Nonlinearities in the Regional Phillips Curve with Labor Market Tightness

Giulia Gitti*

January 11, 2024

[Link to most recent version]

Abstract

This paper is the first to show the presence of nonlinearities in the regional U.S. New Keynesian Phillips curve with labor market tightness as a proxy for economic activity. Such nonlinearities contribute to explaining the unexpected and persistent post-COVID inflation surge and have important implications for monetary policy. The New Keynesian Phillips curve is a structural equation that describes inflation dynamics. It captures the concept that in demand-driven booms, workers ask for higher wages, leading firms to raise prices. Labor market tightness represents labor demand relative to labor supply and is a realistic approximation of labor costs. To guide my empirical exercise, I introduce wage rigidities and search-and-matching frictions in the labor market into a standard multi-sector, two-region New Keynesian model. The model delivers a piecewise log-linear regional Phillips curve, which becomes steeper when labor markets become sufficiently tight. I estimate the Phillips curve using panel variation in core inflation and a newly imputed measure of labor market tightness across U.S. metropolitan areas from December 2000 to April 2023. I instrument labor market tightness with a shift-share instrumental variable to take care of regional supply shocks. The regional Phillips curve has a positive slope that increases almost three times when labor market tightness exceeds the metropolitan area-specific average. This result suggests that if the monetary authority assumes that the Phillips curve is linear, it will risk underestimating inflationary pressures when labor markets run hot, allowing inflation to rise more than expected.

Keywords: Phillips curve, nonlinearities, inflation, labor market tightness, monetary policy

JEL code: E30, J63, J64

*Brown University, Economics Department. Contact: giulia_gitti@brown.edu. I would like to express my deepest appreciation to my committee - Gauti Eggertsson, Amy Handlan, and Pascal Michaillat - for their valuable advice and support. For their helpful comments and suggestions, I also wish to thank Peter Hull, Mary Amiti, Sebastian Heise, Keshav Dogra, as well as seminar participants from Brown University and the Federal Reserve Bank of New York. Finally, I thank the Brown University Merit Dissertation Fellowship and the AEA Summer Fellows Program for their financial support.

1 Introduction

Macroeconomists and policy-makers study inflation dynamics through the New Keynesian Phillips curve, a structural equation that relates inflation to measures of real economic activity, supply shocks, and inflation expectations. The relationship between inflation and real economic activity goes through the labor market. The Phillips curve captures the concept that in demand-driven booms, workers ask for higher wages, leading firms to raise prices. Therefore, tight labor market conditions are relevant indicators of inflationary pressures coming from raising labor costs.

The 20 years leading up to the COVID-19 pandemic, characterized by overall slack labor markets[1] and low and stable inflation, fostered a consensus that the U.S. Phillips curve was linear and flat (Hazell et al., 2022). In other words, fluctuations in economic activity would produce limited effects on inflation. The experience of the post-COVID period, however, challenged this consensus. In the aftermath of the pandemic, the U.S. labor market was the tightest since World War II (Michaillat and Saez, 2022). At the same time, the 12-month core CPI inflation rate in the United States began to rise, reaching a 40-year high at seven percent in September 2022 and remaining high well into the first half of 2023.

In this paper, I investigate whether inflationary pressures in tight labor markets might be stronger than in slack labor markets, potentially leading to a Phillips curve that is nonlinear in the state of the labor market. A major challenge for making this assessment, however, is that time series observations of tight labor markets are limited[2]. For this reason, I turn to panel data, which provides greater variation and more instances of tight labor markets. Figure 1 shows the relationship between the 12-month core CPI inflation rate and the logarithm of labor market tightness[3] across 21 U.S. Metropolitan Statistical Areas (MSAs) from December 2000 to April 2023. The scatter plot illustrates a positive correlation between MSA core inflation and labor market tightness, which becomes steeper around a tightness value of 1 – corresponding to the value of logarithm of tightness equal to 0.

This paper is the first to identify and estimate a nonlinear regional Phillips curve (NRPC from now onward) with labor market tightness as a proxy for real economic activity. To guide my empirical exercise, I introduce two novel features into an otherwise standard multi-sector New Keynesian model of

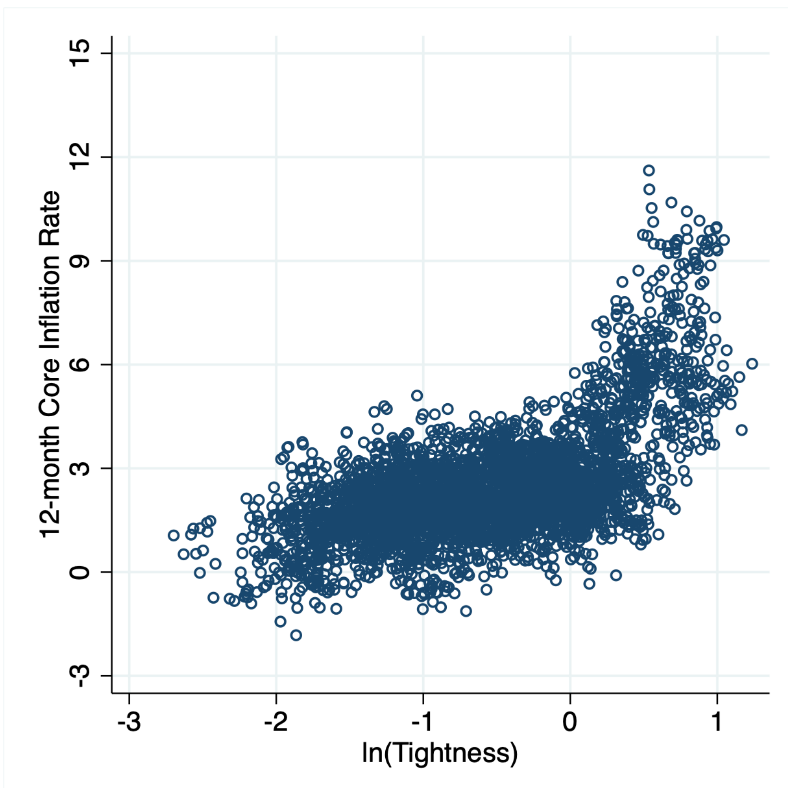
[1] Between 2000 and 2020, the 12-month headline CPI inflation averaged 2.2% in the United States. During the same period, the U.S. labor market was slack, except for a limited period between 2018 and 2019 (Michaillat and Saez, 2022).

[2] Combining data from Barnichon (2010) and Petrosky-Nadeau and Zhang (2021), Michaillat and Saez (2022) show that before 2018-2019 and 2021-2022, the U.S. labor market was inefficiently tight only during World War II, the Korean War, and the Vietnam War.

[3] Labor market tightness is defined as the ratio between vacancies posted by firms and the number of unemployed workers. It measures labor market conditions, taking into account measures of both labor demand (vacancies) and labor supply (unemployed workers looking for a job).

two regions in a monetary union. First, I incorporate search-and-matching frictions in the labor markets, which give rise to a relationship between inflation and tightness. Second, I introduce wage rigidities that generate a kink in this relationship, as in Figure 1. Thanks to these two elements, the model yields a piecewise log-linear regional Phillips curve with labor market tightness as a proxy for real economic activity. To perform the empirical exercise, I impute a novel series of vacancies at the MSA level and use it to measure regional labor market tightness. Thanks to this new variable, I estimate the NRPC derived in the model. Estimation relies on an identification strategy that combines MSA-level panel variation in core inflation and tightness with an instrumental variable approach. I instrument labor market tightness with a shift-share instrumental variable that proxies for sectoral productivity shocks to deal with unobservable regional supply shocks.

Figure 1: Correlation of Inflation and Log of Labor Market Tightness at the MSA Level, Dec00-Apr23



Notes. The figure shows the scatter plot between the 12-month core inflation rate and the logarithm of labor market tightness (vacancy-to-unemployment ratio) across 21 Metropolitan Statistical Areas (MSAs) in the United States from December 2000 to April 2023. Vacancies across metropolitan areas are imputed by combining state-level vacancies collected by BLS JOLTS with population weights from the 2000 and 2010 Census.

I find that the regional Phillips curve is nonlinear in labor market tightness. In particular, I document two novel sets of results. First, I find a positive and significant relationship between inflation and labor market tightness at the regional level from December 2000 to April 2023. Second, I estimate that the slope of the NRPC increases by almost three times when labor market tightness exceeds its MSA-specific average value. This result has important implications for monetary policy. If the central bank assumes that the Phillips curve is linear, it will underestimate inflationary pressures when labor markets become sufficiently tight, allowing inflation to rise more than expected. Moreover, the nonlinearity in the Phillips curve allows the central bank to decrease inflation without causing a significant recession when labor markets are tight, achieving the so-called "soft-landing" during a rate hike.

To guide my empirical exercise, I rely on a New Keynesian general equilibrium model augmented with four key features: two regions in a monetary union, a vertically-linked production structure, search-and-marching frictions, and wage rigidities. The first two components provide the foundation for my empirical strategy, based on panel variation and an instrumental variable approach. MSA-level panel data requires to derive a regional Phillips curve. The instrumental variable exploits productivity shocks in the intermediate-input industries that affect the pricing-setting decisions of final-goods firms.

Search-and-matching frictions in the labor markets introduce into a New Keynesian model the concept of unemployment formally and generate a relationship between inflation and labor market tightness. I model them in the spirit of the Diamond-Mortensen-Pissarides theoretical framework. Households in each region choose their labor force participation, but only a fraction of the labor force is employed. Employment agencies post firms' job vacancies and match them to unemployed workers searching for a job. Crucially for generating labor market frictions, posting a vacancy is costly.

Wage rigidities give rise to nonlinearities in the regional Phillips curve. Based on the evidence in Figure 1, I model a wage-setting mechanism that introduces a kink in the relationship between inflation and labor market tightness. Following Phillips (1958), I assume that during labor market shortages (in sufficiently tight labor markets), firms bid up wages to attract workers, and wages rise fast. In this case, I allow wages to be fully flexible and to be pinned down by the optimal behavior of employment agencies. In normal circumstances (in slack labor markets), workers are reluctant to accept a reduction in their nominal wage rate. In this case, wages fall slowly and constrain firms' labor demand. In sum, the prevailing wage rate depends on the level of labor market tightness in each region. As wages enter the marginal cost structure of price-setting final-goods firms, their nonlinearity generates a kink in the regional Phillips curve.

To bring the structural NRPC curve to the data, I impute a novel measure of vacancies at the metropolitan area level. Measuring labor market tightness at this level of disaggregation is challenging because a public and representative source of MSA-level vacancy data does not exist. I overcome this problem by imputing the number of vacancies at the city level from state-level vacancies recorded by the BLS's Jop Openings and Labor Turnover Survey (JOLTS), using population weights from the 2000 and 2010 Census. The main advantage of this new measure is that it is constructed using vacancies posted by a representative sample of firms. The instrumental variable takes care of the measurement error I introduce in labor market tightness, my main explanatory variable.

To estimate the NRPC, I combine panel variation in core inflation and labor market tightness at the MSA level with an instrumental variable approach. Figure 1 reports a simple correlation between core inflation and tightness, which can be driven by aggregate or regional confounders. Examples of these confounders are aggregate or regional supply shocks, such as a rise in the international price of oil or migration inflows into a regional labor market. A negative supply shock decreases labor market tightness and increases inflation, inducing a downward bias in the relationship between tightness and inflation. Using panel data enables me to incorporate regional and time fixed effects into the empirical model. Time fixed effects capture aggregate confounders, such as aggregate supply shocks, but also long-run inflation expectations that depend on the monetary regime in place (Hazell et al., 2022), and endogenous national monetary and fiscal policies (Fitzgerald and Nicolini 2014, McLeay and Tenreyro 2020).

I employ an instrumental variable approach to take care of regional confounders. Guided by the model, I construct a shift-share instrument that proxies productivity shocks in the tradable intermediate-input industries (Bartik, 1991). These shocks affect labor demand relatively more in those metropolitan areas specialized in such industries. As the labor market is common across sectors within metropolitan areas, a larger change in labor demand will lead to a larger change in wages and, therefore, a larger change in marginal costs of local final-goods firms in those specialized cities, which ultimately will be reflected in higher final-goods prices. For instance, a positive national productivity shock in the manufacturing sector leads to larger cost increases for final-goods firms located in manufacturing-intensive cities like Detroit. The identifying assumption requires that such cost increases are no larger on average for final-goods firms in Detroit than New York[4].

I conduct several robustness checks. First, instead of a single kink, I allow for more flexible nonlinearities and show that my results are unchanged. Second, the estimated nonlinearities are robust to using the most popular measure of economic activity, the unemployment rate, as the main independent variable

[4]Intermediate-input industries' productivity shocks also have a direct impact on intermediate-input prices, affecting local final-goods firms' marginal costs through this channel as well. As intermediate-input prices are observable, I control for this second channel, preventing the invalidation of my instrumental variable strategy.

instead of labor market tightness. In addition, I provide suggestive evidence that the regional Phillips curve was nonlinear in the pre-COVID period as well, indicating that the nonlinearities are not just driven by the post-COVID period[5]. Overall, I conclude that the finding of nonlinearity in the regional Phillips curve with labor market tightness is robust.

Since the seminal work of Phillips (1958), the Phillips curve has been extensively studied theoretically and empirically. In this vast literature, my paper relates to three strands. First, I contribute to an emerging theoretical literature that seeks to model nonlinearities in the aggregate Phillips curve to explain the post-COVID surge in inflation. In particular, guided by the correlation between MSA-level core inflation and labor market tightness in Figure 1, my work builds on Benigno and Eggertsson (2023). With respect to the literature that incorporates search-and-matching frictions and wage rigidities into the New Keynesian theoretical framework (Blanchard and Galí 2010, Christoffel and Linzert 2005, Trigari 2006, Krause and Lubik 2007, Faia 2008, Michaillat 2014), Benigno and Eggertsson (2023) propose a novel wage-setting mechanism based on the theoretical argument of Phillips (1958) that can generate a kink in the aggregate Phillips curve. The contribution of my paper is to embed search-and-matching frictions and Benigno and Eggertsson (2023)'s wage-setting mechanism into a New Keynesian model of two regions in a monetary union. In addition, I model a vertical production structure. This vertical production structure and the regional nature of my model are the key ingredients that allow me to identify the slope of the Phillips curve by combining fixed effects with an instrumental variable approach. Other relevant contributions in this literature are Harding et al. (2022) and Schmitt-Grohé and Uribe (2022). Harding et al. (2022) generate a nonlinear aggregate Phillips curve through a quasi-kinked demand schedule for goods. Schmitt-Grohé and Uribe (2022) employ heterogeneous downward nominal wage rigidity for individual labor varieties to derive a nonlinear aggregate wage Phillips curve.

Second, my paper adds to the growing literature about the cross-sectional identification of the slope of the Phillips curve. Papers such as Mavroeidis et al. (2014), Fitzgerald and Nicolini (2014), McLeay and Tenreyro (2020), Hazell et al. (2022), and Cerrato and Gitti (2022) show how panel variation can help overcoming some of the identification challenges that affect the estimation of the slope of the Phillips curve at the aggregate level. Hazell et al. (2022) and Cerrato and Gitti (2022) are the papers closer to mine. With respect to them, I introduce theoretical and empirical innovations. On the theoretical side, I consider search-and-matching frictions and wage rigidities to derive the NRPC with labor market tightness as an explanatory variable. On the empirical side, I impute a novel measure of vacancies at

[5]As I do not have enough statistical power to estimate the regional Phillips curve with labor market tightness in the pre-COVID period, I run this robustness check using the unemployment rate to proxy for economic activity.

the MSA level to proxy for economic activity with labor market tightness, measured by the ratio of vacancies to unemployment, rather than with the unemployment rate alone. Moreover, I estimate a regional Phillips curve with nonlinearities.

Third, few papers have used regional data to empirically investigate the presence of nonlinearities in the regional Phillips curve before the COVID-19 pandemic. Kiley (2015), Murphy (2017), Babb and Detmeister (2017), Leduc and Wilson (2017), and Hooper et al. (2020) all find evidence of the existence of the Phillips curve at the regional level between 1990 and 2019 and document the presence of nonlinearities. I contribute to this strand of literature by employing an instrumental variable approach to deal with biases coming from regional confounders, which cannot be controlled for by a simple two-way fixed effects model as all the other papers do. Moreover, I am the first to estimate the slope of the regional Phillips curve and to test for nonlinearities using labor market tightness to proxy for economic activity.

The remainder of this paper proceeds as follows. Section 2 describes the model and the derivation of the nonlinear regional Phillips curve. Section 3 details data sources and the construction of the novel measure of MSA-level vacancies. Section 4 discusses the empirical strategy. Section 5 presents the empirical results and their policy implications. Section 6 includes the robustness checks, and Section 7 concludes.

2 Model

I propose a New Keynesian model of two regions in a monetary union, featuring a vertical production structure, search-and-matching frictions in regional labor markets, and wage rigidities. Within the model, I derive an NRPC that I can bring to the data. The model also guides the empirical strategy I use to estimate the NRPC. The regional nature of the model and the vertical production structure are key to developing the empirical strategy. Such empirical strategy is based on two ingredients. First, panel variation across US metropolitan areas requires to estimate a regional Phillips curve. Second, the instrumental variable approach exploits the effects of intermediate-input sectors' labor demand shocks on the pricing-setting decisions of final-goods firms that use the intermediate inputs in their production processes. Search-and-matching frictions in the labor markets formally introduce unemployment into the New Keynesian framework. This feature generates a relation between inflation and labor market tightness. Wage rigidities give rise to the nonlinearity in the Phillips curve.

2.1 Model Setup

The model comprises two regions, i and j, belonging to the same monetary union as in the standard regional New Keynesian framework. The monetary authority sets the common interest rate following a Taylor rule, described in Appendix B. The two regions share the same preferences, market structure, and firm behavior. There is a continuum of representative households of measure ζ in region i and $(1-\zeta)$ in region j. The production side of the economy is composed of three vertically-linked sectors: an international commodity sector, a tradable, perfectly competitive, intermediate-input sector, and a non-tradable, monopolistically competitive, final-goods sector. Labor is immobile across regions and perfectly mobile across sectors within regions.

Differently from the standard regional New Keynesian framework, the labor market is not competitive. I add two frictions: search-and-matching frictions and wage rigidities. The wage-setting mechanism is motivated by the argument developed by Phillips (1958). According to Phillips, wages respond asymmetrically to the state of the labor market: they rise rapidly in tight labor markets and move slowly in slack labor markets. To capture Phillips' idea, I assume that wages are equal to the maximum between the prevailing wage rate and the flexible wage rate. The flexible wage rate is the one that clears the market in the absence of any constraint. To pin it down, I introduce a simple model of employment agencies.

Employment agencies oversee the match-and-searching process between workers and firms in each region. They post firms' vacancies subject to a cost and charge a fee to the workers they match to the vacancies. Employment agencies choose how many vacancies to post in order to maximize real profits. In the optimum, they equate the marginal benefit of posting a vacancy to its marginal cost. The flexible wage rate is pinned down by the problem of the employment agencies. In the following paragraphs, I describe the economy of region i.

2.1.1 Households

The representative household in region i is indexed by h. In each period t, household h chooses how much to consume and how many household members work along the extensive margin in order to maximize the utility flow given by Greenwood–Hercowitz–Huffman (GHH) preferences (Greenwood et al., 1988), defined as

$$u(C_{it}^h, F_{it}^h, \chi_{it}) = \frac{1}{1-\sigma}\left(C_{it}^h - \chi_{it}\int_0^{F_{it}^h} f^{\omega}\, df\right)^{1-\sigma}, \tag{1}$$

where C^h_{it} is total consumption of household h, and F^h_{it} denotes the number of members of household h who decide to participate in the labor market. Each household member is indexed by f and has fixed disutility f^ω from working, as in Galí (2011), with $\omega > 0$. χ_{it} is an exogenous variable governing the intensity of disutility of labor and $\sigma > 0$.

Household members are ordered by their disutility from working, capturing the notion that it may be more costly to have old members in the labor force, rather than young adults. Integrating the cost of labor force participation yields

$$\int_0^{F^h_{it}} f^\omega \, df = \frac{\left(F^h_{it}\right)^{1+\omega}}{1+\omega}. \tag{2}$$

Households decide labor force participation, but not all the labor force is employed due to the presence of search-and-matching frictions in the labor market. The labor force supplied by household h is defined as

$$F^h_{it} = N^h_{it} + U^h_{it},$$

where N^h_{it} denotes employed workers and U^h_{it} represents unemployed workers at the end of period t, after job search and matching took place.

The search-and-matching process is carried out by employment agencies and works as follows. At the beginning of each period t, a fraction $(1-s)$ of the labor force of household h is employed. The remaining fraction, sF^h_{it}, searches for a job. Their ability to enter employment is determined by the employment agencies through the following regional matching function, which determines total employment matches in region i

$$M_{it} = m_{it} U^\eta_{it} V^{1-\eta}_{it}, \tag{3}$$

where $m_{it} > 0$ represents matching efficiency, V_{it} denotes total vacancies posted by employment agencies in region i, U_{it} is total unemployment in region i, and $\eta \in [0, 1]$. Households take V_{it} and U_{it} as given, as they are determined at the regional level. Note that, if $s = 0$, we are back to the standard New Keynesian model with perfectly flexible labor markets. M_{it} are start-of-period unemployed workers screened by employment agencies as suitable for working.

Labor market tightness in region i is defined as $\theta_{it} \equiv \frac{V_{it}}{U_{it}}$, that is, the ratio between vacancies posted by employment agencies and unemployed workers searching for a job. The probability of a job seeker finding a job is $f(\theta_{it}) = \frac{M_{it}}{sF_{it}} = u_{it}\frac{m_{it}\theta^{1-\eta}_{it}}{s}$, where $u_{it} \equiv \frac{U_{it}}{F_{it}}$ is the unemployment rate. Households take as given the probability of a job seeker finding a job, as it depends on regional variables. The number of successful employment matches of the members of household h who were unemployed at the beginning of period t is

$$H_{it}^h = M_{it}^h = f(\theta_{it}) s F_{it}^h = u_{it} m_{it} \theta_{it}^{1-\eta} F_{it}^h, \tag{4}$$

Therefore, the number of members of household h who are employed at the end of period t is:

$$\begin{aligned} N_{it}^h &= (1-s) F_{it}^h + H_{it}^h \\ &= \left(1 - s + m_{it} u_{it} \theta_{it}^{1-\eta}\right) F_{it}^h. \end{aligned} \tag{5}$$

Members of households looking for a job pay to the employment agencies a fraction γ_{it}^b of their income. The representative household is then subject to the following budget constraint

$$P_{it} C_{it}^h + B_{it}^h \leq (1 + i_{t-1}) B_{it-1}^h + \left[(1-s) + (1-\gamma_{it}^b) m_{it} u_{it} \theta_{it}^{1-\eta}\right] F_{it}^h W_{it} + \int_0^1 \Pi_{it}^F(z)\, dz + \int_0^1 \Pi_{it}^E(l)\, dl, \tag{6}$$

where P_{it} is the price index associated with the consumption basket C_{it}^h, B_{it}^h is the quantity of risk-free nominal bond held in region i at time t, paying a nominal national interest rate i_t in period $t+1$, W_{it} denotes the nominal wage rate, $\Pi_{it}^F(z)$ are the profits of the firm producing variety z, and $\Pi_{it}^E(l)$ are the profits of the employment agency l. There is a complete set of financial markets across the two regions.

Household h chooses C_{it}^h, F_{it}^h, and B_{it}^h to maximise utility (1), subject to the budget constraint (6), and given the total cost of labor force participation (2). The representative household take as given all variables not indexed by i. As households behave all the same in equilibrium, I suppress the superscript h going forward.

Households trade off current consumption, C_{it} and current labor force participation, F_{it}. The optimal labor force participation takes the following form:

$$F_{it} = \left[\frac{(1-s) + (1-\gamma_{it}^b) m_{it} u_{it} \theta_{it}^{1-\eta}}{\chi_{it}} w_{it}\right]^{\frac{1}{\omega}}, \tag{7}$$

where $w_{it} \equiv \frac{W_{it}}{P_{it}}$ is the real wage rate. As already stated, I assume that households have GHH preferences. This means that the amount of work the households choose affects the amount of utility they receive from consumption. I make this assumption for tractability, following Hazell et al. (2022). The implication is that income effects are not at play in the optimal choice of labor force participation.

Households optimally trade off consumption in the current and in the next periods, as captured by the following Euler equation:

$$\left(C_{it} - \chi_{it}\frac{F_{it}^{1+\omega}}{1+\omega}\right)^{-\frac{1}{\sigma}} = \beta(1+i_t)E_t\left[\left(C_{it+1} - \chi_{it+1}\frac{F_{it+1}^{1+\omega}}{1+\omega}\right)^{-\frac{1}{\sigma}}\frac{P_{it}}{P_{it+1}}\right]. \tag{8}$$

Furthermore, household optimization implies that a standard transversality condition must hold.

I assume that households have constant elasticity of substitution (CES) preferences over varieties, leading to the following final consumption good aggregator:

$$C_{it} = \left[\int_0^1 C_{it}(z)^{\frac{\epsilon-1}{\epsilon}}\,dz\right]^{\frac{\epsilon}{\epsilon-1}}, \tag{9}$$

where $C_{it}(z)$ denotes consumption of variety z in region i. The parameter $\epsilon > 1$ denotes the elasticity of substitution between different varieties.

Households choose how much to purchase of each variety, $C_{it}(z)$, in order obtain the desired level of consumption C_{it} at a minimal expense. The minimization problem implies the following demand curve for variety z:

$$C_{it}(z) = C_{it}\left(\frac{P_{it}(z)}{P_{it}}\right)^{-\epsilon}, \tag{10}$$

and the following price index:

$$P_{it} = \left[\int_0^1 P_{it}(z)^{1-\epsilon}dz\right]^{\frac{1}{1-\epsilon}}, \tag{11}$$

where $P_{it}(z)$ is the price of variety z.

2.1.2 Firms

The production side comprises a vertical supply chain featuring an international commodity market, a national, perfectly competitive intermediate-input sector, and local, monopolistically competitive final-goods markets. At the first level of the supply chain, commodities are supplied by an international market, according to the following production process:

$$P_t^o = c_t^o O_t, \tag{12}$$

where P_t^o is the international price of commodities, c_t^o is an exogenous marginal cost shock, and O_t is the quantity of commodity produced.

The intermediate-input sector is tradable and is characterized by perfect competition. Hence, the price of intermediate input, P_t^x, is common across the two regions. The representative intermediate-input firm in region i uses commodity O_{it} and labor N_{it}^x to produce a homogeneous good, X_{it}, according to the following production function

$$X_{it} = A_{it}^x N_{it}^{x\rho} O_{it}^{1-\rho}, \tag{13}$$

where A_{it}^x denotes local exogenous technology of the intermediate-input sector and $\rho \in (0,1)$. In every period, the representative firm maximizes its value

$$P_t^x X_{it} - W_{it} N_{it}^x - P_t^o O_{it}, \tag{14}$$

given its production technology.

The final-goods sector is non-tradable and is characterized by monopolistic competition. In region i, there is a continuum of final-goods firms of measure 1 indexed by z. Each firm specializes in the production of a differentiated good consumed locally. The production function is characterized by constant returns to scale

$$Y_{it}(z) = A_{it}^y X_{it}(z)^{1-\phi} N_{it}^y(z)^{\phi}, \tag{15}$$

where A_{it}^y denotes local productivity of the final-goods sector, $X_{it}(z)$ and $N_{it}^y(z)$ denote, respectively, the quantity of intermediate good and labor used by firm z, and $\phi \in (0,1)$. Final-goods firm z maximizes the expected discounted value of profits

$$E_t \sum_{k=0}^{\infty} Q_{it,t+k}[P_{it+k}(z)Y_{it+k}(z) - W_{it+k}N_{it+k}^y(z) - P_{t+k}^x X_{it+k}] \tag{16}$$

subject to the production technology and

$$Y_{it}(z) = Y_{it}\left(\frac{P_{it}(z)}{P_{it}}\right)^{-\epsilon},$$

which denotes the demand for its product. $Q_{it,t+k}$ is the stochastic discount factor between period t and $t+k$ and $P_{it}(z)$ is the price set by firm z for its product. Firm z can set its price freely with probability $(1-\alpha)$ as in Calvo (1983). With probability α, the firm must keep its price unchanged.

2.1.3 Wage Determination

The wage-setting mechanism is motivated by the observation made by Phillips (1958) that the relationship between nominal wage growth and labor market tightness is nonlinear. Phillips (1958) argues that workers are unwilling to take jobs paying below the "prevailing wage rate" even in periods of weak labor demand and high unemployment. On the contrary, workers are perfectly happy to accept jobs paying more than the "prevailing wage rate". Therefore, firms quickly bid up wages to attract workers in periods of sufficiently strong labor demand and low unemployment.

To capture the idea of Phillips (1958), I assume that the wage rate in region i at time t is equal to the maximum between the prevailing wage rate W_{it}^{norm} and the flexible wage rate W_{it}^{flex}, where the flexible wage rate is the one that clears the market in the absence of any constraint. That is:

$$W_{it} = \max\{W_{it}^{norm}, W_{it}^{flex}\}. \tag{17}$$

Such wage-setting mechanism allows for an asymmetric response of wages to the state of the labor market. Consider the case in which labor demand is weak and unemployment is high. In such a case, the prevailing wage rate is greater than the flexible wage rate. The max operator ensures then that $W_{it} = W_{it}^{norm}$. On the other hand, if labor demand is sufficiently strong and unemployment is low, as firms compete with each other to attract workers, the flexible wage rate is higher than the prevailing wage rate. Hence, $W_{it} = W_{it}^{flex}$. In sum, the wage rate in region i rises rapidly in tight labor markets, while it decline slowly in slack labor markets. Equation 17 can be rewritten in real terms:

$$w_{it} = \max\{w_{it}^{norm}, w_{it}^{flex}\}. \tag{18}$$

In the search-and-matching literature the determination of wages is in general not pinned down, since each worker-firm match generates a surplus. How the surplus is divided between the worker and the firm can be done in different ways, the most common assuming Nash bargaining between the employer and the employee. Search-and-matching models incorporating price rigidities typically assume that the real wage is exogenous. To incorporate Phillips' idea of asymmetric wages' response to the state of the labor market, I follow Benigno and Eggertsson (2023). They propose a simple model of employment agencies that oversee the search-and-matching process in the labor market. The optimization problem of the employment agencies will provide the foundation for the flexible wage rate. Once I derive the flexible wage rate, then I show how the prevailing wage rate depends on the flexible wage rate.

2.1.4 Employment Agencies

Employment agencies carry out the process of search and matching in the labor market. There is a continuum of measure one of employment agencies in region i. Each employment agency is indexed by l. Employment agencies match workers with intermediate-input and final-goods firms. They carry out two actions. First, as they have access to the matching technology, they screen workers suitable for employment. Second, they post firms' vacancies. Since each agency is small, they take as given the wage rate and the rate of matches per vacancy posted. The number of matches per vacancy posted is

$$q(\theta_{it}) = \frac{M_{it}}{V_{it}} = \frac{m_{it} U_{it}^{\eta} V_{it}^{1-\eta}}{V_{it}} = m_{it}\theta^{-\eta} \tag{19}$$

The problem of employment agency l is defined as follows. Agency l charges a fee proportional to the real salary of workers screened for employment $\gamma_{it}^{b} w_{it} M_{it}$, and pays a real cost to post a vacancy that amounts to $\gamma_{it}^{c} V_{it}$. Because of such cost, I assume that employment agencies never post a vacancy that cannot be filled by firms. The number of matches agency l generates is given by $q(\theta_{it})V_{it}^{l} = m_{it}\theta^{-\eta} V_{it}^{l}$. By choosing the number of vacancies V_{it}^{l}, the employment agency maximizes real profits

$$Z_{it}^{E}(l) = \gamma_{it}^{b} w_{it} m_{it} \theta^{-\eta} V_{it}^{l} - \gamma_{it}^{c} V_{it}^{l} \tag{20}$$

In the optimum, the employment agency equates the marginal benefit of posting a vacancy to its marginal cost:

$$\underbrace{\gamma_{it}^{b} w_{it}^{flex} m_{it} \theta_{it}^{-\eta}}_{\text{marginal benefit}} = \underbrace{\gamma_{it}^{c}}_{\text{marginal cost}} . \tag{21}$$

As long as the marginal benefit is greater than the marginal cost, the agency will post vacancies. As a consequence, the flexible wage rate will decrease and tightness will increase, lowering the number of matches for each vacancy posted until the equilibrium is reached. In general equilibrium, the flexible wage rate and labor market tightness adjust so that such condition is satisfied. Rearranging equation 21, I provide the expression for the flexible wage rate:

$$w_{it}^{flex} = \frac{1}{m_{it}} \frac{\gamma_{it}^{c}}{\gamma_{it}^{b}} \theta_{it}^{\eta}. \tag{22}$$

Consider now the case in which the prevailing wage rate is higher than the flexible wage rate, that is $w_{it}^{norm} > w_{it}^{flex}$. In this case, $w_{it} = w_{it}^{norm}$ and firms hire less labor than they would have, had the wage rate been flexible. As employment agencies do not post vacancies that will not be filled by firms, they post only the number of vacancies that satisfy firms' constrained labor demand. The optimal condition indicates that, in this case, the marginal value of posting an additional vacancy for the agency remains positive.

I assume that the prevailing wage rate evolves as follows:

$$w_{it}^{norm} = (\bar{w}_i)^{\lambda}(w_{it}^{flex})^{1-\lambda} \tag{23}$$

where $\bar{w}_i$ denotes the steady state level of the real wage in region i and $\lambda \in [0, 1]$. When labor markets are slack, λ determines how quickly the prevailing wage rate adjusts to its flexible rate. The flexible wage rate is an anchor towards which the prevailing wage rate is pulled by a factor of $(1 - \lambda)$. At the extremes, if $\lambda = 0$, the prevailing wage rate is completely flexible. If $\lambda = 1$, the prevailing wage rate is fully rigid and equal to its steady state value. I assume such formulation for the prevailing wage rate to preserve the forward-looking nature of the Phillips curve and for computational straightforwardness. Richer forms of wage rigidities can be considered, but the intuition does not change.

In sum, I can write the behavior of the wage rate in region i at time t as

$$w_{it} = \begin{cases} w_{it}^{flex} & \theta_{it} > \theta_{it}^* \\ (\bar{w}_i)^{\lambda}(w_{it}^{flex})^{1-\lambda} & \theta_{it} \leq \theta_{it}^* \end{cases}, \tag{24}$$

When $\theta_{it} > \theta_{it}^*$, the labor market is sufficiently tight so that firms need to compete among each other to attract workers. In this case, wages are flexible and determined in equilibrium by the optimal behavior of the employment agencies. When $\theta_{it} \leq \theta_{it}^*$, the labor market is slack and the prevailing wage rate is higher than the flexible wage rate. As workers are unwilling to accept wages below the prevailing rate, wages move only gradually towards the flexible wage rate, at a speed that depends on the value of λ.

To close the model, I determine the value of θ_{it}^*. At θ_{it}^*, the prevailing wage rate is equal to the flexible wage rate, yielding:

$$\theta_{it}^* = \left(\frac{1}{m_{it}} \frac{\gamma_{it}^c}{\gamma_{it}^b} \bar{w}_i \right)^{\frac{1}{\eta}}. \tag{25}$$

This formula suggests that θ_{it}^* is region-specific and varies along time. In the empirical exercise I will take a pragmatic approach and approximate it with the region-specific average of labor market tightness.

2.2 Regional Nonlinear Phillips Curve

An equilibrium in this economy is an allocation consistent with optimization choices of households, firms, and employment agencies, the interest rate rule, and market clearing conditions. The definition of the equilibrium can be found in Appendix B.

Log-linearizing the model around a zero-inflation steady state and combining optimal final-goods pricing, households' labor force participation, and the wage-setting mechanism, I obtain the following expression for the regional Phillips curve in region i:

$$\pi_{it} = \begin{cases} \beta E_t \pi_{it+1} + \kappa_\theta^{tight} \hat{\theta}_{it} + \kappa_p \hat{p}^x_{it} + \kappa_\nu^{tight} \hat{\nu}_{it} + \kappa_a \hat{a}^y_{it} & \hat{\theta}_{it} > \hat{\theta}^*_{it} \\ \beta E_t \pi_{it+1} + \kappa_\theta \hat{\theta}_{it} + \kappa_p \hat{p}^x_{it} + \kappa_\nu \hat{\nu}_{it} + \kappa_a \hat{a}^y_{it} & \hat{\theta}_{it} \leq \hat{\theta}^*_{it} \end{cases}, \tag{26}$$

where π_{it} is inflation in region i, $E_t\pi_{it+1}$ denotes regional short-run inflation expectations, and $\hat{\theta}^*_{it}$ is such that when $\hat{\theta}_{it} > \hat{\theta}^*_{it}$, then $\theta_{it} > \theta^*_{it}$. The derivation of equation (26), together with the definitions of the parameters, can be found in Appendix B.

As long as the rate of adjustment of wages when the labor market is slack is positive and less than one, i.e. $\lambda \in (0, 1)$, then the regional Phillips curve is nonlinear and $\kappa_\theta \equiv \kappa_\theta^{tight}(1-\lambda) < \kappa_\theta^{tight}$. This means that labor market tightness exerts higher inflationary pressures when labor markets are tight rather than slack, as we observe in the raw data in Figure 1. When $\lambda = 0$, the real wage is flexible also when labor markets are slack, and the two curves coincide.

Three more terms are present in equation (26) and together they compose the regional cost-push shock. First, $\hat{p}^x_{it} = \widehat{\left(\frac{P^x_t}{P_{it}}\right)}$ denotes the percentage deviation of the regional relative price of intermediate input (i.e., the ratio between the national intermediate-input price, P^x_t, and the regional price level, P_{it}) from its steady-state value. As one of the factors of production, a change in the price of the intermediate input directly affects the marginal costs of final-goods firms and, consequently, their pricing decision. However, the presence of the relative price of intermediate input captures the notion that, as the price of intermediate input is common across regions, an identical absolute intermediate-input price change has a higher (lower) pass-through on regional inflation rates the lower (higher) the regional price level. Second, $\hat{\nu}_{it} \equiv \hat{\gamma}^c_{it} - \hat{\gamma}^b_{it} - \hat{m}_{it}$ represents shocks to the regional labor markets' search and matching process. Note that this shock may have a nonlinear effect on regional inflation as well, with $\kappa_\nu \equiv \kappa_\nu^{tight}(1-\lambda) < \kappa_\nu^{tight}$. Finally, $\hat{a}^y_{it}$ captures local shocks to final-goods sector productivity.

To take equation (26) to the data, I follow Hazell et al. (2022) and solve equation (26) forward. To do so, I make two assumptions. First, I assume that if the labor market is slack in t, then it will remain slack forever. If the labor market is tight in t, then it will remain tight until $t+T-1$. From $t+T$ onward, the labor market becomes slack and will remain slack forever. Second, I assume that $\tilde{\theta}_{it}$, $\hat{p}^x_{it}$, and $\hat{a}^y_{it}$ follow AR(1) processes with autocorrelation coefficients ρ_θ, ρ_p, and ρ_{a^y}.

By doing so, I obtain the following regional Phillips curve:

$$\pi_{it} = \begin{cases} E_t\pi_{t+\infty} + \xi E_t\hat{\theta}_{it+\infty} + \psi_\theta^{tight}\tilde{\theta}_{it} + \psi_p\hat{p}^x_{it} + \psi_a\hat{a}^y_{it} + \varepsilon^{tight}_{it} & \tilde{\theta}_{it} > \tilde{\theta}^*_{it} \\ E_t\pi_{t+\infty} + \psi_\theta^{slack}\tilde{\theta}_{it} + \psi_p\hat{p}^x_{it} + \psi_a\hat{a}^y_{it} + \varepsilon_{it} & \tilde{\theta}_{it} \leq \tilde{\theta}^*_{it} \end{cases}, \tag{27}$$

where $\tilde{\theta}_{it} = \hat{\theta}_{it} + E_t\hat{\theta}_{it+\infty}$ represents the transitory component of the variation in labor market tightness, while $E_t\hat{\theta}_{it+\infty}$ is the permanent component. $E_t\pi_{t+\infty}$ denotes long-run inflation expectations, assumed to be common across regions because they depend on the monetary regime in place. This formulation clarifies how regional data helps in dealing with threats to identification coming from controlling for inflation expectations in the estimation of the Phillips curve. Indeed, common long-run inflation expectations are captured by time fixed effects, while $\xi E_t\hat{\theta}_{it+\infty}$ is captured by region fixed effects.

Crucially, $\psi_\theta^{slack} = \psi^1_\theta$, while $\psi_\theta^{tight} = \psi^1_\theta + \psi^2_\theta$, where the expressions for ψ^1_θ and ψ^2_θ can be found in Appendix B. Hence, the slope of the regional Phillips curve in tight labor markets is larger than the slope in slack labor markets by a measure equal to ψ^2_θ. The formulas for ψ_p and ψ_a can be found in Appendix B, while $\varepsilon_{it} = E_t\sum\limits_{k=0}^{\infty}\beta^k\kappa_\nu\hat{\nu}_{it+k}$ and $\varepsilon^{tight}_{it} = E_t\left[\sum_{k=0}^{T-1}\beta^k\kappa_\nu^{tight}\hat{\nu}_{it+k} + \sum_{k=T}^{\infty}\beta^k\kappa_\nu\hat{\nu}_{it+k}\right]$ denote the expected present discounted value of current and future search-and-matching shocks in slack and tight labor markets, respectively. Appendix B contains the formal derivation of equation (27).

The interpretation of the slope of the Phillips curve differ between equation (26) and equation (27). κ_θ^{tight} and κ_θ denote the effects of current labor market tightness on current inflation, while ψ_θ^{tight} and ψ_θ denote the effects of current and expected future deviations of labor market tightness from its long-run steady state on current inflation. Depending on the degree of persistence of labor market tightness, ψ_θ^{tight} and ψ_θ can be more or less larger than κ_θ^{tight} and κ_θ. In my empirical exercise, I estimate ψ_θ^{tight} and ψ_θ in equation (27) and not κ_θ^{tight} and κ_θ in equation (26) due to sample size limitations. Estimating ψ_θ^{tight} and ψ_θ allows to abstract from empirically modelling future values of labor market tightness, which increases the sample size. This implies greater statistical power for estimating the coefficients of interest.

3 Data

To carry out my empirical exercise, I draw data from different sources from December 2000 to April 2023. The units of observations are 21 Metropolitan Statistical Areas (MSAs) in the United States. The main dependent variable is inflation, constructed as the 12-month percent difference in the Consumer Price Index (CPI). The Bureau of Labor Statistics (BLS) regularly provides CPI data for 21 MSAs on a monthly or bi-monthly basis. Prices from all categories are collected monthly in the metropolitan areas of Chicago, Los Angeles and New York. In the other MSAs, prices for food and energy items are collected monthly, while prices for other categories are collected every two months. I linearly interpolate the bi-monthly CPI time series to maximize the sample size and to fully exploit the variation in labor market and instrumental variables. Interpolation introduces measurement errors in CPI and, consequently, inflation. However, such measurement errors do not lead to an attenuation bias in the estimates, because they affect the dependent variable only. Crucially, they do not affect the exogenous variation provided by the instrumental variable. The initiation date for CPI data collection varies among the included metropolitan areas, starting from January 1986.

My analysis primarily focuses on core inflation, defined as the growth rate of prices of all items excluding food and energy. The literature suggests to use core inflation to avoid the relative volatility of food and energy prices, which are more connected to global factors. Core inflation, on the other hand, is more related to domestic economic activity. Nonetheless, I collect CPI data and construct inflation measures also for other broad categories, such as all items (headline inflation), all items excluding shelter, goods and services. For a more comprehensive understanding of CPI data, I recommend referring to the works of Klenow and Kryvtsov (2008) and Nakamura and Steinsson (2008).

I employ a labor market-based variable to measure economic activity: labor market tightness, also known as vacancy-to-unemployment ratio. Labor market indicators are relevant proxies because the mechanism that the Phillips curve captures relies on demand-driven economic fluctuations that affect labor costs for firms. Labor market tightness is the ratio of two elements: the number of vacancies posted by firms in the numerator provides information on labor demand, while the number of unemployed workers in the denominator tracks the supply of workers available in the market. By including a measure of labor demand, tightness provides more accurate information on labor costs than the unemployment rate (Barnichon and Shapiro, 2022). Nonetheless, researchers have primarily used the unemployment rate to proxy for real economic activity in the estimation of the Phillips curve. This is because the availability of administrative data on the unemployment rate is far broader, both at granular levels and back in time.

Recent evidences provide further reasons for using labor market tightness over the unemployment rate as main independent variable. First, Furman and Powell (2021) show that until the outbreak of the COVID-19 pandemic, the unemployment rate and labor market tightness co-moved well. However, the behavior of the two variables started to diverge in March 2020, with the vacancy-to-unemployment ratio signalling tighter labor markets with respect to the unemployment rate. Hence, an analysis that takes into consideration the COVID and post-COVID periods, such as this one, should be careful on the choice of proxy for economic activity. Furthermore, Barnichon and Shapiro (2022) show that the vacancy-to-unemployment ratio outperforms the unemployment rate in the forecast of both price and wage inflation. For all these reasons, this analysis uses labor market tightness as main explanatory variable, but I check the robustness of my results to the unemployment rate.

As data on vacancies is not publicly available at the MSA level, I impute a new time series of MSA-level vacancies employing state-level data from JOLTS and population weights from the Census. JOLTS is a monthly survey of about 21,000 U.S. business establishments, providing representative data on job openings at the national and state levels since December 2000. Population weights are constructed from the 2000 and 2010 Census. Suppose that metropolitan area i belongs to state x and state y. Vacancies in metropolitan area i in period t are computed as a weighted average of vacancies from states x and y in period t, with weights being the fraction of population of state x living in MSA i and the fraction of population of state y living in MSA i, respectively.

The main advantage of this newly imputed measure of MSA-level vacancies is that it is constructed using publicly available job postings from a representative sample of firms. There exist two private sources of MSA-level vacancies: Lightcast (a merger of Emsi and Burning Glass Technologies) and the Conference Board. Lightcast collects online job postings from 2011. The Conference Board started to collect job advertisements printed in newspapers in the 1950s. This series was discontinued in 2008 and substituted with collection of online job postings since 2005. For a more detailed explanation of the Conference Board data, refer to Barnichon (2010). Since 2019, Lightcast is the only provider of data for the Conference Board. Since 2020, the Conference Board implements an adjustment to bridge the gap between the data from Lightcast and the data from the BLS's Jop Openings and Labor Turnover Survey (JOLTS). The need of such an adjustment underscores the main problem of online job posting data: this data is not representative and might introduce selection bias in the estimates.

The time series of vacancies imputed from JOLTS allows to avoid selection bias, as it comprises job postings from all the sectors of the economy and not from those that recruit online only. However, the distribution of vacancies at the state level might not be similar to the distribution of vacancies in metropolitan areas. The empirical strategy provides for the use of an instrumental variable to instrument labor market tightness. In addition to control for local confounders, the instrumental variable attenuates the measurement error that the imputation of vacancy data introduces in the main independent variable.

In addition to vacancy data, I need to collect unemployment data to construct labor market tightness at the MSA level. I draw monthly MSA-level number of unemployed workers from the BLS's Local Area Unemployment Statistics (LAUS). The LAUS program uses non-survey methodologies to estimate the number of employed and unemployed individuals for sub-national areas, using the national not-seasonally-adjusted estimates from the Current Population Survey as controls. From LAUS I also collect monthly MSA-level unemployment rate to be employed in a robustness check.

To construct the shift-share instrument and controls, I need additional data. For the instrument, I collect two types of data. First, I draw monthly, national employment data by industry from the Current Population Survey (CPS). Second, I collect MSA-level industry employment shares from the 1990 and 2000 Census, and from the 2006, 2011, 2016, and 2021 American Community Survey (ACS). Both types of employment data are taken disaggregated at the level of the three-digit North American Industry Classification System (NAICS) code. Then, I aggregate it at the two-digit level, distinguishing between intermediate-inputs and final-goods industries. Finally, I measure the relative price of intermediate inputs as the ratio between the monthly producer price index (PPI) for the manufacturing sector and the local CPI. Both variables come from the BLS.

The resulting dataset is a panel of MSA-year-month observations, from December 2001 to April 2023. Figure 2 shows the distributions of the dependent and main independent variables. As you can see, there is a significant degree of variation in core inflation and labor market tightness at the metropolitan area level. Furthermore, Figure A.1 in Appendix A shows a key advantage of employing regional data. Comparing the distributions of MSA-level and national labor market tightness, regional data exhibits more variation and a higher number of tight labor markets episodes with respect to national data. Aggregate labor market tightness reaches the maximum level of 2, while MSA-level labor market tightness reaches 3.44. At the national level, we observe a vacancy-to-unemployment ratio greater than one only in 18.22% of the sample. At the MSA-level, the same proportion is 46.33%. The take-away is that regional data provides more variation and episodes of tight labor markets than national data, increasing the statistical power of the empirical exercise.

4 Empirical Strategy

My empirical exercise aims at testing the presence of nonlinearities in the regional New Keynesian Phillips curve, employing labor market tightness to proxy for real economics activity. To do so, I estimate the NRPC – equation (27) – derived in Section 2. The empirical strategy is based on two ingredients. First, panel variation in inflation and tightness at the MSA level provides higher statistical power than time series variation as shown in Section 3, and takes care of aggregate confounders. Second, the instrumental variable approach deals with regional confounders.

To estimate equation (27), I specify the following empirical model:

$$\pi_{it} = c + \alpha_i + \gamma_t + \psi_\theta^1 \ln(\theta_{it}) + \psi_\theta^2 \ln(\theta_{it}) \times I_{\{\theta_{it} > \theta_i^*\}} + \beta I_{\{\theta_{it} > \theta_i^*\}} + \psi_p p_{it}^x + \psi_a z_{it}^y + \varepsilon_{it}, \tag{28}$$

where π_{it} denotes the 12-month, core inflation rate in MSA i and year-month t. Constructing the growth rate of prices over 12 months allows me to reduce seasonality. α_i represents MSA fixed effects, absorbing time-invariant characteristics of metropolitan areas, such as differences in long-run economic fundamentals across cities. γ_t denotes year-quarter fixed effects, absorbing aggregate shocks, such as endogenous fiscal and monetary policies (Fitzgerald and Nicolini 2014, McLeay and Tenreyro 2020), and common beliefs about the long-run monetary policy regime (Hazell et al., 2022).

The nonlinearity of the Phillips curve in labor market tightness is specified as follows. $\ln(\theta_{it})$ is the logarithm of labor market tightness in city i and year-month t. I employ the logarithm of labor market tightness for two reasons. First, it mirrors the log-linearized tightness term in equation (27): $\hat{\theta}_{it}$. Second, it enables me to avoid taking a stance about whether I should measure labor market tightness as the vacancy-to-unemployment ratio or the unemployment-to-vacancy ratio. $I_{\{\theta_{it} > \theta_i^*\}}$ is a dummy variable that takes value of 1 when labor market tightness is greater than a MSA-level threshold, θ_i^*, and 0 otherwise. The threshold is allowed to vary across MSAs and will be determined empirically. The interaction between $\ln(\theta_{it})$ and $I_{\{\theta_{it} > \theta_i^*\}}$ generates the nonlinearity. When labor market tightness is less than θ_i^*, the slope of the Phillips curve is $\psi_\theta = \psi_\theta^1$ and the intercept is c. When labor market tightness is greater than θ_i^*, the slope of the Phillips curve is $\psi_\theta^{tight} = \psi_\theta^1 + \psi_\theta^1$ and the intercept is $c + \beta$.

Finally, guided by equation (27), I add to the main specification two controls. p_{it}^x is the local relative price of intermediate inputs in MSA i and year-month t, measured as the ratio of national manufacturing PPI in year-month t and core CPI in MSA i and year-month t. z_{it}^y denotes the proxy for productivity shocks of non-tradable, final-goods industries in MSA i and year-month t. As explained in Section 2, p_{it}^x,

z_{it}^{y}, and the error term ε_{it} constitute together the local cost-push shock. Through the lenses of the model, the error term ε_{it} contains the shocks to the regional labor markets' search and matching process.

I conduct the empirical exercise employing regional variation for two main reasons. First, regional data provides greater variation and a higher number of episodes of tight labor markets with respect to national data, as shown in Section 3. Second, a growing literature has shown how panel variation helps overcoming three identification problems affecting the estimation of the aggregate Phillips curve. The main challenge for the identification of the slope of the Phillips curve is to distinguish between demand and supply shocks. While demand shocks increase economic activity and inflation, supply shocks depress economic activity and increase inflation, leading to the so-called *simultaneity bias*. The simultaneity bias generates a downward bias in the estimated slope of the Phillips curve. There are two more challenges. First, Fitzgerald and Nicolini (2014) and McLeay and Tenreyro (2020) show that if monetary and fiscal policies react to offset aggregate demand shocks, then the remaining variation in inflation will only be due to supply shocks, leading to a biased estimate of the aggregate slope. Second, the choice of variable to measure inflation expectations affects the estimate of the slope (Mavroeidis et al., 2014).

Panel variation helps in dealing with these aggregate threats to identification because of the introduction of time fixed effects. Time fixed effects absorb any aggregate demand or supply shock, eliminating any bias coming from aggregate fluctuations all at once. Among the aggregate demand shocks, Fitzgerald and Nicolini (2014) and McLeay and Tenreyro (2020) show that time fixed effects capture endogenous monetary and fiscal policies that are set at the national level, solving the problem of omitted variable bias. Hazell et al. (2022) show that time fixed effects are also able to absorb local inflation expectations, if we assume that they are determined by the monetary regime in place[6]. Section 2 and Appendix B discuss and show more in details the derivation of this result.

What is left in the variation of inflation at the regional level comes from both regional demand and regional supply shocks. To solve the simultaneity bias at the regional level, I propose a shift-share instrumental variable that captures productivity shocks in the tradable intermediate-input sectors, similar to Cerrato and Gitti (2022). The shift-share instrument takes the following form:

$$z_{it}^{x} = \sum_{k=1}^{N} e_{ki} \times g_{kt},$$

where e_{ki} is the average employment share of industry k in metropolitan area i, and g_{kt} is the three-

[6]Moreover, Sargent (1982) shows that common beliefs about the long-run monetary regime in place are a major determinant of sudden fluctuations in inflation.

year growth in national employment of industry k at time t. Industries are identified at the level of the two-digit NAICS code, and include: agriculture, mining, manufacturing of durable and non-durable goods, wholesale trade, and intermediate-input sectors in transportation, warehousing and utilities, and in professional and business services. The shifters g_{kt} capture national productivity-driven labor demand shocks in the intermediate-input sectors at business cycle frequencies. The shares e_{ki} measure the regional exposure to such aggregate shocks.

I need an instrumental variable approach to solve the simultaneity bias at the regional level for the following reason. Through the lenses of the model, the regional supply shock is composed of three elements: the regional relative price of intermediate inputs $\hat{p}^x_{it}$, the final-goods sector's productivity $\hat{a}^y_{it}$, and the search and matching shocks $\hat{\nu}_{it}$. While $\hat{p}^x_{it}$ and $\hat{a}^y_{it}$ are observable and can be controlled for, search and matching shocks are not and are contained in the error term. The instrumental variable allows to isolate fluctuations in labor market tightness that are not driven by unobservable search and matching shocks.

The shift-share instrument works as follows. A positive productivity shock in the tradable intermediate-input sector raises labor demand of intermediate-input firms relatively more in those metropolitan areas with a higher degree of specialization in the intermediate-input sector. As a consequence, employment agencies in the specialized metropolitan areas will post more vacancies, leading to higher wages at a speed that depends on labor market tightness. Higher wages represent higher labor costs for final-goods firms, which will ultimately increase final-goods prices. For instance, a positive national productivity shock in the manufacturing sector leads to larger cost increases for final-goods firms located in manufacturing-intensive cities like Detroit. This channel produces an exogenous variation in labor market tightness that the shift-share instrument exploits to identify the slope of the Phillips curve.

The control variables are key to take care of two threats to the exclusion restriction of the shift-share instrument. First, a positive productivity shock in the tradable intermediate-input sector decreases the price of the intermediate input, lowering the cost of production for all final-goods firms that use the intermediate input in their production process. Consequently, final-goods firms lower their prices. Such mechanism captures a channel through which tradable intermediate-input productivity shocks affect inflation that is not through labor market tightness. This represents a violation of the exclusion restriction for the validity of the instrumental variable. However, as prices of intermediate inputs are observable, I can control for their direct incidence on local inflation by including p^x_{it} in the specification.

Second, productivity shocks in the tradable intermediate-input sector captured by the shift-share instrument could be correlated with productivity shocks in the regional final-goods sectors. Such correlation has likely been in place during and after the COVID-19 pandemic (Guerrieri et al., 2022), when local economies experienced robust labor demand recoveries across all sectors. If this is the case, the instrument affects final-goods prices though the correlation with productivity shocks of regional final-goods sectors. To deal with this violatin to the exclusion restriction, I follow Borusyak et al. (2022), and include as control a shift-share variable proxying for productivity shocks in local final-goods sectors, z^y_{it}. The structure of this variable mirrors the of the shift-share instrument, employing two-digit NAICS, non-tradable, final-goods industries [7].

The exogeneity of the instrument, conditional on the controls for the relative price of intermediate inputs and for the final-goods sectors' productivity shocks, stems from the shocks g_{jt}, rather than from the exposure shares e_{ji}. Such a case falls under the framework developed by Borusyak et al. (2022). In this paper, the authors prove that the validity of shift-share instruments can rely on the exogenous variation of the shocks only, allowing the variation in exposure shares to be endogenous. In particular, under some assumptions, shocks can be only as-good-as-randomly assigned, i.e. shocks can be equilibrium objects. An example of this type of shocks are the national industry employment growth rates that this analysis employs, as well as Bartik (1991).

As my instrument is constructed using tradable intermediate-input industries only, I interact the sum of the exposure shares with time fixed effects in the regression. Borusyak et al. (2022) show that failing to control for the sum of exposure shares introduces a bias in the estimated coefficients. The reason being that cities with more diversified economies will tend to have systematically higher shift-shares instruments and they may have systematically different unobservables. For example, cities with more diversified economies might be more resilient to unobserved shocks.

The identifying assumption is that, conditioning on MSA fixed effects, time fixed effects interacted with the sum of the exposure shares, p^x_{it}, and z^y_{it}, industry-level employment growth rates in the intermediate-input sectors capture labor demand shocks plausibly uncorrelated with industry-level aggregates of regional labor supply shocks. Related to the example of a productivity shock in the manufacturing sector illustrated before, the identifying assumption requires that the labor cost increases generated by such shock are no larger on average for final-goods firms in Detroit than New York.

The nonlinearity of my main independent variable requires that I instrument not only the logarithm

[7] Two-digit NAICS, non-tradable, final-goods sectors used to construct z^y_{it}: construction, retail trade, information and communication, finance and insurance, real estate, rental and leasing, educational services, health care and social assistance, arts, entertainment and recreation, accommodation and food services, other services, and final-goods sectors in transportation, warehousing and utilities, and in professional and business services.

of labor market tightness $\ln(\theta_{it})$, but also the interaction term between the logarithm of labor market tightness and the dummy signalling when labor markets are tight $\ln(\theta_{it}) \times I_{\{\theta_{it} > \theta_i^*\}}$. I instrument the interaction term with an interaction between my shift-share instrument and a dummy that takes value of one when the value of the shift-share instrument is above the 75^{th} percentile.

Table A.1 in Appendix A shows the first-stage coefficients and F-statistics. Column 1 shows that the shift-share instrument z_{it}^x strongly predicts $\ln(\theta_{it})$, while column 2 shows that the instrument for the interaction $z_{it}^x \times I_{\{z_{it}^x > z_{75}^x\}}$ strongly predicts the interaction term $\ln(\theta_{it}) \times I_{\{\theta_{it} > \theta_i^*\}}$. Furthermore, the two instruments are strong, as the F-statistics are both greater than 10.

5 Results

5.1 Main Results

I find two novel sets of results. First, a regional Phillips curve in labor market tightness exists from December 2000 to April 2023. This finding confirms that the mechanism captured by the Phillips curve is at work during this period. An increase in labor market tightness induces workers to ask for higher wages and, consequently, firms to raise prices. Second, I find evidence of nonlinearities in the regional Phillips curve. The slope of the regional Phillips curve almost triples when labor market tightness exceeds the MSA-specific average.

Table 1 shows these results. In Table 1, θ_i^* is equal to the average of labor market tightness in MSA i. The distribution of labor market tightness averages across MSAs goes from a minimum of 0.51 to a maximum of 0.9, with a mean value of 0.68. I select these thresholds because they are compatible with the instrumental variables being strong in column (1). Column (1) reports the results of my benchmark specification, where equation (28) is estimated by instrumenting labor market tightness and the interaction term with the shift-share instrument. I estimate the slope of the Phillips curve to be 3.01% below the threshold and 8.4% above the threshold. That is, a 1% increase in labor market tightness leads to a 3% increase in inflation when regional labor markets are slack and an 8.4% increase when regional labor markets become sufficiently tight. In sum, the slope of the NRPC almost triples above the threshold.

Comparing the estimation of equation (28) by IV and OLS is instructive to understanding the role played by the IV. Column (2) of Table 1 reports a simple OLS regression that controls only for the local relative price of intermediate inputs p_{it}^x, the productivity shocks of final-goods sectors z_{it}^y, and MSA-fixed effects. This simple OLS regression yields a nonlinear regional Phillips curve with a slope of 0.98% below the thresholds and 2.26% above the thresholds.

Column (3) shows that the Phillips curve still exists in regional data when I add time fixed effects, but the nonlinearity disappears. Indeed, the coefficient on the interaction term shrinks and becomes statistically not different from zero. Taken together, these results reveal that aggregate shocks play a role in reinforcing stronger inflationary pressures in tight labor markets. When aggregate shocks are absorbed by time fixed effects, the nonlinearity disappears because the presence of regional supply shocks weakens the effect of tight labor markets on inflation. Once I control for them with the IV, the nonlinearity appears again. By running a simple two-way fixed effects model, we could conclude that the Phillips curve is linear in labor market tightness. However, it is important to address the simultaneity of demand and supply shocks at the local level in order to corroborate such result.

The difference in magnitudes between the OLS and IV coefficients is driven by the fact that the instrumental variable corrects for measurement errors. Indeed, as I explain in Section 3, the vacancy variable in the numerator of labor market tightness, my main independent variable, suffers from measurement error. The increase in magnitude goes in the right direction, as the presence of local supply shocks introduces a downward bias in the estimated slope of the Phillips curve. Although my instrumental variables reject the hypotheses of weak instruments and are strongly correlated with the instrumented variables, as Table A.1 shows, the precision of the IV estimates can be improved.

Figure 3 shows the goodness-of-fit of the regional Phillips curve estimated in Table 1, column (1). Inflation deviations are defined as the difference between the 12-month core inflation rates and the estimated controls and fixed effects. What is left is the variation due to labor market tightness and the error terms. Such variation is plotted against the log of labor market tightness, my main independent variable. The red line shows the curve estimated in Table 1, column (1). As you can see, when labor market tightness becomes greater than 0.67, the average of the thresholds across MSAs, the slope of the regional Phillips curve increases, confirming the pattern that we see in the raw data in Figure (1).

5.2 Policy Implications

Ignoring nonlinearities in the Phillips curve can induce the monetary authority to allow inflation to rise more than expected. Figure 4 provides a visual representation of this dynamic. If the central bank believes that the Phillips curve is linear and flat, a given increase in labor market tightness will produce the same small relative increase in inflation, no matter the level of labor market tightness. After the kink, the central bank will think that the economy is on the dashed blue line, where a further increase in labor market tightness produces only limited inflationary pressures. Hence, a central bank running an expansionary monetary policy might have the incentives to keep the accommodative policy stance even if labor markets become tight.

However, if the Phillips curve is nonlinear, after the kink the economy lies on the steeper blue solid line. Hence, a given increase in labor market tightness will spur larger inflationary pressures than expected, leading to an unexpected surge in inflation. The surprise increase in inflation equals to the difference between the solid and dashed line. Through the lenses of the model, the increase in the slope of the Phillips curve in tight labor markets is due to firms bidding up wages to attract workers. Wages become flexible and do not constraint anymore firms' demand, leading to a larger change in marginal costs and consequently prices of final-goods firms.

The steeper slope of the Phillips curve after the kink implies that, as long as labor markets run hot, a restrictive monetary policy stance will reduce inflation without affecting much the labor market. Such possibility of "soft landing" has been much discussed recently, as the central banks around the world are aggressively trying to bring down inflation without causing much economic recession. As I have argued, a nonlinear Phillips curve in labor market tightness can explain this puzzle.

6 Robustness Checks

In this section, I perform four robustness exercises to evaluate the stability of my results. First, instead of a single kink, I show that the nonlinearity of the regional Phillips curve is robust to more flexible functional forms of nonlinearity. Second, I report mixed evidence of nonlinearity in the regional Phillips curve when economic activity is measured by the unemployment rate. This result underlies that the choice of the variable used to proxy for real economic activity in the estimation of the Phillips curve matters. Third, I check whether the nonlinearity in the regional Phillips curve is driven only by the post-COVID period when labor market tightness at the aggregate level increased to levels not seen since the 1940s. I employ the unemployment rate to conduct this robustness check because it ensures enough statistical power. I find that the nonlinearity is still present when I exclude the post-COVID period from the sample. Fourth, I show that the nonlinearity in the regional Phillips curve is robust to introducing a proxy for local inflation expectations.

The first robustness exercise shows that the nonlinearity in the regional Phillips curve is robust to other nonlinear functional forms. I test the following functional forms: piecewise linear, logarithmic, and inverse. All these specifications are linear in parameters and nonlinear in the main independent variables. Table A.2 in Appendix A finds evidence of an NRPC when labor market tightness is subject to logarithmic and inverse transformations. The coefficients on such transformations of labor market tightness, instrumented with the shift-share share, are significantly different from zero.

The second robustness exercise checks whether the existence of the NRPC is robust to the employment of the unemployment rate to measure economic activity. Economists generally use the unemployment rate to estimate the Phillips curve, as this variable is available at the highest frequency and level of disaggregation. Table A.3 in Appendix A shows that my finding is somewhat robust. Using the unemployment rate allows me to have a larger sample size, as the data has been available since January 1990. These additional ten years of data positively affect the instrument's strength, as the F-statistics show. I consider the following functional forms: log, piecewise linear, piecewise log, and inverse. The estimated slope of the Phillips curve is significantly different from zero across all specifications, but the coefficients on the interaction terms are not. The signs of the coefficients all go in the expected direction. An increase in $\ln(u_{it})$ negatively affects core inflation, while an increase in $\frac{1}{u_{it}}$ yields an opposite effect. The coefficients on the interaction terms are negative but not statistically significant. I can conclude that there is mixed evidence of nonlinearities in the Phillips curve when economic activity is measured by the unemployment rate.

The third robustness exercise investigates whether the nonlinearity in the regional Phillips curve is driven only by the dynamics between inflation and economic activity in the aftermath of the pandemic. To do so, I estimate the NRPC on a smaller sample that stops in February 2020. I conduct this robustness check using the unemployment rate as main explanatory variable. As data on the unemployment rate is available for longer, the larger sample size increases the statistical power. With labor market tightness, the instruments are not strong enough when I reduce the sample size.

There is suggestive evidence that the regional Phillips curve with the unemployment rate is nonlinear in the pre-COVID period, as Table A.4 in Appendix A shows. The nonlinearity is present when the unemployment rate is transformed in logarithmic and inverse forms. The coefficients on the interaction terms in the piecewise linear and piecewise log specifications, in line with the results in Table A.3, are not significant. Notice that the estimated coefficients are larger in the pre-COVID period than in the full sample. There might be two reasons for these counter-intuitive results. First, including the COVID period in the full sample might attenuate the coefficients as the slope of the regional Phillips curve dropped to 0 in this period (Cerrato and Gitti, 2022). Second, the F-statistics show that the shift-share instrument is weaker in the pre-COVID sample than in the full sample. Hence, the magnitude of the coefficients might increase because the confidence intervals are wider.

In the last robustness check, I show that the NRPC is robust to the inclusion of a proxy for regional inflation expectations. As my estimates of ψ_θ and ψ_θ^{tight} capture the effect of current and expected future labor market tightness on current inflation, I check that expectations about local future economic conditions do not drive them. If we observed short-run inflation expectations at the local level, we could estimate Equation 26. In that case, the estimated slope of the Phillips curve would only capture the effect of current labor market tightness on current inflation, the parameters κ_θ and κ_θ^{tight}. Coibion and Gorodnichenko (2015) argue that households may form their short-run inflation expectations by observing the changes in prices of salient goods, such as gasoline. Therefore, I proxy for local short-run inflation expectations by using the 12-month MSA-level gasoline inflation rate and include this control in my main specification. My estimates are robust to the inclusion of this proxy for local inflation expectations. Table A.5 in Appendix A shows that the coefficients on labor market tightness and the interaction term decrease, but the increase in the slope of the Phillips curve after the threshold is still of similar size as in Table 1. Local inflation expectations have a null and not significant impact on core inflation.

7 Conclusion

This is the first paper that show the presence of nonlinearities in the regional Phillips curve with labor market tightness as proxy for economic activity. In doing so, this paper provides both theoretical and empirical contributions. From a theoretical point of view, I introduce search-and-matching frictions and wage rigidities in an otherwise standard multi-sector, New Keynesian model of two regions in a monetary union. Search-and-matching frictions give rise to unemployment in the New Keynesian model formally and to a relationship between inflation and labor market tightness. Wage rigidities generate a kink in the regional Phillips curve. Augmented with these features, the model delivers a piecewise log-linear regional Phillips curve in labor market tightness. From an empirical point of view, I impute a novel measure of vacancies across metropolitan areas. This newly imputed variable allows me to measure labor market tightness at the MSA level and to estimate the NRPC derived in the model. Moreover, I enrich the empirical strategy based on MSA-level panel variation with an instrumental variable approach to deal with unobservable regional supply shocks.

I find that the regional Phillips curve has a positive slope that increases almost three times when labor market tightness exceeds the metropolitan area-specific average. Hence, I provide evidence that the Phillips curve is piecewise log-linear at the MSA level. This result has two implications for the monetary authorities. First, if the central bank assumes that the Phillips curve is linear, this might lead them to underestimate the inflationary pressures of sufficiently tight labor markets, allowing inflation to surge more than expected. Second, the nonlinearity of the Phillips curve allows the central bank to bring down inflation without causing a significant economic recession. More work is needed to relate the regional Phillips curve to the aggregate one, to infer more detailed implications for monetary policy. I believe that this is an exciting path for future research.

References

Babb, Nathan R. and Alan K. Detmeister, "Nonlinearities in the Phillips Curve for the United States : Evidence Using Metropolitan Data," Finance and Economics Discussion Series 2017-070, Board of Governors of the Federal Reserve System (U.S.) June 2017.

Barnichon, Regis, "Building a Composite Help-Wanted Index," *Economics Letters*, 2010, *109* (3), 175–178.

_ **and Adam H. Shapiro**, "What's the Best Measure of Economic Slack?," Economic Letter 4, Federal Reserve Bank of San Francisco 2022.

Bartik, Timothy J., "Who Benefits from State and Local Economic Development Policies?," WE Upjohn Institute for Employment Research 1991.

Benigno, Pierpaolo and Gauti B Eggertsson, "It's Baaack: The Surge in Inflation in the 2020s and the Return of the Non-Linear Phillips Curve," Working Paper 31197, National Bureau of Economic Research April 2023.

Blanchard, Olivier and Jordi Galí, "Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment," *American economic journal: macroeconomics*, 2010, *2* (2), 1–30.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-experimental Shift-Share Research Designs," *The Review of Economic Studies*, 2022, *89* (1), 181–213.

Calvo, Guillermo A., "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 1983, *12* (3), 383–398.

Cerrato, Andrea and Giulia Gitti, "Inflation Since COVID: Demand or Supply," *Available at SSRN 4193594*, 2022.

Christoffel, Kai Philipp and Tobias Linzert, "The Role of Real Wage Rigidity and Labor Market Frictions for Unemployment and Inflation Dynamics," 2005.

Coibion, Olivier and Yuriy Gorodnichenko, "Is the Phillips Curve Alive and Well After All? Inflation Expectations and the Missing Disinflation," *American Economic Journal: Macroeconomics*, 2015, *7* (1), 197–232.

Faia, Ester, "Optimal Monetary Policy Rules with Labor Market Frictions," *Journal of Economic dynamics and control*, 2008, *32* (5), 1600–1621.

Fitzgerald, Terry J. and Juan Pablo Nicolini, "Is There a Stable Relationship between Unemployment and Future Inflation?," Federal Reserve Bank of Minneapolis Working Paper 2014.

Furman, Jason and Wilson Powell, "What is the Best Measure of Labor Market Tightness?," Peterson Institute for International Economics 2021.

Galí, Jordi, *Unemployment Fluctuations and Stabilization Policies: a New Keynesian Perspective*, MIT press, 2011.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W Huffman, "Investment, Capacity Utilization, and the Real Business Cycle," *The American Economic Review*, 1988, pp. 402–417.

Guerrieri, Veronica, Guido Lorenzoni, Ludwig Straub, and Iván Werning, "Macroeconomic Implications of COVID-19: Can Negative Supply Shocks Cause Demand Shortages?," *American Economic Review*, 2022, *112* (5), 1437–74.

Harding, Martín, Jesper Lindé, and Mathias Trabandt, "Understanding Post-Covid Inflation Dynamics," 2022.

Hazell, Jonathon, Juan Herreno, Emi Nakamura, and Jón Steinsson, "The Slope of the Phillips Curve: Evidence from US States," *The Quarterly Journal of Economics*, 2022, *137* (3), 1299–1344.

Hooper, Peter, Frederic S. Mishkin, and Amir Sufi, "Prospects for Inflation in a High Pressure Economy: Is the Phillips Curve Dead or is it just Hibernating?," *Research in Economics*, 2020, *74* (2), 26–52.

Kiley, Michael T., "An Evaluation of the Inflationary Pressure Associated with Short- and Long-Term Unemployment," *Economics Letters*, 2015, *137*, 5–9.

Klenow, Peter J. and Oleksiy Kryvtsov, "State-Dependent or Time-Dependent Pricing: Does it Matter for Recent US Inflation?," *The Quarterly Journal of Economics*, 2008, *123* (3), 863–904.

Krause, Michael U and Thomas A Lubik, "The (Ir)relevance of Real Wage Rigidity in the New Keynesian Model with Search Frictions," *Journal of Monetary Economics*, 2007, *54* (3), 706–727.

Leduc, Sylvain and Daniel Wilson, "Has the wage Phillips curve gone dormant?," *FRBSF Economic Letter*, 2017.

Mavroeidis, Sophocles, Mikkel Plagborg-Møller, and James H. Stock, "Empirical Evidence on Inflation Expectations in the New Keynesian Phillips Curve," *Journal of Economic Literature*, 2014, *52* (1), 124–188.

McLeay, Michael and Silvana Tenreyro, "Optimal Inflation and the Identification of the Phillips Curve," *NBER Macroeconomics Annual*, 2020, *34* (1), 199–255.

Michaillat, Pascal, "A Theory of Countercyclical Government Multiplier," *American Economic Journal: Macroeconomics*, 2014, *6* (1), 190–217.

_ **and Emmanuel Saez**, "u*=$\sqrt{uv}$," Working Paper 30211, NBER 2022.

Murphy, Anthony, "Is the US Phillips Curve Convex? Some Metro Level Evidence," *Manuscript*, 2017.

Nakamura, Emi and Jón Steinsson, "Five Facts about Prices: a Reevaluation of Menu Cost Models," *The Quarterly Journal of Economics*, 2008, *123* (4), 1415–1464.

Petrosky-Nadeau, Nicolas and Lu Zhang, "Unemployment crises," *Journal of Monetary Economics*, 2021, *117*, 335–353.

Phillips, Alban W., "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957," *Economica*, 1958, *25* (100), 283–299.

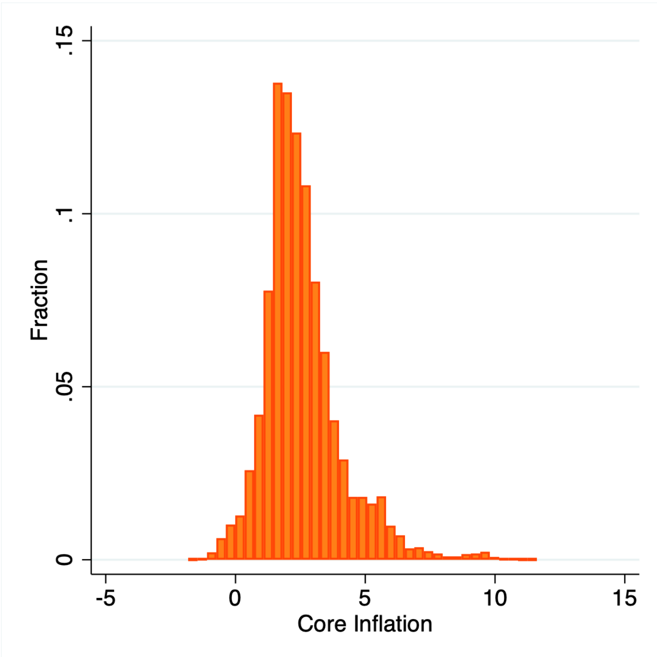
Sargent, Thomas J., "The Ends of Four Big Inflations," in "Inflation: Causes and effects," University of Chicago Press, 1982, pp. 41–98.

Schmitt-Grohé, Stephanie and Martín Uribe, "Heterogeneous Downward Nominal Wage Rigidity: Foundations of a Static Wage Phillips Curve," Technical Report, National Bureau of Economic Research 2022.

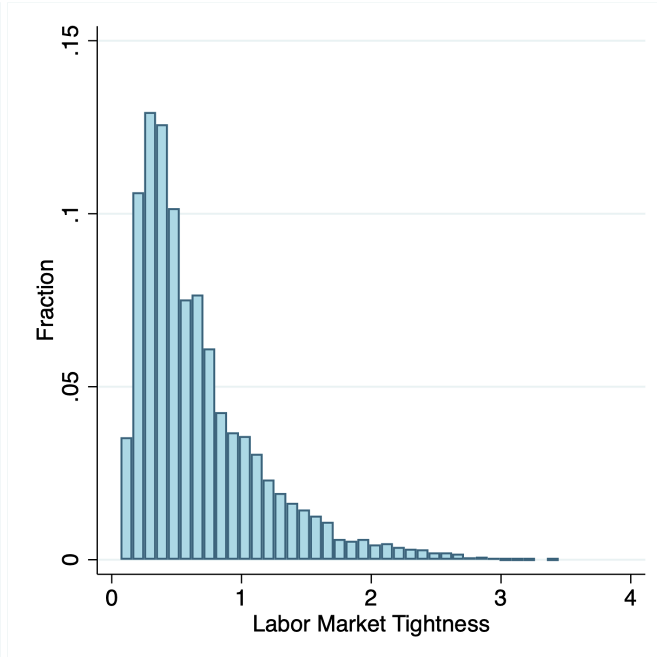
Trigari, Antonella, "The Role of Search Frictions and Bargaining for Inflation Dynamics," 2006.

Main Figures and Tables

Figure 2: Distributions of MSA-level Inflation and Labor Market Tightness, Dec01-Apr23



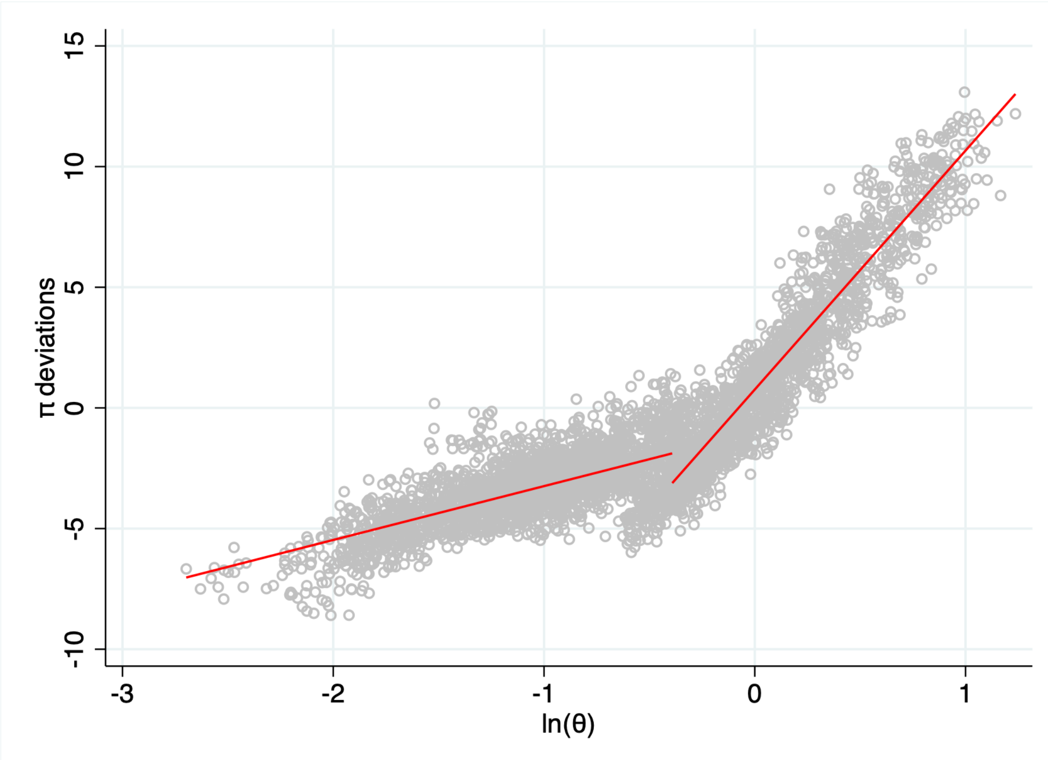
(a) Core Inflation



(b) Labor Market Tightness

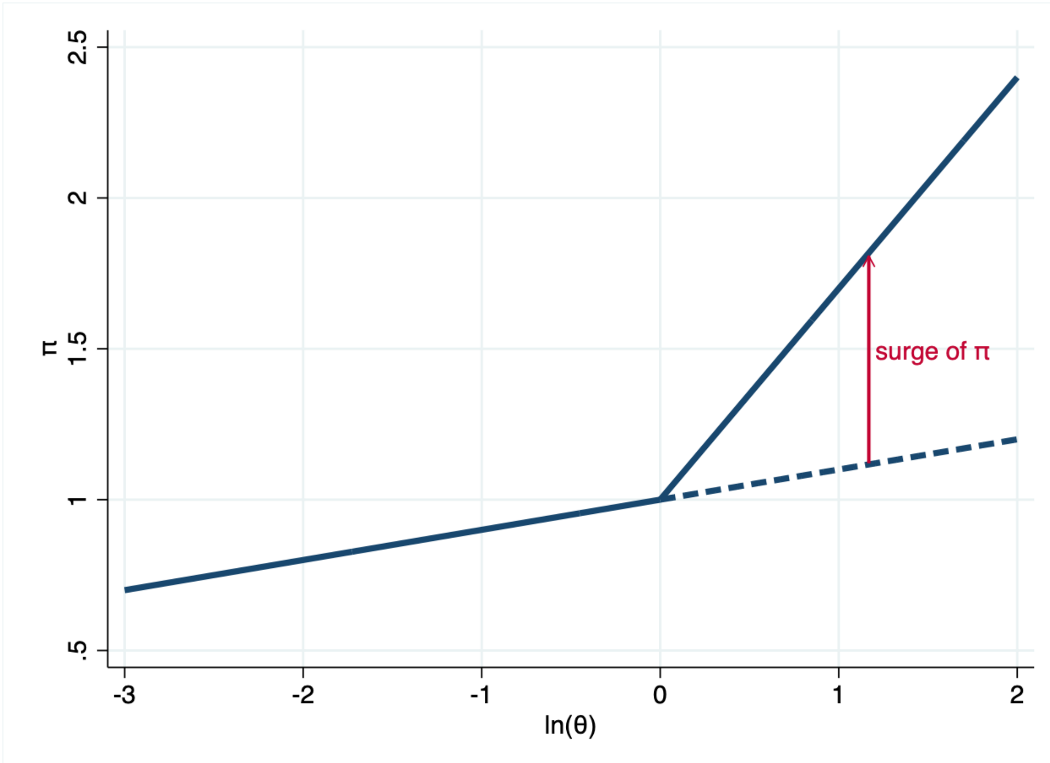
Notes. The figure shows the distribution of core inflation rates (2a) and of labor market tightness (2b) across 21 U.S. metropolitan areas from December 2000 to April 2023. Vacancies across metropolitan areas are imputed combining state-level vacancies collected by BLS JOLTS with population weights from the 2000 and 2010 Census.

Figure 3: Fit of Estimated Phillips Curve



Notes. The figure shows the scatter plot of core inflation deviations and the log of labor market tightness across 21 U.S. metropolitan areas. Inflation deviations are defined as the difference between the 12-month core inflation rate and the estimated controls and fixed effects in column (4) of Table 1. The red lines plot the Phillips curve estimated in column (4) of Table 1.

Figure 4: Implications for Monetary Policy



Notes. The figure shows the implication for monetary policy of the Phillips curve being piecewise log-linear. The blue solid line represents the real Phillips curve, while the blue dashed line denotes the Phillips curve believed to be in place by the monetary authority. The difference between the blue solid and dashed lines after the kink represents the unexpected surge in inflation caused by the central bank ignoring the nonlinearity in the Phillips curve.

Table 1: Estimates of ψ_θ^1 and ψ_θ^2

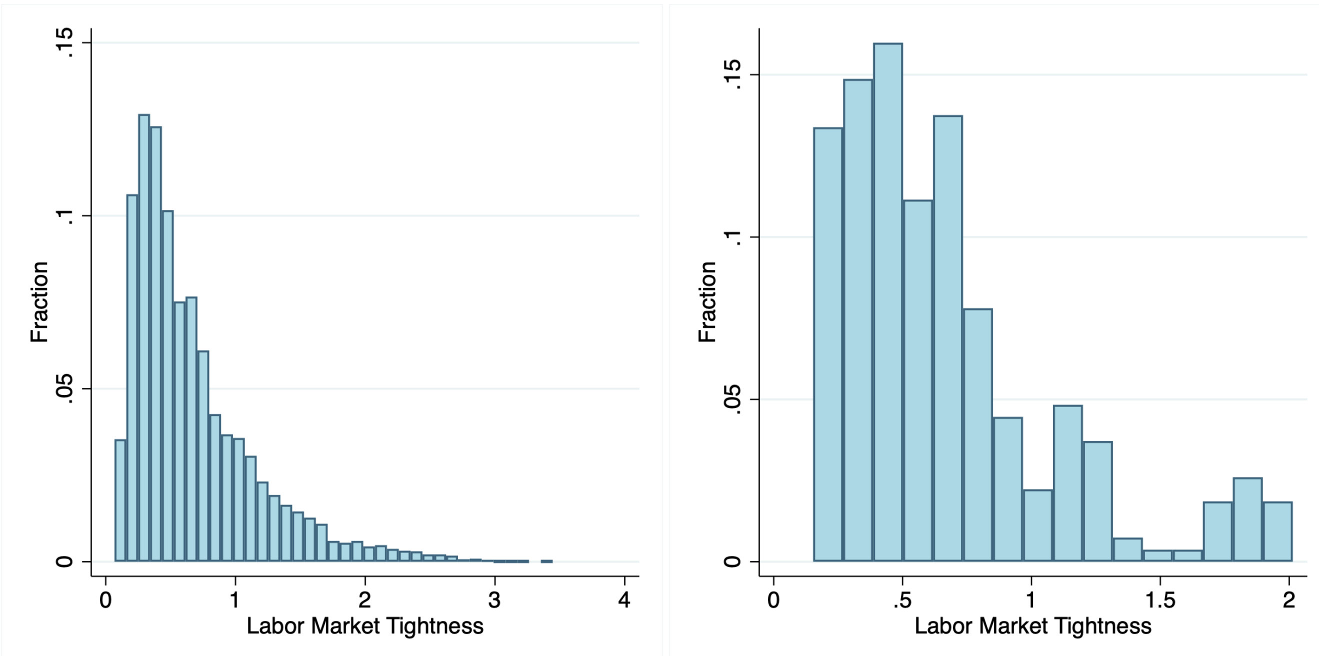
	(1) IV	(2) OLS	(3) OLS
$\ln(\theta_{it})$	3.01**	0.98***	1.21***
	(1.41)	(0.14)	(0.22)
$\ln(\theta_{it}) \times I_{\{\theta_{it}>\theta_i^*\}}$	5.39*	2.26***	0.28
	(2.95)	(0.31)	(0.30)
$I_{\{\theta_{it}>\theta_i^*\}}$	0.57	0.06	0.06
	(1.08)	(0.16)	(0.15)
Observations	4357	4357	4357
MSA FE	✓	✓	✓
Year-Quarter FE			✓
Year-Quarter FE $\times \Sigma_j e_{ji}$	✓		
Controls	✓	✓	✓
F-stat θ	14.64		
F-stat $\theta \times I$	12.65		

Notes. This table presents estimates of ψ_θ^1 and ψ_θ^2 from equation (28) from December 2001 to April 2023. All specifications feature the 12-month core inflation rate as dependent variable and two independent variables: the logarithm of labor market tightness and its interaction with a dummy that takes value of 1 when labor market tightness is greater than θ_i^* and 0 otherwise. θ_i^* represents the average of labor market tightness in MSA i. Column (1) displays IV coefficients, while columns (2) to (3) display OLS coefficients. Columns (1) displays IV estimates of ψ_θ^1 and ψ_θ^2 obtained by instrumenting $\ln(\theta_{it})$ with the shift-share instrument z_{it}^x and the interaction term $\ln(\theta_{it}) \times I_{\theta_{it}>\theta_i^*}$ with the interaction of the shift-share instrument with an dummy that takes value of 1 when the value of the shift-share instrument is above its 75^{th} percentile, $z_{it}^x \times I_{\{z_{it}^x > z_{75}^x\}}$. Column (2) features MSA fixed effects. Column (3) additionally controls for year-quarter fixed effects. All specifications control for the local relative price of intermediate input, p_{it}^x and for the productivity shocks of the non-tradable final-goods sectors, z_{it}^y. Standard errors in parentheses are clustered at the MSA-year level. For each column, first-stage F-statistics from Table A.1 are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix

A Figures and Tables

Figure A.1: Distributions of MSA-level and National Labor Market Tightness, Dec01-Apr23



(a) MSA-level Labor Market Tightness

(b) National Labor Market Tightness

Notes. The figure shows the distribution of labor market tightness across 21 U.S. metropolitan areas (A.1a) and the distribution of labor market tightness at the national level (A.1b) from December 2000 to April 2023. Vacancies across metropolitan areas are imputed combining state-level vacancies collected by BLS JOLTS with population weights from the 2000 and 2010 Census.

Table A.1: First stage coefficients of Equation (28)

	(1) $\ln(\theta_{it})$	(2) $\ln(\theta_{it}) \times I_{\{\theta_{it} > \theta_i^*\}}$
z_{it}^x	5.01***	3.56***
	(1.31)	(0.74)
$z_{it}^x \times I_{\{z_{it}^x > z_{75}^x\}}$	3.28**	-2.48***
	(1.65)	(0.76)
$I_{\{\theta_{it} > \theta_i^*\}}$	0.35***	-0.24***
	(0.02)	(0.02)
Observations	4357	4357
MSA FE	✓	✓
Year-Quarter FE $\times \Sigma_j e_{ji}$	✓	✓
Controls	✓	✓
F-stat	14.64	12.65

Notes. This table presents the first stage regression coefficients for IV estimation of Equation 28. In column (1), the dependent variable is the log of labor market tightness, $\ln(\theta_{it})$. In column (2), the dependent variable is the log of labor market tightness interacted with a dummy that takes value of 1 when labor market tightness is greater than θ_i^* and 0 otherwise, $\ln(\theta_{it}) \times I_{\{\theta_{it} > \theta_i^*\}}$. θ_i^* represents the average of labor market tightness in MSA i. The main independent variables are the shift-share instrument constructed with tradable intermediate-input industries, z_{it}^x, and the interaction between the shift-share instrument and a dummy that takes value of 1 then the value of the shift-share instrument is above the 75^{th} percentile, $z_{it}^x \times I_{\{z_{it}^x > z_{75}^x\}}$. All columns control for MSA fixed effects, year-quarter fixed effects interacted with the sum of intermediate-input industries' exposure shares, the relative intermediate-input prices, and the shift-share proxying for final-goods sectors' productivity shocks. Standard errors in parentheses are clustered at the MSA-year level. F-statistics are reported for each column. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: IV Estimates of Phillips curve with different nonlinear functional forms

	(1) Piecewise Log	(2) Log	(3) Piecewise Lin	(4) Inverse
$\ln(\theta_{it})$	2.08**	3.44***		
	(0.98)	(0.90)		
$\ln(\theta_{it}) \times I_{\{\theta_{it}>\theta_i^*\}}$	3.73*			
	(2.05)			
θ_{it}			3.31*	
			(1.90)	
$\theta_{it} \times I_{\{\theta_{it}>\theta_i^*\}}$			1.23	
			(1.90)	
$\frac{1}{\theta_{it}}$				-2.84***
				(0.74)
Observations	4357	4357	4357	4357
MSA FE	✓	✓	✓	✓
Year-Quarter FE $\times \Sigma_j e_{ji}$	✓	✓	✓	✓
Controls	✓	✓	✓	✓
F-stat $f(\theta)$	14.64	22.84	9.45	24.86
F-stat $f(\theta) \times I$	12.65		11.41	

Notes. This table presents estimates of the slope of the Phillips curve from December 2000 to April 2023. All specifications feature the 12-month, core inflation rate as dependent variable and the standardized vacancy-to-unemployment ratio as the main independent variable, in different functional forms. Column (1) presents my main specification, where labor market tightness is modeled according to a piecewise log-linear function with a kink at θ_i^*. θ_i^* represents the average of labor market tightness in MSA i. Columns (2) to (4) display the following nonlinear transformations: log, piece-wise linear with kink at θ_i^*, and inverse, respectively. All specifications control for MSA fixed effects, year-quarter fixed effects interacted with the sum of intermediate-input industries' exposure shares, the local relative intermediate-input prices, and the shift-share proxy for final-goods sectors' productivity shocks. All columns display IV estimates of the coefficients obtained by instrumenting θ_{it} and its transformations with the shift-share instrument z_{it}^x. The coefficient on the interaction term in column (1) and (3) are obtained by using the interaction of the shift-share instrument with an dummy that takes value of 1 when the value of the shift-share instrument is above its 75^{th} percentile, $z_{it}^x \times I_{\{z_{it}^x > z_{75}^x\}}$. Standard errors in parentheses are clustered at the MSA-year level. For each column, first-stage F-statistics are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: IV Estimates of Phillips curve with unemployment rate

	(1) Log	(2) Piecewise Log	(3) Piecewise Lin	(4) Inverse
$\ln(u_{it})$	-1.44***	-1.17*		
	(0.31)	(0.60)		
$\ln(u_{it}) \times I_{\{u_{it}<5.5\%\}}$		-0.79		
		(1.21)		
u_{it}			-0.95**	
			(0.40)	
$u_{it} \times I_{\{u_{it}<5.5\%\}}$			-1.99	
			(1.44)	
$\frac{1}{u_{it}}$				1.45***
				(0.33)
Observations	6108	5952	5952	6108
MSA FE	✓	✓	✓	✓
Year-Quarter FE $\times \Sigma_j e_{ji}$	✓	✓	✓	✓
Controls	✓	✓	✓	✓
F-stat 1	78.63	48.94	62.11	59.23
F-stat 2		15.57	14.43	

Notes. This table presents estimates of the slope of the Phillips curve from January 1990 to April 2023. All specifications feature the 12-month, core inflation rate as dependent variable and the standardized unemployment rate as the main independent variable, in different functional forms. Columns (2) to (6) display the following nonlinear transformations: log, piecewise log and piecewise lin with kink at 5.5%, and inverse, respectively. All specifications control for MSA fixed effects, year-quarter fixed effects interacted with the sum of intermediate-input industries' exposure shares, the local relative intermediate-input prices, and the shift-share proxy for final-goods sectors' productivity shocks. All columns display IV estimates of the coefficients obtained by instrumenting u_{it} and its transformations with the shift-share instrument z_{it}^x. The coefficient on the interaction terms in column (2) and (3) are obtained by using the interaction between the shift-share instrument and the indicator for the 12-month lag of the unemployment rate being less that 5.5%, $z_{it}^x \times I_{u_{it-12}<5.5\%}$. The coefficient on the quadratic term in column (5) is obtained by using the square of the first-stage predicted unemployment rate as instrument. Standard errors in parentheses are clustered at the MSA-year level. For each column, first-stage F-statistics are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: IV Estimates of Phillips curve with unemployment rate before COVID

	(1) Log	(2) Piecewise Log	(3) Piecewise Lin	(4) Inverse
$\ln(u_{it})$	-1.49***	-1.34**		
	(0.40)	(0.64)		
$\ln(u_{it}) \times I_{\{u_{it}<5.5\%\}}$		-0.62		
		(1.14)		
u_{it}			-1.15**	
			(0.48)	
$u_{it} \times I_{\{u_{it}<5.5\%\}}$			-1.92	
			(1.39)	
$\frac{1}{u_{it}}$				1.48***
				(0.42)
Observations	5310	5154	5154	5310
MSA FE	✓	✓	✓	✓
Year-Quarter FE $\times \Sigma_j e_{ji}$	✓	✓	✓	✓
Controls	✓	✓	✓	✓
F-stat 1	43.27	25.95	34.22	33.17
F-stat 2		12.21	11.71	

Notes. This table presents estimates of the slope of the Phillips curve from January 1990 to February 2020, what I identify as the pre-COVID period. All specifications feature the 12-month, core inflation rate as dependent variable and the standardized unemployment rate as the main independent variable, in different functional forms. Columns (1) to (4) display the following nonlinear transformations: log, piecewise log and piecewise log with kink at 5.5%, and inverse, respectively. All specifications control for MSA fixed effects, year-quarter fixed effects interacted with the sum of intermediate-input industries' exposure shares, the local relative intermediate-input prices, and the shift-share proxy for final-goods sectors' productivity shocks. All columns display IV estimates of the coefficients obtained by instrumenting u_{it} and its transformations with the shift-share instrument z^x_{it}. The coefficient on the interaction terms in column (2) and (3) are obtained by using the indicator for the 12-month lag of the unemployment rate being less that 5.5%, $z^x_{it} \times I_{u_{it-12}<5.5\%}$. Standard errors in parentheses are clustered at the MSA-year level. For each column, first-stage F-statistics are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: IV Estimates of Phillips curve with θ and control for local inflation expectations

	(1) Piecewise Log	(2) Log	(3) Piecewise Lin	(4) Inverse
$\ln(\theta_{it})$	2.15**	3.46***		
	(0.98)	(0.90)		
$\ln(\theta_{it}) \times I_{\{\theta_{it}>\theta_i^*\}}$	3.67*			
	(2.07)			
θ_{it}			3.44*	
			(1.91)	
$\theta_{it} \times I_{\{\theta_{it}>\theta_i^*\}}$			1.14	
			(1.92)	
$\frac{1}{\theta_{it}}$				-2.85***
				(0.74)
π_{it}^{gas}	0.00	0.01**	0.00	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	4353	4353	4353	4353
MSA FE	✓	✓	✓	✓
Year-Quarter FE $\times \Sigma_j e_{ji}$	✓	✓	✓	✓
Controls	✓	✓	✓	✓
F-stat 1	14.71	22.73	9.43	24.76
F-stat 2	12.51		11.28	

Notes. This table presents estimates of the slope of the Phillips curve from December 2000 to April 2023. All specifications feature the 12-month, core inflation rate as dependent variable and the standardized vacancy-to-unemployment ratio as the main independent variable, in different functional forms. Column (1) presents my main specification, where labor market tightness is modeled according to a piecewise log-linear function with a kink at θ_i^*. θ_i^* represents the average of labor market tightness in MSA i. Columns (2) to (4) display the following nonlinear transformations: log, piece-wise linear with kink at θ_i^*, and inverse, respectively. All specifications control for local inflation expectations proxied by the MSA-level growth rate of gasoline prices, together with MSA fixed effects, year-quarter fixed effects interacted with the sum of intermediate-input industries' exposure shares, the local relative intermediate-input prices, and the shift-share proxy for final-goods sectors' productivity shocks. All columns display IV estimates of the coefficients obtained by instrumenting θ_{it} and its transformations with the shift-share instrument z_{it}^x. The coefficient on the interaction term in column (1) and (3) are obtained by using the interaction of the shift-share instrument with an dummy that takes value of 1 when the value of the shift-share instrument is above its 75^{th} percentile, $z_{it}^x \times I_{\{z_{it}^x > z_{75}^x\}}$. Standard errors in parentheses are clustered at the MSA-year level. For each column, first-stage F-statistics are reported. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B Model

B.1 Monetary Authority

The monetary authority implements a common monetary policy across the two regions following the Taylor rule

$$r_t^n = \phi_\pi(\pi_t - \pi_t^*) - \phi_\theta(\hat{\theta}_t - \theta_t^*) + \varepsilon_{rt}, \tag{B.1}$$

where hatted variables represent deviations from a zero-inflation steady state and lower-case variables are logs of upper-case variables. $\pi_t = \zeta\pi_{it} + (1-\zeta)\pi_{jt}$ denotes economy-wide inflation, where $\pi_{it} = p_{it} - p_{it-1}$ is consumer price inflation in region i and π_{jt} is the counterpart in region j. $\hat{\theta}_t = \zeta\hat{\theta}_{it} + (1-\zeta)\hat{\theta}_{jt}$ denotes the deviation of aggregate labor market tightness from its steady-state value. Finally, π_t^* represents a time-varying inflation target. We assume that the monetary authority targets a value for labor market tightness consistent with its long-run inflation target, i.e. $\theta_t^* = \frac{(1-\beta)}{\kappa_\theta}\pi_t^*$. Finally, ϕ_π and ϕ_u ensure a unique locally bounded equilibrium, and ε_{rt} denotes a transitory monetary shock, assumed to follow an AR(1) process. The model in its simplest form abstracts from fiscal policy, as the government does not tax, spend, nor issues debt, and monetary policy has no fiscal implications.

B.2 Derivation of Regional Phillips Curve

From the problem of the final-goods firm in region i described in Section 2, I derive the following optimal pricing condition:

$$\sum_{k=0}^{\infty} \alpha^k E_t \left[Q_{it,t+k} Y_{it+k}(z) \left(\frac{P_{it}(z)}{P_{it-1}} - \frac{\epsilon}{\epsilon - 1} RMC_{it+k} \frac{P_{it+k}}{P_{it-1}} \right) \right] = 0, \tag{B.2}$$

where RMC_{it} denotes firms' real marginal costs, defined as $RMC_{it} \equiv \frac{1}{A_{it}^y} \left(\frac{P_t^x}{P_{it}(1-\phi)} \right)^{1-\phi} \left(\frac{W_{it}}{P_{it}\phi} \right)^{\phi}$. Log-linearazing Equation (B.2) around the zero inflation steady state yields

$$p_{it}(z) - p_{it-1} = (1-\alpha\beta) \sum_{k=0}^{\infty} (\alpha\beta)^\kappa E_t \left[r\hat{m}c_{it+k} - (p_{it+k} - p_{it-1}) \right],$$

where

$$r\hat{m}c_{it} = -\hat{a}_{it}^y + (1-\phi)(p_t^x - p_{it}) + \phi\hat{w}_{it}. \tag{B.3}$$

Rearranging the equation, I obtain

$$p_{it}(z) - p_{it-1} = \alpha\beta E_t \left[p_{it+1}(z) - p_{it}\right] + (1-\alpha\beta) r\hat{m}c_{it} + \pi_{it}, \tag{B.4}$$

where π_{it} is derived from the definition of the price index in Equation (11). Indeed, only $(1-\alpha)$ firms are able to reset their price, and since they are faced by the same probability of changing price in the future and the same current and expected same marginal costs, they will choose the same price P_{it}^*. Hence, the price index becomes

$$P_{it}^{1-\epsilon} = \alpha P_{it-1}^{1-\epsilon} + (1-\alpha) P_{it}^{*1-\epsilon}.$$

Taking a log-linear approximation of this last expression yields

$$p_{it} = \alpha p_{it-1} + (1-\alpha) p_{it}^*,$$

which implies

$$\pi_{it} = (1-\alpha)(p_{it}^* - p_{it}). \tag{B.5}$$

Substituting Equation (B.5) in Equation (B.4), after some manipulations I obtain

$$\pi_{it} = \beta E_t \pi_{it+1} + \delta r\hat{m}c_{it}, \tag{B.6}$$

where

$$\delta = \frac{(1-\alpha\beta)(1-\alpha)}{\alpha}.$$

Combining Equations (B.3) and (B.6), I get

$$\pi_{it} = \beta E_t \pi_{it+1} + \delta(1-\phi)(p_t^x - p_{it}) + \delta\phi\hat{w}_{it} - \delta\hat{a}_{it}^y. \tag{B.7}$$

The log-linearized expression for the wage norm is

$$\hat{w}_{it} = \begin{cases} \hat{w}_{it}^{flex} = \eta\hat{\theta}_{it} - \hat{m}_{it} + \hat{\gamma}_{it}^c - \hat{\gamma}_{it}^b & \hat{\theta}_{it} > \hat{\theta}_i^* \\ 1 - \lambda\hat{w}_{it}^{flex} & \hat{\theta}_{it} \leq \hat{\theta}_i^* \end{cases}. \tag{B.8}$$

Substituting Equation (B.8) into Equation (B.7), I obtain the regional Phillips curve

$$\pi_{it} = \begin{cases} \beta E_t \pi_{it+1} + \kappa_\theta^{tight} \hat{\theta}_{it} + \kappa_p \hat{p}_{it}^x + \kappa_\nu^{tight} \hat{\nu}_{it} + \kappa_a \hat{a}_{it}^y & \hat{\theta}_{it} > \hat{\theta}_i^* \\ \beta E_t \pi_{it+1} + \kappa_\theta \hat{\theta}_{it} + \kappa_p \hat{p}_{it}^x + \kappa_\nu \hat{\nu}_{it} + \kappa_a \hat{a}_{it}^y & \hat{\theta}_{it} \leq \hat{\theta}_i^* \end{cases}, \tag{B.9}$$

where

- $\kappa_\theta^{tight} = \delta\phi\eta$,
- $\kappa_\theta = (1-\lambda)\delta\phi\eta$,
- $\kappa_p = \delta(1-\phi)$,
- $\hat{p}_{it}^x = p_t^x - p_{it}$,
- $\kappa_\nu^{tight} = \delta\phi$,
- $\kappa_\nu = (1-\lambda)\delta\phi$,
- $\hat{\nu}_{it} = -\hat{m}_{it} + \hat{\gamma}_{it}^c - \hat{\gamma}_{it}^b$,
- $\kappa_a = -\delta$.

B.3 Solving Forward

To derive equation 27, I solve equation B.9 forward. I assume that if the labor market is slack in t, then it will remain slack forever. If the labor market is tight in t, it will remain tight until $t+T-1$. From period $t+T$, the labor market becomes slack and will remain slack forever. Hence, the long-run component of the variation of labor market tightness corresponds to a slack labor market. I make this assumption because labor markets in the U.S. are generally slack. The episodes of tight labor markets are indeed quite rare.

Consider the case in which the labor market in region i at time t is slack. Equation B.9 implies that inflation is equal to

$$\pi_{it}^{slack} = \beta E_t \pi_{it+1}^{slack} + \kappa_\theta \hat{\theta}_{it} + \kappa_p \hat{p}_{it}^x + \kappa_a \hat{a}_{it}^y + \kappa_\nu \hat{\nu}_{it}. \tag{B.10}$$

Solving equation B.10 forward, I obtain

$$\pi_{it}^{slack} = E_t \sum_{k=0}^{\infty} \beta^k \kappa_\theta \hat{\theta}_{it+k} + E_t \sum_{k=0}^{\infty} \beta^k \left[\kappa_p \hat{p}_{it+k}^x + \kappa_a \hat{a}_{it+k}^y + \kappa_\nu \hat{\nu}_{it+k} \right]. \tag{B.11}$$

Following Hazell et al. (2022), I decompose the variation of future labor market tightness $\hat{\theta}_{it+k}$ into a transitory and a permanent component. The transitory component is defined as $\tilde{\theta}_{it} = \hat{\theta}_{it} + E_t\hat{\theta}_{it+\infty}$, where $E_t\hat{\theta}_{it+\infty}$ is the permanent component of variation in labor market tightness. Applying the decomposition to equation B.11, I get

$$\pi_{it}^{slack} = E_t \sum_{k=0}^{\infty} \beta^k \kappa_\theta \tilde{\theta}_{it+k} + \frac{\kappa_\theta}{1-\beta} E_t \hat{\theta}_{it+\infty} + E_t \sum_{k=0}^{\infty} \beta^k \left[\kappa_p \hat{p}_{it+k}^x + \kappa_a \hat{a}_{it+k}^y + \kappa_\nu \hat{\nu}_{it+k} \right]. \tag{B.12}$$

Assuming that the shocks $\hat{p}_{it}^x$, $\hat{a}_{it}^y$, and $\hat{\nu}_{it}$ are transitory, and the labor market is slack in the long run, equation B.9 implies that $E_t\pi_{it+\infty} = \frac{\kappa_\theta}{1-\beta}E_t\theta_{it+\infty}$. Moreover, as $E_t\pi_{it+\infty}$ represents the long-run belief about the monetary policy regime and such regime is common between the two regions, then $E_t\pi_{it+\infty} = E_t\pi_{jt+\infty} = E_t\pi_{t+\infty}$. Substituting, I obtain

$$\pi_{it}^{slack} = E_t\pi_{t+\infty} + E_t \sum_{k=0}^{\infty} \beta^k \kappa_\theta \tilde{\theta}_{it+k} + E_t \sum_{k=0}^{\infty} \beta^k \left[\kappa_p \hat{p}_{it+k}^x + \kappa_a \hat{a}_{it+k}^y + \kappa_\nu \hat{\nu}_{it+k} \right]. \tag{B.13}$$

Assuming that $\tilde{\theta}_{it}$, $\hat{p}_{it}^x$, and $\hat{a}_{it}^y$ follow AR(1) processes with autocorrelation coefficients ρ_θ, ρ_p, and ρ_{a^y}, the regional Phillips curve when the labor market is slack at time t takes the following form:

$$\pi_{it}^{slack} = E_t\pi_{t+\infty} + \psi_\theta^1 \tilde{\theta}_{it} + \psi_p \hat{p}_{it}^x + \psi_a \hat{a}_{it}^y + \varepsilon_{it}, \tag{B.14}$$

where

- $\psi_\theta^1 = \frac{\kappa_\theta}{(1-\beta\rho_\theta)}$,
- $\psi_p = \frac{\kappa_p}{(1-\beta\rho_p)}$,
- $\psi_a = \frac{\kappa_a}{(1-\beta\rho_a)}$,
- $\varepsilon_{it} = E_t \sum_{k=0}^{\infty} \beta^k \kappa_\nu \hat{\nu}_{it+k}$.

Let's turn to the case in which the labor market in region i at time t is tight. I assume that the labor market remains tight until $t+T-1$. At $t+T$ the labor market becomes slack and remains slack forever. From equation B.9, inflation in region i at time $t+T$ is equal to

$$\pi_{it+T}^{slack} = \beta E_{t+T}\pi_{it+T+1}^{slack} + \kappa_\theta \hat{\theta}_{it+T} + \kappa_p \hat{p}_{it+T}^x + \kappa_a \hat{a}_{it+T}^y + \kappa_\nu \hat{\nu}_{it+T}. \tag{B.15}$$

Solving equation B.15 forward, I obtain

$$\pi_{it+T}^{slack} = E_{t+T} \sum_{k=T}^{\infty} \beta^{k-T} \kappa_\theta \hat{\theta}_{it+k} + E_{t+T} \sum_{k=T}^{\infty} \beta^{k-T} \left[\kappa_p \hat{p}_{it+k}^x + \kappa_a \hat{a}_{it+k}^y + \kappa_\nu \hat{\nu}_{it+k} \right]. \tag{B.16}$$

Decomposing the variation of labor market tightness in the transitory and permanent components, I get

$$\pi_{it+T}^{slack} = E_{t+T} \sum_{k=T}^{\infty} \beta^{k-T} \left(\kappa_\theta \tilde{\theta}_{it+k} + \kappa_p \hat{p}_{it+k}^x + \kappa_a \hat{a}_{it+k}^y + \kappa_\nu \hat{\nu}_{it+k} \right) + \sum_{k=T}^{\infty} \beta^{k-T} \kappa_\theta E_{t+T} \hat{\theta}_{it+\infty}. \tag{B.17}$$

Finally, I assume that $\tilde{\theta}_{it}$, $\hat{p}_{it}^x$, and $\hat{a}_{it}^y$ follow AR(1) processes with autocorrelation coefficients ρ_θ, ρ_p, and ρ_{a^y}. Equation B.17 becomes

$$\begin{aligned}
\pi_{it+T}^{slack} = & \sum_{k=T}^{\infty} (\beta\rho_\theta)^{k-T} \kappa_\theta \tilde{\theta}_{it+T} + \sum_{k=T}^{\infty} \beta^{k-T} \kappa_\theta E_{t+T} \hat{\theta}_{it+\infty} + \sum_{k=T}^{\infty} (\beta\rho_p)^{k-T} \kappa_p \hat{p}_{it+T}^x \\
& + \sum_{k=T}^{\infty} (\beta\rho_a)^{k-T} \kappa_a \hat{a}_{it+T}^y + E_{t+T} \sum_{k=T}^{\infty} \beta^{k-T} \kappa_\nu \hat{\nu}_{it+k}.
\end{aligned} \tag{B.18}$$

From period t to $t+T-1$, the labor market is tight. From equation B.9, inflation in region i at time $t+T-1$ is equal to

$$\pi_{it+T-1}^{tight} = \beta E_{t+T-1} \pi_{it+T}^{slack} + \kappa_\theta \hat{\theta}_{it+T-1} + \kappa_p \hat{p}_{it+T-1}^x + \kappa_a \hat{a}_{it+T-1}^y + \kappa_\nu^{tight} \hat{\nu}_{it+T-1}. \tag{B.19}$$

Solving backward, I obtain

$$\begin{aligned}
\pi_{it+T-2}^{tight} = & \beta E_{t+T-2} \pi_{it+T-1}^{tight} + \kappa_\theta \hat{\theta}_{it+T-2} + \kappa_p \hat{p}_{it+T-2}^x + \kappa_a \hat{a}_{it+T-2}^y + \kappa_\nu^{tight} \hat{\nu}_{it+T-2} \\
= & \beta E_{t+T-2} \left[\beta E_{t+T-1} \pi_{it+T}^{slack} + \kappa_\theta \hat{\theta}_{it+T-1} + \kappa_p \hat{p}_{it+T-1}^x + \kappa_a \hat{a}_{it+T-1}^y + \kappa_\nu^{tight} \hat{\nu}_{it+T-1} \right] \\
& + \kappa_\theta \hat{\theta}_{it+T-2} + \kappa_p \hat{p}_{it+T-2}^x + \kappa_a \hat{a}_{it+T-2}^y + \kappa_\nu^{tight} \hat{\nu}_{it+T-2} \\
= & \beta^2 E_{t+T-2} \pi_{it+T}^{slack} + \kappa_\theta \left(\beta E_{t+T-2} \hat{\theta}_{it+T-1} + \hat{\theta}_{it+T-2} \right) + \kappa_p \left(\beta E_{t+T-2} \hat{p}_{it+T-1}^x + \hat{p}_{it+T-2}^x \right) \\
& + \kappa_a \left(\beta E_{t+T-2} \hat{a}_{it+T-1}^y + \hat{a}_{it+T-2}^y \right) + \kappa_\nu^{tight} \left(\beta E_{t+T-2} \hat{\nu}_{it+T-1} + \hat{\nu}_{it+T-2} \right) \\
& \ldots \\
\pi_{it}^{tight} = & \beta^T E_t \pi_{it+T}^{slack} + E_t \sum_{k=0}^{T-1} \beta^k \left(\kappa_\theta^{tight} \hat{\theta}_{it+k} + \kappa_p \hat{p}_{it+k}^x + \kappa_a \hat{a}_{it+k}^y + \kappa_\nu^{tight} \hat{\nu}_{it+k} \right).
\end{aligned} \tag{B.20}$$

Decomposing the variation of labor market tightness into the transitory and permanent components, I get

$$\pi_{it}^{tight} = \beta^T E_t \pi_{it+T}^{slack} + E_t \sum_{k=0}^{T-1} \beta^k \left(\kappa_\theta^{tight} \tilde{\theta}_{it+k} + \kappa_p \hat{p}_{it+k}^x + \kappa_a \hat{a}_{it+k}^y + \kappa_\nu^{tight} \hat{\nu}_{it+k} \right) + \sum_{k=0}^{T-1} \beta^k \kappa_\theta^{tight} E_t \hat{\theta}_{it+\infty}. \tag{B.21}$$

Assuming that $\tilde{\theta}_{it}$, $\hat{p}_{it}^x$, and $\hat{a}_{it}^y$ follow AR(1) processes with autocorrelation coefficients ρ_θ, ρ_p, and ρ_{a^y}, I obtain

$$\pi_{it}^{tight} = \beta^T E_t \pi_{it+T}^{slack} + \sum_{k=0}^{T-1} (\beta\rho_\theta)^k \kappa_\theta^{tight} \tilde{\theta}_{it} + \sum_{k=0}^{T-1} \beta^k \kappa_\theta^{tight} E_t \hat{\theta}_{it+\infty} + \sum_{k=0}^{T-1} (\beta\rho_p)^k \kappa_p \hat{p}_{it}^x + \sum_{k=0}^{T-1} (\beta\rho_a)^k \kappa_a \hat{a}_{it}^y + E_t \sum_{k=0}^{T-1} \beta^k \kappa_\nu^{tight} \hat{\nu}_{it+k}. \tag{B.22}$$

Finally, I substitute equation B.18 into equation B.22 and obtain

$$\begin{aligned}
\pi_{it}^{tight} =& \beta^T E_t \Bigg[\sum_{k=T}^{\infty} (\beta\rho_\theta)^{k-T} \kappa_\theta \tilde{\theta}_{it+T} + \sum_{k=T}^{\infty} \beta^{k-T} \kappa_\theta E_{t+T} \hat{\theta}_{it+\infty} + \sum_{k=T}^{\infty} (\beta\rho_p)^{k-T} \kappa_p \hat{p}_{it+T}^x \\
&+ \sum_{k=T}^{\infty} (\beta\rho_a)^{k-T} \kappa_a \hat{a}_{it+T}^y + E_{t+T} \sum_{k=T}^{\infty} \beta^{k-T} \kappa_\nu \hat{\nu}_{it+k} \Bigg] \\
&+ \sum_{k=0}^{T-1} (\beta\rho_\theta)^k \kappa_\theta^{tight} \tilde{\theta}_{it} + \sum_{k=0}^{T-1} \beta^k \kappa_\theta^{tight} E_t \hat{\theta}_{it+\infty} + \sum_{k=0}^{T-1} (\beta\rho_p)^k \kappa_p \hat{p}_{it}^x \\
&+ \sum_{k=0}^{T-1} (\beta\rho_a)^k \kappa_a \hat{a}_{it}^y + E_t \sum_{k=0}^{T-1} \beta^k \kappa_\nu^{tight} \hat{\nu}_{it+k} \\
=& \sum_{k=0}^{T-1} (\beta\rho_\theta)^k \kappa_\theta^{tight} \tilde{\theta}_{it} + \sum_{k=T}^{\infty} (\beta\rho_\theta)^{k-T} \beta^T \kappa_\theta E_t \tilde{\theta}_{it+T} + \left[\kappa_\theta^{tight} \sum_{k=0}^{T-1} \beta^k + \kappa_\theta \sum_{k=T}^{\infty} \beta^k \right] E_t \hat{\theta}_{it+\infty} \\
&+ \sum_{k=0}^{T-1} (\beta\rho_p)^k \kappa_p \hat{p}_{it}^x + \sum_{k=T}^{\infty} (\beta\rho_p)^{k-T} \beta^T \kappa_p E_t \hat{p}_{it+T}^x + \sum_{k=0}^{T-1} (\beta\rho_a)^k \kappa_a \hat{a}_{it}^y \\
&+ \sum_{k=T}^{\infty} (\beta\rho_a)^{k-T} \beta^T \kappa_a E_t \hat{a}_{it+T}^y + E_t \left[\sum_{k=0}^{T-1} \beta^k \kappa_\nu^{tight} \hat{\nu}_{it+k} + \sum_{k=T}^{\infty} \beta^k \kappa_\nu \hat{\nu}_{it+k} \right].
\end{aligned} \tag{B.23}$$

Assuming that $\tilde{\theta}_{it}$, $\hat{p}_{it}^x$, and $\hat{a}_{it}^y$ follow AR(1) processes with autocorrelation coefficients ρ_θ, ρ_p, and ρ_{a^y}, equation B.23 becomes

$$\pi_{it}^{tight} = \left[\frac{1-\beta^T}{1-\beta}\kappa_\theta^{tight} + \frac{\beta^T}{1-\beta}\kappa_\theta\right] E_t\hat{\theta}_{it+\infty} + \left[\frac{1-(\beta\rho_\theta)^T}{1-\beta\rho_\theta}\kappa_\theta^{tight} + \frac{(\beta\rho_\theta)^T}{1-\beta\rho_\theta}\kappa_\theta\right]\tilde{\theta}_{it}$$
$$+ \frac{1}{1-\beta\rho_p}\kappa_p\hat{p}_{it}^x + \frac{1}{1-\beta\rho_a}\kappa_a\hat{a}_{it}^y + \varepsilon_{it}^{tight}, \tag{B.24}$$

where $\varepsilon_{it}^{tight} = E_t\left[\sum_{k=0}^{T-1}\beta^k\kappa_\nu^{tight}\hat{\nu}_{it+k} + \sum_{k=T}^{\infty}\beta^k\kappa_\nu\hat{\nu}_{it+k}\right]$. Manipulating the coefficients on $E_t\hat{\theta}_{it+\infty}$ and $\tilde{\theta}_{it}$, I obtain

$$\pi_{it}^{tight} = \left[\frac{\kappa_\theta}{1-\beta} + \frac{1-\beta^T}{1-\beta}\left(\kappa_\theta^{tight} - \kappa_\theta\right)\right] E_t\hat{\theta}_{it+\infty} + \left[\frac{\kappa_\theta}{1-\beta\rho_\theta} + \frac{1-(\beta\rho_\theta)^T}{1-\beta\rho_\theta}\left(\kappa_\theta^{tight} - \kappa_\theta\right)\right]\tilde{\theta}_{it}$$
$$+ \psi_p\hat{p}_{it}^x + \psi_a\hat{a}_{it}^y + \varepsilon_{it}^{tight}$$
$$= E_t\pi_{t+\infty} + \frac{1-\beta^T}{1-\beta}\lambda\kappa_\theta^{tight}E_t\hat{\theta}_{it+\infty} + \left(\psi_\theta^1 + \frac{1-(\beta\rho_\theta)^T}{1-\beta\rho_\theta}\lambda\kappa_\theta^{tight}\right)\tilde{\theta}_{it} + \psi_p\hat{p}_{it}^x + \psi_a\hat{a}_{it}^y + \varepsilon_{it}^{tight}$$
$$= E_t\pi_{t+\infty} + \xi E_t\hat{\theta}_{it+\infty} + \left(\psi_\theta^1 + \psi_\theta^2\right)\tilde{\theta}_{it} + \psi_p\hat{p}_{it}^x + \psi_a\hat{a}_{it}^y + \varepsilon_{it}^{tight}, \tag{B.25}$$

where

- $\xi = \frac{1-\beta^T}{1-\beta}\lambda\kappa_\theta^{tight}$
- $\psi_\theta^2 = \frac{1-(\beta\rho_\theta)^T}{1-\beta\rho_\theta}\lambda\kappa_\theta^{tight}$

Putting together equations B.14 and B.25 the regional Phillips curve takes the following form:

$$\pi_{it} = \begin{cases} E_t\pi_{t+\infty} + \xi E_t\hat{\theta}_{it+\infty} + \psi_\theta^{tight}\tilde{\theta}_{it} + \psi_p\hat{p}_{it}^x + \psi_a\hat{a}_{it}^y + \varepsilon_{it}^{tight} & \tilde{\theta}_{it} > \tilde{\theta}_{it}^* \\ E_t\pi_{t+\infty} + \psi_\theta^{slack}\tilde{\theta}_{it} + \psi_p\hat{p}_{it}^x + \psi_a\hat{a}_{it}^y + \varepsilon_{it} & \tilde{\theta}_{it} \le \tilde{\theta}_{it}^* \end{cases}, \tag{B.26}$$

where

- $\psi_\theta^{slack} = \psi_\theta^1$
- $\psi_\theta^{tight} = \psi_\theta^1 + \psi_\theta^2$