague these places. Our results illicit questions regarding the relationship between the unemployment rate and LFPRs in West Virginia and the surrounding Appalachian Region. Strong relationships between these economic indicators relate to the discouraged worker phenomenon or how individuals end up dropping out of the labor force. Given historically low unemployment rates in conjunction with the already low LFPRs in these areas, this branch of research is pertinent to economic development in West Virginia. A better understanding of the relationship between these indicators and whether the discouraged worker phenomenon takes place in West Virginia is needed to better implement solutions and see economic prosperity realized in state.

## 2.6 Appendix B

Table B.1: **State LFP Descriptive Statistics**

| Counties | Mean | Median | Min | Max | S.D. |
|---|---|---|---|---|---|
| Barbour County | 63.03 | 62.98 | 57.30 | 70.87 | 2.47 |
| Berkeley County* | 71.90 | 72.00 | 66.00 | 77.25 | 2.47 |
| Boone County* | 53.29 | 54.67 | 45.30 | 58.17 | 3.36 |
| Braxton County | 59.72 | 60.22 | 54.13 | 63.84 | 2.36 |
| Brooke County* | 68.92 | 68.86 | 64.22 | 75.48 | 2.11 |
| Cabell County* | 67.27 | 67.44 | 64.13 | 71.38 | 1.80 |
| Calhoun County | 57.48 | 57.74 | 48.23 | 65.88 | 3.67 |
| Clay County* | 56.16 | 55.97 | 45.50 | 63.89 | 4.25 |
| Doddridge County | 61.79 | 62.80 | 50.52 | 72.16 | 5.90 |
| Fayette County* | 58.13 | 57.99 | 54.23 | 63.71 | 1.55 |
| Gilmer County | 51.02 | 53.11 | 40.97 | 63.32 | 5.96 |
| Grant County | 73.81 | 72.43 | 54.47 | 89.43 | 8.45 |
| Greenbrier County | 68.74 | 67.81 | 64.34 | 77.26 | 2.78 |
| Hampshire County* | 66.24 | 65.88 | 58.55 | 75.33 | 3.69 |
| Hancock County* | 70.78 | 70.63 | 65.71 | 76.17 | 2.29 |
| Hardy County | 75.67 | 77.41 | 59.73 | 93.53 | 9.15 |
| Harrison County | 72.29 | 71.11 | 67.47 | 82.56 | 3.89 |
| Jackson County | 68.57 | 66.57 | 61.43 | 97.80 | 5.97 |
| Jefferson County* | 74.69 | 74.03 | 66.39 | 82.34 | 3.23 |
| Kanawha County* | 73.30 | 72.58 | 69.88 | 80.08 | 2.13 |
| Lewis County | 66.80 | 66.91 | 60.66 | 73.45 | 2.53 |

Continued on next page

**Table B.1 – continued from previous page**

| States | Mean | Median | Min | Max | S.D. |
|---|---|---|---|---|---|
| Lincoln County* | 52.82 | 53.85 | 45.71 | 59.09 | 3.14 |
| Logan County | 52.35 | 52.07 | 49.19 | 57.37 | 1.36 |
| Marion County | 68.54 | 69.36 | 62.25 | 72.92 | 2.91 |
| Marshall County* | 68.52 | 67.86 | 65.59 | 76.60 | 2.34 |
| Mason County | 59.22 | 59.51 | 50.28 | 64.74 | 3.00 |
| McDowell County | 42.35 | 41.47 | 38.82 | 51.13 | 2.59 |
| Mercer County | 62.42 | 61.63 | 56.70 | 69.63 | 3.91 |
| Mineral County* | 69.80 | 68.94 | 63.78 | 76.75 | 3.73 |
| Mingo County | 46.59 | 46.76 | 41.42 | 51.78 | 2.46 |
| Monongalia County* | 66.16 | 65.89 | 61.40 | 72.55 | 2.00 |
| Monroe County | 66.05 | 67.16 | 53.58 | 78.64 | 6.28 |
| Morgan County | 68.01 | 67.72 | 59.48 | 75.71 | 3.85 |
| Nicholas County | 61.45 | 61.42 | 57.53 | 65.68 | 1.58 |
| Ohio County* | 75.50 | 75.89 | 69.46 | 84.73 | 3.52 |
| Pendleton County | 77.15 | 75.02 | 69.42 | 94.68 | 6.22 |
| Pleasants County | 62.08 | 61.58 | 53.67 | 76.46 | 4.11 |
| Pocahontas County | 68.97 | 68.56 | 60.59 | 82.33 | 4.96 |
| Preston County* | 66.93 | 66.69 | 62.78 | 75.42 | 2.38 |
| Putnam County* | 73.40 | 73.55 | 69.52 | 79.35 | 1.99 |
| Raleigh County* | 62.61 | 62.21 | 58.84 | 67.92 | 1.89 |
| Randolph County | 66.86 | 67.04 | 61.52 | 72.47 | 2.55 |
| Ritchie County | 67.87 | 66.83 | 58.85 | 76.80 | 3.97 |
| Roane County | 58.94 | 58.72 | 52.90 | 66.13 | 3.21 |

Continued on next page

**Table B.1 – continued from previous page**

| States | Mean | Median | Min | Max | S.D. |
|---|---|---|---|---|---|
| Summers County | 51.91 | 51.28 | 48.00 | 57.83 | 2.18 |
| Taylor County | 66.97 | 65.99 | 61.65 | 76.41 | 3.63 |
| Tucker County | 71.99 | 71.94 | 60.69 | 81.31 | 5.09 |
| Tyler County | 62.15 | 61.94 | 56.65 | 68.99 | 2.91 |
| Upshur County | 64.98 | 65.08 | 58.81 | 69.74 | 2.45 |
| Wayne County* | 61.22 | 61.14 | 58.39 | 65.52 | 1.58 |
| Webster County | 55.28 | 53.27 | 44.61 | 69.46 | 7.27 |
| Wetzel County | 65.14 | 63.44 | 57.32 | 76.42 | 5.36 |
| Wirt County* | 60.88 | 61.47 | 47.24 | 77.00 | 5.05 |
| Wood County* | 72.01 | 72.13 | 67.29 | 76.68 | 2.53 |
| Wyoming County | 49.57 | 49.98 | 41.92 | 56.49 | 3.81 |

Note: Statistics reflect the county-level labor force participation rates over the sample period January 1990 – July 2020. S.D refers to the standard deviation. * denotes Metro Counties

Table B.2: **County Characteristic Descriptive Statistics**

| Variable | Mean | SD | Min | Max |
|---|---|---|---|---|
| **Metropolitan Counties** | | | | |
| % Government Jobs | 18.82 | 5.63 | 5.63 | 32.60 |
| % Agriculture Jobs | 3.64 | 4.47 | 0.20 | 20.20 |
| % Manufacturing Jobs | 8.41 | 8.03 | 0.00 | 47.20 |
| % Mining Jobs | 3.57 | 7.21 | 0.00 | 40.90 |
| % Non-Farm Jobs | 17.51 | 5.41 | 5.41 | 31.70 |

Continued on next page

**Table B.2 – continued from previous page**

| Variable | Mean | SD | Min | Max |
|---|---|---|---|---|
| Personal Income | 1,636,043.14 | 1,680,666.69 | 54,023.00 | 9,335,977.00 |
| Unemployment Rt | 8.00 | 2.74 | 2.74 | 17.80 |
| Life Expectancy | 75.47 | 1.56 | 1.56 | 79.15 |
| % African American | 2.83 | 2.54 | 0.01 | 8.57 |
| % Female < 25 | 15.41 | 1.81 | 1.81 | 20.65 |
| % Female 25 - 54 | 19.96 | 1.64 | 1.64 | 23.41 |
| % Female 54 - 65 | 34.67 | 25.03 | 6.31 | 64.87 |
| % Another Race | 1.87 | 1.44 | 0.19 | 8.47 |
| % with HS Diploma | 39.57 | 5.55 | 5.55 | 52.70 |
| % with Some College | 20.86 | 5.13 | 5.13 | 33.63 |
| Land Area | 394.35 | 206.66 | 82.61 | 903.17 |
| % Receiving TANF | 0.06 | 0.07 | 0.01 | 0.32 |
| Precipitation | 44.96 | 6.14 | 6.14 | 57.84 |
| Gas Production | 6,963,876.61 | 21,278,811.50 | 0.00 | 131,137,681.00 |
| Number of Gas Wells | 532.48 | 691.02 | 0.00 | 3,295.00 |
| Coal Production | 3,294,927.75 | 6,128,398.42 | 0.00 | 32,446,186.00 |
| **Non-Metropolitan Counties** | | | | |
| % Government Jobs | 18.74 | 4.96 | 4.96 | 35.70 |
| % Agriculture Jobs | 7.10 | 5.15 | 0.00 | 23.70 |
| % Manufacturing Jobs | 8.86 | 7.86 | 0.00 | 44.80 |
| % Mining Jobs | 4.98 | 7.12 | 0.00 | 26.20 |
| %Non-Farm Jobs | 18.55 | 5.84 | 5.84 | 39.90 |
| Personal Income | 515765.63 | 498031.13 | 76819.00 | 3485501.00 |

Continued on next page

Table B.2 – continued from previous page

| Variable | Mean | SD | Min | Max |
|---|---|---|---|---|
| Unemployment | 9.36 | 2.75 | 2.75 | 19.40 |
| Life Expectancy | 75.24 | 1.86 | 1.86 | 79.34 |
| % African American | 1.68 | 2.48 | 0.01 | 13.48 |
| % Female < 25 | 14.52 | 1.82 | 1.82 | 19.82 |
| % Female 25 - 54 | 19.53 | 1.97 | 1.97 | 26.25 |
| % Female 54 - 65 | 35.18 | 24.76 | 6.11 | 66.14 |
| % Another Race | 1.44 | 1.03 | 0.24 | 7.48 |
| % with HS Diploma | 43.13 | 4.98 | 4.98 | 51.70 |
| % with Some College | 17.82 | 4.43 | 4.43 | 29.71 |
| Land Area | 464.08 | 205.88 | 130.10 | 1039.80 |
| % Receiving TANF | 0.07 | 0.11 | 0.01 | 0.80 |
| Precipitation | 47.34 | 7.41 | 7.41 | 66.78 |
| Gas Production | 19,261,843.14 | 75,540,908.33 | 0.00 | 5,97,800,583.00 |
| Number of Gas Wells | 932.43 | 1,077.23 | 0.00 | 5146.00 |
| Coal Production | 2,013,393.13 | 4,142,730.23 | 0.00 | 21,840,363.00 |
| **Total** | | | | |
| % Government Jobs | 18.77 | 5.23 | 5.23 | 35.70 |
| % Agriculture Jobs | 5.78 | 5.18 | 0.00 | 23.70 |
| % Manufacturing Jobs | 8.69 | 7.93 | 0.00 | 47.20 |
| % Mining Jobs | 4.44 | 7.19 | 0.00 | 40.90 |
| % Non-Farm Jobs | 18.15 | 5.70 | 5.70 | 39.90 |
| Personal Income | 943,507.95 | 123,6145.80 | 54,023.00 | 9,335,977.00 |
| Unemployment Rt | 8.84 | 2.82 | 2.82 | 19.40 |

Continued on next page

**Table B.2 – continued from previous page**

| Variable | Mean | SD | Min | Max |
|---|---|---|---|---|
| Life Expectancy | 75.33 | 1.75 | 1.75 | 79.34 |
| % African American | 2.12 | 2.56 | 0.01 | 13.48 |
| % Female < 25 | 14.86 | 1.87 | 1.87 | 20.65 |
| % Female 25 - 54 | 19.69 | 1.87 | 1.87 | 26.25 |
| % Female 56 - 65 | 34.99 | 24.87 | 6.11 | 66.14 |
| % Another Race | 1.60 | 1.22 | 0.19 | 8.47 |
| % with HS Diploma | 41.77 | 5.49 | 5.49 | 52.70 |
| % with Some College | 18.98 | 4.93 | 4.93 | 33.63 |
| Land Area | 437.46 | 208.94 | 82.61 | 1039.80 |
| % Receiving TANF | 0.07 | 0.09 | 0.01 | 0.80 |
| Precipitation | 46.43 | 7.05 | 7.05 | 66.78 |
| Gas Production | 14,566,255.92 | 61,124,387.99 | 0.00 | 597,800,583.00 |
| Number of Gas Wells | 779.72 | 968.21 | 0.00 | 5146.00 |
| Coal Production | 2,502,706.35 | 5,033,593.19 | 0.00 | 32,446,186.00 |

Note: This table presents descriptive statistics for the county characteristics used in the regressions outlined in Section 2.4.4 of this paper. These statistics are aggregated and presented by metropolitan, non-metropolitan, and state totals, respectively. SD refers to standard deviation. For variable descriptions see Table B.3

Table B.3: **Descriptions of Variables Used in Analysis**

| | Variable | Description |
|---|---|---|
| **1** | % Government Jobs | Percentage of county's industry in government for years 1990, 2000, 2010, 2020 from the Bureau of Economic Analysis (BEA). |
| **2** | % Agriculture Jobs | Percentage of county's industry in agriculture for years 1990, 2000, 2010, 2020 from the (BEA). |
| **3** | % Manufacturing Jobs | Percentage of county's industry in manufacturing for years 1990, 2000, 2010, 2020 from the (BEA). |
| **4** | % Mining Jobs | Percentage of county's industry in mining for years 1990, 2000, 2010,2020 from the (BEA). |
| **5** | % Non-Farm Jobs | Percentage of county's industry in non-farming for years 1990, 2000, 2010, 2020 from the (BEA). |
| **6** | Personal Income | County personal income in thousands of U.S. dollars for years 1990, 2000, 2010, 2020. Consists of the income received in return for labor, land, capital use, and other income such as transfer receipts retrieved from the (BEA). |
| **7** | Unemployment Rate | The number unemployed as a percentage of the labor force for the years 1990, 2000, 2010, 2020 from the U.S. Bureau of Labor and Statistics (BLS). |
| **8** | Life Expectancy | Estimates for life expectancy at birth from the Institute for Health Metrics and Evaluation (IHME). 2020 data is unavailable so we use data for years 1990, 2000, 2010, and 2014. |
| **9** | % African American | Percentage of the population that is African American for the years 1990, 2000, 2010, 2020 from the U.S. Census Bureau. |
| **10** | % Female < 25 | Percentage of the working that is Female below the age of 25 for the years 1990, 2000, 2010, 2020 from the U.S. Census Bureau. |
| **11** | % Female 25-54 | Percentage of the working that is Female between the ages 25 - 54 for the years 1990, 2000, 2010, 2020 from the U.S. Census Bureau. |
| **12** | % Female 54- 65 | Percentage of the working that is Female between the ages 54 - 65 for the years 1990, 2000, 2010, 2020 from the U.S. Census Bureau. |

Continued on next page

**Table B.3 – continued from previous page**

| | Variable | Description |
|---|---|---|
| **13** | % Another Race | Percentage of the working that is not white or African American for the years 1990, 2000, 2010, 2020 from the U.S. Census Bureau. |
| **14** | % with HS Diploma | Percentage of adults 25 years or older with a high school diploma or equivalent for years 1990, 2000, 2015 from the United States Department of Agriculture (USDA) and for 2020 from the U.S. Census Bureau. |
| **15** | % with Some College | Percentage of adults 25 years or older with between 1 - 3 years of college for the years 1990, 2000, 2015 from the United States Department of Agriculture (USDA) and for 2020 from the U.S. Census Bureau. |
| **16** | Land Area | Land area of each county in square miles from the U.S. Census Bureau. |
| **17** | % Receiving TANF | Percentage of families with related children who receive public assistance under the Temporary Aid to Needy Families (TANF) program for the years 1990, 2000, 2010, 2015 from the Kids County Data Center at the Annie E. Casey Foundation. |
| **18** | Precipitation | Amount of rain in inches per county for 1990, 2000, 2010, and 2020 from the National Centers for Environmental Information. |
| **19** | Gas Production | Gas production in each county measured in 1000s of cubic feet for the years 1990, 2000, 2010, and 2020 from the West Virginia Geological and Economic Survey. |
| **20** | Number of Gas Wells | Number of gas wells in each county for the years 1990, 2000, 2010, and 2020 from the West Virginia Geological and Economic Survey. |
| **21** | Coal Production | Coal production in each county measured in tonnes for 1990, 2000, 2010, and 2020 from the West Virginia Office of Mines' Health Safety and Training. |
| **22** | Interstate | A dummy variable for whether at least one interstate passes through a given county from the Department of Transportation's Federal Highway Administration Highway Performance Monitoring System (HPMS) |

Note: We use the 2013 USDA RUCC designations in this analysis.

Table B.4: **Panel Regression Results: Variance Decomposition for County LFPR Change**

| Variable | State | Metro | Non-Metro | County |
|---|---|---|---|---|
| % Government Jobs | 0.317 | -2.35** | -0.298 | 0.43 |
| | (0.014) | (0.02) | (0.0009) | (0.01) |
| % Agriculture Jobs | 1.927 | 0.98 | -2.98** | 0.85 |
| | (0.03) | (0.039) | (0.020) | (0.02) |
| % Manufacturing Jobs | -0.558 | -0.75 | -0.133 | 0.81*** |
| | (0.007) | (0.013) | (0.005) | (0.01) |
| %Mining Jobs | -0.176 | -0.533 | -0.299 | 0.39 |
| | (0.008) | (0.013) | (0.004) | (0.01) |
| % Non-Farm Jobs | -0.078 | -0.71 | -0.387 | 0.67 |
| | (0.014) | (0.018) | (0.008) | (0.01) |
| Unemployment Rt | -1.753* | 0.888 | 0.381 | 0.71 |
| | (0.017) | (0.019) | (0.011) | (0.01) |
| Life Expectancy | 7.204 | -7.48 | 0.381 | -7.54* |
| | (0.086) | (0.099) | (0.065) | (0.08) |
| % African American | 0.044 | -1.96 | 6.31 | -1.09 |
| | (0.037) | (0.071) | (0.024) | (0.03) |
| % Female < 25 | -4.559* | 0.66 | 0.487 | 3.87* |
| | (0.048) | (0.054) | (0.029) | (0.04) |
| % Female 25 - 54 | 0.152 | 0.062 | 1.442 | 1.38 |
| | (0.048) | (0.05) | (0.04) | (0.04) |
| % Female 54 - 65 | -0.057 | -0.225 | -0.518 | 0.00 |
| | (0.003) | (0.003) | (0.003) | (0.00) |

Continued on next page

**Table B.4 – continued from previous page**

| Variable | State | Metro | Non-Metro | County |
|---|---|---|---|---|
| % Another Race | 0.889 | -1.35 | -0.003 | -0.18 |
| | (0.053) | (0.053) | (0.044) | (0.05) |
| Personal Income | 0.011*** | 0.003 | 0.548*** | -0.01 |
| | (0.000) | (0.000) | (0.000) | (0.00) |
| % with HS Diploma | 0.287 | -0.838 | -0.011 | 0.33*** |
| | (0.013) | (0.017) | (0.01) | (0.01) |
| % with Some College | -1.375* | 3.327** | -0.671 | 0.84 |
| | (0.14) | (0.022) | (0.01) | (0.01) |
| Land Area | 2.939 | 1.587 | -0.296 | -3.58 |
| | (0.049) | (0.096) | (0.03) | (0.02) |
| % Receiving TANF | -13.359* | -26.733 | 1.75 | 10.48 |
| | (0.37) | (1.03) | (0.14) | (0.32) |
| Precipitation | -0.693 | -0.023 | 1.131 | 0.35 |
| | (0.01) | (0.011) | (0.007) | (0.01) |
| Gas Production | 0.059** | 0.009 | 0.304*** | -0.03* |
| | (0.000) | (0.001) | (0.000) | (0.00) |
| Number of Gas Wells | -.0899* | -0.54 | -0.039 | 1.41*** |
| | (0.012) | (0.034) | (0.02) | (0.01) |
| Interstate | -9.319** | 25.53*** | -0.202 | -4.47** |
| | (0.24) | (0.234) | (0.006) | (0.24) |
| Coal Production | 0.004 | 0.004 | -0.003 | -0.01 |
| | (0.000) | (0.000) | (0.000) | (0.00) |
| Draws | 5000 | 5000 | 5000 | 5000 |

Continued on next page

**Table B.4 – continued from previous page**

| Variable | State | Metro | Non-Metro | County |
|---|---|---|---|---|
| $R^2$ | 0.38 | 0.37 | 0.25 | 0.4 |

Note: This table reports panel regression results with $\theta_{it}^{S}$, $\theta_{it}^{R}$, $\theta_{it}^{i}$ as the regressand in columns 2, 3/4, and 5, respectively. These regressions include all the county characteristics in column 1 jointly as regressors. All regressions include a county specific intercept term. Regressions for each factor (i.e. columns 2-5) are computed 5000 times for 5000 random draws of $\theta_{it}^{S}$, $\theta_{it}^{R}$, $\theta_{it}^{i}$. These draws are distributed evenly over each month of the panel years (1990, 2000, 2010, 2020) and drawn randomly. P-values for each regression are saved. If over 800 of the 5000 saved are statistically significant at the %5 level, we denote with *. If over 1500 or 2000 of the 5000 saved are statistically significant at the %5 level, we denote with ** and ***, respectively.

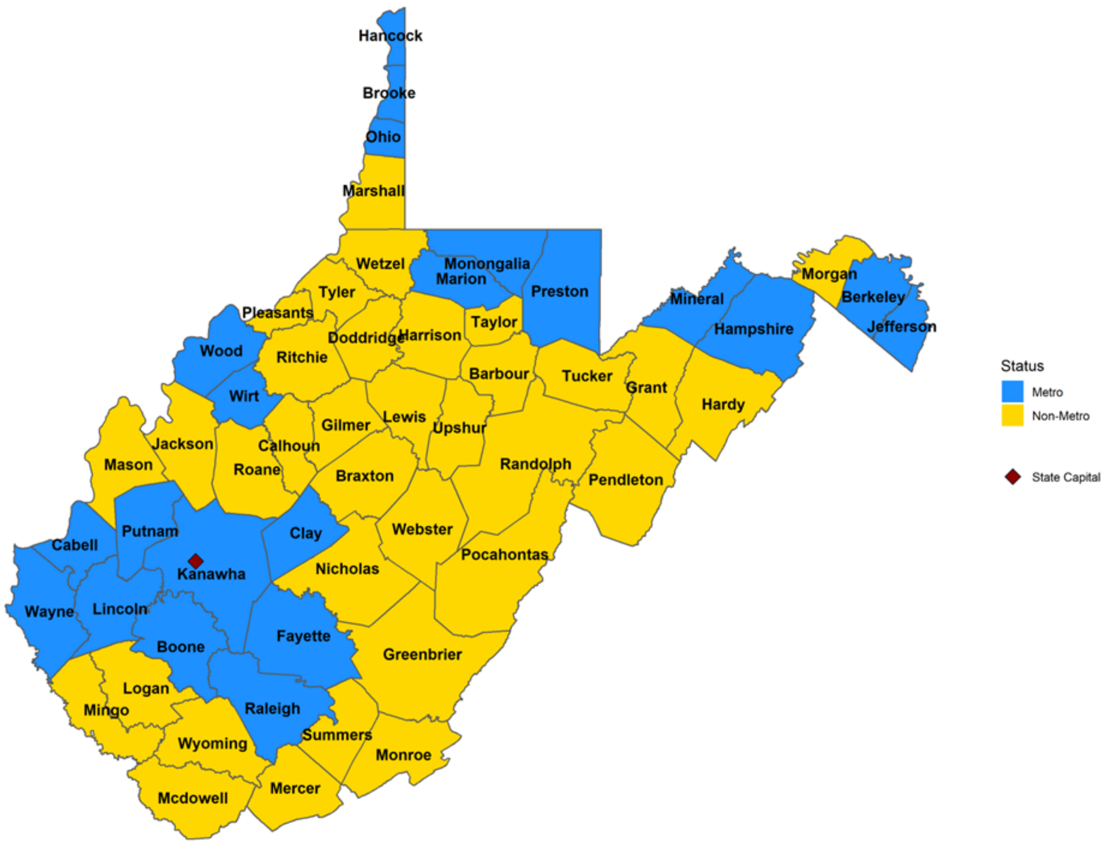


Figure B.1: **West Virginia Counties by RUCC Status**

### 2.6.1 Results Continued

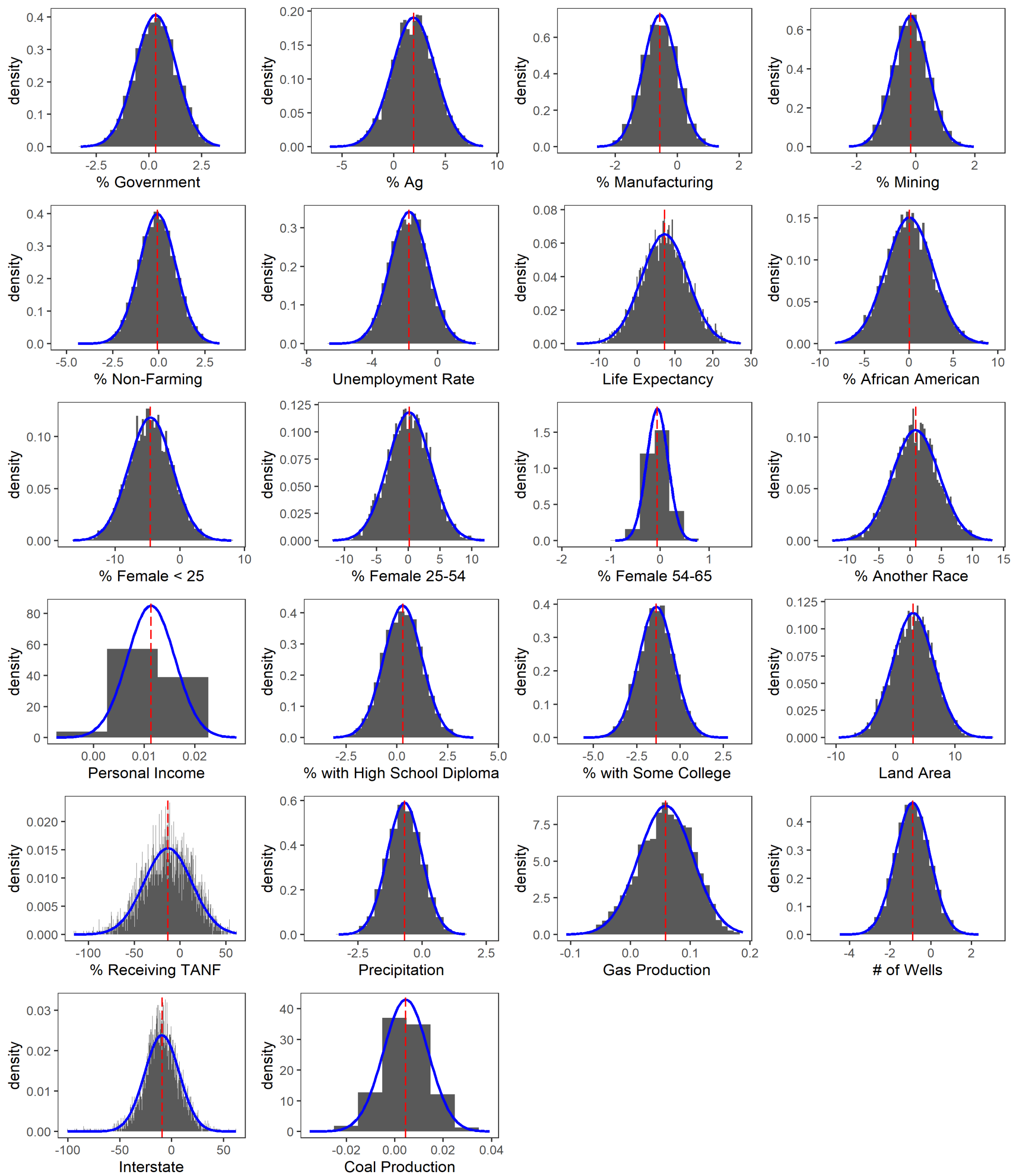


Figure B.2: **Distribution of Regression Coefficients - State Factor**

Note: Shaded bars represent the distribution of coefficients for each variable after the 5000 (State) regressions are run. These Red lines represent the mean of each distribution and blue lines are normal distributions with the same mean and standard deviation as the data mapped over each distribution.

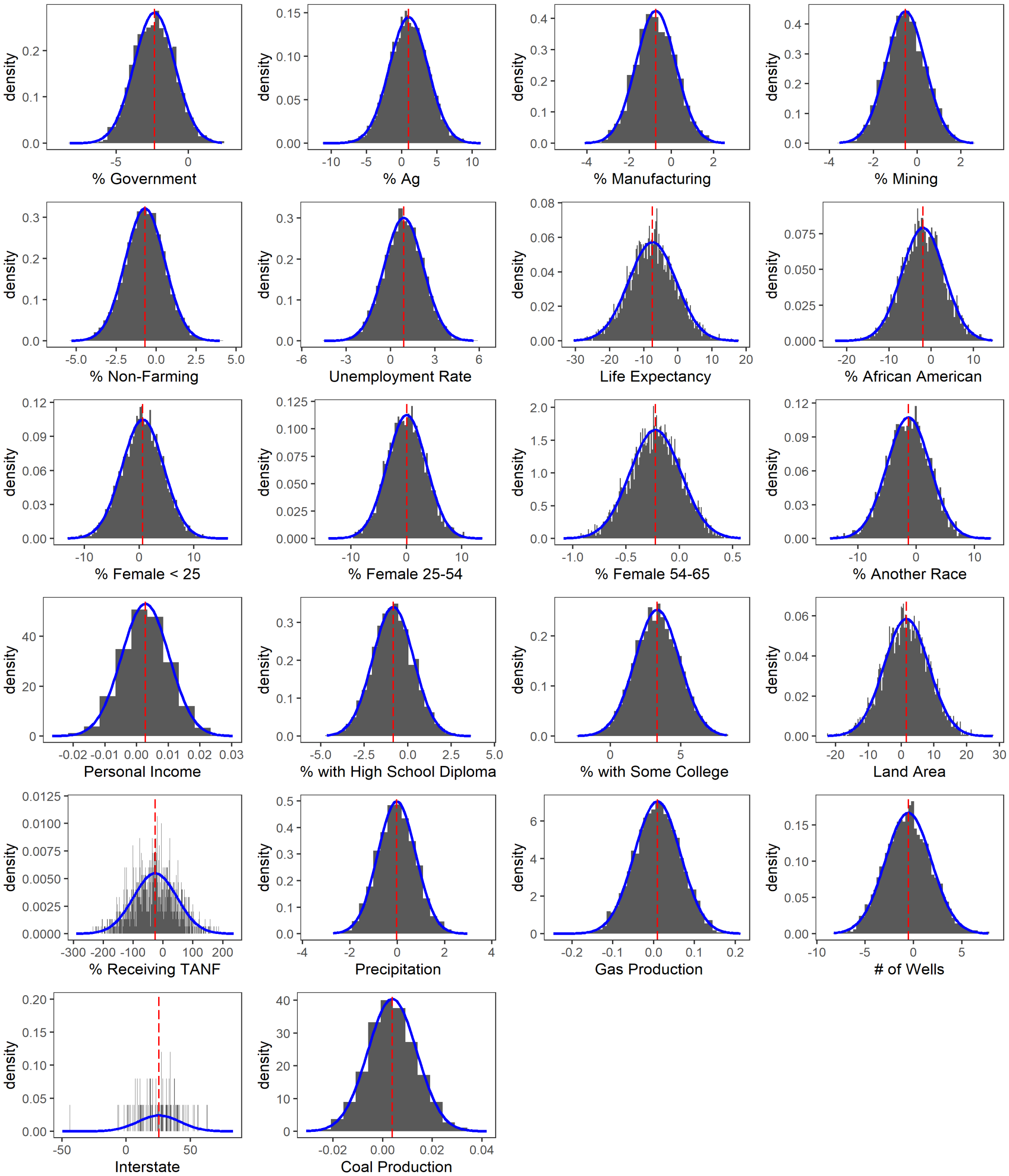


Figure B.3: **Distribution of Regression Coefficients - Metro Factor**

Note:Shaded bars represent the distribution of coefficients for each variable after the 5000 (Metropolitan) regressions are run. These Red lines represent the mean of each distribution and blue lines are normal distributions with the same mean and standard deviation as the data mapped over each distribution.

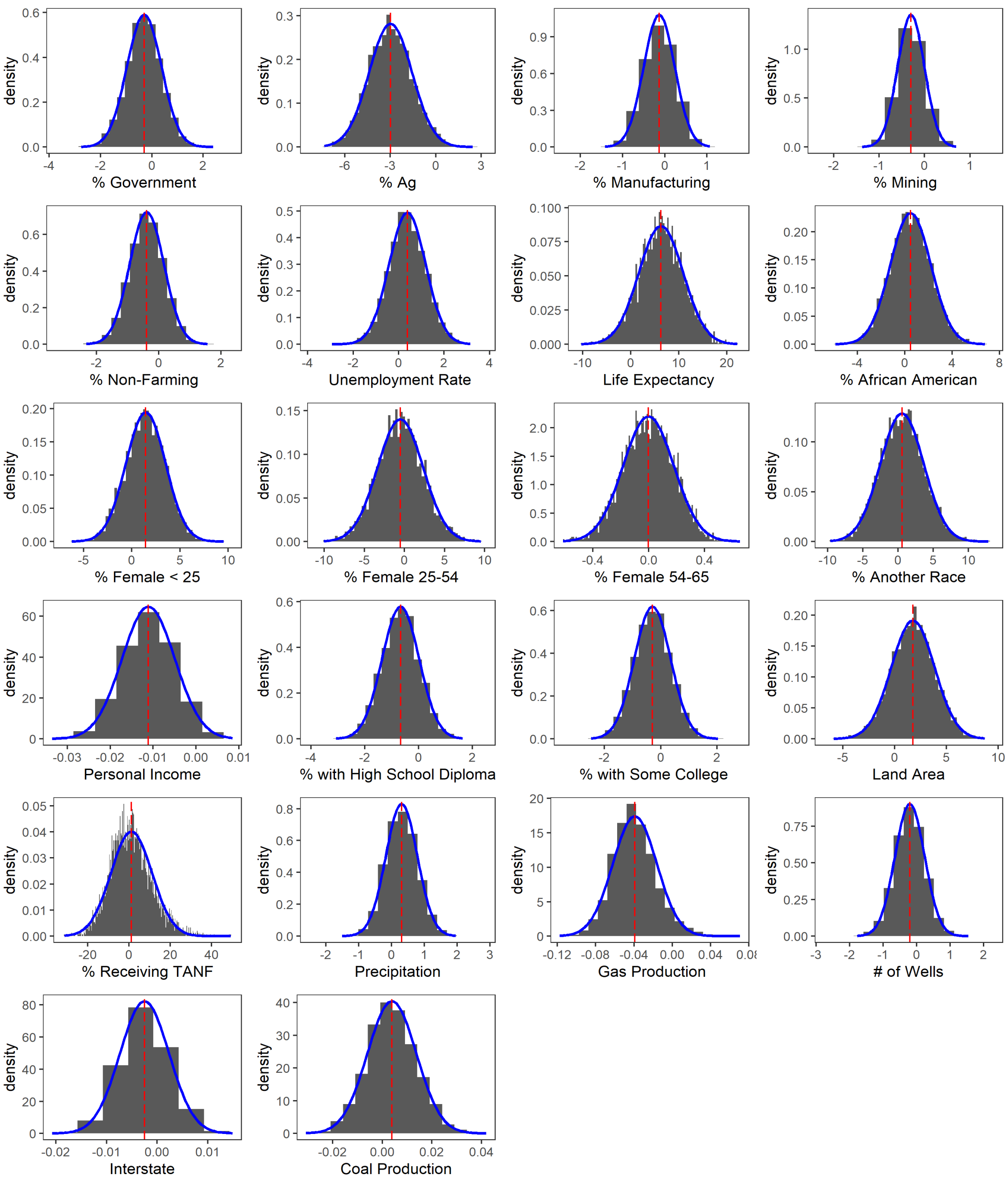


Figure B.4: **Distribution of Regression Coefficients - Non-Metro Factor**

Note: Shaded bars represent the distribution of coefficients for each variable after the 5000 (Non-metropolitan) regressions are run. These Red lines represent the mean of each distribution and blue lines are normal distributions with the same mean and standard deviation as the data mapped over each distribution.

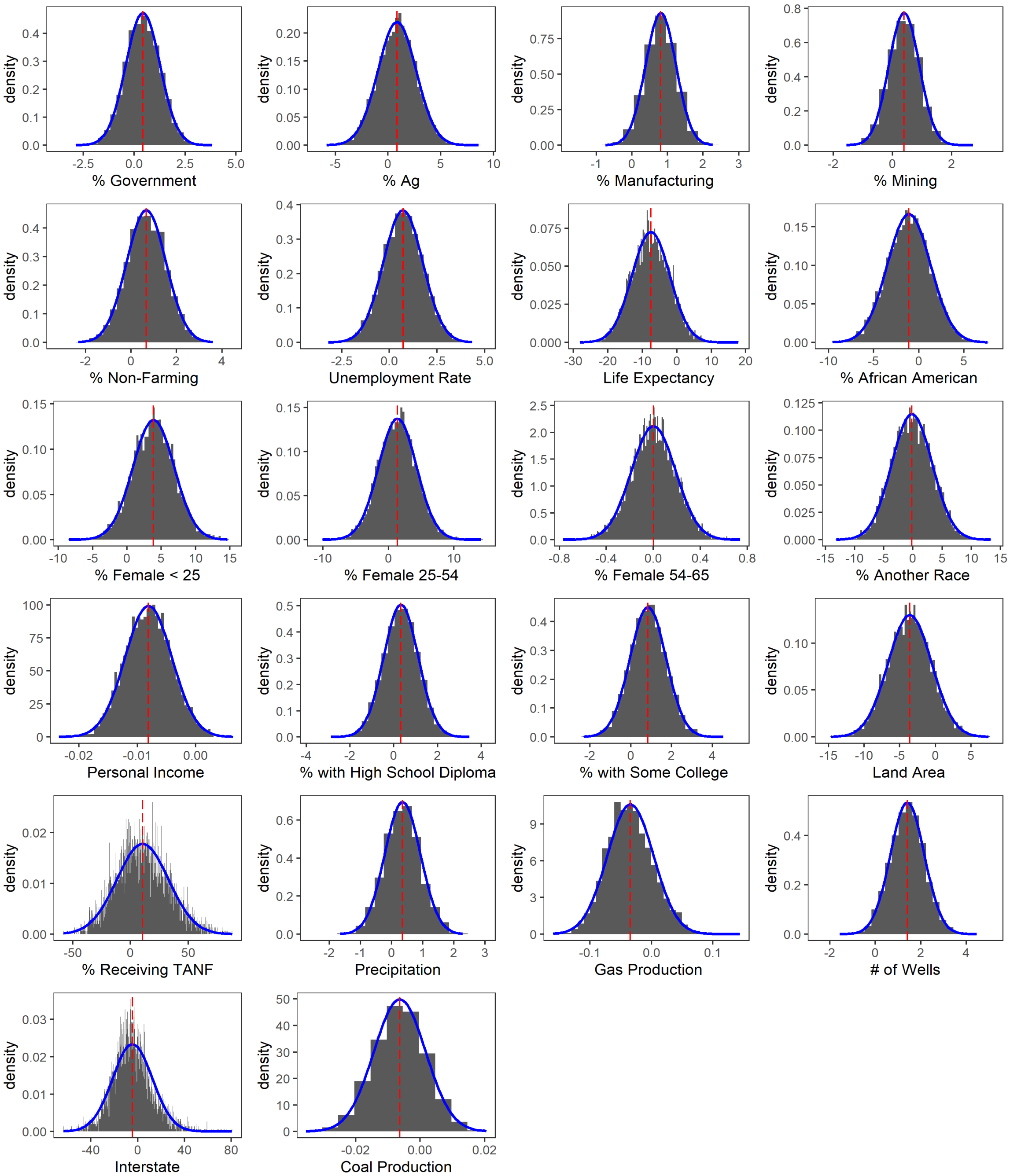


Figure B.5: **Distribution of Regression Coefficients - County Factor**

Note: Shaded bars represent the distribution of coefficients for each variable after the 5000 (County-specific) regressions are run. These Red lines represent the mean of each distribution and blue lines are normal distributions with the same mean and standard deviation as the data mapped over each distribution.

# Chapter 3

# The Unemployment Invariance Hypothesis in West Virginia: A Tale of Two Indicators

## 3.1 Introduction

Soaring unemployment and low labor force participation associated with economic shocks (such as the Great Recession in 2008-09 and the COVID-19 Pandemic) have been particularly concerning for policymakers and labor market participants. The movement of workers between (un)employment and not-in-the-labor-force are impacted by shocks. Key economic indicators such as the unemployment rate (UR) and labor force participation rate (LFPR) are affected by extension. Recent and substantial changes observed in these labor market flows suggest a causal link between the unemployment and labor force participation rates (Blanchard et al., 1990; Burda and Wyplosz, 1994; Bell and Smith, 2002; Gomes, 2012; Lin and Miyamoto, 2012). Alternatively, some researchers posit that while a relationship may exist in the short-run, the UR should be independent of changes in the labor force in the long run (Layard et al., 2005).

Three hypotheses emerge from this debate in the literature to explain the relationship between the LFPR and UR. The first two hypotheses are derived from the two typical responses of households when facing high unemployment. On the one hand, members of the household who are actively searching for work may become “discouraged”, give up their search, and drop out of the labor force. This phenomenon is referred to as the “discouraged worker effect” (DWE). In this context, with a lower labor force (unemployed + employed), falling unemployment rate would signal better labor market (and economic) conditions than the reality would reflect. On the other, household members may engage in searching for additional jobs to compensate for lost wages. These individuals transition from not-in-the-labor-force to either unemployed or employed. This so-called “added worker effect” (AWE) results in increases the LFPR and potentially increases the UR.

The third hypothesis is one where neither of these phenomena (AWE nor DWE) dominates the other. In this case, these two effects cancel each other, resulting in no effect or no relationship between the LFPR and the UR. This is the unemployment invariance hypothesis (UIH).

While initially posited by Layard et al. (2005), empirical testing for the UIH began with Österholm (2010). Österholm (2010) tests the unemployment invariance, DWE, and AWE hypotheses through cointegration testing and with a linear vector error correction model (VECM) on aggregate and gender data in Sweden. While Österholm (2010) rejects the UIH for Sweden, in favor of the DWE, several other authors, following the same approach, find results supporting the UIH. Table 4 shows that for aggregate data in Australia (Nguyen Van, 2016), Romania (Oţoiu and Ţiţan, 2016), Turkey (Tansel et al., 2016b), and parts of Africa (Raifu and Adeboje, 2022), the UIH is supported. However, Table 4 also shows that several authors find results congruent with Österholm (2010). For example, in addition to Sweden, results on aggregate data favoring the DWE are found for Japan (Lin and Miyamoto, 2012), Canada (Tansel et al., 2016a), and parts of Africa (Raifu and Adeboje, 2022). Emerson (2011) also rejects the UIH but finds that the AWE dominates in aggregate U.S. data. Given the lack of consensus in the extant literature, the question of the UIH's validity may not be one of theory but potentially one of the location assessed. Tansel et al. (2016b) suggests that the differences in results across these studies may be due to labor market institution differences across countries.

Table 4: **Survey of Previous Literature Examining the Unemployment Invariance Hypothesis**

| **Author(s)** | **Period** | **Region** | **Conclusion** |
|---|---|---|---|
| Ošterholm (2010) | 1970M1 – 2007M4 | Sweden | DWE |
| Emerson (2011) | 1948M1 - 2010M2 | U.S. | AWE |
| Lin and Miyamoto (2012) | 1980M1 - 2010M12 | Japan | DWE |
| Nguyen Van (2016) | 1978M2 – 2014M12 | Australia | UIH |
| Oţoiu and Ţiţan (2016) | 1996Q1 – 2012Q4 | Romania | UIH |
| Tansel et al. (2016a) | 1976M1 – 2015M12 | Canada | DWE |
| Tansel et al. (2016b) | 1983Q3 - 2013Q4 | Turkey | UIH |
| Altuzarra et al. (2019) | 1987Q2 - 2016Q4 | Spain | UIH |
| Congregado et al. (2021) | 1995M1 - 2016:M4 (3 sub-samples) | Poland | AWE & DWE & UIH |
| Raifu and Adeboje (2022) | 1991 - 2018 (Annual) | Africa (5 Regions) | DWE & UIH |

Note: UIH indicates the Unemployment Invariance Hypothesis. AWE indicates the Added Worker Effect and DWE indicates the Discouraged Worker Effect.

Admittedly, the location may not be the only factor affecting results in these studies. Congregado et al. (2021) suggests that the results of these previous studies may be biased by virtue of failing to account for instabilities or structural breaks in the long-run relationship between the LFPR and the UR. Congregado et al. (2021) contend that the relationship between the LFPR and the UR may change over a long period of time. If such is the case, estimating linear cointegration relations between the LFPR and the UR without accounting for the structural changes would lead to spurious inference. Extending the approach of Österholm (2010) done for Sweden, Congregado et al. (2021) investigate the unemployment invariance, DWE, and AWE hypotheses across different regimes/periods surrounding calculated structural breaks in Poland. As shown in Table 4, Congregado et al. (2021) finds evidence for all three hypotheses depending on the sub-sample studied.

In this paper, we follow the approach of Congregado et al. (2021) when testing for the unemployment invariance, DWE, and AWE hypotheses across vintages for West Virginia. We first examine the time-series properties of monthly seasonally-adjusted labor force in West Virginia over the period 1976:1-2022:6 and testing for unit roots. We proceed to formally determining potential structural breaks and instability across the sample. Using a Vector Error Correction model (VECM) we test the validity of our three hypotheses our the full sample and sub-samples.

Our traditional unit root measures suggest that West Virginia's unemployment and labor force participation rates are first difference stationary (i.e., I(1)). When accounting for structural breaks, we confirm the existence of three breaks, which implies four regimes over the 47-years span. Further, our Johansen's cointegration methodology (Johansen, 1988, 1991) reveals the existence of a single cointegrating relationship LFPR and UR for three of the four regimes. Our cointegration and VECM estimation results demonstrate that whether the DWE, AWE, or UIH prevails depends on the regime studied. For the full sample and the third regime (2010:M05 - 2022:M06), we find that the discouraged worker effect (DWE) prevails in West Virginia. However, for the other three regimes (1976:M01 - 1989:M08, 1989:M09 - 1998:M04, and 2010:M05 - 2022:M06), we find the unemployment invariance hypothesis to hold. The UIH for these regimes indicates that the DWE and AWE cancel each other resulting in no relationship between the LFPR and UR in West Virginia.

These findings imply that policies and programs that encourage labor force growth and economic development should be tailored to the prevailing phenomenon of the current regime. In the case of the DWE, workers may become discouraged and drop out of the labor force due to frictional or structural unemployment. We suggest policymakers design policies and implement programs to inform workers of job vacancies, train and re-train individuals, and provide incentives to participate in the workforce. Incentives may include flexible work schedules, broadband infrastructure investments to facilitate remote work, and childcare options or reimbursements (O'Leary and Boettner, 2015). Accepting a propensity for the DWE in West Virginia, policies should be designed to create jobs and to mitigate unemployment in the short and long run. Accounting for future implications of current policies may be a vital first step to ensuring future economic prosperity in economically discouraged areas like West Virginia.

In the case of the AWE, the UR may still increase if the added workers disproportionately join the unemployed rather than the employed. To avoid the potential increase in the UR caused by added workers transitioning from not-in-the-labor-force to unemployed, we again suggest efforts to reduce frictional and structural unemployment. Spletzer (1997) notes that given lower discount factors, lower search costs, and lower wage offer averages, job search for added workers is typically shorter for unskilled jobs. Therefore, policymakers should emphasize expanding industry and creating skilled positions to shorten the job search of skilled laborers joining the market. In the case of the UIH, Layard et al. (2005) posits that the labor market works efficiently to self-stabilize, resulting in no effects on long-run unemployment. If we accept that the theory holds, policymakers may focus on increasing productivity, increasing the labor force participation rate, and supporting economically struggling areas without fear of inadvertent or adverse effects in the long run.

Our focus on West Virginia is two-fold. First, compared to the rest of the U.S., economic disparity in that state has persisted for decades. This disparity extends to measures of poverty, education, health care, and labor markets, including both the LFPR and the UR (Billings, 1974; Stephens and Deskins, 2018; Isserman and Rephann, 1993; Behringer and Friedell, 2006; Muntaner and Barnett, 2000). Several counties in West Virginia remain some of the most economically distressed areas in the U.S. and the Appalachian Region (Appalachian Regional Commission, 2020). Given the wealth of natural resources in the state, the West Virginia labor market has relied heavily on the coal and natural gas industries for decades. As policymakers continue to air-mark funding to boost the LFPRs in economically distressed regions, like West Virginia, it is important to understand which phenomenon (UIH, DWE, or AWE) holds. Evidence for the UIH in West Virginia would imply that economic development efforts and policies designed to encourage labor force participation will have negligible effects on the unemployment rate in the long run. However, evidence for the DWE would imply larger effects on long-run unemployment in the state, potentially perpetuating high levels of unemployment. A better understanding of the relationship between UR and LFPR could therefore help to break the cycle of despair for these economically struggling areas and jump-start economic prosperity.

Secondly, we study the relationship between the LFPR and the UR in West Virginia because of the sparsity in economic research for the state. The historic low unemployment rate (3.5 percent) attained by West Virginia in May 2022 received much media attention. However, the reliability of the unemployment rate as an economic indicator depends on its relationship to LFPR (Gustavsson and Österholm, 2006). In the case of the DWE, low unemployment rates results from individuals dropping out of the labor force which falsely signals economic prosperity. Despite this media attention and the persistent disparity in economic indicators in West Virginia, ours is the first study to investigate the validity of the UIH and the existence of the DWE or AWE in the state. As such, we believe West Virginia provides an interesting addition to the debate in the literature discussed above.

As Österholm (2010) and others point out, studying the relationship between the LFPR and UR has implications for economic theory, modeling, and labor market policies. Our study contributes to economic theory by comparing West Virginia results to other regions and aggregate levels to help determine if labor market institution differences and aggregation levels lead to differences in results. This study also contributes to modeling the relationship between labor force participation and unemployment rates by extending the approach to consider structural breaks in the data in the spirit of Congregado et al. (2021). Lastly, we suggest that performing the analysis on state-level data better informs local economies' labor policies by shifting the focus away from national intervention and towards actionable policies at the local level. This may help distressed areas like West Virginia reverse persistent economic under-performance in the future

The remainder of the paper proceeds as follows. We review the conceptual framework for the UIH in Section 3.2. We describe the data in Section 3.3. In Section 3.4, we discuss the empirical methodology and present our results in Section 3.5. We conclude and discuss policy implications for our findings in Section 3.6.

## 3.2 Conceptual Framework and Background

In this section, we discuss the conceptual framework of the UIH and assumptions made on a basic unemployment model imposed by Layard et al. (2005). To start, consider a policy or shock that restricts workers' ability to retire early, effectively increasing the working-age population. Alternatively, suppose a policy promotes research and development, which, in turn, increases productivity, effectively increasing labor demand. All else being equal, with increases in the working-age population or demand for labor, one may ask if these policies will affect unemployment in the long-run. In response to questions like these, Layard et al. (2005) developed the UIH. The hypothesis posits that regardless of what may occur in the short run, the unemployment rate should be independent of changes in the labor force (capital stock and productivity growth) in the long run. Therefore, policies that increase the working population or productivity would not affect on long-run unemployment. Layard et al. (2005) proposed that the labor market self-stabilizes back to equilibrium to ensure that long-run unemployment remains unaffected. Any effects on unemployment will be eventually counteracted by labor demand, labor supply, wage setting, or other labor market responses.

Figure 17 demonstrates the theoretical effect on long-run unemployment after an increase in productivity, assuming the UIH holds. As shown in the figure, an increase in productivity leads to an increase in the labor demand curve (LD) which specifies aggregate employment for given real wages. This increase is indicated by the shift from LD to $LD_2$. The UIH posits, in this case, that the wage-setting curve– the wages associated with given unemployment rates– will shift by the same amount in the long run. This is indicated by the shift from $WS_1$ to $WS_2$. The equilibrium employment level (E*) and equilibrium real wage ($W_2$*) are determined by the intersection of the labor demand curve and the wage-setting curve. Shifts by the same amount results in no change in the equilibrium employment level. Since the equilibrium unemployment level, U*, is determined by the difference between the labor supply curve and the equilibrium employment level, there is no change in long-run unemployment.

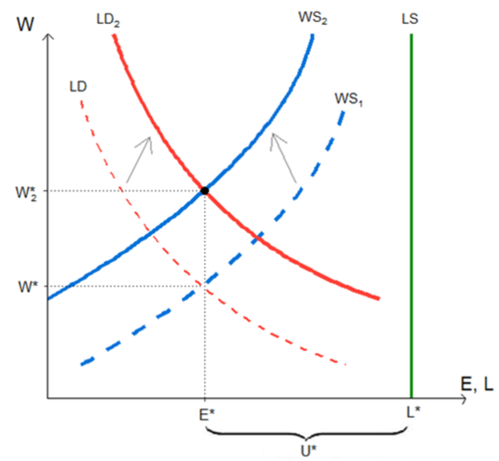


Figure 17: **The Effect of a Productivity Increase on Long-Run Unemployment Under the Unemployment Invariance Hypothesis**

Note: LD denotes the initial labor demand curve and $LD_2$ denotes the labor demand curve after a shift caused by increased productivity. WS denotes the initial wage setting curve and $WS_2$ denotes the wage setting curve after a shift in response to labor demand increase. LS denotes the labor supply curve. U* represents unemployment or the difference between the labor force (L*) and the employment level (E*). W* and $W^*_2$ denote the equilibrium wage before and after the shift in the wage setting curve, respectively.

To offer another example, Figure 18 illustrates the theoretical effect on long-run unemployment after an increase in the labor force, assuming the UIH holds. An increase in the labor force leads to an increase in the labor supply curve (LS). The labor supply curve shows the size of the labor force at given wages. The increase in labor supply is indicated by the shift from LS to $LS_2$, in Figure 18. After this increase, the UIH posits that the wage-setting curve increases by the same amount. This is indicated by a right shift from $WS_1$ to $WS_2$. Again, the equilibrium employment levels (E* and $E^*_2$) and equilibrium real wages (W* and $W^*_2$) are determined by the intersection of the labor demand curve (LD) and the wage-setting curves ($WS_1$ and $WS_2$). The equilibrium unemployment levels (U* and $U^*_2$) are determined by the respective differences between the labor supply curves and the equilibrium employment levels (L* - E* and $L^*_2$ - $E^*_2$). Since the labor supply curve and the wage-setting curve increase proportionally, the long-run unemployment rate remains unchanged. This is denoted by U* and $U^*_2$ where $U^* = U^*_2$.

The debate regarding the validity of the UIH revolves around the assumptions and restrictions made

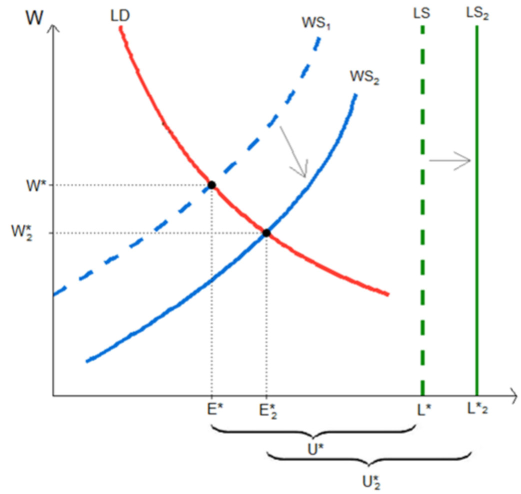


Figure 18: **Labor Force Size Increase on Long-Run Unemployment under the Unemployment Invariance Hypothesis**

Note: LD denotes the labor demand curve. LS denotes the initial labor supply curve and $LS_2$ is the labor supply curve after a shift caused by an increase in the labor force (L* and $L*_2$, respectively). $WS_1$ denotes the initial wage setting curve and $WS_2$ denotes the wage setting curve after the shift. U* and $U*_2$ represents the initial and final unemployment, respectively. These are the differences in the initial and new labor force (L*, $L*_2$) and the initial and new employment levels (E*, $E*_2$),(L* - E*, $L*_2$ - $E*_2$) respectively. The equilibrium wage is denoted W* before and $W*_2$ after the shift.

in Layard et al. (2005) on the basic model describing equilibrium unemployment. As Table 4 shows, the UIH, as proposed by Layard et al. (2005), is often rejected by the data. However, the model and restrictions used to develop the UIH are still widely used (Karanassou and Snower, 2004). Since this is the case, this paper provides further tests of the UIH and its validity. However, since the theoretical model assumptions and restrictions fueling the contention have policy implications, like in the examples we have previously discussed, we believe it prudent to set up the basic model describing equilibrium unemployment and highlight the restrictions imposed. To do this, we start with price and wage decisions by agents in the market. The interaction between price and wage mark-up equations define the unemployment equilibrium (Blanchard, 1986).

Prices are set by firms according to profit-maximizing behavior and as a mark-up on expected wages as follows:[1]

$$P_t - W_t^e = \lambda_0 - \lambda_u u_t \tag{19}$$

where $P_t$ is the logged price level, $W_t^e$ is the logged expected wage level, and $u_t$ is the unemployment rate. It is assumed that the coefficients $\lambda_0$ and $\lambda_u$ are positive.

Wages are set by workers through wage efficiency or union bargaining as a mark-up on expected prices as follows:

$$W_t - P_t^e = \delta_0 - \delta_u u_t \tag{20}$$

where $W_t$ is the logged wage level, $P_t^e$ is the logged expected price level, and $\delta_0$ and $\delta_u$ are assumed to be positive. In equilibrium, expected price and wage are the actual price and wage. So, $P_t^e = P_t$ and $W_t^e = W_t$. Substituting these into Equations 19 and 20 and solving for $u*$, the unemployment rate is determined as:

$$u* = \frac{\lambda_0 + \delta_0}{\lambda_u + \delta_u} \tag{21}$$

To see the necessary restrictions for the UIH to hold, this model is extended to include capital stock, labor force, and productivity. It is first assumed that the price mark-up equation in Equation 19 depends positively on employment ($E_t$), due to diminishing returns on labor. All else equal, since an increase in the labor force reduces a firm's search cost, the price mark-up over the expected wage is assumed to depend negatively on the labor force ($L_t$). Since an increase in the capital stock ($K_{t-1}$) or technology ($\tau_t$), increases productivity and permits a firm to reduce the price relative to the wage, the price mark-up over the expected wage is assumed depend negatively on capital stock and technology. Including all of these in the model, the extended price mark-up equation becomes:

$$P_t - W_t^e = \lambda_0 + \lambda_E E_t - \lambda_L L_t - \lambda_K K_{t-1} - \lambda_\tau \tau_t \tag{22}$$

where $\lambda_0$, $\lambda_E$, $\lambda_L$, $\lambda_K > 0$.

[1] All variables are expressed in natural logs in this model construct.

To similarly extend Equation 20, the wage mark-up is assumed to depend positively on the capital stock and technology, $K_{t-1}$ and $\tau_t$ respectively. Also, since productivity increases enable workers to claim higher wages, the wage mark-up is assumed to depend positively on capital stock and technology. The extended wage mark-up equation is as follows:

$$W_t - P_t^e = \delta_0 - \delta_u u_t + \delta_K K_{t-1} + \delta_\tau \tau_t \tag{23}$$

where $\delta_0$, $\delta_u$, $\delta_K$, $\delta_\tau > 0$. Since, $L_t$ and $E_t$ are logged labor force and logged employment, respectively, the unemployment rate can be approximated as:

$$u_t = L_t - E_t \tag{24}$$

Using the fact that in equilibrium, $P_t = P_t^e$ and $W_t = W_t^e$, we combine Equations 22 and 23 and solve for equilibrium unemployment rate, $u^*$, as follows:

$$W_t - \lambda_0 - \lambda_E E_t + \lambda_L L_t + \lambda_K K_{t-1} + \lambda_\tau \tau_t - W_t = \delta_0 - \delta_u u_t + \delta_K K_{t-1} + \delta_\tau \tau_t$$

$$\lambda_L L_t + \lambda_K K_{t-1} + \lambda_\tau \tau_t - \lambda_0 - \lambda_E E_t = \delta_0 - \delta_u u_t + \delta_K K_{t-1} + \delta_\tau \tau_t$$

$$\delta_u u_t - \lambda_E E_t = (\lambda_0 + \delta_0) + (\delta_K - \lambda_K) K_{t-1} + (\delta_\tau - \lambda_\tau)\tau_t - \lambda_L L_t$$

Adding $\lambda_E L_t$ to both sides we get:

$$\delta_u u_t + \lambda_E L_t - \lambda_E E_t = (\lambda_0 + \delta_0) + (\delta_K - \lambda_K) K_{t-1} + (\delta_\tau - \lambda_\tau)\tau_t + (\lambda_E - \lambda_L) L_t$$

$$\delta_u u_t + \lambda_E (L_t - E_t) = (\lambda_0 + \delta_0) + (\delta_K - \lambda_K) K_{t-1} + (\delta_\tau - \lambda_\tau)\tau_t + (\lambda_E - \lambda_L) L_t$$

$$\delta_u u_t + \lambda_E (u_t) = (\lambda_0 + \delta_0) + (\delta_K - \lambda_K) K_{t-1} + (\delta_\tau - \lambda_\tau)\tau_t + (\lambda_E - \lambda_L) L_t$$

$$u_t(\delta_u + \lambda_E) = (\lambda_0 + \delta_0) + (\delta_K - \lambda_K) K_{t-1} + (\delta_\tau - \lambda_\tau)\tau_t + (\lambda_E - \lambda_L) L_t$$

Therefore,

$$u_t^* = \frac{(\lambda_0 + \delta_0) + (\delta_K - \lambda_K) K_{t-1} + (\delta_\tau - \lambda_\tau)\tau_t + (\lambda_E - \lambda_L) L_t}{\delta_u + \lambda_E} \tag{25}$$

Therefore, for the UIH to hold and the unemployment rate to be unaffected by capital stock, labor force, and productivity in the long-run, then:

$$\lambda_E = \lambda_L = \lambda_u, \quad \lambda_K = \delta_K, \quad \lambda_\tau = \delta_\tau$$

These restrictions imply that the labor market contains all the equilibrating mechanisms that guarantee that the long-run unemployment remains unaffected by changes in the labor force, capital stock, and productivity growth. Any implications for policy regarding the DWE, AWE, and UIH are subject to these assumptions and restrictions.

## 3.3 Data

To test the UIH for West Virginia, we use seasonally adjusted monthly West Virginia labor force participation rate and unemployment rate data obtained from the Bureau of Labor Statistics (BLS). This data spans January 1976 to June 2022. The BLS defines the labor force participation rate as the percentage of the civilian and non-institutional population between the ages of 16-64 that are in the labor force (Bureau of Labor Statistics, 2021a). The labor force consists of the sum of the number of individuals who are either employed or unemployed. Individuals in the labor force are either working or not working but searching for work. The BLS defines unemployment rate as the percentage of the labor force who are unemployed, which, again, are individuals who are not working but are searching for work (Bureau of Labor Statistics, 2021a).

Figure 19 shows the West Virginia LFPR and UR compared to the U.S. averages over time. For over two-thirds of the sample period, West Virginia's UR exceeded the U.S. average, reaching over 18% at its peak. After briefly surpassing West Virginia unemployment rate during and following the Great Recession (2008–2010), the U.S. average unemployment rate gradually declined to fall below West Virginia levels again. In that state, the 1980s was marked by high unemployment (averaging more than 15%) and low LFPRs (peaking at around 54%). The labor growth observed in the late 1980s was buoyed by a restructuring of the West Virginia labor market away from the coal industry (Stevens, 1986). West Virginia's LFPR has remained relatively stable throughout the 1990s and 2000s, albeit far below the U.S. average.

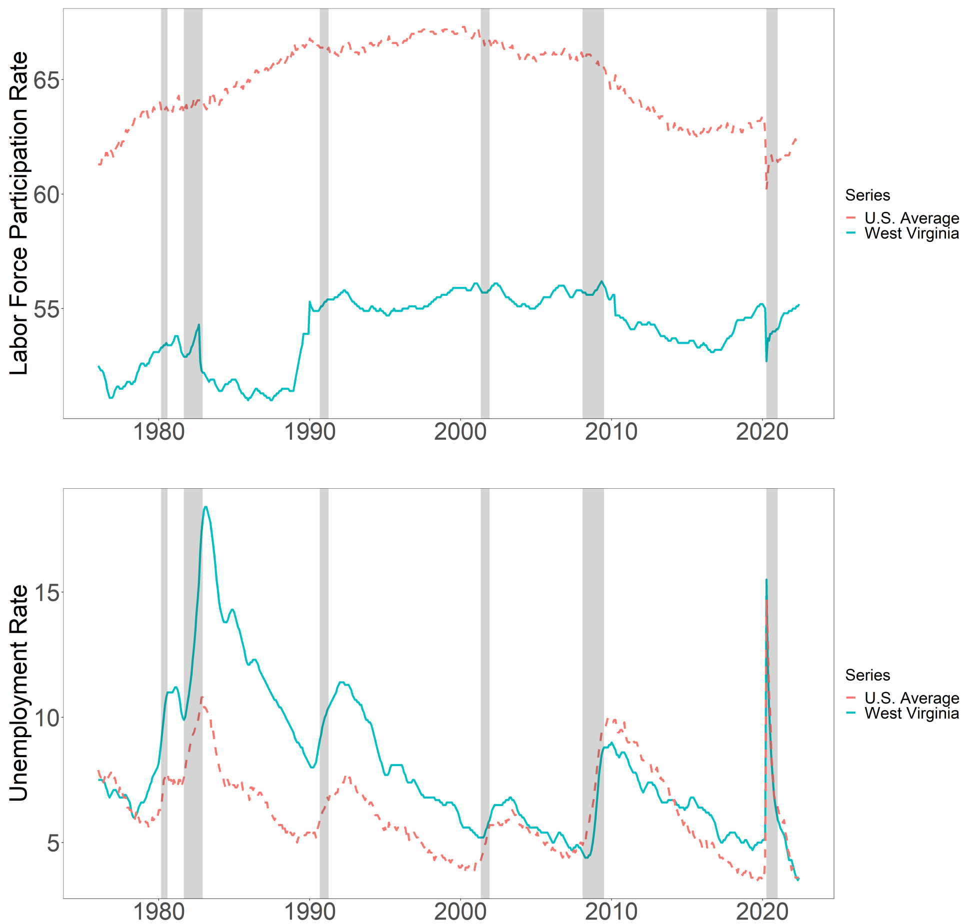


Figure 19: **Labor Force Participation and Unemployment**

Note: Shaded regions are the NBER-dated recessions.

## 3.4 Methodology

To examine whether a long-run relationship exists between the LFPR and the UR in West Virginia, we closely follow the approach of Congregado et al. (2021). This approach investigates the existence of such a relationship by applying the methodology initially developed by Österholm (2010) while also considering the possibility of structural breaks in the data. The procedure is as follows:

### 3.4.1 Unit Root Tests

First, to determine whether a long-run relationship between the West Virginia LFPR and UR may exist, we establish the time-series properties of our variables through a series of unit root tests. If the LFPR and UR are integrated of different orders, no cointegration or long-run relationship can exist. Therefore, to test for non-stationarity (unit root), we employ the Augmented Dickey-Fuller (ADF) test with GLS detrending (Elliott et al., 1996), the Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test (Kwiatkowski et al., 1992), and the Phillips-Perron (PP) test (Phillips and Perron, 1988). The null hypotheses for the ADF-GLS and PP tests are that the data has a unit root, while the null hypothesis of the KPSS test is that the data is stationary. For robustness, we also apply the Fourier Lagrange Multiplier (LM) and Fourier Augmented-Dickey-Fuller (ADF) unit root tests (Enders and Lee, 2012a). Since the Fourier approximation allows for multiple unknown break dates, unspecified break forms, and a reduction in the number of parameters to estimate, Fourier LM and ADF tests have an advantage over other unit root tests (Enders and Lee, 2012b). In both tests, the null hypothesis is a unit root with an unknown number of breaks. The alternative hypothesis is a stationary process with an unknown number of breaks. In applying these unit root tests, we determine the order of integration of the West Virginia LFPR and the UR and whether a long-run relationship between the two variables is possible.

### 3.4.2 Structural Break Tests

Next, in the spirit of Congregado et al. (2021), we proceed by testing for multiple unknown breakpoints in the West Virginia LFPR and the UR. As Congregado et al. (2021) points out, it is conceivable that the LFPR, the UR, and the relationship between the two, may change over a long period (47 years of data in our case). In the case of one or more structural breaks, the results for cointegrating relationships may differ across periods between the breaks. To check the stability of the series, we apply four structural break tests outlined in Bai and Perron (2003). In applying these tests, we allow for serial correlation in the errors by specifying a quadratic spectral kernel based heteroskedasticity and autocorrelation consistent (HAC) covariance matrix using prewhitened residuals. As suggested by Bai and Perron (2003), we determine the kernel bandwidth using the Andrews AR(1) method (Andrews, 1991; Andrews and Monahan, 1992). The first of these tests we employ is the Sup F Wald Test, which tests for $l$ breaks against the null hypothesis of no breaks. Secondly, we apply the double maximum tests (UDmax and WDmax), which test the null hypothesis of no breaks against the alternative of an unknown number of breaks. For the WDmax test, weights are applied to the individual tests so that the marginal p-values are equal across the number of breakpoints. For the UDmax test, the weights are distributed equally.

The third structural break test we use is a sequential procedure (Seq) to test the null hypothesis of $l$ breaks against the alternative of $l+1$ breaks. Lastly, we also consider two information criteria to select the number of breaks. We use the modified Schwarz criterion (LWZ), proposed by Liu et al. (1997), and the Bayesian Information Criterion (BIC), suggested by Yao (1988). In all of these tests, we use 15% trimming with a maximum number of breaks equal to four. We arrive at these parameters after inspecting the National Bureau of Economic Research (NBER) U.S. business cycle dates. On average, over our sample period, business cycles last about eight years in the U.S. Each cycle represents approximately 17% of the total observations we have in our data set (558). This would allow for approximately five regimes or periods surrounding four potential breakpoints. Since rounding up and setting trimming to 20% would restrict the maximum number of breakpoints to three, we set trimming equal to 15% and restrict the maximum number of breakpoints to four. In testing for breakpoints, we allow the slope and intercept to change.

### 3.4.3 Vector Auto-Regressive (VAR) Model Estimation and Cointegration Tests

After finding the structural breaks, we follow Congregado et al. (2021) by applying the approach for testing cointegration in Österholm (2010) to each regime. Österholm (2010) tests for cointegration between the LFPR and UR in Sweden using Johansen's methodology (Johansen, 1988, 1991). This first requires estimating a finite order Vector Auto-Regressive (VAR) Model. This unrestricted VAR can be written as:

$$y_t = \eta + \sum_{i=1}^{p} A_i y_{t-i} + \mu_t \tag{26}$$

where $y_t = [\text{lfpr}_t,\ \text{ur}_t]'$ is a vector of the non-stationary variables, the labor force participation rate ($\text{lfpr}_t$) and unemployment rate ($\text{ur}_t$) in the model. In our case, $\eta$ represents the constant in the model, $A_i$ represents a 2x2 matrix of parameters, and $\mu_t$ is a 2x1 vector of residuals. We use the Schwarz information criterion (SIC) to choose the lag length in this VAR model. We examine the cointegrating relationship between the West Virginia LFPR and UR in the context of the VAR using the Johansen (1988, 1991) trace and maximum eigenvalue tests. Using the notation of Österholm (2010) and Hjalmarsson and Österholm (2007), we denote and define the trace and maximum eigenvalue test statistics as $J_{trace}$ (Equation 27) and $J_{max}$ (Equation 28), respectively.

$$J_{trace} = -T \sum_{i=r+1}^{k} \ln(1 - \hat{\lambda}_i) \tag{27}$$

$$J_{max} = -T \ln(1 - \hat{\lambda}_{r+1}) \tag{28}$$

where we denote the sample size as T, and the $i$th largest canonical correlation as $\hat{\lambda}_i$ The null hypothesis tested by the trace test is that the number of cointegrating vectors is $r$. The alternative is that the number of cointegrating vectors is $k$. The maximum eigenvalue test has the same null hypothesis as the trace test, but the alternative hypothesis is that the number of cointegrating vectors is $r + 1$. In both cases, testing continues for increasing values of $r$ until the test results in a non-rejection. The corresponding value for the first non-rejection is then used as the estimator for $r$.

### 3.4.4 Vector Error Correction Model (VECM) Estimation

Given the existence of at least one cointegrating relationship, we proceed by estimating a Vector Error Correction (VECM) model which is a restricted (cointegrated) VAR Model. In this model, any cointegrating relationships are built into the specification. This allows for short-run adjustment dynamics as the endogenous variables are restricted to converge to the long-run equilibrium. Subtracting $y_{t-1}$ from both sides of Equation 26, we can rewrite the unrestricted VAR in its VECM representation:

$$\begin{aligned}
\Delta y_t &= \eta + A_1 y_{t-1} + A_2 y_{t-2} - y_{t-1} + \mu_t \\
&= \eta + (A_1 - I) y_{t-1} + A_2 y_{t-2} + \mu_t \\
&= \eta + (A_1 - I) y_{t-1} - (A_1 - I) y_{t-2} + (A_1 - I) y_{t-2} + A_2 y_{t-2} + \mu_t \\
&= \eta + (A_1 - I) \Delta y_{t-1} + (A_1 + A_2 - I) y_{t-2} + \mu_t
\end{aligned}$$

Therefore,

$$\Delta y_t = \eta + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-1} + \Pi y_{t-p} + \mu_t \tag{29}$$

where

$$\Pi = \sum_{i=1}^{p} A_i - I \quad \text{and} \quad \Gamma_i = \sum_{i=1}^{p-1} A_i - I$$

If the coefficient matrix $\Pi$ has rank equal to the number of cointegrating relationships, then $\Pi$ can be expressed as $\Pi = \boldsymbol{\alpha}\boldsymbol{\beta}'$, where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are both 2x1 vectors with rank equal to one. The adjustment parameter vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ contains the adjustment parameters that represent the speed of adjustment back to the long-run equilibrium in the model for the LFPR and UR respectively. The cointegrating vector, $\boldsymbol{\beta} = (1, -\beta_2)$ contains the cointegrating coefficients. The coefficient for the LFPR is normalized to 1. The system can be written as:

$$
\begin{aligned}
\Delta \text{lfpr}_t &= \Gamma(L)\Delta \text{lfpr}_{t-1} + \alpha_1 \left(\text{lfpr}_{t-1} - \beta_2 \text{ur}_{t-1}\right) + \mu_t^{\text{lfpr}} \\
\Delta \text{ur}_t &= \Gamma(L)\Delta \text{ur}_{t-1} + \alpha_2 \left(\text{lfpr}_{t-1} - \beta_2 \text{ur}_{t-1}\right) + \mu_t^{\text{ur}}
\end{aligned}
\tag{30}
$$

where $\Gamma(L)$ denotes the lag operator and $\left(\text{lfpr}_{t-1} - \beta_2 \text{ur}_{t-1}\right)$ is the long-run equilibrium.

## 3.5 Results

### 3.5.1 Unit Root Test Results

To determine whether a long-run relationship between the West Virginia LFPR and UR is possible we report the results of the unit root tests in Table 5. As indicated by the decisions columns of the table, we fail to reject the null hypothesis for both the ADF-GLS and PP tests. This suggests that the West Virginia LFPR and UR have unit roots. Additionally, we reject the null hypothesis for the KPSS test at the 1% level which suggests that the data are, again, not stationary (have unit roots). To test the robustness of these results, we re-apply the tests to first differences of the data and conclude that the raw LFPR and UR are non-stationary and have the same order of integration, I(1).

In addition, we apply the Fourier LM and ADF unit root tests as a preliminary test for possible structural breaks. The results are reported in Table 6. Critical values are taken from Enders and Lee (2012a). With trimming again set to 15% we find that the test statistics for the LFPR and the UR in both tests are smaller than the 10% level critical values. Therefore, we cannot reject the null hypothesis that the series have unit roots with an unknown number of structural breaks. The results of the unit root tests imply that structural breaks and a cointegrating relationship may exist in the data.

Table 5: Unit Root Tests on Individual Series

| | **ADF-GLS** | **Decision** | **KPSS** | **Decision** | **PP** | **Decision** |
|---|---|---|---|---|---|---|
| **LFPR** | -1.733 | Fail to Reject | 1.19*** | Reject | -1.62 | Fail to Reject |
| **ΔLFPR** | -6.06*** | Reject | 0.08 | Fail to Reject | -23.4*** | Reject |
| **UR** | -1.761 | Fail to Reject | 1.426*** | Reject | -1.799 | Fail to Reject |
| **ΔUR** | -24.98*** | Reject | 0.08 | Fail to Reject | -24.94*** | Reject |

Note: ADF-GLS refers to the test statistics from the augmented Dickey-Fuller test with GLS detrending. The optimal lag length was selected using the Schwarz information criterion. KPSS refers to the Kwiatkowski, Phillips, Schmidt and Shin test. PP refers to the Phillips-Perron test. LFPR and UR refer to the labor force participation rate and unemployment rate, respectively. $\Delta(\cdot)$ is the 1st-differenced data.

*** Indicates significance at the 1% level.

### 3.5.2 Structural Break Point Test Results

To examine the number and dates of possible structural breaks, we present the results of the Bai and Perron (2003) structural break tests in Table 7. We set trimming equal to 15% and the maximum number of breaks to four. Under these conditions, the Sup F Wald test, double maximum tests, and the information criteria select four breaks for the LFPR. The sequential procedure (Seq) selects three. For the UR, the Sup F Wald test and the double maximum test again select four breaks. However, for unemployment, the information criteria select three breaks and the Seq test selects one. Bai and Perron (2003) point out that in with serial correlation, the LWZ and BIC information criteria choose a value much higher than the truth. In this case, Bai and Perron (2003) suggests using the Seq test to choose the number of structural breaks. Following this recommendation, we choose the structural break dates corresponding to 1989:M08, 1998:M04, and 2010:M04. This suggest the following four regimes: Regime 1: 1976:M01 - 1989:M08, Regime 2: 1989:M09 - 1998:M04, Regime 3: 1998:M05 - 2010:M04, Regime 4: 2010:M05 - 2022:M06.

Table 6: **Fourier Unit Root Tests Allowing for Multiple Unknown Structural Breaks**

| | **Labor Force Participation Rate** | | | **Unemployment Rate** | | |
|---|---|---|---|---|---|---|
| **Trimming** | 15% | | | 15% | | |
| **Fourier (LM)** | -1.863 | | | -2.715 | | |
| Critical Values | 10% | 5% | 1% | 10% | 5% | 1% |
| | -2.83 | -3.13 | -3.74 | -2.83 | -3.13 | -3.74 |
| **Fourier (ADF)** | -2.479 | | | -3.061 | | |
| Critical Values | 10% | 5% | 1% | 10% | 5% | 1% |
| | -4.00 | -4.27 | -4.80 | -3.67 | -3.98 | -4.58 |

Note: Column 1 shows the type of test used followed by the critical values for test acquired from Enders and Lee (2012a). Column 2 shows the trimming setting and the test statistics for each test for the labor force participation rate data. Column 3 shows the trimming setting and the test statistics for the unemployment rate data. Fourier (LM) refers to the Fourier Lagrange Multiplier unit root test. Fourier (ADF) refers to the Fourier Augmented-Dickey-Fuller unit root test.

### 3.5.3 Cointegration Test Results

Having determined the structural break dates, we present the cointegration test results. Table 8 shows that for the full sample and Regimes 2, 3, and 4, the hypothesis that there is no cointegrating relationship ($r = 0$) between the West Virginia LFPR and UR is rejected at the 5% level or below. Also, the hypothesis that there is one cointegrating relationship is not rejected for these samples. This implies that there is one and only one cointegrating relationship between the LFPR and the UR for these periods. However, for Regime 1, the period between 1976:M01 and 1989:M08, we fail to reject the hypothesis that there is no cointegrating relationship between the LFPR and the UR in West Virginia. Therefore, we cannot reject the UIH in this case and conclude that the AWE and DWE cancel each other during this period.

Table 7: Tests for Multiple Structural Breaks in West Virginia Data

| **Variable** | **Labor Force Participation Rate** | **Unemployment Rate** |
|---|---|---|
| **Trimming** | 15% | 15% |
| Sup F(1) | 23.68* | 10.74* |
| Sup F(2) | 31.54* | 8.78* |
| Sup F(3) | 63.05* | 7.72* |
| Sup F(4) | 79.57* | 6.80* |
| UDmax | 79.57* | 10.73* |
| WDmax | 136.82* | 11.69* |
| Seq | 3 | 1 |
| BIC | 4 | 3 |
| LWZ | 4 | 3 |
| Number of Breaks Allowed | 4 | 4 |
| Break Dates | 1989:08<br>1998:04<br>2010:04 | 1994:09 |

Note: Column 1 shows the test used to determine the number and significance of structural breaks. Sup F(.) refers to the supreme F Wald test. UDmax and WDmax are the two double maximum test. Seq refers to the sequential procedure test. LWZ and BIC refer to the Schwarz and Bayesian information criteria, respectively. More information on the test can be found in Bai and Perron (1998). Column 2 and 3 show the test statistic, number of breaks, and break dates for the labor force participation rate and unemployment rate data, respectively. * indicates significance at the 5% level. The critical values are taken from Bai and Perron (1998)

### 3.5.4 VECM Results

While we find evidence for cointegrating relationships for the full sample, and Regimes 2, 3, and 4, Hjalmarsson and Österholm (2010) point out that the Johansen tests may have size distortions if LFPR and UR do not have exact unit roots. Following Österholm (2010), we test the cointegrating relationship restrictions ($\beta = (0,1)'$ and $\beta = (1,0)'$) in the context of the VECM to ensure that the cointegrating relationships we found in Table 8 are not due a single stationary variable. Rejecting these restrictions provides support for the cointegrating relationship results in Table 8. However, failing to reject either restriction implies that the relationship found in Table 8 is due to either a stationary LFPR variable (in the first restriction) or a stationary UR variable (in the second

Table 8: Johansen (1988, 1991) Cointegration Tests

| | $J_{trace}$ | $J_{max}$ |
|---|---|---|
| *Full Sample: 1976:1 - 2022:6* | | |
| $H_0 : r = 0$ | 19.02* | 15.35* |
| $H_0 : r = 1$ | 3.67 | 3.67 |
| *Regime 1: 1976:1 - 1989:8* | | |
| $H_0 : r = 0$ | 12.39 | 7.61 |
| $H_0 : r = 1$ | 4.78 | 4.78 |
| *Regime 2: 1989:9-1998:4* | | |
| $H_0 : r = 0$ | 28.51* | 26.31* |
| $H_0 : r = 1$ | 2.21 | 2.21 |
| *Regime 3: 1998:5 - 2010:4* | | |
| $H_0 : r = 0$ | 31.88* | 27.50* |
| $H_0 : r = 1$ | 4.38 | 4.38 |
| *Regime 4: 2010:5 - 2022:6* | | |
| $H_0 : r = 0$ | 15.38* | 15.34* |
| $H_0 : r = 1$ | 0.034 | 0.034 |

Note: Column 1 shows the hypothesis for the number of cointegrating relationships, $r$. Column 2 and 3 show the test statistics for the trace and maximum eigenvalue tests from Johansen (1988, 1991), respectively. We choose the lag length in the VAR based on the Schwarz information criterion (SIC).

* indicates significance at the 5% level

restriction). Table 9 displays the results of these likelihood ratio tests. For the full sample and Regime 3, both restrictions are rejected at 1% or better. This indicates that a long-run relationship between the West Virginia LFPR and UR exists and we can rule out that these results are due to a single stationary variable. However, in the second regime we fail to reject the restriction, $\beta = (0, 1)'$, and in the fourth regime we fail to reject the restriction, $\beta = (1, 0)'$. This implies that only the UR enters the cointegrating vector in the former case, and only the LFPR enters the cointegrating vector in the latter. This indicates that there is no long relationship between the West Virginia LFPR and UR and we cannot reject the UIH for these periods.

Additionally, we investigate the error-correction terms of the VECM model. We follow (Österholm, 2010; Emerson, 2011; Kakinaka and Miyamoto, 2012, and others) to test if West Virginia LFPR and

UR error-correct to each other. We test for weak exogeneity or the hypothesis that both variables adjust to deviations from the long-run relationship. We verify this assumption by imposing the restrictions, $\alpha = (\alpha_1, 0)'$ and $\alpha = (0, \alpha_2)'$. The results of these tests are reported below our $\beta$ restriction results for each regime in Table 9. We see that the first restriction cannot be rejected in Regime 2. This implies that the WV UR does not adjust to the long-run relationship or error correct, but LFPR does. For the full sample, we find the opposite. We fail to reject the second restriction, which implies that LFPR does not error correct to the long-run relationship, but the UR does. For Regimes 3 and 4, we reject both restrictions at or below the 1% level indicating that neither the UR nor LFPR adjusts to the long-run relationship.

Given the results in Table 9, we estimate our VECM for the sub-samples where at least one of the restrictions on the cointegrating relationship is rejected. The results for the cointegration vector, $\beta$ from Equation 30 are reported in Table 10. The results for the full sample and Regime 3 show a rejection of the UIH, indicating a long-run relationship between the West Virginia LFPR and UR. For these two periods, the positive coefficients indicate that the DWE prevails over the AWE in West Virginia. Lower labor force participation rates are associated with higher unemployment rates. Our results show that a 1% change in the unemployment rate induces about a 0.70% decrease in the LFPR for the full sample and a 0.65% decrease for the 1998 - 2010 sample. Interestingly, our full sample results differ from the findings on U.S. aggregate data. Emerson (2011) finds the AWE to prevail over the DWE. However, as many authors point out, West Virginia has a unique labor history connected to a dependence on natural resources (see Dorsey, 1991; Lewis, 1993; Boal and Pencavel, 1994, for example). The period 1998:M05 - 2010:M04, specifically, is characterized by stagnant and slow growth in the natural resource industry in West Virginia (O'Leary and Boettner, 2011). Consequently, West Virginia experienced some of the lowest coal mining employment and earning levels in the state's history (O'Leary and Boettner, 2011). Also, this regime includes the 2001 recession, the Great Recession of 2008 and 2009, and the retirement of the Post-War and Baby Boomer generations. Our DWE results for Regime 3 imply that net labor market exits such as aging of the population, and emigration or dropping out of the labor force due to poor labor market conditions impact long-run unemployment.

For the remaining sub-samples, we cannot reject the UIH, which implies that during these other periods, West Virginia seems to have a self-adjusting labor market. The low labor force participation rate does not translate to lower long-run unemployment. Regimes 1, 2, and 4 have characterizations that may explain this result. In the 1970's, due to a reduction in the supply of natural resources caused by the OPEC oil embargo, West Virginia experienced a coal boom. This increased the LFPR, earnings, and overall economic activity in the state (Juhn, 1992; Black et al., 2002; Van Zandweghe et al., 2017). Later, in the 1980s, even though West Virginia experienced a coal bust, the decline in coal was simultaneous with a labor market restructuring into other industries (Stevens, 1986; Black et al., 2002). Also, in the last decade (Regime 4), the natural gas industry has experienced significant growth, replacing coal as the primary industry in the state. Notably, throughout Regimes 1, 2, and 4, labor unions have significantly impacted benefits, pay, and other incentives in major industries in the state. Examples include the Bituminous coal strike of 1977-1978, the Pittson coal strike of 1989-1990, and the Public School Teacher's strikes in 1990 and 2018-2019. With these examples, it is apparent that the West Virginia labor market can self-adjust in the long run.

## 3.6 Conclusion

In this study, we examine the relationship between LFPR and UR in West Virginia to determine whether the unemployment invariance hypothesis (UIH), added worker effect (AWE), or the discouraged worker effect (DWE) prevails. We confirm that, against the background of structural breaks in the data, the conclusion is driven by the choice of our estimation period. Specifically, we observe that the DWE prevails over the full-sample (1976:M01 – 2022:M06), as well as the 1998:M05 – 2010:M04 vintage. However, we find no evidence of a cointegrating relationship between the LFPR and UR during the 1976:M01 – 1989:08, 1989:M09 – 1998:M04, and 2010:M05 – 2022:M06 sub-periods. This would indicate that during these periods, the AWE and DWEs offset each other, providing validity to the UIH.

We suggest that changes in labor market conditions over time explain these results, validating the incorporation of structural breaks into our analysis. For example, in West Virginia, we see labor market expansions and restructuring, increased earnings, growth in major industries, and increased work incentives through labor union efforts throughout the 1970s, 80s, 90s, and 2010s. Consequently, we find that the persistently low LFPR in West Virginia does not translate to lower unemployment rates during these times (UIH). However, in the 2000s, West Virginia experienced record low employment and earnings levels in the state's primary industry, coal mining. Additionally, multiple recessions and labor market exits, such as the aging of the population, took place at this time. Our results ascribe the DWE to this period in West Virginia.

In the main, these findings highlight the importance of accounting for potential structural breaks when conducting analyses and suggest a non-linear relationship between the LFPR and UR (Congregado et al., 2021). Moreover, we are able to add to the extant literature and fill a gap in analyzing the state of West Virginia, one of the U.S.'s most economically depressed states. Our findings also support that varying regional results may be due to labor market differences. This is buoyed by the fact that our results for West Virginia differ from other regions, including the aggregate U.S. (Emerson, 2011). Given that labor market shocks, recessionary periods, and major labor market restructuring within a region's labor market may change the relationship between the LFPR and the UR over time, it follows that different labor market institutions may as well.

Our findings have wider implications for future labor policy in West Virginia (through the lense of LFPR and UR). In particular, the COVID-19 pandemic has resulted in an uptick in the number of discouraged workers and voluntary unemployment across the U.S. If we accept this as the case in West Virginia as well, current labor policy aimed at increasing the labor force participation may inadvertently increase long-run unemployment in the state. To avoid perpetuating cycles of high unemployment and low LFPR, labor market policy design should include measures to mitigate the DWE and stimulate the LFPR. To that end, we contend that strategies aimed at decreasing unemployment and increasing the LFPR should be underpinned by a focus on job creation and encouraging participation of those who have left the labor force. Additionally, policymakers should focus on implementing policies and programs to train and/or re-train existing unemployed and not-in-the-labor-force residents. In addition to these, policies that support flexible work schedules, child care, and tax credits may be helpful to increase LFPR in West Virginia and mitigate the discouraged worker effects (O'Leary and Boettner, 2015)

### 3.6.1 Avenues for Future Work

Lastly, we recognize that our study has some limitations. Some of the previous studies, briefly discussed in Section 3.1, include sub-population data such as gender and age groups. If an interested researcher were able to overcome the lack of availability of high-frequency state-level time series data for gender or other demographics characteristics, we contend that a much richer set of policy prescriptions could be offered for disadvantaged/vulnerable groups in West Virginia.

In addition, West Virginia is the only state in the U.S., with every county designated as a part of the Appalachian Region. Beverly et al. (2022) note a unique relationship between West Virginia and the Appalachian Region. We suggest that expanding our study to include the other states in Appalachia could shed light on the relationship between LFPR and UR in the region and provide another interesting angle to the debate regarding the UIH. Since the Appalachian Region is defined at the county level, we suspect some creativity may be needed to investigate this question for the region. Overall, more research is needed to better understand the LFPR and UR relationship and help provide a future with the hope of economic prosperity in economically discouraged areas like West Virginia.

Table 9: Cointegration VAR Restriction Tests

| **Restriction** | **P-Value** | **Decision** |
|---|---|---|
| | *Full Sample:1976:1-2022:6* | |
| $\beta = (1, 0)'$ | 0.01 | Rejected |
| $\beta = (0, 1)'$ | 0.001 | Rejected |
| $\alpha = (\alpha_1, 0)'$ | 0.007 | Rejected |
| $\alpha = (0, \alpha_2)'$ | 0.98 | Fail to Reject |
| | *Regime 1:1976:1 – 1989:8* | |
| $\beta = (1, 0)'$ | | |
| $\beta = (0, 1)'$ | No Cointegrating Relationship | |
| $\alpha = (\alpha_1, 0)'$ | | |
| $\alpha = (0, \alpha_2)'$ | | |
| | *Regime 2:1989:9 – 1998:4* | |
| $\beta = (1, 0)'$ | 0.19 | Fail to Reject |
| $\beta = (0, 1)'$ | 0.000 | Rejected |
| $\alpha = (\alpha_1, 0)'$ | 0.419 | Fail to Reject |
| $\alpha = (0, \alpha_2)'$ | 0.000 | Rejected |
| | *Regime 3:1998:5 – 2010:4* | |
| $\beta = (1, 0)'$ | 0.000 | Rejected |
| $\beta = (0, 1)'$ | 0.001 | Rejected |
| $\alpha = (\alpha_1, 0)'$ | 0.000 | Rejected |
| $\alpha = (0, \alpha_2)'$ | 0.000 | Rejected |
| | *Regime 4:2010:5 – 2022:6* | |
| $\beta = (1, 0)'$ | 0.002 | Rejected |
| $\beta = (0, 1)'$ | 0.872 | Fail to Reject |
| $\alpha = (\alpha_1, 0)'$ | 0.008 | Rejected |
| $\alpha = (0, \alpha_2)'$ | 0.048 | Rejected |

Note: Column 1 lists the regime and restrictions tested. Column 2 reports p-values from the likelihood ratio tests for the restrictions listed. Column 3 shows the test result based on the p-values.

Table 10: Estimated Cointegration Vectors

| | Coefficient | Conclusion |
|---|---|---|
| | *Full Sample: 1976:1 – 2022:6* | |
| LFPR | 1 | |
| UR | 0.704 (0.159) | **DWE** |
| Constant | -59.743 | |
| | *Regime 1: 1976:1 – 1989:8* | |
| LFPR | | |
| UR | No Cointegrating Relationship | **UIH** |
| Constant | | |
| | *Regime 2: 1989:9 – 1998:4* | |
| LFPR | 1 | |
| UR | $-0.05$ (0.034) | **UIH** |
| Constant | -54.68 | |
| | *Regime 3: 1998:5 – 2010:4* | |
| LFPR | 1 | |
| UR | 0.645 (0.131) | **DWE** |
| Constant | -60.145 | |
| | *Regime 4: 2010:5 – 2022:6* | |
| LFPR | 1 | |
| UR | $-8.803$ (0.436) | **UIH** |
| Constant | 0.592 | |

Note: LFPR and UR, refer to the labor force participation rate and unemployment rate, respectively. The estimated coefficients are from the cointegration vector, $\beta$, in Equation 30. Standard errors are reported in parentheses.

# Bibliography


Aaronson, S., Cajner, T., Fallick, B., Galbis-Reig, F., Smith, C., and Wascher, W. (2014). Labor force participation: recent developments and future prospects. *Brookings Papers on Economic Activity*, 2014(2):197–275.

Aastveit, K. A., Bjørnland, H. C., and Thorsrud, L. A. (2015). What drives oil prices? emerging versus developed economies. *Journal of Applied Econometrics*, 30(7):1013–1028.

Andrews, D. W. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica: Journal of the Econometric Society*, pages 817–858.

Andrews, D. W. and Monahan, J. C. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica: Journal of the Econometric Society*, pages 953–966.

Appalachian Regional Commission (2020). ARC's History and Work in Appalachia. https://www.arc.gov/arcs-history-and-work-in-appalachia/. Accessed: 2020-11-01.

Azzoni, C. R. and Silveira-Neto, R. M. (2005). Decomposing regional growth: labor force participation rates, structural changes, and sectoral factor reallocation. *The Annals of Regional Science*, 39(2):221–239.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, pages 47–78.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of applied econometrics*, 18(1):1–22.

Behringer, B. and Friedell, G. H. (2006). Appalachia: where place matters in health. *Preventing chronic disease*, 3(4).

Bell, B. and Smith, J. M. (2002). On gross worker flows in the united kingdom: evidence from the labour force survey. *Bank of England Working Paper*.

Beverly, J., Stewart, S., and Neill, C. L. (October 22, 2022). The Dynamics of Labor Force Participation: Is All Quiet on the Appalachian Front? Available at SSRN: "https://ssrn.com/abstract=4275976".

Bhatt, V., Kishor, N. K., and Ma, J. (2017). The impact of emu on bond yield convergence: Evidence from a time-varying dynamic factor model. *Journal of Economic Dynamics and Control*, 82:206–222.

Bian, Z., Ma, J., Ni, J., and Stewart, S. (2020). Synchronization of regional growth dynamics in china. *China Economic Review*, 61:101226.

Billings, D. (1974). Culture and poverty in appalachia: A theoretical discussion and empirical analysis. *Social Forces*, 53(2):315–323.

Birecree, A. M. (1996). The importance and implications of women's participation in the 1989-90 pittston coal strike. *Journal of Economic Issues*, 30(1):187–210.

Black, D., Daniel, K., and Sanders, S. (2002). The impact of economic conditions on participation in disability programs: Evidence from the coal boom and bust. *American Economic Review*, 92(1):27–50.

Black, D., McKinnish, T., and Sanders, S. (2005). The economic impact of the coal boom and bust. *The Economic Journal*, 115(503):449–476.

Blanchard, O. J. (1986). The wage price spiral. *The Quarterly Journal of Economics*, 101(3):543–565.

Blanchard, O. J., Diamond, P., Hall, R. E., and Murphy, K. (1990). The cyclical behavior of the gross flows of us workers. *Brookings papers on economic activity*, 1990(2):85–155.

Boal, W. M. and Pencavel, J. (1994). The effects of labor unions on employment, wages, and days of operation: coal mining in west virginia. *The Quarterly Journal of Economics*, 109(1):267–298.

Bollinger, C., Ziliak, J. P., and Troske, K. R. (2011). Down from the mountain: Skill upgrading and wages in appalachia. *Journal of Labor Economics*, 29(4):819–857.

Bowen, W. G. and Finegan, T. A. (1966). Educational attainment and labor force participation. *The American Economic Review*, 56(1/2):567–582.

Bradley, D. H., Herzenberg, S., and Wial, H. (2001). *An assessment of labor force participation rates and underemployment in Appalachia*. Keystone Research Center.

Brainard, L. et al. (2017). Why persistent employment disparities matter for the economy's health: a speech at" disparities in the labor market: What are we missing?" a research conference sponsored by the board of governors of the federal reserve system, washington, dc, september 26, 2017. Technical report.

Bryant, J., Jacobsen, V., Bell, M., and Garrett, D. (2004). Labour force participation and gdp in new zealand. Technical report, New Zealand Treasury Working Paper.

Bullard, J. et al. (2014). The rise and fall of labor force participation in the united states. *Federal Reserve Bank of St. Louis Review, First Quarter*, 96(1):1–12.

Burda, M. and Wyplosz, C. (1994). Gross worker and job flows in europe. *European economic review*, 38(6):1287–1315.

Bureau of Economic Analysis (2021). GDP by State. https://www.bea.gov/data/gdp/gdp-state. Accessed: 2020-12-11.

Bureau of Labor Statistics (2021a). Concepts and Definitions. https://www.bls.gov/cps/definitions.htm. Accessed: 2021-04-17.

Bureau of Labor Statistics (2021b). Multifactor Productivity. https://www.bls.gov/mfp/. Accessed: 2021-04-17.

Bustelo, M., Flabbi, L., Piras, C., and Tejada, M. (2019). Female labor force participation, labor market dynamic, and growth. Technical report, IDB Working Paper Series.

Cai, F. and Lu, Y. (2013). Population change and resulting slowdown in potential gdp growth in china. *China & World Economy*, 21(2):1–14.

Cajner, T., Radler, T., Ratner, D., and Vidangos, I. (2017). Racial gaps in labor market outcomes in the last four decades and over the business cycle.

Chalmers, J. A. and Greenwood, M. J. (1985). The regional labor market adjustment process: Determinants of changes in rates of labor force participation, unemployment, and migration. *The annals of regional science*, 19(1):1–17.

Chib, S. and Greenberg, E. (1994). Bayes inference in regression models with arma (p, q) errors. *Journal of Econometrics*, 64(1-2):183–206.

Chinitz, B. (1961). Contrasts in agglomeration: New york and pittsburgh. *The American Economic Review*, 51(2):279–289.

Chung, H. T., Fallick, B., Nekarda, C. J., and Ratner, D. (2014). Assessing the change in labor market conditions.

Compton, J. and Pollak, R. A. (2014). Family proximity, childcare, and women's labor force attachment. *Journal of Urban Economics*, 79:72–90.

Congregado, E., Gałecka-Burdziak, E., Golpe, A. A., and Pater, R. (2021). Unemployment invariance hypothesis and structural breaks in poland. *The Journal of Economic Asymmetries*, 24:e00198.

Council of Economic Advisors (2014). The labor force participation rate since 2007: Causes and policy implications.

Daly, M. C., Hobijn, B., Kwok, J., et al. (2009). Labor supply responses to changes in wealth and credit. *FRBSF Economic Letter*.

Del Negro, M. and Otrok, C. (2008). Dynamic factor models with time-varying parameters: measuring changes in international business cycles. *FRB of New York Staff Report*, (326).

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.

Dorsey, S. (1991). The Strange Case of the Missing West Virginia Labor Force. *Growth and Change*, 22(3):49–65.

Elhorst, J. P. (2003). The mystery of regional unemployment differentials: Theoretical and empirical explanations. *Journal of economic surveys*, 17(5):709–748.

Elliott, G., Rothenberg, T. J., and Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, 64(4):813–836.

Elsby, M. W., Hobijn, B., and Şahin, A. (2013). Unemployment dynamics in the oecd. *Review of Economics and Statistics*, 95(2):530–548.

Elsby, M. W., Hobijn, B., Şahin, A., Valletta, R. G., Stevenson, B., and Langan, A. (2011). The labor market in the great recession—an update to september 2011 [with comment and discussion]. *Brookings Papers on Economic Activity*, pages 353–384.

Elsby, M. W., Michaels, R., and Solon, G. (2009). The ins and outs of cyclical unemployment. *American Economic Journal: Macroeconomics*, 1(1):84–110.

Emerson, J. (2011). Unemployment and labor force participation in the united states. *Economics Letters*, 111(3):203–206.

Enders, W. and Lee, J. (2012a). The flexible fourier form and dickey–fuller type unit root tests. *Economics Letters*, 117(1):196–199.

Enders, W. and Lee, J. (2012b). A unit root test using a fourier series to approximate smooth breaks. *Oxford bulletin of Economics and Statistics*, 74(4):574–599.

Epstein, B. (2018). Labor Force Participation Dynamics. https://ideas.repec.org/p/pra/mprapa/88776.html.

Erceg, C. J. and Levin, A. T. (2014). Labor force participation and monetary policy in the wake of the great recession. *Journal of Money, Credit and Banking*, 46(S2):3–49.

Fernández, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review*, 103(1):472–500.

Fernandez, R. and Fogli, A. (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American economic journal: Macroeconomics*, 1(1):146–77.

Fogli, A. and Veldkamp, L. (2011). Nature or nurture? learning and the geography of female labor force participation. *Econometrica*, 79(4):1103–1138.

Foster, K. A., Havenner, A. M., and Walburger, A. M. (1995). System theoretic time-series forecasts of weekly live cattle prices. *American Journal of Agricultural Economics*, 77(4):1012–1023.

Fujita, S. and Ramey, G. (2006). The cyclicality of job loss and hiring.

Gomes, P. (2012). Labour market flows: Facts from the united kingdom. *Labour Economics*, 19(2):165–175.

Grossman, D. A. and Levin, M. R. (1961). The Appalachian Region: A National Problem Area. *Land Economics*, 37(2):133.

Gustavsson, M. and Österholm, P. (2006). The informational value of unemployment statistics: a note on the time series properties of participation rates. *Economics Letters*, 92(3):428–433.

Hamrick, K. S. (1997). Rural labor markets often lead urban markets in recessions and expansions. *Rural America/Rural Development Perspectives*, 12(2221-2019-2639):11–17.

Hjalmarsson, E. and Österholm, P. (2007). Testing for cointegration using the johansen methodology when variables are near-integrated. *Available at SSRN 1007890*.

Hjalmarsson, E. and Österholm, P. (2010). Testing for cointegration using the johansen methodology when variables are near-integrated: size distortions and partial remedies. *Empirical Economics*, 39(1):51–76.

Hotchkiss, J. L. and Rios-Avila, F. (2013). Identifying factors behind the decline in the us labor force participation rate. *Business and Economic Research*, 3(1):257.

Howe, W. J. and Parks, W. I. (1989). Labor market completes sixth year of expansion in 1988. *Monthly Lab. Rev.*, 112:3.

Isserman, A. M. and Rephann, T. J. (1993). Geographical and Gender Differences in Labor Force Participation: Is there an Appalachian Effect? *Growth and Change*, 24(4):539–578.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, 12(2-3):231–254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica: journal of the Econometric Society*, pages 1551–1580.

Jones, T. (2020). Labor force participation rate mystery: Why have so many americans stopped working? https://www.investors.com/news/labor-force-participation-rate-low/. Accessed: 2020-11-04.

Juhn, C. (1992). Decline of male labor market participation: The role of declining market opportunities. *The Quarterly Journal of Economics*, 107(1):79–121.

Juhn, C. and Potter, S. (2006). Changes in labor force participation in the united states. *Journal of Economic Perspectives*, 20(3):27–46.

Kakinaka, M. and Miyamoto, H. (2012). Unemployment and labour force participation in japan. *Applied Economics Letters*, 19(11):1039–1043.

Karanassou, M. and Snower, D. J. (2004). Unemployment invariance. *German Economic Review*, 5(3):297–317.

Kilkenny, M. and Huffman, S. K. (2003). Rural/urban welfare program and labor force participation. *American Journal of Agricultural Economics*, 85(4):914–927.

Kim, C.-J., Nelson, C. R., et al. (1999). State-space models with regime switching: classical and gibbs-sampling approaches with applications. *MIT Press Books*, 1.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The review of economic studies*, 65(3):361–393.

Kose, M. A., Otrok, C., and Whiteman, C. H. (2003). International business cycles: World, region, and country-specific factors. *american economic review*, 93(4):1216–1239.

Kottis, A. P. (1990). Shifts over time and regional variation in women's labor force participation rates in a developing economy: The case of greece. *Journal of development economics*, 33(1):117–132.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1-3):159–178.

Layard, P. R. G., Layard, R., Nickell, S. J., and Jackman, R. (2005). *Unemployment: macroeconomic performance and the labour market*. Oxford University Press on Demand.

Lewis, R. L. (1993). Appalachian restructuring in historical perspective: coal, culture and social change in west virginia. *Urban Studies*, 30(2):299–308.

Lin, C.-Y. and Miyamoto, H. (2012). Gross worker flows and unemployment dynamics in japan. *Journal of the Japanese and International Economies*, 26(1):44–61.

Liu, J., Wu, S., and Zidek, J. V. (1997). On segmented multivariate regression. *Statistica Sinica*, pages 497–525.

Ma, J., Vivian, A., and Wohar, M. E. (2018). Global factors and equity market valuations: Do country characteristics matter? *International Journal of Finance & Economics*, 23(4):427–441.

Marston, S. T. (1985). Two views of the geographic distribution of unemployment. *The Quarterly Journal of Economics*, 100(1):57–79.

Maurer-Fazio, M., Hughes, J. W., and Zhang, D. (2005). Economic reform and changing patterns of labor force participation in urban and rural china.

Mehrotra, S. and Parida, J. K. (2017). Why is the labour force participation of women declining in india? *World Development*, 98:360–380.

Mryyan, N. (2014). Demographics, labor force participation, and unemployment in jordan. *The Jordanian labour market in the new millennium*, pages 39–63.

Mumtaz, H. and Surico, P. (2012). Evolving international inflation dynamics: world and country-specific factors. *Journal of the European Economic Association*, 10(4):716–734.

Muntaner, C. and Barnett, E. (2000). Depressive symptoms in rural west virginia: labor market and health services correlates. *Journal of health care for the poor and underserved*, 11(3):284–300.

National Land Cover Dataset (2000). Multi-resoultion land characteristics consortium (u.s.).

Nguyen Van, P. (2016). Examining the unemployment invariance hypothesis: The case of australia. *Australian Economic Review*, 49(1):54–58.

O'Leary, S. and Boettner, T. (2011). Booms and Busts, The Impact of West Virginia's Energy Economy. https://wvpolicy.org/wp-content/uploads/2018/5/BoomsBusts072111.pdf. Accessed: 2022-10-11.

O'Leary, S. and Boettner, T. (2015). State of Working West Virginia 2015. https://wvpolicy.org/state-of-working-west-virginia/. Accessed: 2022-8-30.

Österholm, P. (2010). Unemployment and labour-force participation in sweden. *Economics Letters*, 106(3):205–208.

Oţoiu, A. and Ţiţan, E. (2016). Does the unemployment invariance hypothesis hold for romania? *Applied Economics Letters*, 23(12):884–887.

Partridge, M. D., Betz, M. R., and Lobao, L. (2013). Natural resource curse and poverty in appalachian america. *American Journal of Agricultural Economics*, 95(2):449–456.

Petrongolo, B. and Pissarides, C. A. (2008). The ins and outs of european unemployment. *American economic review*, 98(2):256–62.

Phillips, P. C. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346.

Phimister, E., Vera-Toscano, E., and Weersink, A. (2002). Female participation and labor market attachment in rural canada. *American Journal of Agricultural Economics*, 84(1):210–221.

Raghupathi, W. and Raghupathi, V. (2018). An Empirical Study of Chronic Diseases in the United States: A Visual Analytics Approach to Public Health. *International Journal of Environmental Research and Public Health*, 15(3):431.

Raifu, I. A. and Adeboje, O. M. (2022). Labour force participation and unemployment rate: does discouraged worker effect hypothesis or unemployment invariance hypothesis hold in africa? *African Journal of Economic and Management Studies*.

Rickman, D. S. and Wang, H. (2017). Us regional population growth 2000–2010: Natural amenities or urban agglomeration? *Papers in Regional Science*, 96:S69–S90.

Rogers, C., Mencken, K., and Mencken, F. C. (1997). Female labor force participation in central appalachia: A descriptive analysis. *Journal of Appalachian studies*, 3(2):189–210.

Sakanishi, A. (2020). Spatial analysis of female labor force participation rates in japan. *International Journal of Economic Policy Studies*, 14(2):351–361.

Schmitt, E. R. (2009). The appalachian thread in the antipoverty politics of robert f. kennedy. *The Register of the Kentucky Historical Society*, 107(3):371–400.

Shoven, J. B. (2007). New age thinking: Alternative ways of measuring age, their relationship to labor force participation, goverment policies and gdp. Technical report, National Bureau of Economic Research.

Spletzer, J. R. (1997). Reexamining the added worker effect. *Economic inquiry*, 35(2):417–427.

Stephens, H. M. and Deskins, J. (2018). Economic Distress and Labor Market Participation. *American Journal of Agricultural Economics*, 100(5):1336–1356.

Stevens, W. K. (1986). Economy reviving in west virginia. *The New York Times*, Section A:18–18.

Stock, J. H. and Watson, M. (2011). *Dynamic factor models*. Oxford University Press.

Stock, J. H. and Watson, M. W. (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In *Handbook of macroeconomics*, volume 2, pages 415–525. Elsevier.

Tansel, A., Ozdemir, Z. A., and Aksoy, E. (2016a). Does the unemployment invariance hypothesis hold for canada?

Tansel, A., Ozdemir, Z. A., and Aksoy, E. (2016b). Unemployment and labour force participation in turkey. *Applied Economics Letters*, 23(3):184–187.

Taylor, J. B. (2016). Can we restart the recovery all over again? *American Economic Review*, 106(5):48–51.

Topel, R. H. (1986). Local labor markets. *Journal of Political economy*, 94(3, Part 2):S111–S143.

USDA (2013). United States Department of Agriculture Economic Research Service, Rural Urban Continuum Codes (RUCC). Available online: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/. Accessed: 2022-06-12.

Van Zandweghe, W. (2012). Interpreting the recent decline in labor force participation. *Economic Review-Federal Reserve Bank of Kansas City*, page 5.

Van Zandweghe, W. et al. (2017). The changing cyclicality of labor force participation. *Federal Reserve Bank of Kansas City Economic Review Q*, 3.

Veracierto, M. (2008). On the cyclical behavior of employment, unemployment and labor force participation. *Journal of Monetary Economics*, 55(6):1143–1157.

Voicu, A. (2001). *Labor force participation dynamics in the Romanian labor market: A Bayesian Markov chain Monte Carlo approach*. Rutgers The State University of New Jersey-New Brunswick.

Walburger, A. M. and Foster, K. A. (1998). Determination of focal pricing regions for us fed cattle. *American journal of agricultural economics*, 80(1):84–95.

Weingarden, A. E. et al. (2017). Labor market outcomes in metropolitan and non-metropolitan areas: Signs of growing disparities. Technical report, Board of Governors of the Federal Reserve System (US).

West, K. D. and Wong, K.-F. (2014). A factor model for co-movements of commodity prices. *Journal of International Money and Finance*, 42:289–309.

World Bank (2021). World bank, labor force participation rate. https://data.worldbank.org/indicator/SL.TLF.CACT.ZS?most_recent_value_desc=true. Accessed: 2021-02-04.

Yao, Y.-C. (1988). Estimating the number of change-points via schwarz'criterion. *Statistics & Probability Letters*, 6(3):181–189.