ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

The Evolution of Regional Beveridge Curves

Authors	Michael T. Owyang, Hannah G. Shell, and Daniel Soques
Working Paper Number	2022-037B
Revision Date	October 2025
Citable Link	https://doi.org/10.20955/wp.2022.037
Suggested Citation	Owyang, M.T., Shell, H.G., Soques, D., 2025; The Evolution of Regional Beveridge Curves, Federal Reserve Bank of St. Louis Working Paper 2022-037. URL https://doi.org/10.20955/wp.2022.037

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

The Evolution of Regional Beveridge Curves*

Michael T. Owyang†

Hannah G. Shell‡

Daniel Soques§

keywords: mismatch, vacancies, clustering

October 24, 2025

Abstract

The slow recovery of the labor market in the aftermath of the Great Recession highlighted mismatch, or the misallocation of workers across space or across industries. We consider the historical evolution of regional mismatch. We construct MSA-level unemployment rates and vacancy data using techniques similar to Barnichon (2010) and a new dataset of online help-wanted ads by MSA. We estimate regional Beveridge curves, identifying the slopes by restricting them to be equal across locations with similar labor market characteristics. We find that the 51 U.S. cities in our sample have four groupings which are influenced by industry classification, union membership, and geographic proximity. Additionally, allowing for a structural break suggests match elasticity decreased across regions over time.

[JEL codes: C33; J63]

*The authors benefitted from discussions with David Andolfatto, Mitch Downey, Victoria Gregory, Thomas Lubik, Howard Wall and David Wiczer. Seminar participants at Warwick Business School, the University of York, the University of Indiana, Michigan State University, and UCSD provided comments and suggestions. Julie K. Bennett, Brooke Hathhorn, Marie Hogan, Constanza Liborio, Ashley Stewart, and Kate Vermann provided research assistance. Views expressed here are the authors' alone and do not reflect the opinions of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

†Research Division. Federal Reserve Bank of St. Louis. Corresponding author: Michael.T.Owyang@stls.frb.org

‡Research Division. Federal Reserve Bank of St. Louis.

§Department of Economics and Finance. University of North Carolina Wilmington

1 Introduction

The slow recovery of the labor market after the Great Recession led some researchers to investigate the potential misallocation of workers across industries and regions. For example, Valletta (2013) argues that the elevated levels of unemployment that persisted after the Great Recession may be related to workers' inability to relocate to areas with more jobs. The primary rigidity arises from a decline in real estate prices that prevents workers from selling their houses and moving to areas with more fertile labor markets.[1] Şahin, Song, Topa, and Violante (2014, SSTV) further argue that skill mismatch—workers allocated to the wrong industry—became more prevalent after the Great Recession. SSTV develop a method to measure the misallocation of workers across industries or across space but found no significant role for regional mismatch at the state or county level post-Great Recession.

The SSTV measure of mismatch relies on an estimate of the slope of the Beveridge curve, an empirical relationship between unemployment and vacancies. More than just a reduced-form construct, the Beveridge curve can be obtained from any number of matching models of the labor market and can have implications for estimates of match efficiency, mismatch, hysteresis, etc. In particular, if one assumes a Cobb-Douglas matching function in steady state with a constant separation rate, the slope of the Beveridge curve returns the match elasticities. These elasticities capture the matching technologies, which, in principle, could vary by location and over time.[2]

One of the main issues with the estimation of the Beveridge curve is differentiating between *shifts* in the Beveridge curve from *movements along* the Beveridge curve. Previous papers have identified shifts in a variety of ways, including eyeballing changes, allowing time-varying parameters, or imposing deterministic trends. A relatively small number of papers have used panel data, constructing regions by imposing that the Beveridge curves of the cities/counties/states within a region have the same slopes and experience the same deterministic trends [Courtney (1991) for the

[1] Ferreira, Gyourko, and Tracy (2010) found that negative housing equity reduced labor mobility during the Great Recession; Schulhofer-Wohl (2011), however, found no relationship. Sterk (2015) and Head and Lloyd-Ellis (2012) construct theoretical models where the inability to sell one's house is an essential friction for cross-regional heterogeneity in the unemployment rate. Alvarez and Shimer (2011) differentiate between search and rest unemployment, where the former is modeled as an active search in another industry but could be viewed as search outside one's home city.

[2] As noted by Petrongolo and Pissarides (2001), a number of issues arise when using estimated Beveridge curves to infer features about the underlying matching function. Direct estimation of the matching function using flow data would be preferred but these data are currently limited or unavailable.

U.S. and Jones and Manning (1992) for the U.K.] or the same time fixed effects [Börsch-Supan (1991) and Wall and Zoëga (2002)]. Most of these papers construct regions geographically. In the U.S., these types of restrictions may not be optimal as geographic proximity does not guarantee that cities or states will have similar business cycle characteristics [see Owyang, Piger, and Wall (2005) and Owyang, Piger, Wall, and Wheeler (2008)].

Instead, we can define regions endogenously, as in Hamilton and Owyang (2012), based on similarities of their Beveridge curves. We use the Metropolitan Statistical Area (MSA) as our preferred geographical unit of analysis.[3] While we make similar identifying assumptions as Börsch-Supan (1991) and Wall and Zoëga (2002) about the within-region slopes and fixed effects, we relax the restriction that the regions are known *ex ante*. Instead, we allow the data to group the cities into regions. Moreover, we can determine whether geography plays a role in forming regions by adopting a logistic prior on regional membership. Thus, we can test whether factors such as Census region, industrial composition, industrial concentration, and the city's level of unionization in its labor force are determinants in the responsiveness of unemployment to changes in vacancies.

While our analysis is in the similar vein as previous papers quantifying mismatch, we do not limit our analysis to the Great Recession and its aftermath. Instead, we construct MSA-level vacancy, unemployment, and labor force data for a time sample of 1978 to 2021 using the 1990 MSA definitions. We compile vacancies from multiple sources, including both print advertisements and new, proprietary data on internet job postings. Unemployment and labor force data prior to the 1990 MSA reclassification are constructed using microdata from the CPS. We estimate the Beveridge curves for the MSAs for the full sample to determine the number of regions, the composition of the regions, and the matching elasticities across cities.

We find evidence suggesting that, for the full sample, the 51 cities we analyzed are best grouped into 4 regions. Interestingly, our regions do not appear to be geographically-based; instead, characteristics of the cities' industries (unionization, industry concentration) appear to influence the slope and position of cities' Beveridge curves. In particular, each region's composition depends on a different aspect of the city-level labor market. For example, higher union membership is correlated

[3] Other papers [e.g., SSTV and Anderson and Burgess (2000)] use state-level data. In the context of Alvarez and Shimer (2011), the cost of searching outside of one's city would likely come with a relocation cost. Such costs might not be incurred while searching across states (e.g., cities that overlap several states) or counties (e.g., cities that encompass several counties). For these reasons, we contend that the proper level of disaggregation is at the city level.

with a flatter Beveridge curve; larger shares of manufacturing and/or financial and professional business services employment steepens the Beveridge curve. We also consider a model with an endogenous break in the number and compositions of the regions. The break model suggests a decrease in match elasticity over time.

The balance of the paper is outlined as follows: Section 2 describes both the national Beveridge curve and our version of the regional Beveridge curve. Section 3 describes the data, prior, and estimation, paying particular attention to the determination of the regions. Sections 4 and 5 describe the full sample and endogenous break results. Section 6 summarizes and concludes.

2 Model

In order to evaluate the misallocation of workers across cities, we estimate city-level Beveridge curves grouped into regions and describe the estimation of the clustered panel model.

2.1 The National Beveridge Curve

The Beveridge curve characterizes the empirical relationship between the unemployment rate (U_t) and vacancies (V_t):

$$u_t = \alpha_t + \beta v_t + \varepsilon_t, \tag{1}$$

where $u_t = \ln U_t$, $v_t = \ln V_t$, α_t is a (possibly time-varying) intercept, β is a time-invariate slope coefficient, and $\varepsilon_t \sim N\left(0, \sigma^2\right)$. One can derive a Beveridge curve type relationship from a simple matching framework (e.g., a Cobb-Douglas matching function with constant separation), where the slope of the Beveridge curve is a function of the parameters in the matching function.

The critical issue when estimating the relationship is identifying shifts of versus movements along the Beveridge curve. The former are the typical characterization of the business cycle effects on labor markets, reflecting the trade-off between unemployment and vacancies. The latter is often thought to represent structural changes in labor markets that raise or lower both unemployment and vacancies.

Past studies have identified shifts in the Beveridge curve by plotting the data and "eyeballing" the timing of the shifts [Jackman, Pissarides, and Savouri (1990); Bean (1994); Gregg and Petrongolo (1997); and Gartner (1997)]. Once the timing is identified, the size of the shift can be estimated using additional dummy variables in the Beveridge curve equation. Alternatively, shifts can be identified by imposing structure on the time variation of the intercept (e.g., random walk time-varying parameters [Benati and Lubik (2014)]) or imposing a deterministic—either linear or quadratic—trend [Courtney (1991) and Jones and Manning (1992)]. After the trend is extracted, the movements along the Beveridge curve can be estimated and the resulting model parameters are identified.

An alternative method for identifying the slope of the Beveridge curve exploits the cross-section of city labor markets. Labor markets are localized objects, with only small labor flows from one city to another.[4] City-level data is used to estimate the aggregate Beveridge curve by imposing cross-equation restrictions on the slopes but allowing for city-level variation in the intercepts. The cross-equation restriction that identifies shifts-in from movements-along the Beveridge curve obtains from a matching technology that is common across the cities, making the slopes of the city Beveridge curves identical.

In addition to imposing that they have the same slope, some other previous papers [e.g., Börsch-Supan (1991) and Wall and Zoëga (2002, WZ)] impose that the shifts in the city-level Beveridge curves occur simultaneously. The idea here is that structural innovations to the matching markets occur homogeneously across the country, either because of the business cycle [as in Blanchard and Diamond (1989)] or because of mismatch, search effectiveness, or hysteresis. The simultaneous shifting of the Beveridge curve is captured by time fixed effects.

Then, city n's Beveridge curve has the form:

$$u_{nt} = \alpha_n + \tau D_t + \beta v_{nt} + \varepsilon_{nt}, \tag{2}$$

where α_n is a time-invariant city-specific intercept, τ is a vector of time fixed effects, D_t is a vector of dummies picking out time t (that is, D_t is a $(T-1) \times 1$ vector of 0's with a 1 at time t, where we have normalized around period T), β is a time–and-region invariant slope coefficient, and $\varepsilon_{nt} \sim N\left(0, \sigma_n^2\right)$ assumes that the errors are uncorrelated across the cities.[5]

The national Beveridge curve is obtained by estimating (2) using a panel of city-level data. This

[4] A number of studies [see Valletta (2013)] argue that the workers were unable to obtain new jobs during the Great Recession because of house-lock, the inability to move because of an underwater mortgage.

[5] The spatially-uncorrelated error assumption is straightforward to relax.

method also produces a set of city-level Beveridge curves with common slopes and common shifts but with intercepts that vary by location and different idiosyncratic innovations. The common slope is the slope of the national Beveridge curve (i.e., all cities and the nation have the same sloped Beveridge curve) and the time fixed effects represent the shifts in the national Beveridge curve.

As discussed by Borowczyk-Martins, Jolivet, and Postel-Vinay (2013) and SSTV, a key econometric concern in estimating eq. (2) is the potential simultaneity between unemployment and vacancies. Specifically, current vacancy levels, v_{nt}, may be correlated with the unobserved error, ε_{nt}, due to shocks that jointly affect both vacancies and unemployment. To address this, we use a Bayesian instrumental variable approach, where v_{nt} is jointly determined with u_{nt} but is instrumented using its own lag, v_{nt-1}.

Similar to Borowczyk-Martins, Jolivet, and Postel-Vinay (2013) and SSTV, we model the relationship between current and lagged vacancies as an AR(1) process:

$$v_{nt} = \rho_n v_{nt-1} + \xi_{nt}, \quad |\rho_n| < 1, \tag{3}$$

that allows the structural error terms to be correlated and city-specific:

$$\begin{pmatrix} \xi_{nt} \\ \varepsilon_{nt} \end{pmatrix} \sim \mathcal{N}\left(\mathbf{0}, \Sigma_n\right), \quad \Sigma_n = \begin{pmatrix} \sigma^2_{\xi,n} & \sigma_{\xi\varepsilon,n} \\ \sigma_{\xi\varepsilon,n} & \sigma^2_{\varepsilon,n} \end{pmatrix}. \tag{4}$$

Identification of the slope coefficient, β, in the Beveridge curve equation relies on the assumption that v_{nt-1} is correlated with v_{nt} but uncorrelated with the idiosyncratic unemployment shock, ε_{nt}. In this framework, all parameters—including the city-specific covariance matrices, Σ_n—are estimated jointly, improving the model's robustness to endogeneity bias in v_{nt}.

2.2 Regional Beveridge Curves

The use of a panel of city-level data to identify a national Beveridge curve relies on the assumption that the matching function is invariant to location. This assumption may be viewed as rather strong, especially given heterogeneity in the industrial composition, population, and demographics of cities. A perhaps less strong identifying assumption is that some cities—in particular, those that have similar labor markets or industrial compositions—may have similar matching functions and, hence, similar Beveridge curves. We can think of these groups or clusters of cities as being "regions" defined by the commonality of their Beveridge curve slopes.

WZ estimated UK regional Beveridge curves using county-level data and the identifying restriction that the slopes of Beveridge curves for counties in the same region are identical and shifts of the Beveridge curve within a region are common.WZ choose their regions exogenously (defined by the government) and to be geographically continuous. For the UK, geographic regions might be justified; however, proximity between MSAs in the U.S. does not necessarily guarantee similarity in the labor markets. Thus, the assumption of a common slope coefficient across proximate MSAs may be too strong.[6] In addition to affecting the estimate of the Beveridge curve, the common-slope-across-all-cities assumption would affect our estimates of mismatch.

Our model adopts a similar design as WZ with the additional feature that we allow the data to determine which cities belong in the same region. Our approach has the additional advantage of allowing us to test which city-level characteristics might affect the form of the matching function.

We adapt model (2), allowing the data to indicate how labor market similarities can drive commonality in city-level Beveridge curves. We retain the assumption that cities within a region have a common slope but relax the assumption that we know, *ex ante*, the composition of the regions. To do this, we create an indicator that will allocate the city to a region. In the model described in this section, the number of regions K is assumed to be set *ex ante*; however, the choice of K can be treated as a model selection problem and is discussed below.

Let $n = 1, ..., N$ index the cities and $k = 1, ..., K$ index the regions. Define $\gamma_n = \{1, ..., K\}$ indicating to which region city n belongs and define

$$\gamma_{nk} = \begin{cases} 1 & \text{if } \gamma_n = k \\ 0 & \text{otherwise} \end{cases},$$

where $\sum_{k=1}^{K} \gamma_{nk} = 1$, imposing that an MSA belongs to only one region. Further, define

[6] Other papers have found that proximity does not necessarily beget similarity in business cycles or labor markets. Crone and Clayton-Matthews (2005) finds that Arizona may be more similar to states in the northeast. Hamilton and Owyang (2012) find states form regions based on industrial similarity, which can be related to proximity (e.g., energy and mining) or not at all (e.g., finance and real estate).

$$g_{nk} = \begin{cases} 1 & \text{if } k \leq \gamma_n \\ 0 & \text{otherwise} \end{cases}, \tag{5}$$

so that if MSA n belongs to region k, $\gamma_{nk} = 1$, and $g_{nm} = 1$ for $m \leq k$. We can then write the Beveridge curve for city n as

$$u_{nt} = \alpha_n + \tau D_t + \sum_{k=1}^{K} g_{nk} \beta_k v_{nt} + \varepsilon_{nt}, \tag{6}$$

where α_n is a city-specific fixed effect, τ is the vector of time fixed effects, D_t is the vector of time dummies, and $\beta_k < 0$ for all k. In this formulation, β_1 represents the common within-region slope of the Beveridge curve for region 1, and β_k, $k > 1$, reflects the difference between the slopes of the Beveridge curves of region k and region $k-1$. The slope for region $k > 1$ is the sum of the β_ks. We retain the assumption that ε_{nt} and ξ_{nt} are uncorrelated across cities. Consistent with the identification discussion above, we allow for within-city correlation between these shocks to account for the endogeneity of vacancies.

Collecting all of the cities in the panel, we can write (6) in matrix notation as

$$\mathbf{u}_t = \alpha + \tau D_t + \mathbf{G} \beta \mathbf{v}_t + \varepsilon_t \tag{7}$$

where $\mathbf{u}_t = [u_{1t}, ..., u_{Nt}]'$, $\mathbf{v}_t = [v_{1t}, ..., v_{Nt}]'$, $\alpha = [a_1, ..., \alpha_N]'$, $\mathbf{\Gamma} = [\Gamma_1', ..., \Gamma_N']'$, $\Gamma_n = [\gamma_{n1}, ..., \gamma_{nK}]$, $\mathbf{G} = [g_1', ..., g_N']'$, $g_n = [g_{n1}, ..., g_{nK}]$, $\beta = [\beta_1, ..., \beta_K]'$, and $\varepsilon_t = [\varepsilon_{1t}, ..., \varepsilon_{Nt}]'$.

Notice that the interpretation for the regional Beveridge curve is similar to the interpretation of the aggregate Beveridge curve. Cities within a region have a common slope but experience similar national shifts. In principle, we could set $\mathbf{\Gamma}$ ex ante (as in WZ) if we knew it with certainty. On the other hand, if we are unsure about the composition of the clusters, our best bet is to rely on the data to determine which factors influence the likelihood that series cluster together.[7] In our case, cities are allocated to regions based on the similarities in the slopes of their Beveridge curves.

[7]Francis, Owyang, and Savascin (2017) show, in a clustered factor setting, the pitfalls of imposing even slightly misspecified clusters.

2.3 Measuring Mismatch

Once we have obtained the estimates of the slopes of the regional Beveridge curves, we can compute mismatch, or the misallocation of labor across regions. Jackman, Layard, and Savouri (1991) develop an index describing regional mismatch. They compare the region or industry level of unemployment to a steady-state value that, under their assumptions, would be equal across units. SSTV construct their own indices for mismatch which rely on both the slope and intercept of the Beveridge curve that also accounts for vacancies.

Their simple index is based on a theoretical matching model with a Cobb-Douglas matching function with homogenous matching elasticities across entities. For robustness, SSTV extend their model to allow for heterogenous matching elasticies across sectors. We apply this framework to consider different matching elasticities across regions.

In their model, the matching function for region n is given by

$$h_{nt} = V_{nt}^{1-\beta_n} U_{nt}^{\beta_n} \tag{8}$$

where β_n is match elasticity (i.e., slope of the Beveridge curve) for region n. Given total unemployment U_t and regional vacancies V_{nt}, the social planner chooses the optimal allocation of unemployed for each region U_{nt}^* to maximize aggregate output.[8] The mismatch index is

$$M_t = 1 - \frac{h_t}{h_t^*} \tag{9}$$

where $h_t = \sum_{n=1}^{N} h_{nt}$ measures the actual number of matches whereas $h_t^* = \sum_{n=1}^{N} h_{nt}^* = V_{nt}^{1-\beta_n} {U_{nt}^*}^{\beta_n}$ measures the optimal number of matches from the planner's allocation. Thus, M_t measures the fraction of matches lost due to a suboptimal allocation of unemployed workers across regions. Because we do not have total unemployment and vacancies for the whole country, we compute mismatch only over the available cities in our sample.

[8]The SSTV measure embeds a number of simplifying assumptions. Workers are equally productive regardless of location, match efficiency and job destruction is homogenous across locations, and the reallocation of labor in the planner problem is costless. Appendix A.7 of SSTV outlines the solution to the planner's problem.

3 Econometric Implementation

As shown in Frühwirth–Schnatter and Kaufmann (2008) and others, the model is straightforward to estimate in a Bayesian environment. The following sections describe the data, provide our assumptions about the prior distributions, and outline the sampler. Details for construction of the data and the sampler are provided in a technical appendix.

3.1 Data

Our model requires three sets of data: (1) a panel of MSA-level unemployment rate and labor force data; (2) a panel of MSA-level vacancy data and (3) the cross-sectional data that determines the priors for defining the regions. The MSA-level unemployment rate and the size of the labor force are available from the BLS starting in 1990. To obtain the data prior to that, we extract the data from the CPS. The MSA-level vacancy data must be constructed from multiple sources. The cross-sectional data is also obtained from a number of sources. These datasets are explained in detail in the following subsections.

3.1.1 MSA-level Unemployment and Labor Force

The unemployment rate and the size of the labor force at the MSA level are obtained from the Bureau of Labor Statistics' Local Area Unemployment Statistics (LAUS). These data are available monthly from July 1989 through the present. They are seasonally adjusted by the BLS using the SEATS program.

To measure the unemployment rate and labor force before 1989, we use the change in unemployment and labor force calculated from the Basic Monthly Current Population Survey (CPS) Public Use Microdata. The CPS data for January 1978 through December 1990 are accessed from Unicon CPS Utilities. Not all MSAs are defined before 1985, so we have an unbalanced panel with some MSAs missing data between 1978 and 1985. Also, MSA definitions change over time, so we use a series of crosswalks to estimate the data for MSAs consistent with the 1990 definitions. Details for these and other data transformations are in the Appendix.

3.1.2 MSA-level Vacancies

To measure vacancies, some past studies have used the Conference Board's Help-Wanted Index (HWI) that measures advertisements for job openings in major newspapers in 51 MSAs back through 1955. Beginning with the onset of the internet in the 1990s, help wanted advertising in newspapers became underrepresentative of total advertising. Increasingly, firms post open positions on the internet, resulting in a secular decline in the national composite measure of newspaper help wanted advertising starting in the mid-1990s. In June 2005, the Conference Board created the Help-Wanted Online Index (HWOL) that measures vacancies from online job search websites. Soon after the Conference Board started publishing the HWOL, it stopped reporting the HWI, which ends in May 2008.

To construct a harmonized measure of vacancies, we need to account for the onset and diffusion of online advertising and the decline of the print advertising. The problem is that we are missing HWOL data during the initial period of online advertising and we are missing HWI data after 2008. Barnichon (2010) proposed a method of combining the HWI and HWOL series, resulting in a national-level composite help-wanted index that captures total advertising (print + online) from 1955 through 2021. Barnichon's method requires estimating the print share of advertising, $S_t^p = P_t/(P_t + O_t)$, from 1995 forward, where P_t and O_t are print and online advertising, respectively.

To construct his national-level index, Barnichon assumes that $S_t^p = 1$ prior to 1995 and that the diffusion of online advertising follows a Mixed-Information Source Model (MISM), which implies the print share of advertising decreases over time along the path

$$S_t^p = 1 - \frac{1 - e^{-\lambda_1 t}}{1 + \lambda_2 e^{-\lambda_3 t}}, \tag{10}$$

where $\lambda_1, \lambda_2, \lambda_3 > 0$. The MISM allows for an asymmetric adoption rate that is faster during the initial stages. Barnichon first constructs estimates of S_t^p for the post 1995 period using various techniques and then estimates $\lambda_1, \lambda_2, \lambda_3$ from (10) through nonlinear least squares.[9] After 2008 when the HWI is no longer available, Barnichon propagates S_t^p based on (10) and the estimated parameters, then backs out the change in total advertising.

The assumption that the diffusion of online advertising follows an MISM is borne out in the

[9] The technical appendix contains details of how Barnichon obtains estimates of the print share from 1995 to 2008.

aggregate data as the share of print advertising does appear to decline monotonically over time. For some of the cities—perhaps because of slower (or later starting) diffusion or business cycle effects, the estimates of the print share do not decline monotonically and the estimation of (10) produces a unit future print share. In adherence to the idea that print share should decrease over time, we impose a prior on the parameters $\lambda_1, \lambda_2, \lambda_3$ and estimate (10) using Bayesian methods. We first estimate the model for national data with a flat prior and then impose that the cities diffuse the internet similarly (but perhaps with different starting points) to the nation. Thus, the priors for λ_1, λ_3 are independent gamma centered on the national estimates and λ_2 has a normal prior centered on the national estimate.

3.1.3 Covariate Data

The third set of data that we require is the covariates, $\mathbf{x}_n$, in the logistic prior that inform us about whether a city belongs in a region. We include only a small set of covariates that answer specific questions about what factors might influence the nature of the matching function and the Beveridge curve.[10] First, we include a set of dummies for the census region of the city. These dummies will allow us to tell whether geography plays a role in the clustering of cities into regions. Second, we include a few measures of industrial composition consisting of the share of manufacturing employment, the share of finance and professional business employment, and a measure of industrial concentration. The latter may determine how easily workers can flow across firms: more concentrated cities might offer more opportunities to workers without much retraining. We also include union concentration, which may inhibit the flow of workers across industries.

One limitation of the clustering algorithm that we use is that it requires static covariates. Thus, we cannot allow the clusters to evolve slowly across time as the covariates evolve. Because of this restriction, we use the 1990 vintage of covariate data for the full sample; however, using the 2000 vintage does not change the results substantially.

[10] Glaeser, Kallal, Scheinkman, and Shleifer (1992) and Glaeser, Scheinkman, and Shleifer (1995), among many others, consider the factors that might induce variation in the output and employment growth rates across cities.

3.2 Priors and the Sampler

We estimate the model using the Gibbs sampler. The Gibbs sampler forms the joint posterior distribution of the full set of model parameters from draws from each of the conditionals. The parameters are split into five blocks: (1) the coefficients $\Psi = \{\alpha, \beta, \tau\}$, (2) the city-specific variance-covariance matrices Σ_n, (3) the vacancy persistence parameters $\rho = [\rho_1, ..., \rho_N]'$, (4) the multinomial logit parameters δ, (5) and the regional indicators $\mathbf{\Gamma}$. The sampler iterates through draws of one block, conditional on the past draw from the other blocks. The resulting set of draws approximates the full joint posterior distribution after discarding the first 20,000 draws to eliminate initial conditions. We use 10,000 draws to form the joint posterior distribution.

Table 1: **Prior Hyperparameters** – This table shows the prior distributions used to estimate the model.

Parameter	Prior
α	$N(\boldsymbol{a}_0, \boldsymbol{A}_0)$ $\boldsymbol{a}_0 = 0$ $\boldsymbol{A}_0 = I_N$
β	$N(\boldsymbol{b}_0, \boldsymbol{B}_0)$ $\boldsymbol{b}_0 = \boldsymbol{c}$ $\boldsymbol{B}_0 = I_N$
τ	$N(\boldsymbol{h}_0, \boldsymbol{H}_0)$ $\boldsymbol{z}_0 = 0$ $\boldsymbol{Z}_0 = I_N$
Σ_n	$\mathcal{IW}(\nu_0, S_0)$ $\nu_0 = 30$ $S_0 = I_2$
ρ_n	$N(r_0, \boldsymbol{R}_0)$ $r_0 = 0.7$ $\boldsymbol{R}_0 = 0.1^2$
δ_k	$N(\boldsymbol{d}_0, \boldsymbol{D}_0)$ $\boldsymbol{d}_0 = 0$ $\boldsymbol{D}_0 = I_N$ $\forall k$

The prior hyperparameters are given in Table 1. We impose a set of relatively standard priors on the main Beveridge curve parameters as well as the vacancy persistence parameter: $\alpha \sim N(\mathbf{a}_0, \mathbf{A}_0)$; $\beta \sim N(\mathbf{b}_0, \mathbf{B}_0)$; $\tau \sim N(\mathbf{h}_0, \mathbf{H}_0)$; $\rho_n \sim N(r_0, \boldsymbol{R}_0)$; and $\Sigma_n \sim \mathcal{IW}(\nu_0, S_0)$. We set the prior mean $\mathbf{b}_0 = \mathbf{c}$ for the slope coefficients β so that the slopes are larger (in absolute value) for higher k, matching the model assumptions. The prior for the regional indicators $\mathbf{\Gamma}$ is multinomial logistic. Suppose there exists a vector, $\mathbf{x}_n$, of series $n-$specific covariates that may influence whether series n belongs to cluster k. We could then model the prior probability that city n belongs to region k as a multinomial logit, where

$$p(\gamma_{nk}) = \begin{cases} \exp\left(\mathbf{x}'_{nk}\delta_k\right) / \left[1 + \sum_j \exp\left(\mathbf{x}'_{nj}\delta_j\right)\right] & k = 1, ..., K-1 \\ 1/\left[1 + \sum_j \exp\left(\mathbf{x}'_{nj}\delta_j\right)\right] & k = K \end{cases} \tag{11}$$

for $n = 1, ..., N$ and where we have normalized $\delta_K = 0$ for identification. Note also that the vector,

$\mathbf{x}_{nk}$, need not be composed of the same variables for each cluster. As discussed in Hamilton and Owyang (2012), we can treat the δ_k as population parameters to be estimated with all the other model parameters. We draw the hyperparameters δ_k using the difference in random utility model (dRUM) method from Frühwirth–Schnatter and Frühwirth (2010) and Kaufmann (2015).

3.3 Choosing the Number of Regions

In the model above, the choice of K—the number of regions—is taken as given. Each of the sampler steps is made conditional on the number of regions. One method of selecting the number of regions could be to build in an indicator for the number of regions that is drawn at each iteration of the sampler. Because the number of regions is assumed to be small compared to the number of cities, we can instead treat the choice of K as a model selection problem. We estimate the model for each of $K = 2, ..., 7$ and choose the optimal K by minimizing the Bayesian information criterion (BIC). [11]

4 Results

In this section, we describe the full sample results using the unbalanced panel of vacancy rates and unemployment rates and a fixed cross-section of covariate data.[12] For the full sample, the data favor $K = 4$ regions.

4.1 Characterizing the Beveridge Curve

The steepness of the Beveridge curve reflects the strength of the relationship between the unemployment rate and vacancies: The steeper the Beveridge curve, the larger the unemployment rate changes associated with a given change in vacancies. The top panel of Table 2 shows the slopes ($\sum_{m\leq k} \beta_m$) of the estimated Beveridge curves for each of region k. By construction, the regions are identified by ordering them in increasing steepness of their Beveridge curves. Region 1's Beveridge curve is nearly flat, while Region 4 is more than twice as steep as Region 2.

[11] See Ando (2007).

[12] A previous version of this paper used a different, now-defunct MSA HWOL dataset from the Conference Board. In 2019, the Conference Board updated their methodology and data gathering. Results here are consistent with the older dataset, although the number of estimated regions differ slightly.

Table 2: **Estimated Slopes and Covariate Marginal Effects (Full Sample)** – This table shows the estimated cluster-specific parameters for the baseline model without a structural break. The table reports the posterior median along with the 68% posterior coverage interval. The top row shows the estimated slope coefficient for the Beveridge curve. The bottom panel shows the estimated marginal effects of each covariate on the prior probability of cluster membership. Bold indicates the 68% posterior coverage for the marginal effect does not include zero. The marginal effects can be interpreted as the difference in the prior probability when the respective covariate is relatively high and low.

	K = 1	K = 2	K = 3	K = 4
Slope Coefficient (β_k)	−0.03 [−0.04, −0.02]	−0.15 [−0.16, −0.15]	−0.22 [−0.22, −0.21]	−0.36 [−0.36, −0.35]
Union Members (% of emp)	0.08 [−0.01,0.19]	0.13 [−0.02,0.31]	−0.10 [−0.24,0.03]	−0.12 [−0.26,0.01]
Manufacturing (%of emp)	0.03 [−0.06,0.14]	**−0.22** [−0.38, −0.05]	0.11 [−0.03,0.27]	0.06 [−0.08,0.20]
Industrial Concentration	0.05 [−0.05,0.17]	−0.04 [−0.21,0.13]	0.04 [−0.11,0.19]	−0.05 [−0.20,0.09]
Finance/Prof. Bus. (% of emp)	0.00 [−0.11,0.12]	0.16 [−0.01,0.34]	**−0.22** [−0.40, −0.05]	0.05 [−0.10,0.20]
Northeast	**0.21** [0.05,0.38]	−0.16 [−0.47,0.12]	−0.06 [−0.36,0.19]	0.04 [−0.16,0.23]
Midwest	−0.08 [−0.32,0.07]	0.11 [−0.17,0.33]	−0.04 [−0.34,0.18]	0.11 [−0.08,0.28]
West	−0.13 [−0.39,0.03]	−0.19 [−0.49,0.06]	**0.36** [0.13,0.56]	0.04 [−0.14,0.21]

Figure 1: **Time Fixed Effects** – This figure shows the estimated year fixed effects (blue line) and the separation rate (red line). The solid blue line shows the posterior median for the fixed effects and the dashed lines show the 68% posterior coverage interval. The separation rate is constructed by Fujita and Ramey (2009) from the Current Population Survey. The gray bars show the NBER recession dates.

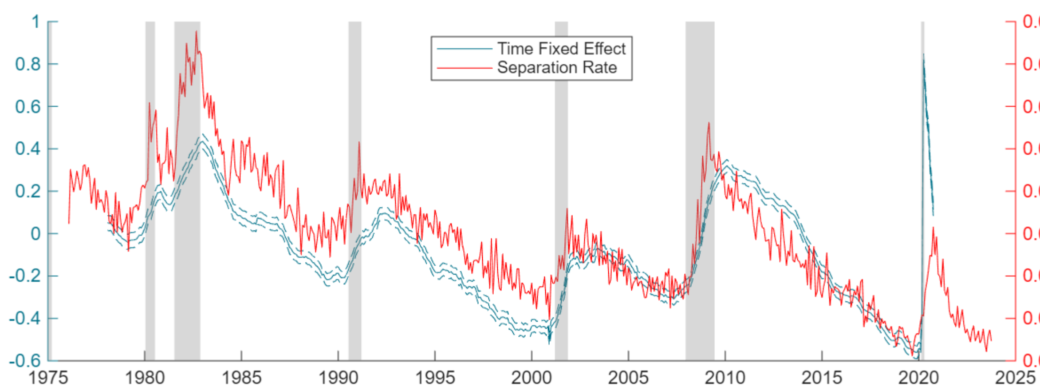


Figure 1 shows the values of the year-fixed-effects for the full sample (blue line) as well as the separation rate (red line) from Fujita and Ramey (2009). This result is consistent with those of WZ for the UK and Benati and Lubik (2014), who find that the shifts in the Beveridge curve appear to be correlated with the business cycle. The correlation between the estimated fixed effects and the separation rate is 0.78. The time-varying intercept appears to be capturing a higher separation rate during economic downturns, which leads to a higher unemployment rate for a given vacancy rate.

4.2 Regional Composition

Figure 2: **Posterior Regions – Full Sample.** The map shows the four regions for the MSAs based on the modal posterior draw. Membership is determined by similarity in Beveridge curve slopes as well as country-specific factors influencing the prior. Areas in white are not included in the sample.

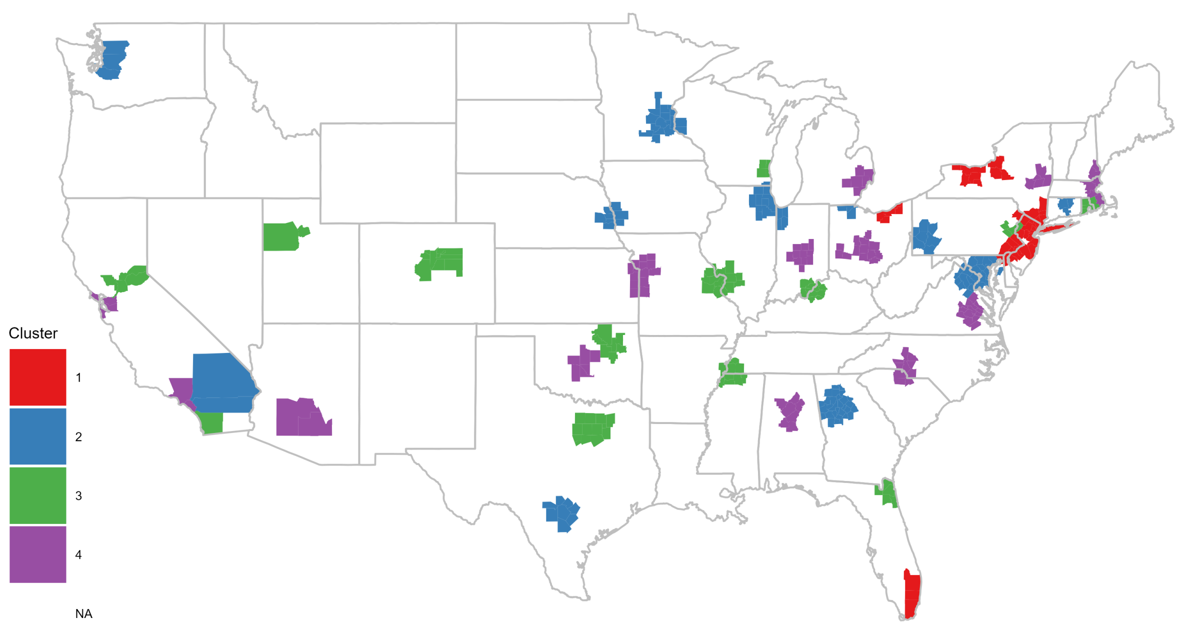


Figure 2 shows the composition of the regions for the full sample estimation in the form of a map. The stark result is that—with the exception of Region 1—the composition of the regions does not appear to follow a geographic pattern. Cities that are located in close proximity are not necessarily in the same region. Three of the four Beveridge curve regions are geographically dispersed across the country, each having member cities in all four Census regions. While some geographically-close cities (e.g., New York and Philadelphia; Rochester and Syracuse; or Columbus and Dayton) are grouped into the same region, some cities that share borders (e.g., D.C. and Richmond; or Riverside and San Diego) are not grouped together.

Similarly, the city's population does not seem to be important to grouping the cities. One might have hypothesized that the size of the labor market—proxied by population—might effect the efficiency of the labor markets. This does not appear to be the case. While some large cities (DC, Atlanta, and Chicago) are grouped together, other larger cities like Dallas, Los Angeles, and New York (in the top 5 by population) are in different regions.

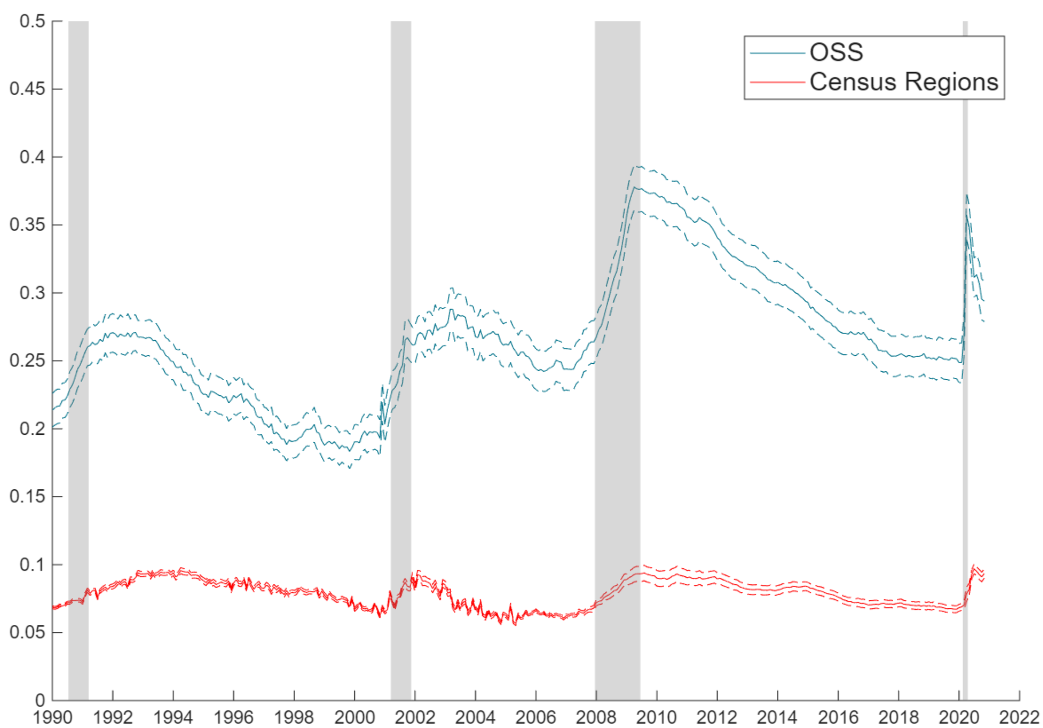
We investigate what factors influence regional composition by computing the "marginal effect" of a change in the covariates—the change in the inclusion probability, γ_{nk}, resulting from increasing the ith covariate. To facilitate exposition, we compute the marginal effect of increasing the ith covariate from one-standard-deviation below its mean to one-standard-deviation above its mean from eq. (11), holding all other covariates at their cross-city means. Positive (negative) marginal effects are interpreted as an increase in the covariate raising (lowering) a city's prior probability of being included in that region.

The bottom panel of Table 2 contains the marginal effects of the covariates in the logistic prior for the optimal number of regions using the full sample of data. For the full sample, we use the 1990 vintage of covariate data; however, using the 2000 vintage does not change the results substantially. Based on the estimated hyperparameters for the prior, increasing the share of manufacturing employment increases the probability of being in Region 2. Conversely, Region 3 cities are more likely to have a larger share of employment in financial and professional business services. Region 1 and Region 3 are somewhat geographically-defined. Cities in the Northeast are more likely to be grouped into Region 1, while cities in the West are more likely in Region 3. Region 4, in contrast, does not exhibit a clear relationship with the included covariates.

4.3 Mismatch

The blue lines in Figure 3 shows the regional mismatch measure [eq. (9)] for the full (balanced) sample along with its 68-percent posterior coverage. We have previously discussed some of the papers that suggest that mismatch was a contributing factor in slow recovery of the labor market following the Great Recession. We similarly find that regional mismatch rises subsequent to the

Figure 3: **Mismatch** – This figure shows the regional mismatch for the full sample using the baseline endogenous clusters (blue) and the geographic exogenous clusters (red). The solid line represents the posterior median and the dashed lines show the 68% posterior coverage interval. The gray bars show the NBER recession dates.



financial crisis, and the increase is persistent and substantially larger than for the previous two recessions. This finding contrasts with previous work by SSTV which found that regional mismatch was not a major factor in the slow labor market recovery after the Great Recession. [13] Moreover, our finding is in-line with the hypothesis that house lock—the inability to relocate to a more fertile job market because of a negative mortgage balance—may have played a substantial role as previously documented in some studies.

While it is reassuring that our results are consistent with what others find for the Great Recession period, our focus is on examining the properties of the longer time series of mismatch. While the early 1990s recession and subsequent recovery period is accompanied by a rise in mismatch, it appears that regional mismatch fluctuated around a constant trend before 2000. Regional mismatch then appears to fall consistently, except during recessions and the short periods that directly follow recessions. During recessions and subsequent recovery periods, unproductive labor is shed but cannot immediately reallocate to regions where it might be needed and mismatch rises.

Finally, we consider how estimating the regions might change how mismatch is measured versus pre-defined region such as the four Census regions. The red lines show regional mismatch when using the fixed regions. While many of the fluctuations in mismatch are similar across these two specifications, fixed regions substantially *underestimates* mismatch, and the Great Recession does not have an outsize effect compared to the first two recessions in our sample.

5 Breaks

The nature of the Beveridge curve may have changed over time for a myriad of reasons, including the widespread adoption of internet job postings [Faberman and Kudlyak (2019), among others]; demographic changes [Aaronson et al. (2015), among others]; sectoral reallocation from manufacturing to services [Groshen and Potter (2003), among others]; and a shift from temporary to permanent layoffs [Groshen and Potter (2003), among others]. We examine this possibility by allowing an endogenously-timed structural break in the model parameters. The timing of the break will be suggestive, but we draw no firm causal conclusions. If, for example, sectoral reallocation made

[13] The key explanation for this difference is that the regional mismatch index of SSTV only considers differences in productivity across four broad Census regions, whereas our measure is based on differences in match elasticity across endogenous clusters of cities.

cities more similar by making jobs more homogeneous, the number of regions might decline over the sample period. This could also change the Beveridge curve slopes by altering the parameters of the matching function.

We begin by estimating the model, allowing a break in both the Beveridge curve slopes and the number of regions at an endogenously determined date, τ. At each Gibbs iteration, we draw the break date, τ, based on the ratio of the posterior likelihood for each date in the interior 60 percent of the sample. Previously, the number of regions were chosen as a model selection problem. We continue that approach here by estimating the break model for combinations of regions, $K_{pre}, K_{post} = \{2, ..., 7\}$, and choosing the region combination based on the BIC. We can also compare the break model to the no-break model based on the BIC. We use 1990 values for the covariates pre-break and 2000 values after the break.

We find that the pre-break sample favors $K_{pre} = 2$ regions, while the post-break sample favors $K_{post} = 6$ regions with a break date at $\tau = 1990{:}12$, suggesting more cross-MSA heterogeneity post-break.[14] This break specification has a BIC of $-252,490$, compared with $-245,890$ for the optimal no-break model. This difference in the BIC exceeds conventional thresholds [e.g., Kass and Raftery (1995)], thereby providing very strong evidence in favor of the break model.

The two panels of Table 3 show the pre- and post-break results for the preferred specification of $K_{pre} = 2$ (top) and $K_{post} = 6$ (bottom). Within each panel, the top subpanel shows the slope of each region's Beveridge curve (ordered by increasing steepness) and the bottom panel shows the marginal effects of the covariates for each region. Figure 4 shows the cluster membership pre- and post-break.

One stark result from Table 3 is that, in the post-break period, the regional Beveridge curves are uniformly flatter than in the pre-break period.[15] In fact, the flattest regional Beveridge curve in the pre-break period is steeper than the steepest regional Beveridge curve in the post-break period. Larger magnitude (steeper) slopes suggest lower matching elasticity, meaning that firms must post more vacancies to bring the unemployment rate down. This result then suggests one possible interpretation: after the break, internet job posting decreased match elasticity.

[14] One might be concerned that the timing of the break might be influenced by how we combine the HWI to HWOL data. However, the timing of the break suggests this is not the issue.

[15] The slopes in the pre-break sample are more consistent with the literature estimating the national Beveridge curve. Michaillat and Saez (2021) find that "the Beveridge elasticity fluctuated between 0.84 and 1.02, averaging 0.91" between 1951 to 2019. The post-break estimates are similar to those found by WZ for Britain (-0.196).

Table 3: **Estimated Slopes and Covariate Marginal Effects (Pre- and Post-Break)** – This table shows the estimated cluster-specific parameters for the model with a structural break. We estimated the structural break that occurred in December 1990. The table reports the posterior median along with the 68% posterior coverage interval. The top panels show the estimated slope coefficient for the Beveridge curve. The bottom panels show the estimated marginal effects of each covariate on the prior probability of cluster membership. Bold indicates the 68% posterior coverage for the marginal effect does not include zero. The marginal effects can be interpreted as the difference in the prior probability when the respective covariate is relatively high and low.

(a) Pre-Break

	K = 1	K = 2
Slope Coefficient (β_k)	−0.40 [−0.41,−0.39]	−0.60 [−0.62,−0.59]
Union Members (% of emp)	**-0.43** [−0.58,−0.28]	**0.43** [0.28,0.58]
Manufacturing (% of emp)	0.07 [−0.12,0.25]	−0.07 [−0.25,0.12]
Industrial Concentration	0.12 [−0.07,0.31]	−0.12 [−0.31,0.07]
Finance/Prof. Bus. (% of emp)	0.13 [−0.08,0.32]	−0.13 [−0.32,0.08]
Northeast	0.01 [−0.30,0.32]	−0.01 [−0.32,0.30]
Midwest	**0.42** [0.11,0.65]	**-0.42** [−0.65,−0.11]
West	0.05 [−0.27,0.36]	−0.05 [−0.36,0.27]

(b) Post-Break

	K = 1	K = 2	K = 3	K = 4	K = 5	K = 6
Slope Coefficient (β_k)	−0.0005 [−0.0014,−0.0001]	−0.06 [−0.07,−0.06]	−0.13 [−0.14,−0.12]	−0.19 [−0.20,−0.18]	−0.28 [−0.29,−0.27]	−0.37 [−0.38,−0.37]
Union Members (% of emp)	0.09 [−0.01,0.24]	−0.11 [−0.23,0.01]	0.28 [0.14,0.46]	**-0.15** [−0.28,−0.03]	−0.09 [−0.20,0.01]	−0.06 [−0.14,0.02]
Manufacturing (% of emp)	**-0.39** [−0.57,−0.22]	**0.19** [0.05,0.37]	0.01 [−0.08,0.12]	−0.01 [−0.14,0.13]	**0.11** [0.01,0.26]	0.04 [−0.03,0.13]
Industrial Concentration	−0.10 [−0.25,0.01]	**0.16** [0.03,0.33]	−0.00 [−0.11,0.10]	0.00 [−0.14,0.13]	−0.03 [−0.17,0.08]	0.00 [−0.09,0.08]
Finance/Prof. Bus. (% of emp)	**-0.30** [−0.49,−0.14]	**0.44** [0.27,0.62]	−0.02 [−0.13,0.08]	**-0.15** [−0.32,−0.01]	0.04 [−0.07,0.16]	0.01 [−0.07,0.09]
Northeast	**0.22** [0.05,0.42]	0.19 [−0.01,0.41]	**-0.22** [−0.51,−0.04]	−0.05 [−0.33,0.13]	−0.01 [−0.22,0.12]	−0.02 [−0.12,0.06]
Midwest	−0.18 [−0.47,0.00]	**0.24** [0.05,0.43]	0.02 [−0.16,0.14]	−0.00 [−0.24,0.16]	0.02 [−0.15,0.15]	0.04 [−0.06,0.13]
West	−0.14 [−0.42,0.03]	−0.19 [−0.48,0.01]	**0.23** [0.08,0.42]	0.19 [−0.01,0.40]	−0.02 [−0.21,0.12]	0.02 [−0.07,0.10]

Although the direction of change is consistent across cities, the magnitude of flattening varies considerably. Figure 5 illustrates this by plotting the change in Beveridge curve slopes, measured as the post-break slope minus the pre-break slope. To explore these differences, we regress the change in slope on the same MSA-level characteristics used in the cluster prior. The regression results shown in Table 4 indicate that MSAs with higher union membership experienced a larger flattening of their Beveridge curves, consistent with slower vacancy adjustment or greater wage rigidity in more unionized labor markets. In contrast, Midwestern cities experienced a significantly smaller flattening, implying that matching efficiency declined less sharply in that region. This regional resilience may reflect the persistence of traditional manufacturing networks and firm-based search channels that were less affected by the diffusion of online job posting in the 1990s.

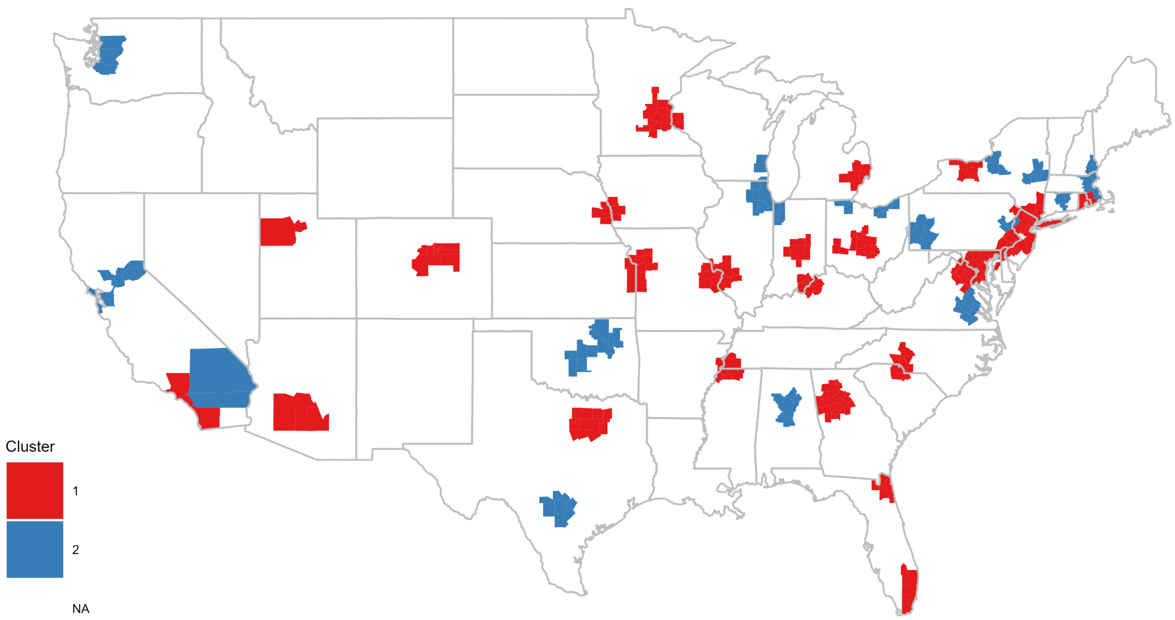
Table 4: **Explaining the Change in Slope** – This table shows the posterior means and 90% posterior coverage from regressing the change in slope, measured as post-break β minus pre-break β, on the MSA-specific characteristics. Bold indicates the 90% posterior coverage for the coefficient does not include zero.

Variable	Estimate
Intercept	0.41 [0.30,0.51]
Union Members (% of emp)	**0.10** [0.04,0.16]
Manufacturing (% of emp)	−0.04 [−0.09,0.02]
Industrial Concentration	−0.03 [−0.09,0.02]
Finance/Prof. Bus. (% of emp)	−0.02 [−0.09,0.04]
Northeast	−0.01 [−0.18,0.15]
Midwest	**−0.16** [−0.31,−0.01]
West	−0.11 [−0.25,0.04]

Table 3 reports the marginal effects of the covariates in the logistic prior for the pre- and post-break samples. The results show that the determinants of cluster membership changed somewhat after the break. Before 1990, higher union membership was associated with a greater probability of belonging to Region 2, which corresponds to a steeper Beveridge curve. However, this relationship reverses in the post-break period as more unionized MSAs are more likely to be members of regions with flatter Beveridge curves. One potential explanation is that unionized labor markets became

Figure 4: **Posterior Regions Pre- and Post-Break** – The top map shows the two regions for the MSAs based on the modal posterior draw prior to the break in December 1990, while the bottom map shows the six regions after the break. Membership is determined by similarity in Beveridge curve slopes as well as country-specific factors influencing the prior. Areas in white are not included in the sample.

(a) Cluster Composition Pre-Break



(b) Cluster Composition Post-Break

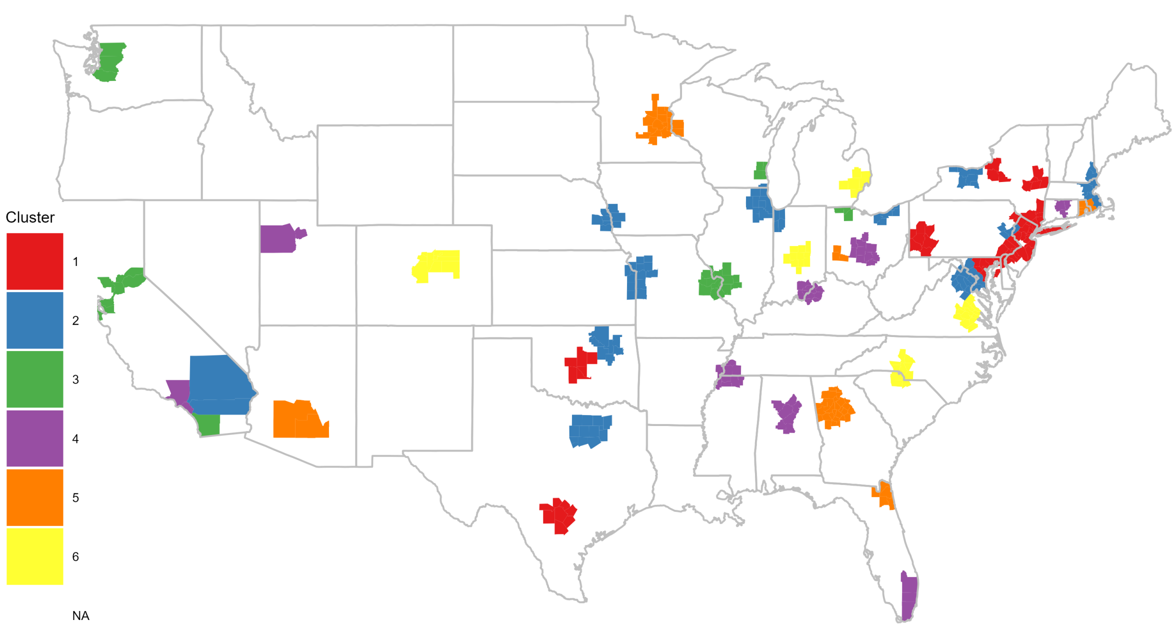
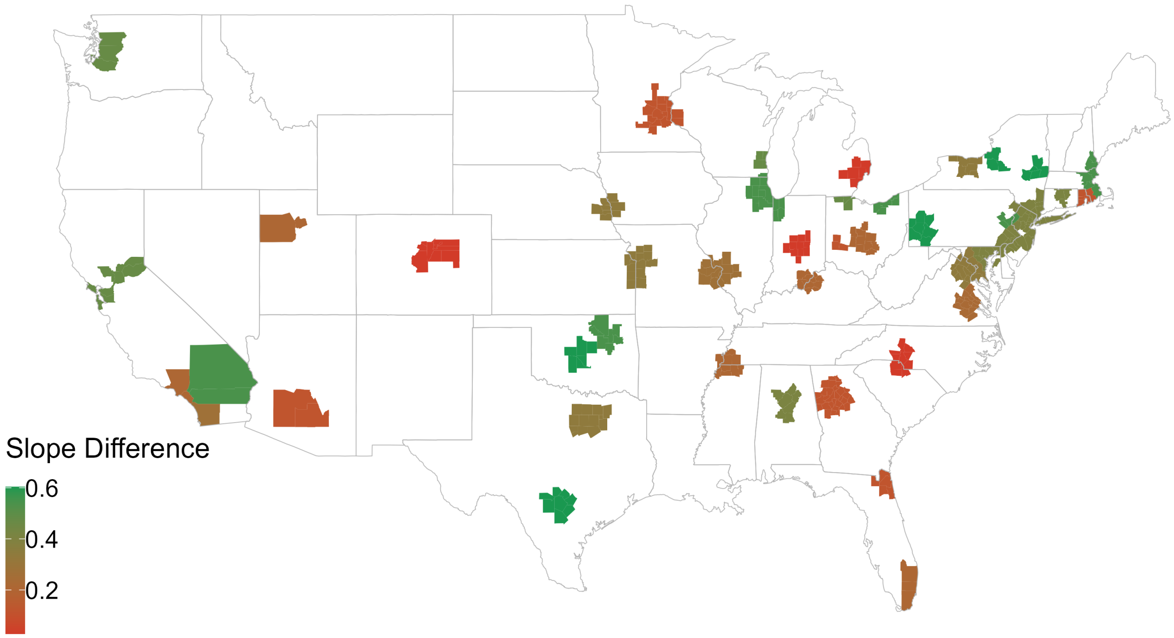


Figure 5: **Change in Beveridge Curve Slopes Pre- vs. Post-Break** – This map shows the post-break minus pre-break Beveridge curve slopes by region. A positive value indicates a flatter Beveridge curve after the break. Areas in white are not included in the sample.



relatively less responsive to vacancy fluctuations after the structural break. Cities located in the Midwest consistently tend to belong to flatter-slope clusters, suggesting a more stable relationship between vacancy and unemployment in that region. By contrast, the post-break relationships between industrial composition and Beveridge-curve slopes are more difficult to interpret. The marginal effects of manufacturing and financial employment shares are no longer monotonic as higher manufacturing employment does not uniformly increase the probability of belonging to a steeper region.

6 Conclusion

We use proprietary data on internet job postings to estimate regional Beveridge curves from MSAs. We combine the internet data with help-wanted data from newspapers to form a longer sample. Regional Beveridge curves are estimated under the assumption that cities in the same region have the same sloped Beveridge curve. This assumption allows us to form regions endogenously.

We find evidence for four regions in the full sample. Each region appears to have some city-level characteristic common to its MSAs. For example, the region with the flattest Beveridge curve (in the full sample) is comprised of cities with high shares of financial and professional business services employment and high industry concentration.

We also find evidence of a sample break in the early 1990s, after the time the internet was adopted. The break changes the number and composition of the regions. In the post-break period, Beveridge curves became flatter across the board, suggesting a fall in matching elasticity.

References

[1] Aaronson, Daniel, Luojia Hu, Arian Seifoddini, and Daniel G. Sullivan. "Changing labor force composition and the natural rate of unemployment." *Chicago Fed Letter* 338 (2015).

[2] Alvarez, Fernando, and Robert Shimer. "Search and rest unemployment." *Econometrica* 79, no. 1 (2011): 75-122.

[3] Anderson, Patricia M., and Simon M. Burgess. "Empirical Matching Functions: Estimation and Interpretation Using State-Level Data." *The Review of Economics and Statistics* 82, no. 1 (2000): 93-102.

[4] Ando, Tomohiro. "Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models." *Biometrika* 94, no. 2 (2007): 443-458.

[5] Barnichon, Regis. "Building a composite Help-Wanted Index." *Economics Letters* 109, no. 3 (2010): 175-178.

[6] Bean, Charles R. "European Unemployment: A Survey." *Journal of Economic Literature* 32, no. 2 (1994): 573-619.

[7] Benati, L., and T. A. Lubik. "The Time-varying Beveridge Curve." In *Advances in Nonlinear Economic Modeling*. (2014): 167-204.

[8] Blanchard, Olivier Jean and Peter Diamond. "The Beveridge Curve." *Brookings Papers on Economic Activity* 1989, no. 1 (1989): 1-76

[9] Borowczyk-Martins, Daniel, Gregory Jolivet, and Fabien Postel-Vinay. "Accounting for endogeneity in matching function estimation." *Review of Economic Dynamics* 16, no. 3 (2013): 440-451.

[10] Börsch-Supan, Axel H. "Panel Data Analysis of the Beveridge Curve: Is There a Macroeconomic Relation between the Rate of Unemployment and the Vacancy Rate?." *Economica* 58, no. 231 (1991): 279-297.

[11] Chib, Siddhartha. "Estimation and comparison of multiple change-point models." *Journal of Econometrics* 86, no. 2 (1998): 221-241.

[12] Courtney, Hugh George. "The Beveridge Curve and Okun's Law: a re-examination of fundamental macroeconomic relationships in the United States." PhD diss., Massachusetts Institute of Technology, 1991.

[13] Crone, Theodore M., and Alan Clayton-Matthews. "Consistent economic indexes for the 50 states." *Review of Economics and Statistics* 87, no. 4 (2005): 593-603.

[14] Faberman, R. Jason, and Marianna Kudlyak. "The Intensity of Job Search and Search Duration. " *American Economic Journal: Macroeconomics* 11, no. 3 (2019): 327-357.

[15] Ferreira, Fernando, Joseph Gyourko, and Joseph Tracy. "Housing busts and household mobility." *Journal of Urban Economics* 68, no. 1 (2010): 34-45.

[16] Francis, Neville, Michael T. Owyang, and Özge Şavascin. "An endogenously clustered factor approach to international business cycles." *Journal of Applied Econometrics* 32, no. 7 (2017): 1261-1276.

[17] Frühwirth-Schnatter, Sylvia, and Rudolf Frühwirth. "Data Augmentation and MCMC for Binary and Multinomial Logit Models." In *Statistical Modelling and Regression Structures*, pp. 111-132. Physica-Verlag HD, 2010.

[18] Frühwirth-Schnatter, Sylvia, and Sylvia Kaufmann. "Model-based clustering of multiple time series." *Journal of Business & Economic Statistics* 26, no. 1 (2008): 78-89.

[19] Fujita, Shigeru, and Garey Ramey. "The Cyclicality of Separation and Job Finding Rates." *International Economic Review* 50, no. 2 (2009): 415-430.

[20] Gärtner, Manfred. *A primer in European macroeconomics*. Prentice Hall, 1997.

[21] Glaeser, Edward L., Hedi D. Kallal, José A. Scheinkman, and Andrei Shleifer. "Growth in cities." *Journal of Political Economy* 100, no. 6 (1992): 1126-1152.

[22] Glaeser, Edward L., José Scheinkman, and Andrei Shleifer. "Economic growth in a cross-section of cities." *Journal of Monetary Economics* 36, no. 1 (1995): 117-143.

[23] Gregg, Paul, and Barbara Petrongolo. "Random or Non-Random Matching? Implications for the Use of the UV Curve as a Measure of Matching Performance. " Centre for Economic Performance & Institute of Economics, 1997.

[24] Groshen, Erica L., and Simon M. Potter. "Has structural change contributed to a jobless recovery? " *Current Issues in Economics and Finance* 9, no. 8 (2003).

[25] Hamilton, James D., and Michael T. Owyang. "The propagation of regional recessions." *The Review of Economics and Statistics* 94, no. 4 (2012): 935-947.

[26] Head, Allen, and Huw Lloyd-Ellis. "Housing Liquidity, Mobility, and the Labour Market." *The Review of Economic Studies* 79, no. 4 (2012): 1559-1589.

[27] Jackman, Richard., Richard Layard, and Savvas Savouri. "Mismatch: A Framework for Thought." In *Mismatch and Labour Mobility*. Cambridge University Press, 1991.

[28] Jackman, Richard., Christopher Pissarides, and Savvas Savouri. "Labour Market Policies and Unemployment in the OECD." *Economic Policy* 5, no. 11 (1990): 449–90.

[29] Jones, David R., and D. N. Manning. "Long Term Unemployment, Hysteresis and the Unemployment–Vacancy Relationship: A Regional Analysis." *Regional Studies* 26, no. 1 (1992): 17-29.

[30] Kass, Robert E., and Adrian E. Raftery. "Bayes Factors." *Journal of the American Statistical Association* 90, no. 430 (1995): 773-795.

[31] Kaufmann, Sylvia. "K-state switching models with time-varying transition distributions—Does loan growth signal stronger effects of variables on inflation?." *Journal of Econometrics* 187, no. 1 (2015): 82-94.

[32] Michaillat, Pascal, and Emmanuel Saez. "Beveridgean unemployment gap." *Journal of Public Economics Plus* 2 (2021): 100009.

[33] Owyang, Michael T., Jeremy Piger, and Howard J. Wall. "Business Cycle Phases in US States." *The Review of Economics and Statistics* 87, no. 4 (2005): 604-616.

[34] Owyang, Michael T., Jeremy M. Piger, Howard J. Wall, and Christopher H. Wheeler. "The economic performance of cities: A Markov-switching approach." *Journal of Urban Economics* 64, no. 3 (2008): 538-550.

[35] Petrongolo, Barbara, and Christopher A. Pissarides. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39, no. 2 (2001): 390-431.

[36] Şahin, Ayşegül, Joseph Song, Giorgio Topa, and Giovanni L. Violante. "Mismatch unemployment." *American Economic Review* 104, no. 11 (2014): 3529-3564.

[37] Schulhofer-Wohl, Sam. "Negative Equity Does Not Reduce Homeowners' Mobility." No. w16701. *National Bureau of Economic Research*, 2011.

[38] Sterk, Vincent. "Home equity, mobility, and macroeconomic fluctuations." *Journal of Monetary Economics* 74 (2015): 16-32.

[39] Valletta, Robert G. "House lock and structural unemployment." *Labour Economics* 25 (2013): 86-97.

[40] Wall, Howard J., and Gylfi Zöega. "The British Beveridge curve: A tale of ten regions." *Oxford Bulletin of Economics and Statistics* 64, no. 3 (2002): 257-276.

A Vacancies Data

Prior to 1995, vacancies can be proxied by the HWI, which reflects the number of job openings advertised in newspapers. After the advent of the internet, job postings were increasingly posted online (or sometimes both online and in print advertisements). Thus, after 1995, the HWI alone cannot be viewed as a proper representation of vacancies. Ideally, we could combine the HWI and the HWOL to produce a composite index; however, both indices are only available for a few years of the post-1995 sample.

Barnichon (2010) suggests how to combine the HWI and HWOL—with some additional information and assumptions—into a composite index. The key component in constructing the composite index is determining the share of print advertising, S_t^p; then, combining this share with either HWI or HWOL, whichever is available, can yield the composite index. Barnichon assumes that the true (unobserved) print share is 1 prior to 1995 and declines monotonically as internet advertising diffuses. He assumes that the true print share follows a diffusion model. From 1995 to 2008, he computes an estimate of the print share, $\widehat{S}_t^p$. The techniques used for constructing the estimate $\widehat{S}_t^p$ varies over the subperiods where the data availability changes.

He then estimates the diffusion model parameters from the estimated print share, $\widehat{S}_t^p$, from 1995 to 2008 to obtain the true latent print share, S_t^p. Next, he generates values of the true latent print share, S_t^p, post-2008 by propagating the diffusion model using the estimated parameters. The following subsections provide details of Barnichon's procedure using aggregate data; the final two subsections show how the procedure produces problematic results for some MSAs and how we adapt the procedure using Bayesian methods to correct for these MSA-specific issues.

A.1 Computing $\widehat{S}_t^p$ from 1995 to 2008

From 1995 to 2008, the HWI existed but the HWOL did not. Barnichon's suggested approach is to fit a polynomial trend to print advertising from the beginning of the sample (in our case, 1955) to 2008. He then estimates the print share $\widehat{S}_t^p$ for the years 1995 to 2000 by setting $\widehat{S}_t^p = \widetilde{P}_t/\widetilde{P}_{1994}$, where $\widetilde{P}_t$ is the value of the polynomial computed at time t and $\widetilde{P}_{1994}$ is the value of the polynomial computed in 1994.[16]

[16] Barnichon also suggests estimating a Beveridge curve over 1990-1994, where $P_t = \alpha + \beta u_t + \eta_t$. He then uses α and β to forecast total advertising from 1995-2000 by assuming that, before 1995, online advertising had not yet

From 2000 to 2008, Barnichon also uses the JOLTS data to generate his estimate of the print share to check his estimates. The key assumption is that total vacancies comoves with the job openings series, J_t, from the JOLTS. Then, starting with the estimated print share from 2000 estimated from the polynomial trend, the change in the log of in the estimated print share can be computed from the difference between the changes in the log levels of print advertising and job openings:

$$\Delta \ln \widehat{S}_t^p = \Delta \ln P_t - \Delta \ln J_t.$$

A.2 Computing S_t^p from $\widehat{S}_t^p$

While the methods in the previous subsection yield estimates of the print share, the actual print share is considered to be a latent variable approximated by $\widehat{S}_t^p$ that has unit value in 1994 and then follows the MISM diffusion:

$$S_t^p = 1 - \frac{1 - e^{-\lambda_1 t}}{1 + \lambda_2 e^{-\lambda_3 t}}, \tag{12}$$

where the parameters λ_1, λ_2, and λ_3 are fit with $\widehat{S}_t^p$ estimates from 1995 to 2008 using nonlinear least squares with the additional restrictions that $\lambda_1, \lambda_2, \lambda_3 > 0$. After 2008, S_t^p is taken directly from (12) using the estimated parameters $\widehat{\lambda}_1$, $\widehat{\lambda}_2$, and $\widehat{\lambda}_3$.

A.3 Computing the Composite Index from S_t^p

Once the print shares are obtained from the methods above, constructing the composite index, v_t, is straightforward. However, the construction of the index again varies across subperiods as the data availability changes. From 1950 to 1995, composite index is identically HWI:

$$v_t = P_t.$$

From 1995 to 2005, we still have HWI but the print share is declining. We can compute the change in log v_t based on the change in the ratio of log HWI to the print share:

affected P_t. He backs out S_t^p for 1995-2000 using the forecasted values of print advertising to total advertising.

$$\Delta \ln v_t = \Delta \ln \frac{P_t}{S_{t-1}^p}, \tag{13}$$

starting with the last value of v_t in 1994. From 2005 to 2008, we have both HWI and HWOL and we can construct the index from:

$$\frac{\Delta v_t}{v_{t-1}} = S_{t-1}^p \frac{\Delta P_t}{P_{t-1}} + (1 - S_{t-1}^p) \frac{\Delta O_t}{O_{t-1}}. \tag{14}$$

Finally, after 2008, the index is constructed from the level of the online index and the print share:

$$\Delta \ln v_t = \Delta \ln \frac{O_t}{1 - S_{t-1}^p}. \tag{15}$$

A.4 Bayesian Estimation of the Diffusion Model Parameters

Computation of the indices for the MSAs is complicated by the idiosyncrasies in timing of the global maxima of the MSA-level print share series. One of the issues is that 1994 may be a global max for aggregate P_t but may not be a global max for the MSA. Thus, estimating the print share by normalizing the polynomial trend around 1994 may result in estimated print shares that are greater than one for periods subsequent to 1995. Moreover, for some of the MSAs, the estimated print shares $\widehat{S}_t^p$ are trending upward from 1995 to 2008, causing the estimates of the MISM parameters to yield a process for which online advertising never diffuses (i.e., the print share is identically 1 for all time).

One solution is to set the print share to one for all values up to the global max for the city. This implies that the internet did not diffuse in the city until the time of the global max, which seems a bit unlikely. Moreover, if the global maximum is too late in the sample, there may not be enough data to estimate the MISM parameters, and the diffusion process will be biased. The second solution is to estimate the print share normalizing the polynomial trend as before, which implies that the estimated print share could be greater than one (or, equivalently, set to one) for some periods in between 1995 and 2000. Again, if the global maximum is too close to 2000 (or even after 2000), there may not be sufficient information to estimate MISM parameters.

To resolve this issue, we employ Bayesian methods to ensure that the MISM print share S_{nt}^p eventually declines over time. To do this, we first estimate the national print share, $\widehat{S}_t^p$, and the

MSA-level print share $\widehat{S}^p_{nt}$ as documented above and then fit the MISM to the national data, obtaining the national level, $\lambda_1, \lambda_2, \lambda_3$, using Bayesian methods and a relatively uninformative prior. We then estimate the MSA-level MISM using the national parameters to center the prior distributions of the MSA parameters.

Because the MISM is nonlinear in the parameters, we form the posterior distributions for $\lambda_1, \lambda_2, \lambda_3$ using Metropolis-Hastings (MH). Given some starting values $\widetilde{\lambda}_1, \widetilde{\lambda}_2, \widetilde{\lambda}_3$, the MH algorithm generates the unknown posterior distribution (the target density) by drawing a candidate vector of parameters, $\lambda_1^*, \lambda_2^*, \lambda_3^*$, from a joint proposal distribution and accepts the candidate with probability proportional to the ratio of the likelihoods of the proposal to the last accepted candidate. The closer the proposal density is to the target density, the more efficient the algorithm will be. Because of the restrictions on the model parameters, we choose independent random walk gamma distributions as the candidate densities for λ_1 and λ_3 and a random walk truncated normal distribution for λ_2.

A.5 Computing Vacancies

We start at the end of our time sample computing vacancies as $V_t = \frac{O_t}{1-S_t^p}$ from June 2008 forward. We then project the index backwards using the inverse of the calculations from Barnichon, (13), (14), and (15). This method results in V_t being in units of O_t, which is an outright measure of advertisements.

B Unemployment Rate and Labor Force Data

The unemployment rate and the size of the labor force at the MSA level are obtained from the Bureau of Labor Statistics' Local Area Unemployment Statistics (LAUS). These data are available monthly from July 1989 through the present. They are seasonally adjusted by the BLS using the SEATS program.

To measure the unemployment rate and labor force before 1989, we use the change in unemployment and labor force calculated from the Basic Monthly Current Population Survey (CPS) Public Use Microdata. The CPS data for January 1978 through December 1990 are accessed from Unicon CPS Utilities. The specific Unicon variables used for MSA identification are smsark (1978-1985) and msark (1985-1990). Not all MSAs are defined before 1985, so we have an unbalanced panel with some MSAs missing data between 1978 and 1985 (see Table 1 for exact MSAs). The only other variables we use in the CPS microdata are mlr for labor force status and the basic monthly CPS weights to calculate aggregate numbers.

After calculating the time series data for each MSA, we apply seasonal adjustment using the Census X-13 ARIMA seasonal adjustment program. The seasonally adjusted data are then smoothed using a Christiano-Fitzgerald band pass filter to remove cycles less than 2 or greater than 12 months for the unemployment rate and less than 2 or greater than 60 months for the labor force. The filter is applied to extract a smoother trend from the otherwise extremely volatile sub-national data CPS data. To account for changing MSA definitions over time, we wrote a crosswalk between the 1970, 1980, and 2009 MSA definitions with the goal to match the 2009 MSA definitions as closely as possible (crosswalk files available upon request). Because discrete jumps at definition changes still remain in the data, we use backwards changes in the CPS data and back out the levels based from the first observation of the LAUS data back to either 1978 or 1986 depending on the MSAs availability.

$$UR_{t-1} = UR_t^{LAUS} * \left(1 + \frac{(UR_{t-1}^{CPS} - UR_t^{CPS})}{UR_t^{CPS}}\right) \tag{16}$$

C The Sampler

The model can be estimated via the Gibbs sampler using four blocks: (1) $\Psi = \{\alpha, \beta, \tau\}$, (2) Σ, (3) the multinomial logit parameters δ, and (4) $\mathbf{\Gamma}$. We impose a set of standard priors on the parameter blocks: $\alpha \sim N(\mathbf{a}_0, \mathbf{A}_0)$; $\beta \sim N(\mathbf{b}_0, \mathbf{B}_0)$; $\sigma_{nn}^{-2} \sim \text{Gamma}(\omega_0, \Omega_0)$; and $\delta \sim N(\mathbf{d}_0, \mathbf{D}_0)$. The prior hyperparameters are given in Table 1. Let $\mathbf{\Theta}$ represent the full set of parameters and $\mathbf{\Theta}_{-m}$ represent the full set of parameters with m excluded. We let $\mathbf{Y} = \{\mathbf{U}, \mathbf{V}\}$ collect the time series data and $\mathbf{X} = \{\mathbf{x}_n\}_{n=1}^N$ collect the cross sectional covariate data. The Gibbs sampler forms the joint posterior distribution of the full set of model parameters from draws of the conditionals for each. The following sections describe the draws from the conditional distributions.

C.1 Generating $\Psi \mid \mathbf{\Theta}_{-\Psi}$, $\mathbf{Y}$, $\mathbf{X}$

The set of coefficients Ψ is sampled from a straightforward posterior:

$$\Psi \sim N\left(\mathbf{m}, \mathbf{M}\right),$$

where

$$\mathbf{M} = \left(\mathbf{M}_0 + \zeta_T' \zeta_T\right)^{-1},$$

$$\mathbf{m} = \mathbf{M}\left(\mathbf{M}_0^{-1}\mathbf{m}_0 + \sigma_n^{-2}\zeta_T' U_T\right),$$

$\zeta_T = [\zeta_1, ..., \zeta_T]'$ and $\zeta_t = [I, \mathbf{\Gamma} \odot D, \mathbf{G}V_t]'$.

C.2 Generating $\Sigma \mid \mathbf{\Theta}_{-\Sigma}$, Y, X

The prior for elements of Σ is inverse gamma. Conditional on their cluster affiliations, the other model parameters, and the data, we can sample σ_n^{-2} from a Wishart posterior

$$\sigma_n^{-2} | \mathbf{Y}, \varsigma_K, \Theta_{-\sigma_n^2} \sim \mathbf{W}\left(\omega_n, \Omega_n\right),$$

where $\omega_n = \left(\omega_0 + T\right)/2$ represents the degrees of freedom and the scale parameter is

$$\Omega_n = \frac{1}{2}\left(\Omega_0 + \varepsilon_T \varepsilon_T'\right).$$

C.3 Generating $\delta \mid \mathbf{\Theta}_{-\delta}$, Y, X

We draw the prior hyperparameters using the difference in random utility model method from Frühwirth–Schnatter and Frühwirth (2010) and Kauffman (2015). Under this model, the posterior for $\boldsymbol{\delta}_k$ follows a normal distribution

$$\boldsymbol{\delta}_k \sim N(\boldsymbol{d}_k, \boldsymbol{D}_k)$$

where

$$d_k = D_k \left(\sum_{n=1}^{N} \frac{\boldsymbol{x}_{k,n}\left[\omega_{k,n} + \log(\chi_{-k,n})\right]}{s^2_{R_{k,n}}} + \boldsymbol{D}_0^{-1} d_0 \right),$$

$$D_k = \left(\sum_{n=1}^{N} \frac{\boldsymbol{x}_{k,n}\boldsymbol{x}_{k,n}'}{s^2_{R_{k,n}}} + \boldsymbol{D}_{0k}^{-1} \right)^{-1},$$

and the differences in utility $\omega^h_{k,n}$, constants $\chi^h_{-k,n}$, and standard deviations $s_{R^h_{k,n}}$ are defined as in Frühwirth–Schnatter and Frühwirth (2010).

C.4 Generating $\Gamma \mid \boldsymbol{\Theta}_{-\Gamma}$, Y, X

We draw the cluster membership indicator h_{nk} city-by-city by combining the conditional likelihood and logistic prior:

$$\Pr\left(h_{nk} = 1 | \Theta_{-\boldsymbol{\Gamma}}, \boldsymbol{U}, \boldsymbol{X}\right) = \frac{p(\boldsymbol{U}_n | h_{nk} = 1, \Theta_{-\boldsymbol{\Gamma}}) p(h_{nk} = 1 | \boldsymbol{\delta}, \boldsymbol{X})}{\sum_{j=1}^{K} p(\boldsymbol{U}_n | h_{nj} = 1, \Theta_{-\boldsymbol{\Gamma}}) p(h_{nj} = 1 | \boldsymbol{\delta}, \boldsymbol{X})}.$$