The Cyclical Component of Labor Market Polarization and Jobless Recoveries in the US

Paul Gaggl, and Sylvia Kaufmann

Working Paper 14.03

The Cyclical Component of Labor Market Polarization and Jobless Recoveries in the US

April 28, 2014

Abstract: We analyze quarterly occupation-level data from the US Current Population Survey for 1976-2013. Based on common cyclical employment dynamics, we identify two clusters of occupations that roughly correspond to the widely discussed notion of "routine" and "non-routine" jobs. After decomposing the cyclical dynamics into a cluster-specific ("structural") and an occupation-specific ("idiosyncratic") component, we detect significant structural breaks in the systematic dynamics of both clusters around 1990. We show that, absent these breaks, employment in the three "jobless recoveries" since 1990 would have recovered significantly more strongly than observed in the data, even after controlling for observed idiosyncratic shocks.

JEL: J21, E32, E24
Keywords: employment polarization, jobless recoveries, factor model

Paul Gaggl[1]

University of North Carolina at Charlotte
Belk College of Business
Department of Economics
9201 University City Blvd
Charlotte, NC 28223-0001
Email: pgaggl@uncc.edu

Sylvia Kaufmann[1]

Study Center Gerzensee
Dorfstrasse 2, P.O. Box 21
CH-3115 Gerzensee
Switzerland

Email: sylvia.kaufmann@szgerzensee.ch

[1]We are grateful to Nir Jaimovich, Jim Nason, Travis Berge as well as participants of the 2013 SEA meetings for extremely valuable comments and feedback. We would further like to thank Debapriti Chakraborty and Jonathan Viscount for their invaluable help with data collection.

1. Introduction

The "recoveries" after the last three recessions in the US were "jobless". That is, while output was recovering, jobs were not—or at least at a much slower pace. To this date (April 2014), it remains unclear whether this phenomenon is simply due to a sequence of unfavorable shocks or whether the US labor market is undergoing systematic structural change. Following Jaimovich and Siu (2012), we consider labor market "polarization" as one possible structural cause and develop a statistical framework to disentangle it from idiosyncratic shocks.

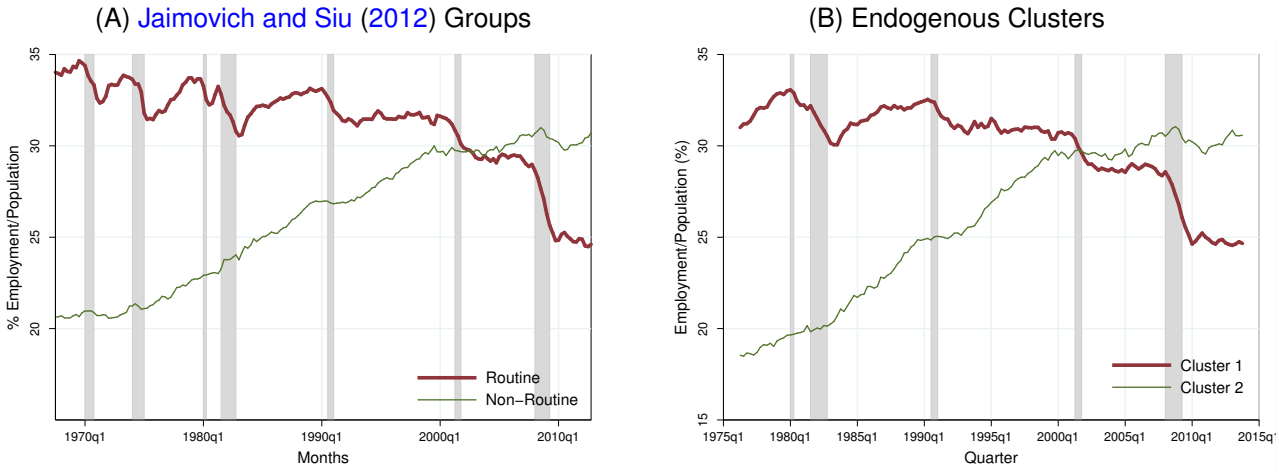
A number of recent studies document that, over the past several decades, the employment share of the lowest and highest skilled occupations increased, while it declined for middle skilled jobs. Over the same period, wages for middle skilled occupations grew substantially less than wages at the tail ends of the skill distribution. These trends are in large part attributed to the widespread adoption of computing technology and the rising importance of offshoring, which substitutes for tasks performed by middle skilled workers.[2]

Much less is known about the cyclical aspects of this apparent trend. In pioneering work, Jaimovich and Siu (2012) use a crude and highly aggregated mapping of skills into jobs and document that 92% of the decline in "routine" jobs in the US—ones that are considered easily replaceable by technology and require "middle" skills—occurred within a 12 month window of NBER dated recessions since the mid 1980s.[3] Moreover, as is immediately apparent from panel A of Figure 1, "routine" (middle skill) occupations used to strongly rebound after recessions prior to 1990, while these swift rebounds were absent in the last three recessions. To the contrary, "non-routine" occupations—ones considered to directly or indirectly complement technology and comprising both low and high skilled workers—appear to be fairly immune to recessions and do not seem to have experienced a marked change in employment dynamics around 1990. Jaimovich and Siu

[2] Acemoglu (1999) was the first to document employment polarization in the US over the period 1983–1993. For more recent periods, Goos and Manning (2007) find similar patterns in the UK, Goos et al. (2009) for 16 EU countries, and Autor et al. (2008) as well as Autor and Dorn (2013) for the US. Autor and Dorn (2013) further show compelling evidence that PC adoption was more prevalent in areas with a historical abundance of workers performing "routine tasks".

[3] The crude mapping of aggregate job categories into three broad skill groups is based on the work surveyed in Acemoglu and Autor (2011).

Figure 1: Group Employment Trends



Notes: Panel A plots employment trends as reported in Jaimovich and Siu (2012). Data prior to 1983 are taken from the US Department of Labor's *Employment & Earnings* publications and from FRED thereafter. The occupations are grouped as suggested in Acemoglu and Autor (2011). Panel B illustrates the cumulative growth of employment/population in each occupation assigned to factors 1 and 2 in model (1), which are listed in Table 3. Data for this graph are directly constructed from the monthly basic CPS files for the consistent panel of occupations compiled by Dorn (2009). The levels in both figures are imputed from quarterly growth rates and start with the level of employment/population at the beginning of each sample. We seasonally adjusted all time series from both data sources using the US Census X11 method.

(2012) then provide a counterfactual based on descriptive statistics to illustrate that the change in the cyclical pattern of routine jobs alone may be an important driver for the jobless recoveries since 1990.

If these stark cyclical patterns are truly due to a distinguishing feature, that is common to occupations within each broad category considered by Jaimovich and Siu (2012), then we should be able to identify this group-specific component as well as potential group-specific structural breaks form high frequency employment dynamics in the underlying detailed occupations.

To accomplish this, we estimate a dynamic factor model with latent clusters based on detailed occupation level data from the US Current Population Survey (CPS) for the period 1976-2013. Our model is designed to capture three important aspects of the data: We estimate and thereby identify occupation clusters displaying common cyclical dynamics in the employment to population ratio, we allow for group-specific structural breaks in the common cyclical dynamics, and we ultimately disentangle structural (cluster-specific) dynamics from idiosyncratic shocks.

In particular, our model identifies two occupation clusters that broadly coincide with groups that have previously been labeled "routine" and "non-routine" jobs, respectively (see Acemoglu

and Autor, 2011, for a survey). Moreover, a comparison of panels A and B in Figure 1 illustrates that the aggregate dynamics of these clusters strongly resemble the dynamic patterns of the crude aggregates considered by Jaimovich and Siu (2012).[4]

Our estimates further deliver a fairly agnostic measure of "exposure to polarization" at the occupation level. Most previous studies classify occupations into "routine" and "non-routine" groups using cross-sectional information on the "task content" of each occupation (see Acemoglu and Autor, 2011, for a survey). This approach faces numerous challenges, including the lack of high quality longitudinal information, difficult to interpret ordinal metrics, as well as the necessity to make ad hoc assumptions about these ordinal metrics (see Autor, 2013, for a detailed discussion of these difficulties). Our approach, on the other hand, is based on common cyclical dynamics across individual occupations and thereby circumvents these issues. Nevertheless, the estimated classification maps surprisingly well into groupings found in previous research. In particular, the estimated occupation clusters suggest that traditional blue collar jobs as well as sales and administrative support are most strongly associated with the gradually disappearing occupation group (cluster 1), while managerial and service jobs—such as child and health care—are most representative for the strongly growing occupation group (cluster 2).

To capture the group-specific business cycle dynamics, we adopt a Markov switching structure that allows for asymmetric dynamics across expansions and recessions and for a potential break in these dynamics. Based on this flexible setup, we find a significant structural break in the group-specific cyclical dynamics for both "routine" *and* "non-routine" occupations around 1990. In fact, our analysis suggests that the break in "non-routine" occupations was at least as pronounced as that in "routine" occupations.

Finally, we construct counterfactuals to assess whether the break in the structural components was large enough to explain a significant portion of the low aggregate job growth during the last three recoveries in the US. As descriptively suggested by Jaimovich and Siu (2012), we find that the aggregate employment/population ratio would have recovered significantly more strongly

[4]In Section 3 we discuss in detail why the levels in panels A and B of Figure 1 don't match up. We argue that the information about the dynamics of aggregate employment contained in either aggregation scheme is essentially the same.

in the absence of the observed structural change around 1990. This is an important finding as it suggests scope for directed policy interventions—in addition to undirected policy measures like monetary easing—that may directly affect the structural aspect of jobless recoveries.

While our findings are generally in line with Jaimovich and Siu (2012) as well as Cortes et al. (2014), they contrast the main conclusions of Foote and Ryan (2012). These authors argue that large aggregate shocks caused jobless recoveries in the US and that labor market polarization played only a minor role in this context. They reach this conclusion by observing that the variance share explained by industry-skill-specific variation is low compared to the variance share explained by a single unobserved common component. Our results refine this picture. We identify two distinct clusters of occupations with common cyclical dynamics, which jointly explain at least 60% of the variation in the data. The two common components display vastly different dynamic patterns over the business cycle and each experiences a component-specific structural break in the cyclical employment dynamics after 1990. Furthermore, our endogenously determined clusters are in line with the grouping of occupations suggested in previous work (Acemoglu and Autor, 2011) and the dynamic patterns of the common components reflect the stylized features observed in panel A of Figure 1.

2. Empirical Model

Let y_{it} denote the growth in the employment/population ratio of occupation $i = 1, \ldots, N$ in period $t = 1, \ldots, T$. We capture the notion of "common dynamics" within a set of K "occupation clusters" by specifying a factor model with K factors. To identify distinct occupation clusters, each occupation, i, is exclusively assigned to one factor. We write the model compactly as

$$
\begin{aligned}
y_{it} &= \sum_{k=1}^{K} \lambda_{ik} f_{kt} + \varepsilon_{it} && (1)\\
&= \lambda_{i\delta_i} f_{\delta_i t} + \varepsilon_{it} && (2)\\
\phi_k(L) f_{kt} &= \mu_{kS_t} + \nu_{kt},\ \nu_{kt} \sim N(0,1) && (3)\\
\psi_i(L)\varepsilon_{it} &= \epsilon_{it},\ \epsilon_{it} \sim N\left(0, \sigma_i^2\right) && (4)
\end{aligned}
$$

where going from (1) to (2) reflects the fact that the latent classification indicator $\delta_i \in \{1, \ldots, K\}$ exclusively assigns one factor to each occupation: $\lambda_{ik} \neq 0$ if $\delta_i = k$ and 0 otherwise, with $k = 1, \ldots, K$. To capture remaining idiosyncratic dynamics, the idiosyncratic shocks ε_{it} follow an autoregressive (AR) process of order q, $\psi_i(L) = 1 - \psi_1 L - \cdots - \psi_q L^q$.

Ex post, the latent classification indicator δ_i is inferred from the co-movement in y_{it} across occupations, yet our framework allows to incorporate occupation-specific, cross-sectional information information into the prior classification probability distribution:

$$P(\delta_i = k | Z_i, \beta) = \frac{\exp(Z_i \beta_k)}{1 + \sum_{j=2}^{K} \exp(Z_j \beta_j)} \quad k = 1, \ldots, K \text{ and } \beta_1 = 0 \tag{5}$$

where Z_i' is a vector containing fixed effects and occupation-specific characteristics. By default, the reference effects are the ones for the probability to load on the first factor, $\beta_1 = 0$. All our results in Section 4 are based on the special case of a discrete uniform prior classification probability of $P(\delta_i = k) = 1/K$, in which we incorporate no additional prior information, and which is nested in (5) when setting $Z_i = 1$ and $\beta_k = 0, \forall k$.

The factor equation (3) is dynamic of order p and restricted to the stationary region, i.e., we assume the roots of the characteristic equation of $\phi_k(L)$ to lie outside the unit circle. Period-specific factor growth, μ_{kS_t}, is determined by the second latent indicator, $S_t \in \{1, \ldots, 4\}$, which takes one of four values. In all our empirical applications, the four possible states have the following ex-post interpretation: 1 = pre-break recession, 2 = pre-break expansion, 3 = post-break recession, 4 = post-break expansion.

We assume that the factor processes are independent of each other. However, since each occupation is governed by a single factor, we could relax this independence assumption and allow for a general vector autoregressive dynamic structure driving the factors. Since our goal is to provide a first set of empirical results with a clear-cut interpretation, we leave this generalization for a future investigation.

We specify a time-varying transition distribution for S_t, in which real GDP growth, x_t, determines the transition distributions between recessions and expansions. We define the state transi-

tion probability matrix as

$$\boldsymbol{\xi}_t = \begin{bmatrix} \xi_{11,t} & \xi_{12,t} & 0 & \xi_{14,t} \\ \xi_{21,t} & \xi_{22,t} & \xi_{23,t} & 0 \\ 0 & & \xi_{33,t} & \xi_{34,t} \\ & & \xi_{43,t} & \xi_{44,t} \end{bmatrix}, \tag{6}$$

where $\xi_{jl,t} = P\left(S_t = l | S_{t-1} = j\right)$ is the probability to switch to state l if state j prevailed in period $t-1$. To ensure identification, and inspired by the stylized patterns in Figure 1, we put two sets of restrictions on the transition probabilities. First, we only allow a structural break to happen once. That is, once states 3 or 4 are reached, the economy cannot go back to either 1 or 2. This is implemented by setting the lower-left block of transition probabilities to 0. Second, a potential structural break must happen in transition from a recession to an expansion, or vice versa. For example, we don't allow the economy to switch from a pre-break expansion into a post-break expansion. This is implemented by setting the upper-right diagonal elements of the transition distribution to 0.

To specify the period specific transition probabilities in $\boldsymbol{\xi}_t$ we use a logit prior and assume state 1 to be the reference state, $\gamma_{j1,\cdot} = 0$, $j = 1, 2$:

$$\boldsymbol{\xi}_t = \begin{bmatrix} \frac{1}{1+\sum_{s=\{2,4\}} \exp(X_t'\gamma_{1s})} & \frac{\exp(\gamma_{12,0}+\gamma_{12,1}x_t+\gamma_{12,2}t)}{1+\sum_{s=\{2,4\}} \exp(X_t'\gamma_{1s})} & 0 & \frac{\exp(\gamma_{14,2}t)}{1+\sum_{s=\{2,4\}} \exp(X_t'\gamma_{1s})} \\ \frac{1}{1+\sum_{s=\{2,3\}} \exp(X_t'\gamma_{2s})} & \frac{\exp(\gamma_{22,0}+\gamma_{22,1}x_t+\gamma_{22,2}t)}{1+\sum_{s=\{2,3\}} \exp(X_t'\gamma_{2s})} & \frac{\exp(\gamma_{23,2}t)}{1+\sum_{s=\{2,3\}} \exp(X_t'\gamma_{2s})} & 0 \\ 0 & 0 & \frac{\exp(\gamma_{33,0}+\gamma_{33,1}x_t)}{\sum_{s=3}^{4} \exp(\gamma_{3s,0}+\gamma_{3s,1}x_t)} & \frac{\exp(\gamma_{34,0}+\gamma_{34,1}x_t)}{\sum_{s=3}^{4} \exp(\gamma_{3s,0}+\gamma_{3s,1}x_t)} \\ 0 & 0 & \frac{\exp(\gamma_{43,0}+\gamma_{43,1}x_t)}{\sum_{s=3}^{4} \exp(\gamma_{4s,0}+\gamma_{1,4s}x_t)} & \frac{\exp(\gamma_{44,0}+\gamma_{44,1}x_t)}{\sum_{s=3}^{4} \exp(\gamma_{4s,0}+\gamma_{4s,1}x_t)} \end{bmatrix}, \tag{7}$$

in which $X_t = (1, x_t, t)'$ contains a constant, GDP growth, x_t, which determines the transition probabilities between recessions and recoveries, and a time trend, t, which is used to introduce prior information on the break date. The parameters $\gamma_{jl,m}$, with $j, l = 1, \ldots, 4$, and $m = 0, 1, 2$, correspond to the state-dependent, state-specific effects of the variables in X_t. So, $\gamma_{jl,m}$ is either the mean ($m = 0$) effect, the effect of GDP growth ($m = 1$) or the trend effect ($m = 2$) on the probability of switching from state j to state l. The denominators are written in a general form,

but note that appropriate elements of γ_{14} and γ_{23} are restricted to zero.

Time explicitly enters the transition distribution of the states 1 and 2, to include prior information on the break date around 1990. We normalize t to be zero in the third quarter of 1990, which corresponds to the peak of the 1980s expansion. The effect of time t should be decreasing for $\xi_{j2,t}$, $j = 1, 2$ and increasing for $\xi_{14,t}$ and $\xi_{23,t}$. Therefore, we expect to estimate $(\gamma_{12,2}, \gamma_{22,2}) \leq 0$ and $(\gamma_{14,2}, \gamma_{23,2}) > 0$.

The prior specification on the latter parameters is currently quite informative and favors a break after the expansion of the 1980s going into the early 1990s recession. We explicitly make this choice based on the stylized patterns in Figure 1 and because our prime interest is to identify the existence and potential *effects* of a structural break around 1990, rather than the timing of the break itself.[5] The current specification provides a convenient framework to conduct posterior inference on the dynamics of the structural component before and after 1990. Yet, we set out to construct a framework that is general enough to conduct inference on the break date itself in future research.

2.1. Bayesian Posterior Inference: Likelihood and Priors

We adopt a Bayesian approach to draw posterior inference on the latent factors $\boldsymbol{f} = (f_1, \ldots, f_T)$, on the values of both the unobserved latent indicators, $\boldsymbol{S} = (S_1, \ldots, S_T)$ and $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_N)$, and all population parameters, $\theta = \{\boldsymbol{\lambda}, \boldsymbol{\psi}, \boldsymbol{\phi}, \boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\gamma}\}$. The complete data likelihood of the model is

$$L\left(\boldsymbol{y}|\boldsymbol{f}, \boldsymbol{\delta}, \theta\right) = \prod_{i=1}^{N} \prod_{t=q+1}^{T} l\left(y_{it}|f_{\delta_i t}, \theta\right) \tag{8}$$

where the period t density contribution of unit i, $l\left(y_{it}|\cdot\right)$, is normal

$$l\left(y_{it}|f_{\delta_i t}, \theta\right) = \frac{1}{\sqrt{2\pi}\sigma_{\epsilon_i}} \exp\left\{-\frac{1}{2\sigma_{\epsilon_i}^2}\left(y_{it}^* - \sum_{j=0}^{q} \psi_j \lambda_{i\delta_i} f_{\delta_i, t-j}\right)^2\right\} \tag{9}$$

[5]Note that, the prior is not dogmatic in the sense that the break could still be estimated to occur after the 1990s recession, however.

and y_{it}^* represents the filtered value $y_{it}^* = y_{it} - \psi_{i1} y_{i,t-1} - \dots \psi_{iq} y_{i,t-q}$. Our prior for the unobserved factors is

$$\pi\left(\boldsymbol{f}|\mathbf{S},\theta\right) = \prod_{t=p+1}^{T} \pi\left(f_t|f^{t-1},S_t,\theta\right)\pi\left(f^p|S^p,\theta\right) \tag{10}$$

where $\pi\left(f^p|S^p,\theta\right)$ is the density of the initial states. The prior for the unobserved state indicator factorizes to

$$\pi\left(\mathbf{S}|\mathbf{x},\theta\right) = \prod_{t=p+1}^{T} \pi\left(S_t|S^{t-1},x_t,\theta\right) \tag{11}$$

The prior for the classification indicator $\boldsymbol{\delta} = (\delta_1,\dots,\delta_N)$ is assumed to be uniform discrete, i.e. $\boldsymbol{\beta} = 0$, $P\left(\delta_i = k\right) = 1/K,\ \forall i$.

To complete the model, we assume that the parameters are block-independent a priori, which allows us to write their joint distribution as

$$\pi\left(\theta\right) = \pi(\boldsymbol{\lambda}|\boldsymbol{\delta})\pi(\boldsymbol{\psi})\pi(\boldsymbol{\phi})\pi(\boldsymbol{\mu})\pi(\boldsymbol{\sigma})\pi(\boldsymbol{\gamma}) \tag{12}$$

and to specify the following individual priors:

1. $\pi(\boldsymbol{\lambda}|\boldsymbol{\delta}) = \prod_{i=1}^{N}\pi(\lambda_{i\delta_i}) = \prod_{i=1}^{N} N(\mathrm{l}_0,\mathrm{L}_0)$
2. $\pi(\boldsymbol{\psi}) = \prod_{i=1}^{N}\pi(\psi_{i1},\dots,\psi_{iq}) = \prod_{i=1}^{N} N(q_0,Q_0)I_{\{Z(\psi_i)>1\}}$
   where $I_{\{\cdot\}}$ is the indicator function and $Z(\varphi) > 1$ means that the characteristic roots of the process $\varphi(L)$ lie outside the unit circle.
3. $\pi(\boldsymbol{\phi}) = \prod_{k=1}^{K}\pi(\phi_{k1},\dots,\phi_{kp}) = \prod_{k=1}^{K} N(p_0,P_0)I_{\{Z(\phi_k)>1\}}$
4. $\pi(\boldsymbol{\mu}) = \prod_{k=1}^{K}\pi(\mu_{k1},\dots,\mu_{k4}) = \prod_{k=1}^{K} N(m_0,M_0)$
5. $\pi(\boldsymbol{\sigma}) = \pi(\sigma_1^2,\dots,\sigma_N^2) = \prod_{i=1}^{N} IG(\mathrm{e}_0,\mathrm{E}_0)$
6. $\pi(\boldsymbol{\gamma}) = \prod_{s=2}^{S}\pi(\boldsymbol{\gamma}_s) = \prod_{s=2}^{S} N(g_{0s},G_{0s})$
   where $\gamma_s = (\gamma_{1s},\dots,\gamma_{Ss})$, $s = 2,\dots,4$, and g_{0s} and G_{0s} have appropriate dimensions.

2.2. Posterior Sampler

We briefly outline the sampling steps here and refer the interested reader to Appendix A for more details.

1. Draw the latent factors from $\pi\left(\boldsymbol{f}|\boldsymbol{y},\mathbf{S},\boldsymbol{\delta},\theta\right)$ using the efficiency-based sampler developed by Chan and Jeliazkov (2009) (see also Kaufmann and Schumacher, 2013).
2. Given the factors, we use multi-move sampling (adjusted for the restricted transition distribution (7)) to draw the states from $\pi\left(\mathbf{S}|\mathbf{x},\boldsymbol{f},\boldsymbol{\mu},\boldsymbol{\gamma},\boldsymbol{\phi}\right)$. For details see Chib (1996) or Frühwirth-Schnatter (2010).
3. Given a draw for the states we draw the parameters governing the transition distribution from $\pi\left(\boldsymbol{\gamma}|\mathbf{S},\mathbf{x}\right)$. Introducing two layers of data augmentation, the non-linear, non-normal model (7) becomes linear-normal and we can sample from posterior conditional normal distributions. For details see Frühwirth-Schnatter and Frühwirth (2010), and Kaufmann (2011) for the extension to sampling state-dependent parameters.
4. To draw the latent classification indicator from $\pi\left(\boldsymbol{\delta}|\boldsymbol{y},\boldsymbol{f},\boldsymbol{\psi},\boldsymbol{\sigma}\right)$ we sample independently over N from the discrete distribution

$$\pi\left(\delta_i=k|\boldsymbol{y}^i,\boldsymbol{f},\psi_i,\sigma_i^2\right)\propto\exp\left\{-\sum_{t=q+1}^{T}\left(y_{it}^*-\frac{\sum y_{it}^*f_{ikt}^*}{\sum {f_{ikt}^*}^2}f_{ikt}^*\right)^2\right\}P\left(\delta_i=k\right),\quad k=1,\ldots,K \tag{13}$$

   where y_{it}^* and f_{ikt}^* represent the filtered values $y_{it}^*=y_{it}-\psi_{i1}y_{i,t-1}-\ldots\psi_{iq}y_{i,t-q}$ and $f_{ikt}^*=f_{kt}-\psi_{i1}f_{k,t-1}-\ldots\psi_{iq}f_{k,t-q}$, respectively.
5. Given draws for $\boldsymbol{f}$, $\boldsymbol{S}$ and $\boldsymbol{\delta}$, it is then straightforward to sample the remaining parameters from their posterior distributions (see Appendix A.4 for the derivation of the posterior moments):
   (a) $\pi(\boldsymbol{\lambda}|\boldsymbol{y},\boldsymbol{f},\boldsymbol{\delta},\boldsymbol{\sigma},\boldsymbol{\psi})=\prod_{i=1}^{N}N(\mathrm{l}_i,\mathrm{L}_i)$
   (b) $\pi(\boldsymbol{\psi}|\boldsymbol{y},\boldsymbol{f},\boldsymbol{\delta},\boldsymbol{\sigma},\boldsymbol{\lambda})=\prod_{i=1}^{N}N(q_i,Q_i)I_{\{Z(\psi_i)>1\}}$
   (c) $\pi(\boldsymbol{\phi},\boldsymbol{\mu}|\boldsymbol{f},\mathbf{S})\prod_{k=1}^{K}N(p_k,P_k)I_{\{Z(\phi_k)>1\}}$
   (d) $\pi(\boldsymbol{\sigma}|\boldsymbol{y},\boldsymbol{f},\boldsymbol{\delta},\boldsymbol{\psi},\boldsymbol{\lambda})=\prod_{i=1}^{N}IG(\mathrm{e}_i,\mathrm{E}_i)$

3. Data

Our main data source are detailed individual level data from the basic US Current Population Survey (CPS) covering the period 1976m1-2013m12. Based on CPS sampling weights we estimate

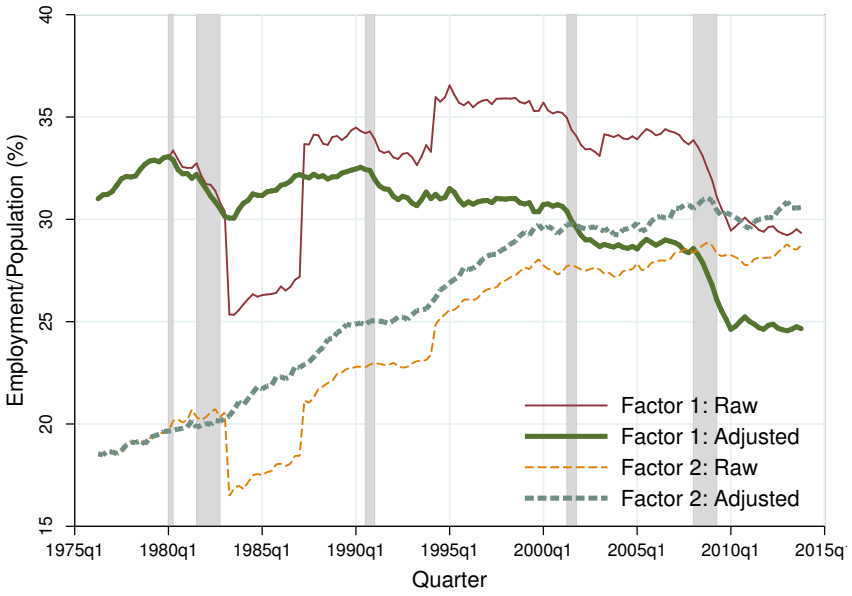
employment levels at the detailed occupation level on a monthly frequency throughout the entire sample. Since the US Department of Labor's (DOL) classification of occupations changes several times during our sample period, we aggregate individuals into a panel of 330 consistent occupations as designed by Dorn (2009).[6] For our baseline specification, summarized in Table 1, we further aggregate the detailed occupations into 21 groups, as also designed by Dorn (2009). Following Jaimovich and Siu (2012) we use the share of employment within the US population of age 16 and older to obtain each occupation's labor market dynamics over time.

To compare our results to Jaimovich and Siu (2012) we also replicate their dataset, spanning the period 1967-2012, for which data prior to 1983 are taken from the DOL's *Employment & Earnings* publications and from the Federal Reserve Bank of St. Louis' FRED database thereafter. These data contain the level of employment and are already aggregated to about 10 broad occupation groups. Unfortunately, the group definitions are neither fully consistent over time (especially prior to 1983) nor between the aggregates in FRED and in the *Employment & Earnings* publications. However, consistent with Jaimovich and Siu (2012), we are able to group occupations into the four broad occupation groups suggested by Acemoglu and Autor (2011): non-routine cognitive (professional, managerial, and technical occupations), routine cognitive (clerical, support, and sales occupations), routine manual (production and operative occupations), non-routine manual (service occupations). We further seasonally adjust all time series (from both data sources) using the US Census X11 method. Based on this dataset, panel A of Figure 1 displays employment as a share of population for the two groups of non-routine and routine occupations. As expected, this figure resembles Figure 4 in Jaimovich and Siu (2012).

One of the biggest challenges in working with the detailed CPS data are the frequent changes in the DOL's system for classifying occupations. Even Dorn's (2009) consistent panel features many jumps in the level of employment since various occupations "jump" from one group to another, new occupations are introduced, or old ones disappear. These jumps are not readily visible in long run comparisons (e.g. across decades) but they become immediately apparent at

[6]The DOL has implemented the latest change in their occupation classification system in 2011, and we thank Nir Jaimovich for providing a crosswalk (as used in Cortes et al., 2014) to extend Dorn's (2009) panel of occupations beyond this newest change of occupation classifications.

Figure 2: Employment Trends Based on Growth Rates



Notes: The figure illustrates the cumulative growth of employment/population in each occupation assigned to factors 1 and 2 in model (1), which are tabulated in Table 3. The imputed level series start with the level of employment/population in 1976q1 and illustrate the variation in growth rates used in our estimation procedure. The series labeled "raw" are are based on the unadjusted growth rates in the occupation level series while those labeled "adjusted" are based on a growth rate series in which the administrative "jumps" were interpolated based on the median January growth (all administrative jumps happen in January). Data for this graph are directly constructed from the monthly basic CPS files for the consistent panel of occupations compiled by Dorn (2009).

higher frequencies. To avoid this problem, Foote and Ryan (2012), who also study the cyclicality of labor market polarization, decide to use industry-skill cells as a proxy for jobs/tasks instead of occupations specified by the DOL.

However, since the level jumps are due to purely administrative changes, they always happen in a single month. Therefore, one way to accommodate the level jumps, is to use growth rates instead of levels and average out the jumps for the occupations in which administrative changes happen. Figure 2 shows the levels implied by our adjusted growth rate series. It is obvious that any adjustment procedure introduces some measurement error, but Figure 1 illustrates that the dynamic patterns in the level of routine and non-routine jobs implied by these adjusted growth rates is virtually the same as in the level series employed by Jaimovich and Siu (2012). In fact, our

Table 1: Preferred Model Specification

A. Specification		
Number of Factors	K	2
Number of States	S	1 = pre-recession, 2 = pre-expansion, 3 = post-recession, 4 = post-expansion
Factor AR lags	p	2
Idiosyncratic AR lags	q	2
B. Sample		
Employment Variable	$y_{i,t}$	Quarterly growth in employment/population (CPS basic files, SA X11)
Aggregate Variable	x_t	Quarterly growth in real GDP
Estimation Sample	T	150 quarters: 1976q2 – 2013q3
Occupation Groups	N	21 Dorn (2009) detailed occupation groups
B. Posterior Sampler		
Total Posterior draws		500,000
Burn-in		300000
Retained Observations		50,000 (every fourth draw after burn-in)

approach to adjust in growth rates is very similar in spirit to the "flows approach" of Cortes et al. (2014).

Ultimately, it should be clear that all four approaches, crude aggregation as in Jaimovich and Siu (2012), forming industry-skill cells as in Foote and Ryan (2012), the "flows approach" by Cortes et al. (2014), and our adjustment in growth rates, are an imperfect solution and introduce some form of measurement error. However, given the nature of administrative changes in the DOL's definition of occupations, these are the best options available.[7]

Finally, we employ real GDP growth as our aggregate measure to help identify business cycles and we draw these data from FRED.

4. Empirical Analysis

We estimate model (1) using the Markov Chain Monte Carlo (MCMC) posterior sampler described in Section 2.2. In total, we draw 500,000 times out of the posterior distribution and discard the first 300,000 as burn-in. To remove autocorrelation across draws, we retain every fourth of

[7] We obtain the same qualitative results when we estimate our model with 9 occupation groups assembled as in Jaimovich and Siu (2012).

Table 2: Model Selection

		$Var(\hat{Y}_k)/Var(Y_k)$		
q	p	Cluster 1	Cluster 2	Avg.
1	1	0.345	0.353	0.349
1	2	0.286	0.216	0.251
2	1	0.138	0.138	0.138
2	2	**0.607**	0.588	**0.597**
2	3	0.306	0.213	0.260
3	2	0.484	**0.620**	0.552
3	3	0.223	0.199	0.211

Notes: The table reports the fraction of the variation in aggregate employment/population that is explained by common cluster dynamics, conditional on the median factor assignment. Maxima are highlighted.

the remaining 200,000 draws.[8] Our preferred specification is summarized in Table 1. Note that we obtain the most precise factor assignment (see Section 4.1) when we set $K = 2$ and since the ultimate goal of this study is to analyze aggregate labor market dynamics, we choose the specification for which the variance share explained by cluster specific variation is largest. In particular, Table 2 lists this statistic for alternative AR lag lengths, p and q, and shows that a specification with $p = q = 2$ performs best on average according to this metric.

4.1. Occupation Clusters

Our model identifies two clusters of occupations with distinct cyclical patterns in employment dynamics. Panel B of Figure 1 illustrates that the identifying feature of cluster 2 is the relatively steady average growth in the employment/population ratio throughout the entire period 1976-2013. The employment share of this group grew from less than 20% in 1976 to more than 30% in 2013. Moreover, employment of this group does not appear particularly "cyclical".

On the other hand, cluster 1 groups occupations with employment patterns that differ dramatically from those of cluster 2. First, the employment/population ratio of this group has declined from around 33% at its peak in 1980 to about 25% at its trough in 2013. Second, employment in

[8] In Appendix B, we graphically illustrate the retained draws for selected model parameters after conditioning on the specific cluster assigment tabulated in Table 3.

these occupations appears highly "cyclical". Growth rates obviously differ between recessions and expansions, and there seems to be a change in these growth rates around 1990.

Overall, the groups identified by our preferred model specification resemble the patterns in employment dynamics presented by Jaimovich and Siu (2012), as displayed in panel A of Figure 1.[9] While the aggregate levels in panels A and B of Figure 1 don't perfectly match up, the long run growth patterns are virtually the same. Both aggregation schemes indicate roughly a 50% increase in the employment/population ratio for cluster 2 (non-routine in Panel A) and about a 25% decrease for cluster 1 (routine) over the period 1980 to 2013. There are two major reasons for why the aggregate level series don't match up exactly. First, our consistent panel does not cover all occupations, since some occupations are simply not available consistently over the entire sample period (see Dorn, 2009, for details). Second, our model is based on adjusted growth rates instead of the actual level series. This comes at the expense of losing some information about the "true" level of employment but allows for a straightforward way to accommodate the administrative changes in the occupation classification. However, since the focus of our analysis are the *dynamics* in group-specific employment, this choice does not dramatically affect our inference or our conclusions (see the strong similarity in the dynamic patterns between panel A and panel B of Figure 1).

Our clustering approach at the disaggregated level allows us to further analyze the composition of the two identified occupation groups. The first two columns of Table 3 tabulate the posterior assignment probabilities for the two factors. Notice that almost all of the 21 Dorn (2009) occupation groups are nearly perfectly assigned to one of the two clusters. Only a handful of occupations have an assignment probability of less than 2/3. Panels A and B respectively group occupations for which the posterior probability of being determined by factors 1 and 2 is larger than 50%.[10]

Notice that there are two service occupations (C.35 and C.1) that essentially have a 50/50

[9] Note that both panels in Figure 1 simply plot the data, but for different aggregation schemes. Panel A is based on the same data and aggregation as in Jaimovich and Siu (2012), while in panel B we use our detailed CPS dataset and aggregate the employment/population ratios in the two clusters of occupations listed in Table 3.

[10] Except for the particular assignemnt of the few occupaions that are not decisevly associated with either cluster, this classification is robust across all other model specifications with $K = 2$ factors that we considered. As discuessed in the text, the "unassigned" occupations are quantitatively small and show little growth throughout the entire sample. Therefore, they are not likely to have much influence on aggregate employment dynamics.

Table 3: Cluster Analysis: Factor Assignment

	Assignment δ_i		Factor Loading $\lambda_{ik}\|\boldsymbol{\delta}^{50}$			
21 Occupation Groups (Dorn, 2009)	$Pr[\delta_i = 1\|\boldsymbol{y}]$	$Pr[\delta_i = 2\|\boldsymbol{y}]$	Mean	Median	68% Coverage	
A. Factor 1 (Routine)						
F.1 Machine Operators, Assemblers, and Inspectors	0.999	0.001	1.069	1.072	0.950	1.191
E.2 Construction Trades	0.999	0.001	0.883	0.890	0.770	0.991
E.4 Precision Production	0.999	0.001	0.840	0.844	0.750	0.939
F.2 Transportation and Material Moving	1.000	0.000	0.686	0.684	0.617	0.754
E.1 Mechanics and Repairers	1.000	0.000	0.626	0.622	0.552	0.695
B.2 Sales	0.975	0.025	0.466	0.463	0.413	0.516
B.3 Administrative Support	0.760	0.240	0.243	0.250	0.170	0.322
C.31 Food Preparation and Service	0.608	0.392	0.168	0.167	0.100	0.240
B. Factor 2 (Non-Routine)						
E.3 Extractive	0.229	0.771	0.941	0.940	0.544	1.361
A.2 Management Related	0.000	1.000	0.818	0.821	0.729	0.918
C.37 Misc. Personal Care and Service	0.031	0.969	0.800	0.785	0.574	1.013
A.1 Executive, Administrative, and Managerial	0.001	0.999	0.599	0.598	0.532	0.675
C.36 Child Care Workers	0.130	0.870	0.586	0.567	0.411	0.750
C.32 Healthcare Support	0.076	0.924	0.444	0.439	0.359	0.536
A.3 Professional Specialty	0.005	0.995	0.374	0.379	0.329	0.425
C.2 Protective Service	0.335	0.665	0.329	0.325	0.256	0.411
C.33 Building/Grounds Cleaning/Maintenance	0.105	0.895	0.296	0.290	0.198	0.388
C.34 Personal Appearance	0.326	0.674	0.241	0.245	0.128	0.363
B.1 Technicians and Related Support	0.144	0.856	0.239	0.240	0.171	0.307
C.35 Recreation and Hospitality	0.476	0.524	-0.061	-0.061	-0.251	0.142
C.1 Housekeeping and Cleaning	0.418	0.582	-0.257	-0.259	-0.376	-0.141

Notes: The first two columns report the fraction of posterior draws that classify each occupation into either factor $k = 1$ or $k = 2$. Panel A groups occupations with $Pr[\delta_i = 1|\boldsymbol{y}] > 1/2$ while panel B collects those with $Pr[\delta_i = 2|\boldsymbol{y}] > 1/2$ based on $50,000$ retained posterior draws. The last three columns report the posterior mean, median, as well as the upper and lower bound of the 68% posterior coverage region for the factor loading $\lambda_{ik}|\boldsymbol{\delta}^{50}$, conditional on the median factor assignment, $\boldsymbol{\delta}^{50}$. Within each panel, the occupations are sorted in decreasing order by their conditional factor loading $\lambda_{ik}|\boldsymbol{\delta}^{50}$.

chance of belonging to either factor. It turns out that employment in these occupations is essentially constant (in levels) throughout the entire sample and that their share in total employment is very small. Therefore, these occupations do not contain much information about the factor dynamics, which is also reflected in an insignificant factor loading for the recreation and hospitality group (C.1).

In addition to the assignment probabilities, the last four columns of Table 3 report the posterior mean and median factor loading, $\lambda_{ik}|\boldsymbol{\delta}^{50}$, as well as the associated 68% posterior coverage

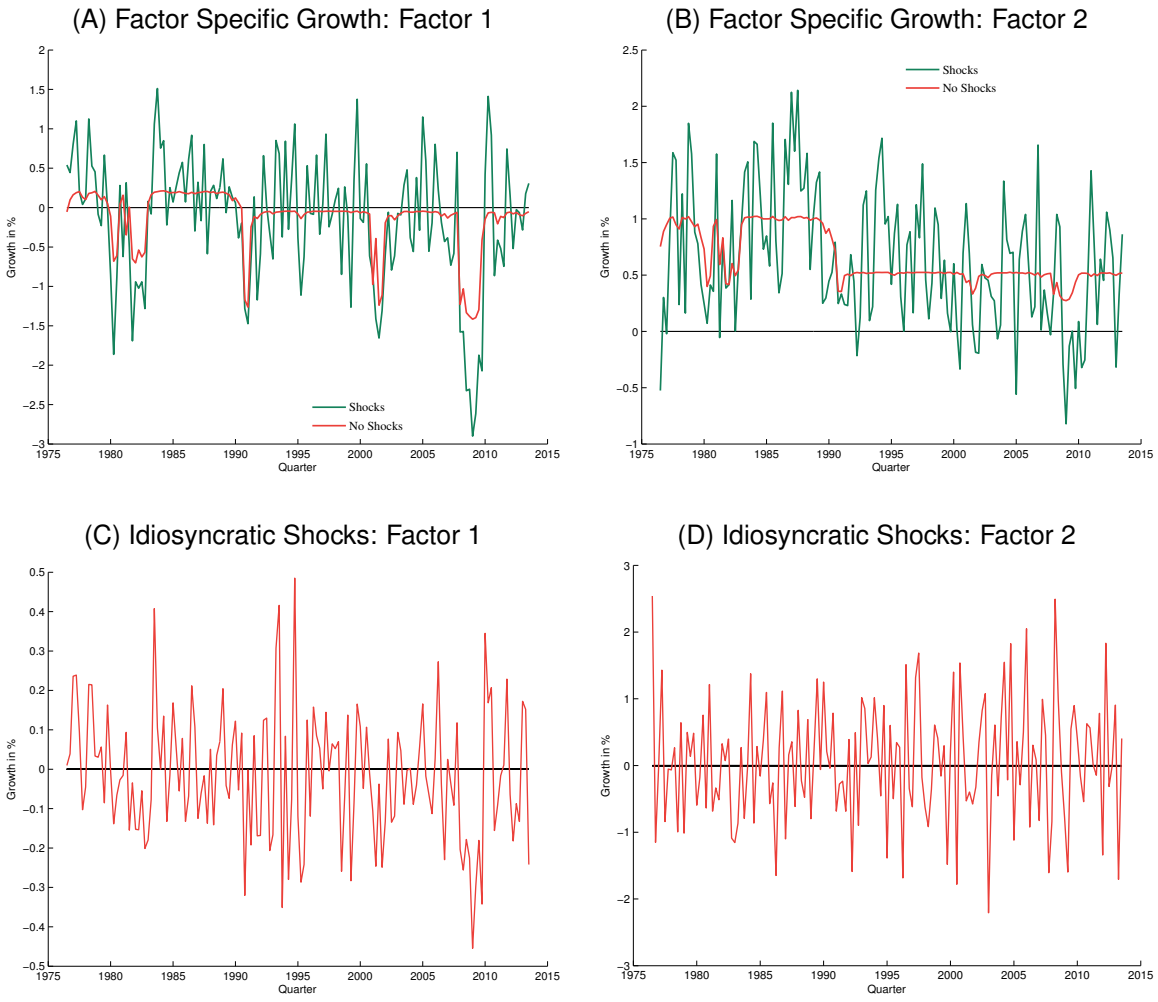
interval, conditional on the median factor assignment, $\boldsymbol{\delta}^{50}$.[11] Notice that all but one of these intervals exclude zero. The occupation with an insignificant factor loading is precisely one of the two service occupations for which the assignment probability is not decisive. If we estimate specifications with $K > 2$, the assignment classification deteriorates considerably for all occupations, which further suggests that the "unassigned" occupations simply experience very idiosyncratic employment dynamics rather than being influenced by an additional factor.

Within panels A and B of Table 3, we sort occupations in decreasing order of the posterior mean/median factor loading. This provides a measure of the "intensity" with which a given occupation is influenced by the common factor dynamics and we observe that a considerable amount of heterogeneity is present across the factor loadings. This measure indicates that employment dynamics for an occupation with a factor loading close to zero are mostly driven by idiosyncratic dynamics, captured by the mean zero AR(q) process ε_{it} in our model.

The long run patterns (see Figure 1) as well as the composition (see Table 3) of the two identified clusters are consistent with recent evidence presented in Autor and Dorn (2013). They show that polarization in the US labor market over the period 1980-2005 is mainly driven by the growth in service and in "abstract cognitive" occupations combined with the decline in "routine" occupations, which are easily replaceable by technology or offshoring. In line with their evidence, panel B of Table 3 illustrates that cluster 2 largely consists of managerial and professional specialty occupations and of a number of service occupations. However, our classification also suggests that some service occupations, like "recreation and hospitality" and "housekeeping and cleaning", are not decisively assigned to either factor, and that one service occupation ("food preparation and service") is assigned to cluster 1 with a 60% probability. On the other hand, one traditional "blue collar" occupation ("extractive") is decisively assigned to factor 2. This may very well be driven by a rising demand for highly skilled engineers in areas such as "fracking", off-shore drilling, and other recent high-tech resource extraction techniques.

[11] That is, these moments are computed from joint posterior draws for which $Pr[\delta_i = 2|\boldsymbol{y}] > 1/2$ *and* for which each occupation, i, is assigned to one of the two clusters exactly as in Table 3.

Figure 3: Cluster Specific vs. Idiosyncratic Dynamics (Full Sample)



Notes: Panels A and B illustrate the median of $\hat{f}_{kt}$ (green) as well as $\hat{f}_{kt} - \hat{\nu}_{kt}$ (red) for the two factors $k \in \{1, 2\}$. Panels C and D depict the occupation specific idiosyncratic shocks aggregated over all occupations according to the median assignment probability within each cluster, $\hat{\varepsilon}_{kt} = \sum_{i|\hat{\delta}_i=k} \hat{\varepsilon}_{it}$ for $k \in \{1, 2\}$.

4.2. Cluster-Specific Cyclical Dynamics

Conditional on the median factor assignment from Table 3, we now shift our focus to the decomposition of employment dynamics into a "structural" (factor specific) and an "idiosyncratic" (occupation-specific) component. Starting with the structural component, we illustrate the estimated factor specific growth in the employment/population ratio, $\hat{f}_{tk}$ for $k \in \{1, 2\}$, in panels A and B of Figure 3 ("shocks"). These panels also graphically illustrate the importance and statistical

properties of the factor specific residuals, ν_{kt}, by plotting $\hat{f}_{kt} - \hat{\nu}_{kt} = \hat{\mu}_{kS_t} + \hat{\phi}_{k1}\hat{f}_{k,t-1}$ for $k \in \{1, 2\}$ ("no shocks").

Two important features of the factor-specific dynamics are immediately apparent from the "filtered" series, labeled "no shocks": first, both factors feature substantial systematic cyclical variation, which is not immediately visible in Figure 1. Second, factor 1 appears to have become "more cyclical" after 1990, while factor 2 became "less cyclical".[12] Finally, growth rates in both clusters are lower after 1990 both in expansions *and* recessions. Specifically, for factor 1 the decrease in growth rates has been more pronounced in recessions, while for factor 2 the decrease has been more pronounced in expansions. We will further investigate these factor and state-specific breaks in section 4.3.

To illustrate the "idiosyncratic" component of employment dynamics, panels C and D of Figure 3 plot the estimated idiosyncratic shocks, aggregated over all occupations according the median posterior assignment probability within each cluster, $\hat{\varepsilon}_{kt} = \sum_{i|\hat{\delta}_i=k} \hat{\varepsilon}_{it}$ for $k \in \{1, 2\}$. One interesting feature of this illustration is that negative idiosyncratic shocks during the most recent recessions for factor 1 have been substantially more pronounced than in the past. We will discuss this aspect further below in our counterfactual analysis.

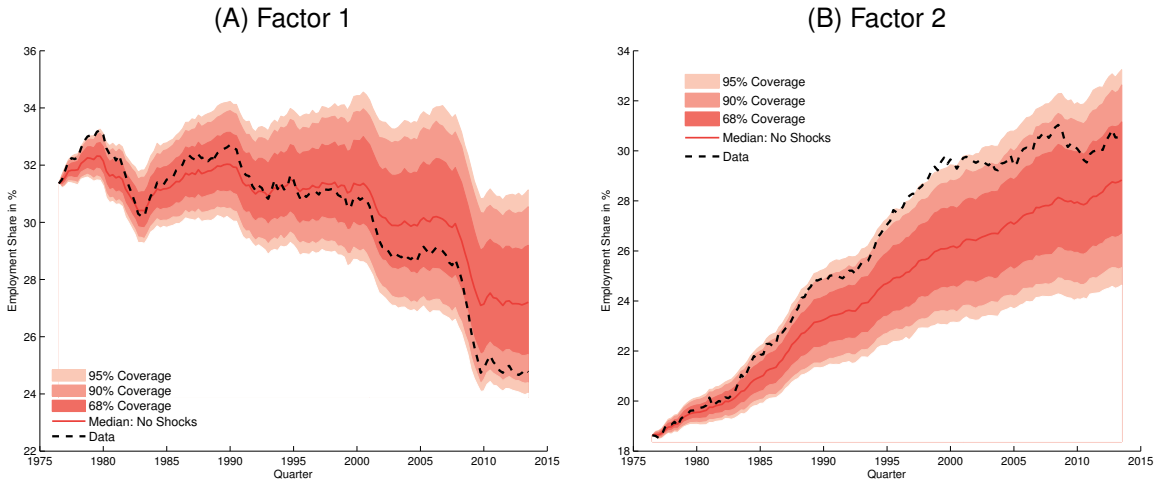
Since we are ultimately interested in aggregate employment dynamics, Figure 4 illustrates what the estimated dynamics imply for cluster-specific employment levels. To construct these graphs, we first recursively compute the occupation-specific implied level series (setting $\varepsilon_{it} = 0$ for all t) for each of the retained joint posterior draws, conditional on the median posterior factor assignment probability $\boldsymbol{\delta}^{50} = \{\hat{\delta}_1, \ldots, \hat{\delta}_N\}$:

$$\tilde{y}_{i,t}^{(m)} = \left(1 + \hat{\lambda}_{\hat{\delta}_i}^{(m)} \hat{f}_{\hat{\delta}_i,t}^{(m)}\right) \tilde{y}_{i,t-1}^{(m)} \quad \text{for all draws } m|\boldsymbol{\delta}^{50}, \tag{14}$$

where m is the respective MCMC draw, $\tilde{y}_{i,0}^{(m)} = y_{i,0}$ is the observed initial level of the employment/population ratio for occupation i, and $\tilde{y}_{i,t}^{(m)}$ denotes the implied level in period t. Based on these occupation-specific implied level series we then aggregate across occupations within each

[12] When we say "more cyclical" we mean that the difference between the median growth rates of expansions and recessions has increased.

Figure 4: Structural Employment Dynamics (Full Sample)



Notes: Conditional on $Pr[\hat{\delta}_i|\boldsymbol{y}] > 0.5$, the figure shows the cluster-specific observed aggregate employment/population ratio and the estimated implied aggregate employment/population ratio, setting $\varepsilon_{it} = 0$. We construct these aggregates by summing occupation-specific implied levels. The posterior coverage region is obtained from quantiles of the empirical posterior implied by all MCMC draws, conditional on the median factor assignment, $\boldsymbol{\delta}^{50}$.

cluster:

$$\hat{Y}_{kt}^{(m)} = \sum_{i|\hat{\delta}_i=k} \tilde{y}_{i,t}^{(m)} \quad \text{for all } m = 1, \ldots, M \text{ and } k = 1, 2, \tag{15}$$

These implied cluster-specific level series capture the predictive ability of the common factor dynamics, while the deviations from observed employment levels are explained by the accumulation of the occupation-specific idiosyncratic component, ε_{it}. We utilize the quantiles of the empirical posterior distribution of the cluster-specific level series to compute posterior coverage regions.[13]

Figure 4 shows that, for factor one, the data lie within the 95% posterior coverage region almost throughout the entire sample. This implies that the common dynamics of occupations in cluster 1 capture a large portion of the aggregate dynamics in employment. In fact, Table 2 reveals that the structural component of factor 1 explains about 61% of the variation in the data. The structural component of factor 2 captures less of the observed aggregate employment dynamics, given that the data series lies marginally outside the 95% coverage region for some quarters in the second half of the 1990s (see panel B of Figure 4). Nevertheless, the explained variance share of the

[13]Note that, while inference for this simulation is conditional on the median factor assignment, $\boldsymbol{\delta}^{50}$, we fully account for conditional uncertainty in all other parameters.

structural component in cluster 2 amounts to about 59% (see Table 2). Therefore, as suggested by Jaimovich and Siu (2012), a significant structural break in the common component of employment dynamics has the potential to significantly affect aggregate employment dynamics. In fact, our model identifies a significant structural break in the factor dynamics around the NBER recession in 1990 and we will discuss its properties and implications in Sections 4.3 and 4.4.

Despite the importance of the structural component, Figure 4 also reveals that there is a significant amount of variation originating from the idiosyncratic component, ε_{it}, especially since the recession in 2001. Evidently, both the 2001 and 2008 recessions were "unusually" severe for occupations associated with factor 1, given that the common factor dynamics over-predict the data development. On the other hand, common dynamics in factor 2 significantly under-predict the data in a number of periods, especially throughout 1995-2000, while developing largely "in parallel" thereafter.

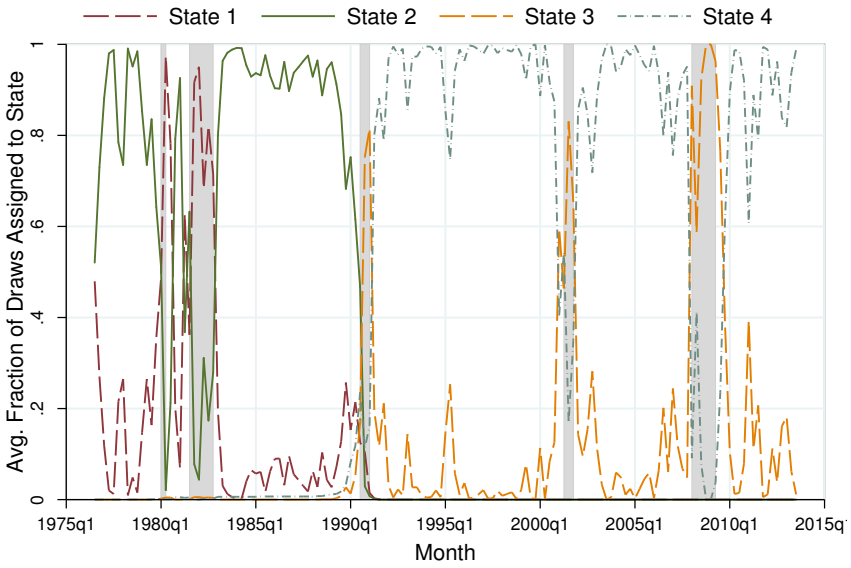
This illustrates that our model provides a convenient framework to disentangles the factor-specific from the idiosyncratic component of employment dynamics. In Section 4.4, we will utilize this aspect of our model to draw inference about the importance of a structural break in the common component for explaining aggregate employment dynamics after the 1990 recession, while simultaneously controlling for the idiosyncratic variation in ε_{it}.

4.3. A Structural Break in the Cluster Dynamics

Inspired by panel A of Figure 1, Jaimovich and Siu (2012) have recently hypothesized that the marked change in the business cycle dynamics of "routine" occupations around 1990 may constitute a structural break. They then provide a counterfactual based on descriptive statistics to illustrate that the apparent break in the cyclical dynamics of "routine" jobs alone has the potential to explain the three "jobless recoveries" since 1990.

Our econometric approach allows us to address this question more formally. Specifically, the Markov switching specification allows for a structural break in the cyclical dynamics of both factors and our posterior estimates provide evidence for such a break at the end of the long-lasting expansion during the 1980s (see Figure 5).

Figure 5: Markov Switching

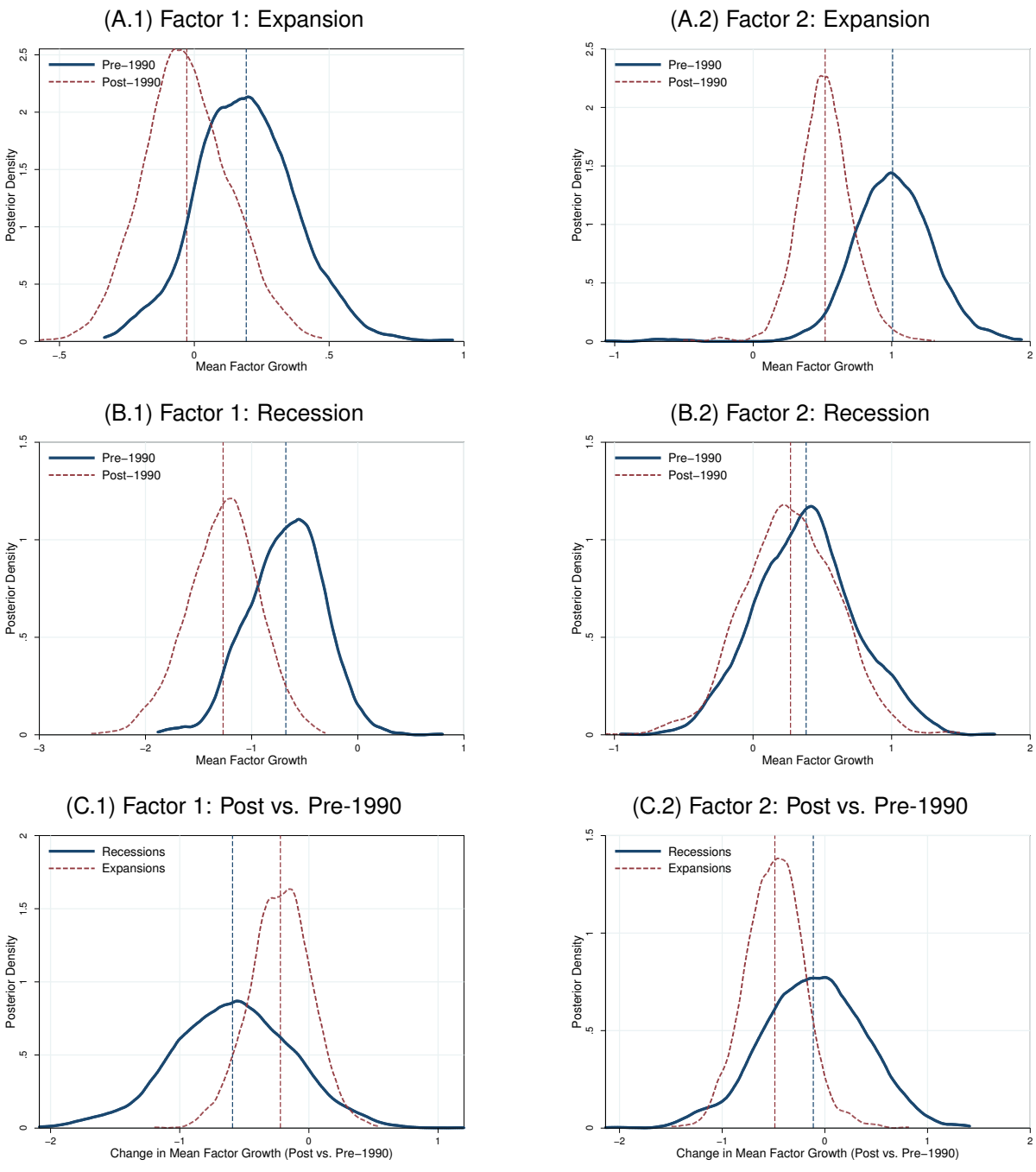


Notes: The figure illustrates the estimated assignemnt probabilities (mean of retained MCMC draws) for the four latent states of the economy: pre-break recession (state 1), pre-break expansion (state 2), post-break recession (state 3), post-break expansion (state 4).

Figure 5 depicts the estimated posterior state probabilities and reveals an almost perfect match with the NBER's business cycle dating committee's classification. The state probabilities are jointly identified from variation in the employment/population ratios and variation in real GDP growth. In particular, GDP growth mainly helps to identify expansions and recessions, while the factor dynamics—inferred from dynamics in the individual employment/population ratios—identify the structural break. Specifically, Figure 5 reveals that the structural break occurred during the 1990/91 recession, as our posterior estimates assign this recession to state 3. The break date is obviously influenced by our informative prior on timing, but the differences in the estimates of several structural parameters confirm the existence of a break.

To discuss the source of the structural break we plot the posterior distributions of the estimated factor- and state-specific means, μ_{kS_t}, in Figure 6. Panels A.1 and B.1 compare the posterior distribution of pre-1990 and post-1990 growth rates for factor 1 during expansions and recessions, respectively. Panels A.2 and B.2 plot the same comparisons for factor 2. Finally, panels C.1 and C.2 show the posterior distributions of the difference between post- and pre-1990 state-specific

Figure 6: Structural Breaks in Estimated Factor Means



Notes: The graphs illustrate kernel density estimates based on our sample of posterior draws for factor and state specific growth rates in employment/population.

growth rates in factor 1 and 2, respectively..

Several insights are worth noting: First, after 1990, both clusters of occupations experienced lower average growth during recessions *and* expansions. Second, and somewhat surprisingly, factor 2 (non-routine) occupations experienced the most pronounced decrease in employment growth during expansions after 1990, while factor 1 occupations (routine) experienced the largest decline in growth during post-1990 recessions. In particular, the state-specific growth rates in panel A.1 and B.1 suggest that the more pronounced employment losses during recessions after 1990 were no longer offset by positive growth rates during expansions. Furthermore, panel A.2 and B.2 show that, although employment in factor 2 occupations was still fairly immune to recessions, average employment growth during expansions was significantly weaker after 1990.

4.4. Can the Structural Break Explain Jobless Recoveries?

We use our model to formally assess whether the structural break in the cluster-specific structural component—discussed in Section 4.3—has the potential to explain the "jobless recoveries" in the aftermath of the last three US recessions. To assess this hypothesis, we consider the following thought experiment: If state specific average factor growth after 1990 (μ_{k3} and μ_{k4}) had remained the same as before 1990 (μ_{k1} and μ_{k2}), how would the employment/population ratio have evolved after 1990?

While this thought experiment is similar to the one analyzed by Jaimovich and Siu (2012), ours differs in several respects. First, Jaimovich and Siu (2012) consider a structural break in post-1990 average employment growth for routine jobs and during recoveries only. In light of the estimates displayed in Figure 6, we find it more natural to allow for a structural break in the dynamics for *both* occupation groups and in *both* states of the economy. Job polarization is a phenomenon that works jointly through the dynamics of routine and non-routine occupations. Thus, if forces like routine biased technical change or a trend toward offshoring are truly the ultimate cause for the observed structural break—and in turn job polarization—, then it seems important to account for structural changes in the dynamics of both routine and non-routine jobs, as one may offset the other.

Second, although we are ultimately interested in the effect of a structural change in the *common* component of routine and non-routine jobs, we also detect a substantial amount of idiosyncratic variation in employment growth (see Figure 4), captured by ε_{it} in our model. We believe that a clean test should control for this idiosyncratic variation and we therefore isolate common from idiosyncratic components which jointly drive aggregate employment dynamics. Finally, our empirical framework allows us to draw formal posterior inference about all estimated effects.

To construct our counterfactual, we simulate factor dynamics for the post-1990 period under the assumption that $\tilde{\mu}_{k3} = \hat{\mu}_{k1}$ (recessions) and $\tilde{\mu}_{k4} = \hat{\mu}_{k2}$ (expansions), where $\hat{\mu}_{ki}$, $i = 1, 2$, denotes the factor-specific growth in the respective business cycle states prior to 1990.[14] Based on the resulting hypothetical factor series, $\tilde{f}_{kt} = \tilde{\mu}_{kS_t} + \hat{\phi}_k \tilde{f}_{kt-1} + \hat{\nu}_{kt}$, and conditional on the estimated occupation classification, we then compute two versions of the occupation-specific employment growth: one in which we assume that $\varepsilon_{it} = 0$ for all t ("no shocks"),

$$y_{it}^{NS} = \hat{\lambda}_{it} \tilde{f}_{\hat{\delta}_i t} \text{ for all } i,$$

and one in which we postulate that $\varepsilon_{it} = \hat{\varepsilon}_{it}$ ("shocks"), that is

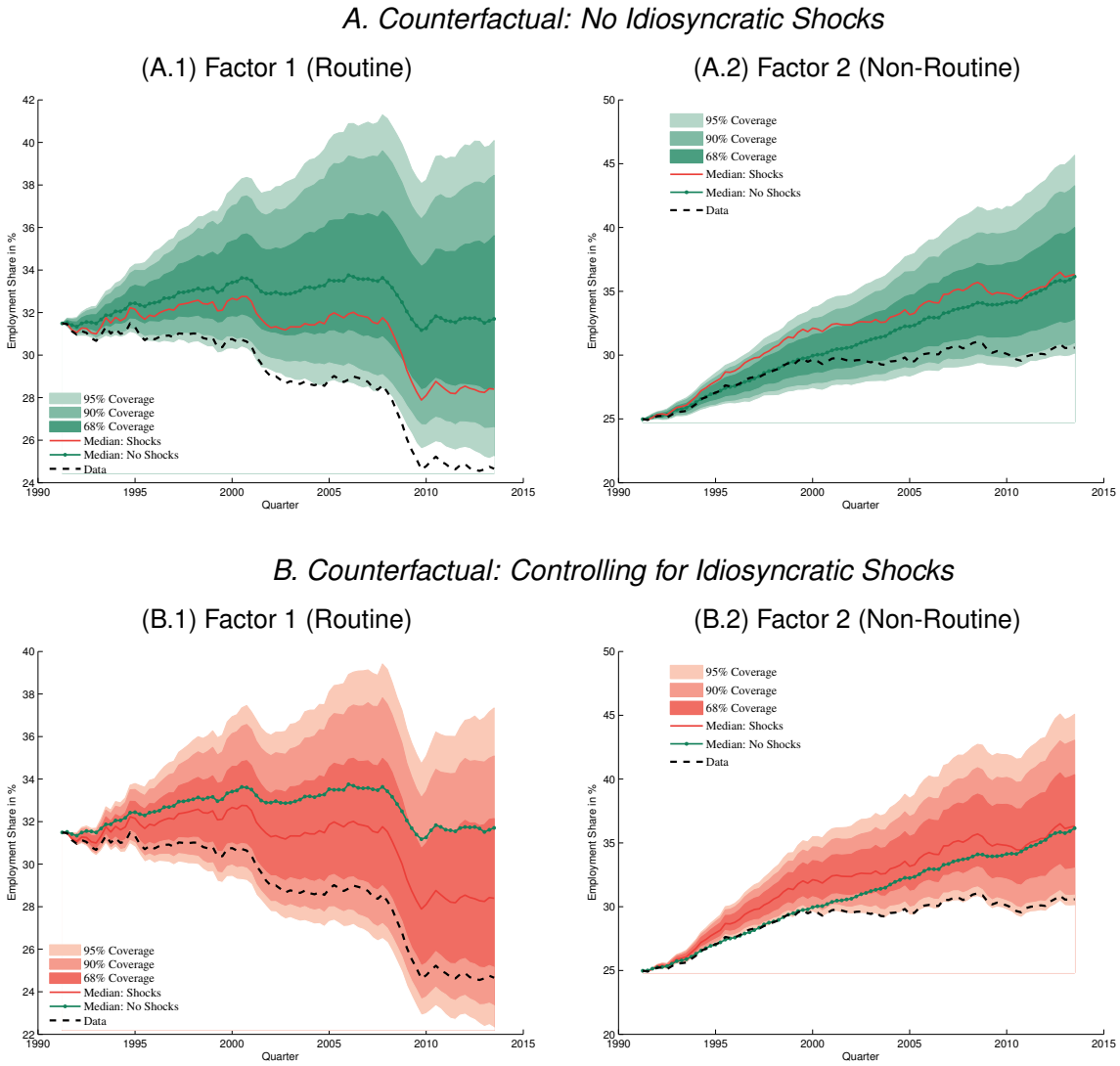
$$y_{it}^{S} = y_{it}^{NS} + \hat{\varepsilon}_{it} \text{ for all } i,$$

where $\hat{\varepsilon}_{it} = y_{it} - \hat{\lambda}_{it} \hat{f}_{\hat{\delta}_i t}$ are the idiosyncratic shocks implied by the estimated factors. The cluster-specific aggregate of these shocks is illustrated in panels C and D of Figure 3.

Panel A of Figure 7 illustrates the counterfactual implied cluster-specific employment level without idiosyncratic shocks (based on y_{it}^{NS} with $\varepsilon_{it} = 0$), while panel B depicts the implied cluster-specific employment when we control for idiosyncratic shocks (based on y_{it}^{S} with $\varepsilon_{it} = \hat{\varepsilon}_{it}$). In both thought experiments we start the simulation at the trough of the 1990/91 recession (1991m3). Panel A shows that the data in cluster 1 lie outside the 95% coverage region during

[14] Generally, we denote all MCMC estimates for the pre and post-break period with a "hat". To draw proper inference on the counterfactuals, we compute implied employment levels for each of the MCMC draws, $m|\boldsymbol{\delta}^{\mathbf{50}}$, and use the resulting empirical posteriors to construct coverage regions. For the ease of notation we omit m for the rest of this section and we ask the reader to keep this in mind.

Figure 7: Counterfactual Experiment: Individual Groups



Notes: Panel A illustrates inference for a counterfactual experiment, in which the factors f_{kt} are governed by the systematic component of the the pre-break dynamics ($\mu_{k3} = \hat{\mu}_{k1}$ during recessions and $\mu_{k4} = \hat{\mu}_{k2}$ during expansions) with $\varepsilon_{i,t} = 0$ for all t. In panel B we show inference for a thought experiment in which we recompute the counterfactual from panel A but under the assumption that the idiosyncratic shocks are those observed since 1990, $\varepsilon_{it} = \hat{\varepsilon}_{it}$. The aggregate versions of these estimated shocks across occupations within each cluster are illustrated in panels C and D of Figure 3.

most of the sample period, while the counterfactual prediction for cluster 2 only starts to diverge after 2000.[15]

[15] In fact, this reveals an interesting insight, and a possible direction for future work. Visual inspection of Figure 1 suggests that the two factors may in fact have experienced structural breaks at different points in time. Factor 1 around 1990 and factor 2 around 2000. This would explain why our counterfactual prediction in panel B of Figure 7 tracks the data perfectly until 2000 and starts to deviate thereafter. While we consider this an interesting possibility to improve the model fit, this does not invalidate our thought experiment, especially after we control for the variation absorbed in the idiosyncratic shocks. Therefore, we leave this modification for future research.

This suggests that, in the absence idiosyncratic shocks, employment growth in both routine and non-routine jobs would have likely been significantly stronger than observed in the data, at least until the onset of the great recession in 2008. However, as the red line in all panels indicates, controlling for idiosyncratic variation ($\hat{\varepsilon}_{it}$) substantially changes the median predictions for employment in both clusters. In particular, the recessions starting in 2001 and in 2008 were significantly more severe than predicted by pre-1990 common factor dynamics. In contrast, controlling for idiosyncratic variation in non-routine jobs reveals that positive idiosyncratic shocks partially offset the negative common growth effect until 2000.[15]

Panel B shows the posterior coverage regions for cluster-specific implied levels when we control for idiosyncratic shocks (based on y_{it}^{S} with $\varepsilon_{it} = \hat{\varepsilon}_{it}$). Despite the substantial downward correction of the median employment level for routine jobs, the data still lie outside the 68% posterior coverage region. Since idiosyncratic shocks were mostly favorable for non-routine occupations, the data now lie outside the 95% coverage region almost throughout the entire prediction horizon in panel B.2.

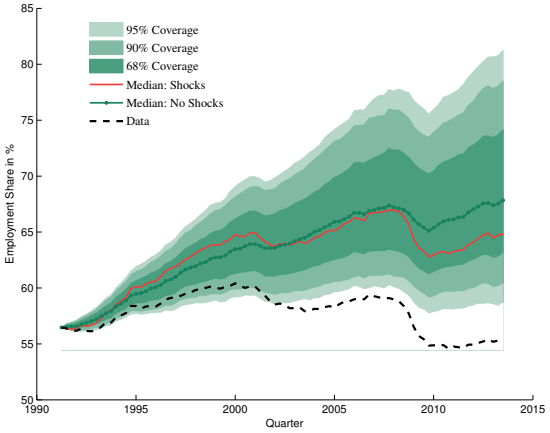
These alternative thought experiments reveal that it is important to both account for structural changes in factor dynamics and to control for idiosyncratic shocks. Ignoring the structural change employment dynamics of non-routine occupations biases the aggregate effects of the structural change downward. Failing to control for idiosyncratic variation biases the estimated effect on routine jobs upward and the effect on non-routine jobs downward. Thus, depending on the relative magnitude of these two distortions, the aggregate bias may be positive or negative.

While these decompositions uncover the cluster-specific effects, we ultimately care about the aggregate effect on employment growth in the aftermath of the last three recessions. Therefore, Figure 8 investigates the aggregate employment effects of the structural changes in the cluster-specific factor dynamics. While panel A is the aggregate version of the underlying effects illustrated in Figure 7, panels B and C alternatively start the simulations at the trough of the 2001 and the 2008/2009 recessions, respectively. These alternative thought experiments are initialized with the observed level for the employment/population ratio at the troughs of each recession. Without controlling for idiosyncratic shocks (panels A.1, B.1, and C.1) the data lie outside the 68% posterior

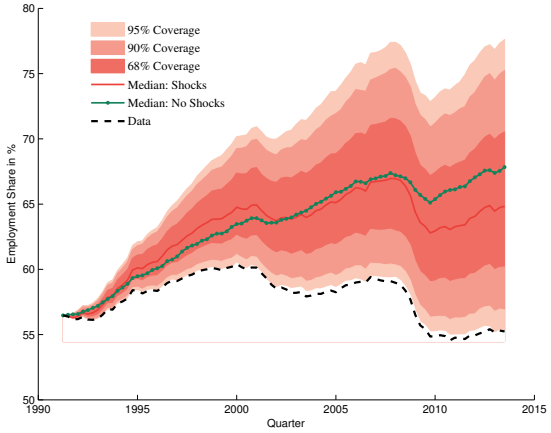
Figure 8: Counterfactual Experiment: Aggregate

A. Trough of the 1990/91 Recession (1991m3)

(A.1) No Shocks

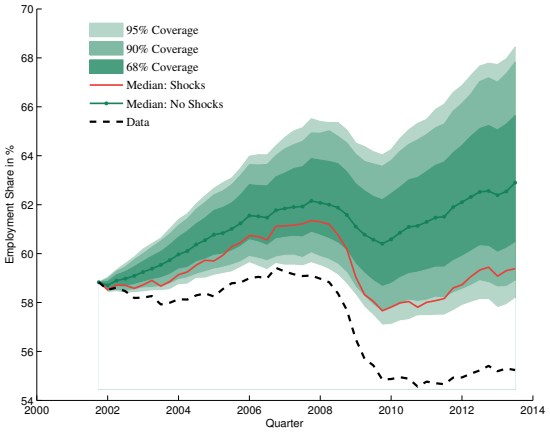


(A.2) Idiosyncratic Shocks

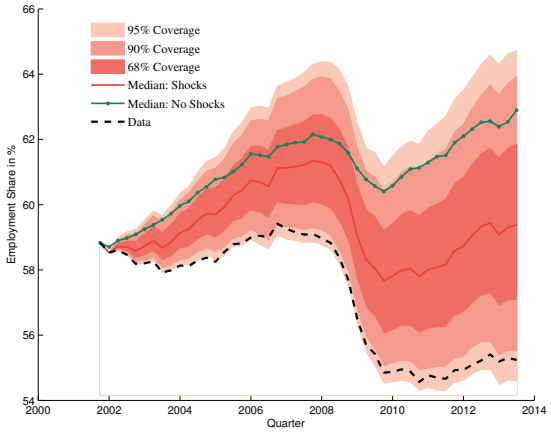


B. Trough of the 2001 Recession (2001m11)

(B.1) No Shocks

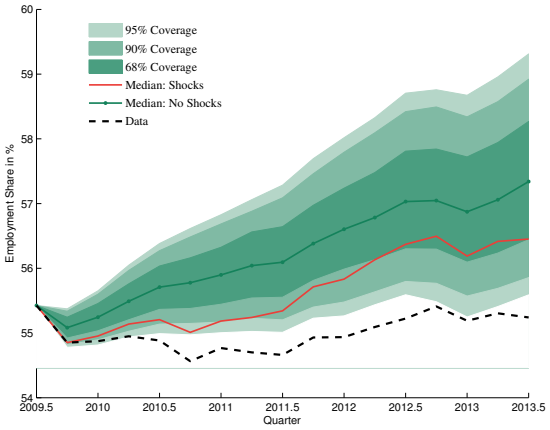


(B.2) Idiosyncratic Shocks

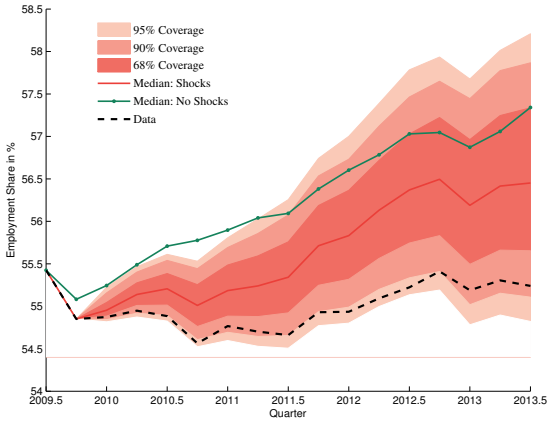


C. Trough of the 2008/09 Recession (2009m6)

(C.1) No Shocks



(C.2) Idiosyncratic Shocks



Notes: Panesl A.1, B.1, and C.1 illustrate counterfactual experiments, in which the factors f_{kt} are governed by the systematic component of the the pre-break dynamics but with $\varepsilon_{i,t} = 0$ for all t. In panels A.2, B.2, and C.2 we recompute these counterfactuals but assuming that the idiosyncratic residuals, ε_{it}, are those observed in the post-break period.

coverage region during all three recoveries and are outside the 95% coverage region after both the 2001 and the 2008/2009 recessions. After controlling for idiosyncratic shocks (panels A.2, B.2, and C.2) the data still lie outside at least the 90% coverage region for all three simulations. In our view, these thought experiments provide formal evidence for the relevance of structural change in explaining the unusually slow job growth during the "jobless recoveries" since 1990.

5. Conclusions

We provide a framework to disentangle group-specific (common) employment dynamics from occupation-specific (idiosyncratic) ones. Our approach simultaneously identifies clusters of jobs that share common cyclical patterns and estimates an empirical model for the group-specific dynamic patterns. In particular, we employ a Markov switching structure with which we distinguish expansions from recessions and allow for a potential structural break.

Using detailed occupation level data from the CPS we detect two clusters of occupations that roughly coincide with occupation groups that Autor et al. (2003) label "routine" and "non-routine" jobs, respectively. Moreover, we find a significant structural break in the cluster-specific dynamics of both routine and non-routine occupations around the 1990/91 recession. Motivated by Jaimovich and Siu (2012), we then assess the impact of this structural break in the common group dynamics on employment growth in the three recoveries since 1990. We find that, in the absence of this structural break, aggregate employment in the US would have recovered significantly more strongly than observed in the data during these "jobless recoveries".

While our analyses do not provide any new insight into the causes of the structural break itself, they nevertheless highlight that the observed "jobless recoveries" are at least in part due to structural change. The particular form of structural change that we document here is closely related to the widely discussed phenomenon of labor market polarization (e.g. Autor and Dorn, 2013), which itself is likely caused by routine biased technological change (RBTC) or the rising importance of offshoring and international trade (e.g. Autor et al., 2013). Therefore, our results suggest that undirected countercyclical policy measures alone—like expansive monetary policy—might not be sufficient to ensure a sustainable recovery in the US labor market after the great

recession of 2008/2009.

References

Acemoglu D. 1999. Changes in unemployment and wage inequality: An alternative theory and some evidence. *American Economic Review* **89**: 1259–1278.
URL http://ideas.repec.org/a/aea/aecrev/v89y1999i5p1259-1278.html

Acemoglu D, Autor D. 2011. *Skills, Tasks and Technologies: Implications for Employment and Earnings*, volume 4 of *Handbook of Labor Economics*, chapter 12. Elsevier, 1043–1171.
URL http://ideas.repec.org/h/eee/labchp/5-12.html

Autor D, Dorn D, Hanson GH. 2013. Untangling Trade and Technology: Evidence from Local Labor Markets. IZA Discussion Papers 7329, Institute for the Study of Labor (IZA).
URL http://ideas.repec.org/p/iza/izadps/dp7329.html

Autor DH. 2013. The "task approach" to labor markets : an overview. *Journal for Labour Market Research* **46**: 185–199.
URL http://ideas.repec.org/a/iab/iabjlr/v2013i3p185-199.html

Autor DH, Dorn D. 2013. The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review* **103**: 1553–97.
URL http://ideas.repec.org/a/aea/aecrev/v103y2013i5p1553-97.html

Autor DH, Katz LF, Kearney MS. 2008. Trends in u.s. wage inequality: Revising the revisionists. *The Review of Economics and Statistics* **90**: 300–323.
URL http://ideas.repec.org/a/tpr/restat/v90y2008i2p300-323.html

Autor DH, Levy F, Murnane RJ. 2003. The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics* **118**: 1279–1333.
URL http://ideas.repec.org/a/tpr/qjecon/v118y2003i4p1279-1333.html

Chan JC, Jeliazkov I. 2009. Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation* **1**: 101–120.

Chib S. 1996. Calculating posterior distributions and modal estimates in markov mixture models. *Journal of Econometrics* **75**: 79–97.
URL http://ideas.repec.org/a/eee/econom/v75y1996i1p79-97.html

Cortes GM, Jaimovich N, Nekarda CJ, Siu HE. 2014. The micro and macro of disappearing routine jobs: A flows approach. Unpublished manuscript, Duke Univeristy.

Dorn D. 2009. Essays on inequality, spatial interaction, and the demand for skills. Dissertation 3613, University of St. Gallen.

Foote CL, Ryan RW. 2012. Labor-market polarization over the business cycle. Public Policy Discussion Paper 12-8, Federal Reserve Bank of Boston.
URL http://ideas.repec.org/p/fip/fedbpp/12-8.html

Frühwirth-Schnatter S. 2010. *Finite Mixture and Markov Switching Models*. Springer Series in Statistics. Springer. ISBN 9781441921949.
URL http://books.google.com/books?id=WvJ5cgAACAAJ

Frühwirth-Schnatter S, Frühwirth R. 2010. Data augmentation and mcmc for binary and multinomial logit models. In Kneib T, Tutz G (eds.) *Statistical Modelling and Regression Structures*. Physica-Verlag HD. ISBN 978-3-7908-2412-4, 111–132.
URL http://dx.doi.org/10.1007/978-3-7908-2413-1_7

Goos M, Manning A. 2007. Lousy and lovely jobs: The rising polarization of work in britain. *The Review of Economics and Statistics* **89**: 118–133.
URL http://ideas.repec.org/a/tpr/restat/v89y2007i1p118-133.html

Goos M, Manning A, Salomons A. 2009. Job polarization in europe. *American Economic Review* **99**: 58–63.
URL http://ideas.repec.org/a/aea/aecrev/v99y2009i2p58-63.html

Jaimovich N, Siu HE. 2012. The trend is the cycle: Job polarization and jobless recoveries. Working Paper 18334, National Bureau of Economic Research.
URL http://www.nber.org/papers/w18334

Kaufmann S. 2011. K-state switching models with endogenous transition distributions. Working Papers 2011-13, Swiss National Bank.
URL http://ideas.repec.org/p/snb/snbwpa/2011-13.html

Kaufmann S, Schumacher C. 2013. Bayesian estimation of sparse dynamic factor models with order-independent identification. Working Papers 13.04, Swiss National Bank, Study Center Gerzensee.
URL http://ideas.repec.org/p/szg/worpap/1304.html

Appendix A. Posterior Distributions

To derive the posterior distributions, we first condense the variables in model (1)-(4) to:

$$\begin{aligned}\Psi(L)y_t = y_t^* &= \lambda f_t - \lambda \odot (\psi_{\cdot 1} \otimes \mathbf{1}_{1\times k})\, f_{t-1} - \cdots - \lambda \odot (\psi_{\cdot q} \otimes \mathbf{1}_{1\times k})\, f_{t-q} + \varepsilon_t \qquad \text{(A.1)}\\ &\quad \varepsilon_t \sim N(0, \Sigma_\varepsilon)\,,\ \Sigma_\varepsilon \text{ diagonal}\\ f_t &= \mu_{S_t} + \Phi_1 f_{t-1} + \cdots + \Phi_p f_{t-p} + \eta_t, \quad \eta_t \sim N(0, I_k) \qquad \text{(A.2)}\end{aligned}$$

where $\odot$ and $\otimes$ represent the Hadamar and the Kronecker product, respectively. The row vector $\mathbf{1}_{1\times k}$ contains as elements k ones. Note that the matrices Φ_j are diagonal and that each row of λ, λ_i contains only one non-zero element, i.e. $\lambda_{ik} \neq 0$ if $\delta_i = k$ and $\lambda = 0$ otherwise. We stack the observations to obtain the matrix representation:

$$\begin{aligned}\mathbf{Y} &= \mathbf{\Lambda F} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N(0, I_{T-q} \otimes \Sigma_\varepsilon) \qquad \text{(A.3)}\\ \mathbf{\Phi F} &= \boldsymbol{\mu} + \boldsymbol{\eta} \quad \boldsymbol{\eta} \sim N(0, \mathbf{\Omega}) \qquad \text{(A.4)}\end{aligned}$$

where $\mathbf{Y} = \left(y_{q+1}^{*\prime}, \ldots, y_T^{*\prime}\right)'$ contains all observed data, $\mathbf{F} = \left(f'_{q+1-\max(p,q)}, \ldots, f'_{q+1}, \ldots, f'_T\right)'$ stacks all unobserved factors, including initial states. The matrices $\mathbf{\Lambda}$ and $\mathbf{\Phi}$ are respectively of dimension $(T-q)N \times (T+d)\,k$ and square $(T+d)k$, with $d = (p-q)I_{\{p>q\}}$. Typically, these matrices are sparse and banded around the main diagonal:

$$\mathbf{\Lambda} = \left[\begin{array}{c|cccccc} & -\lambda \odot (\psi_{\cdot q} \otimes \mathbf{1}_{1\times k}) & \ldots & \lambda & 0 \ldots & & 0 \\ \mathbf{0}_{(T-q)N\times dk} & & \ddots & \ddots & & \ddots & \vdots \\ & 0 \ldots & & 0 & & -\lambda \odot (\psi_{\cdot q} \otimes \mathbf{1}_{1\times k}) & \ldots & \lambda \end{array}\right]$$

$$\mathbf{\Phi} = \left[\begin{array}{cccccc} I_p \otimes I_k & & 0 & \ldots & & \\ \hline -\Phi_p & \ldots & -\Phi_1 & I_k & 0 & \ldots \\ & & & & \ddots & \\ & \ldots & 0 & -\Phi_p & \ldots & -\Phi_1 \quad I_k \end{array}\right], \quad \mathbf{\Omega} = \left[\begin{array}{ccc} I_p \otimes \Sigma_\eta^0 & 0 & \ldots \\ 0 & & \\ \vdots & & I_{T+d-p} \otimes I_k \end{array}\right]$$

where Σ_η^0 represents the variance of the initial states (see below). The vector $\boldsymbol{\mu}$ includes the state-dependent intercept:

$$\boldsymbol{\mu} = \begin{bmatrix} \mathbf{0}_{pk\times 1} \\ \mu_{S_{q+1}} \\ \vdots \\ \mu_{S_T} \end{bmatrix}$$

Appendix A.1. Sampling the Factors $\boldsymbol{F}$

We adapt the sampler proposed in Chan and Jeliazkov (2009) to the present setup, which allows to sample f^T in one sweep. Given the representation in (A.3)-(A.4), the complete data likelihood has a normal density

$$f\left(\mathbf{Y}|\mathbf{F},\theta\right) \sim N\left(\mathbf{\Lambda F}, I_{T-q}\otimes\Sigma_\varepsilon\right) \tag{A.5}$$

For the unobserved factors, from (A.4) we obtain the following prior:

$$\begin{aligned} \mathbf{F}|\mathbf{S},\theta &\sim N\left(\mathrm{f}_0, F_0^{-1}\right) \\ &\mathrm{f}_0 = \mathbf{\Phi}^{-1}\boldsymbol{\mu},\ F_0 = \mathbf{\Phi}'\mathbf{\Omega}^{-1}\mathbf{\Phi} \end{aligned} \tag{A.6}$$

In $\mathbf{S}$, the variance of the initial states, Σ_η^0, may be chosen to be diffuse. Here, we will choose Σ_η^0 to be a multiple of the identity matrix, $\Sigma_\eta^0 = \kappa I_k$.

Combining the prior with the likelihood, the posterior is:

$$\begin{aligned} \mathbf{F}|\mathbf{Y},\mathbf{S},\boldsymbol{\delta},\theta &\sim N\left(\mathrm{f}, F^{-1}\right) \\ F &= F_0 + \mathbf{\Lambda}'\left(I_{T-q}\otimes\Sigma_\varepsilon^{-1}\right)\mathbf{\Lambda} \\ \mathrm{f} &= F^{-1}\left(\mathbf{\Lambda}'\left(I_{T-q}\otimes\Sigma_\varepsilon^{-1}\right)\mathbf{Y} + F_0\mathrm{f}_0\right) \end{aligned} \tag{A.7}$$

To avoid the full inversion of F we take the Cholesky decomposition, $F = LL'$, then $F^{-1} = L^{-1}L^{-1'}$. We obtain a draw $\mathbf{F}$ by setting $\mathbf{F} = \mathrm{f} + L^{-1}\boldsymbol{\nu}$, where $\boldsymbol{\nu}$ is a $(T+d)k$ vector of independent draws from the standard normal distribution.

Appendix A.2. Sampling the States

To derive the sampling scheme for S^T, recall that the time-varying matrix $\boldsymbol{\xi}_t$, with elements $\xi_{lk,t}$, $l,k = 1,\ldots,K$ representing the transition probability from state l in $t-1$ to state k in t is subject to zero restrictions:

$$\boldsymbol{\xi}_t = \begin{bmatrix} \xi_{11,t} & \xi_{12,t} & 0 & \xi_{14,t} \\ \xi_{21,t} & \xi_{22,t} & \xi_{23,t} & 0 \\ 0 & & \xi_{33,t} & \xi_{34,t} \\ & & \xi_{43,t} & \xi_{44,t} \end{bmatrix} \tag{A.8}$$

where the elements $\xi_{lk,t}$ are given by (7).

We express the posterior $\pi(\mathbf{S}|\boldsymbol{f},\mathbf{x},\theta)$ as $\pi(\mathbf{S}|\boldsymbol{f},\boldsymbol{\xi},\theta_{-\gamma})$ and factorize it

$$\pi(\mathbf{S}|\boldsymbol{f},\boldsymbol{\xi},\theta_{-\gamma}) = \pi(S_T|f_T,\boldsymbol{\xi}_T,\theta_{-\gamma}) \prod_{t=1}^{T-1} \pi(S_t|f_t,\boldsymbol{\xi}_t,\theta_{-\gamma})\, \pi\left(S_{t+1}|S_t,\boldsymbol{\xi}_{t+1}\right)$$

The filter density $\pi(S_t|f_t,\boldsymbol{\xi}_t,\theta_{-\gamma})$ is obtained by iterating forward through $t = 1,\ldots,T$

$$\begin{aligned} \pi(S_t|f_t,\boldsymbol{\xi}_t,\theta_{-\gamma}) &\propto f(f_t|S_t,\theta_{-\gamma})\, \pi(S_t|f_{t-1},\boldsymbol{\xi}_t,\theta_{-\gamma}) \\ \pi(S_t|f_{t-1},\boldsymbol{\xi}_t,\theta_{-\gamma}) &= \boldsymbol{\xi}_t' \pi\left(S_{t-1}|f_{t-1},\boldsymbol{\xi}_{t-1},\theta_{-\gamma}\right) \end{aligned}$$

The prior distribution of the initial state $\pi(S_0)$ is assumed to be uniform over the first two states: $P(S_0 = k) = 1/2$, $k = 1,2$ and 0 for $k = 3,4$.

State S_T is sampled out of $\pi(S_T|f_T,\boldsymbol{\xi}_T,\theta_{-\gamma})$. We proceed backwards $t = T-1,\ldots,0$ and draw from the posterior sampling density

$$\pi(S_t|f_t,S_{t+1},\boldsymbol{\xi}_t,\theta_{-\gamma}) \propto \pi(S_t|f_t,\boldsymbol{\xi}_t,\theta_{-\gamma})\, \xi_{S_t S_{t+1},t+1}$$

where $\xi_{S_t S_{t+1},t+1}$ extracts the column S_{t+1} of the matrix $\boldsymbol{\xi}_{t+1}$.

Appendix A.3. Sampling the Parameters of the State Transition Distribution

The steps described in the following yield a draw from $\pi(\boldsymbol{\gamma}|\mathbf{S}, \mathbf{x})$ based on the difference in random utilities model (dRUM) representation.

Step (iii.a): Sample the $(S-1)T$ *utility differences from* $\pi(\boldsymbol{\omega}|\mathbf{S}, \gamma) = \prod_{s=2}^{S} \pi(\omega_{s2}, \ldots, \omega_{sT}|\mathbf{S}, \gamma)$

The dRUM extension expresses the multinomial logit model as differences in the latent utilities S_{st}^u associated to state s in period t, where

$$\begin{aligned} S_{st}^u &= \mathbf{X}_t'\gamma_s + \nu_{st}, \text{ for } s = 2, \ldots, S \\ S_{1t}^u &= \nu_{1t}, \text{ from the identification restriction } \gamma_1 = 0, \end{aligned} \tag{A.9}$$

and ν_{st} follows a Type I extreme distribution. The extended vector $\mathbf{X}_t' = \left(X_t D_{t-1}^{(1)}, \ldots, X_t D_{t-1}^{(S)}\right)$, with $D_t^{(s)} = 1$ if $S_t = s$ and 0 otherwise, includes past state-specific information. The parameter $\gamma_s = (\gamma_{1s}', \ldots, \gamma_{Ss}')'$ stacks all s-state relevant, state-dependent parameters, see (7). Forming the differences $w_{st} = S_{st}^u - S_{1t}^u$, given that the parameters of the reference transition are zero, $\gamma_1 = 0$, we obtain

$$w_{st} = \mathbf{X}_t'\gamma_s + \epsilon_{st}, \ \epsilon_{st} \sim \text{Logistic}, \ s = 2, \ldots, S \tag{A.10}$$

with $\epsilon_{st} = \nu_{st} - \nu_{1t}$. Working with this representation would be quite involving because, in contrast to the error terms ν_{st} in (A.9), the error terms ϵ_{st} in (A.10) are not independent any more across states. Therefore, Frühwirth-Schnatter and Frühwirth (2010) consider a partial representation of the dRUM model, which relies on the observation that

$$S_t = s \Leftrightarrow S_{st}^u > S_{-s,t}^u, \ S_{-s,t}^u = \max_{j \in \mathcal{S}_{-s}^*} S_{jt}^u \tag{A.11}$$

i.e. that state s is observed if S_{st}^u is larger than the maximum of all other relevant utilities collected in $\mathcal{S}_{-s}^*$. The relevant utilities are those which are related to states that are possible given the restrictions of the transition matrix in (7). For example, if $S_t = 2$ given state $S_{t-1} = 1$, then the relevant utilities would be the ones associated to the states 1 and 4, i.e. $\mathcal{S}_{-s}^* = \{1, 4\}$, because these are the

Table A.4: Relevant states in $\mathcal{S}^*_{-s}$ for $S_t = s$ given S_{t-1}

$S_t =$	2		3			4		
S_{t-1}=	1	2	2	3	4	1	3	4
$\mathcal{S}^*_{-s} =$	{1,4}	{1,3}	{1,2}	{4}	{4}	{1,2}	{3}	{3}

only states that could prevail in t beside state 2, see also table A.4.

For all states but the reference state we define the latent difference utilities ω_{st} and the binary observation $D_t^{(s)}$:

$$\omega_{st} = S^u_{st} - S^u_{-s,t},\ D_t^{(s)} = I\{S_t = s\},\ \forall s \in \mathcal{S}^*_{-1} \tag{A.12}$$

Given the multinomial logit model for S_t, ω_{st} has an explicit distributional form. Recall that

$$\exp\left(-S^u_{-s,t}\right) \sim \mathcal{E}\left(\sum_{j \in \mathcal{S}^*_{-s}} \lambda_{jt}\right) \tag{A.13}$$

where $\lambda_{jt} = \exp\left(X_t'\gamma_j\right)$ and define $\lambda_{-s,t} = \sum_{j\in\mathcal{S}^*_{-s}} \lambda_{jt}$. We then can write $S^u_{-s,t} = \log\left(\lambda_{-s,t}\right) + \nu_{-s,t}$, where $\nu_{-s,t}$ follows an extreme value distribution. Thus, the multinomial logit model has the partial dRUM representation

$$\begin{aligned} \omega_{st} &= S^u_{st} - S^u_{-s,t} = \mathbf{X}_t'\gamma_s - \log\left(\lambda_{-s,t}\right) + \nu_{s,t} - \nu_{-s,t} \\ &= \mathbf{X}_t'\gamma_s - \log\left(\lambda_{-s,t}\right) + \epsilon_{s,t},\ D_t^{(s)} = I\{S_t = s\} \end{aligned} \tag{A.14}$$

where $\nu_{s,t}$ and $\nu_{-s,t}$ are i.i.d. and follow an EV distribution, and $\epsilon_{s,t}$ follows a logistic distribution. The constant $-\log\left(\lambda_{-s,t}\right)$ in (A.14) depends only on the parameters γ_{-s}. Therefore, given $\omega_s^T = (\omega_{s1}, \ldots, \omega_{sT})$ and γ_{-s}, we obtain a linear regression with parameter γ_s and logistic errors.

The sub-sampling steps can now be outlined explicitly. For each state s, we first sample the latent utility differences ω_s^T from logistic distributions.[16] Across s, we sample independently T

[16] $\omega_{st}|S^T, \gamma_s$ follows a logistic distribution truncated to $[0,\ \infty)$ if $S_t = s$, and truncated to $(-\infty,\ 0]$ if $S_t \neq s$.

values W_{st} from a uniform distribution $W_{st} \sim U\,[0,\ 1]$ and obtain

$$\omega_{st} = \mathbf{X}_t'\gamma_s - \log\left(\lambda_{-s,t}\right) + F_\epsilon^{-1}\left(D_t^{(s)} + W_{st}\left(1 - D_t^{(s)} - \pi_{st}\right)\right) \tag{A.15}$$

where $\pi_{st} = P\left(D_t^{(s)} = 1|\gamma\right) = 1 - F_\epsilon\left(-\mathbf{X}_t'\gamma_s + \log\left(\lambda_{-s,t}\right)\right) \propto \lambda_{st}/\lambda_{-s,t}$; $F_\epsilon\left(p\right)$ represents the cumulative distribution function of the logistic distribution, and $F_\epsilon^{-1}\left(p\right) = \log\left(p\right) - \log\left(1-p\right)$ its inverse.

Step (iii.b) Sample the components $R^{(S-1)T}$ from $\pi\left(R^{(S-1)T}|\omega^{(S-1)T}, \gamma\right)$

Given $\omega^{(S-1)T}$, the posterior of γ_s is derived based on (A.14), approximating the logistic distribution of the errors ϵ_{st} by a mixture of normal distributions with M components. The components R_{st} are drawn from a multinomial distribution

$$P\left(R_{st} = r|\omega_{st}, \gamma_s\right) \propto \frac{w_r}{d_r}\exp\left\{-\frac{1}{2}\left(\frac{\omega_{st} + \log\left(\lambda_{-s,t}\right) - \mathbf{X}_t'\gamma_s}{d_r}\right)^2\right\} \tag{A.16}$$

where $r = 1, \ldots, 6$, and the respective component's standard deviation d_r and weight w_r, are taken from Frühwirth-Schnatter and Frühwirth (2010), Table 1.

Step (iii.c): Sample γ from $\pi\left(\gamma|\omega^{(S-1)T}, R^{(S-1)T}\right)$

Conditional on the components R_s^T, model (A.14) becomes normal in γ_s:

$$\tilde{\omega}_{st} = \omega_{st} + \log\left(\lambda_{-s,t}\right) \quad = \quad \mathbf{X}_t'\gamma_s + \epsilon_{st}, \quad \epsilon_{st}|R_{st} \sim N\left(0, d_{R_{st}}^2\right) \tag{A.17}$$

Assuming a normal prior for γ_s, $\pi\left(\gamma_s\right) = N(g_0, G_0)$, conditional on ω_s^T and R_s^T the posterior is normal, too:

$$\pi\left(\gamma_s|\omega_s^T, R_s^T\right) \quad = \quad N\left(g_s, G_s\right) \tag{A.18}$$

$$G_s = \left(\sum_{t=1}^{T}\mathbf{X}_t\mathbf{X}_t'/d_{R_{st}}^2 + G_0^{-1}\right)^{-1} \tag{A.19}$$

$$g_s = G_s\left(\sum_{t=1}^{T}\mathbf{X}_t\tilde{\omega}_{st}/d_{R_{st}}^2 + G_0^{-1}g_0\right) \tag{A.20}$$

Appendix A.4. Sampling the classification indicator and the rest of the parameters

1. To draw from $\pi(\boldsymbol{\delta}|\boldsymbol{y}, \boldsymbol{f}, \boldsymbol{\psi}, \boldsymbol{\sigma})$ we sample independently over N from the discrete distribution in (13). The indicator δ_i is set equal to

$$k = \left(\sum_{l=1}^{K} I \left\{ \left(\sum_{j=1}^{l} P(\delta_i = j|\cdot) \right) \leq U \right\} \right) + 1$$

where $I\{\cdot\}$ is the indicator function, $P(\delta_i = j|\cdot)$ are the normalized posterior indicator probabilities obtained from (13) and $U \sim U(0,1)$ is a draw from the uniform distribution.

2. The posterior distributions of the remaining parameters are standard distributions:

   (a) $\pi(\boldsymbol{\lambda}|\boldsymbol{y}, \boldsymbol{f}, \boldsymbol{\delta}, \boldsymbol{\sigma}, \boldsymbol{\psi}) = \prod_{i=1}^{N} N(\mathrm{l}_i, \mathrm{L}_i)$, where

$$\mathrm{L}_i = \left(\sigma_i^{-2} \sum {f^*_{i\delta_i t}}^2 + \mathrm{L}_0^{-1} \right)^{-1}, \quad \mathrm{l}_i = \mathrm{L}_i \left(\sigma_i^{-2} \sum y^*_{it} f^*_{i\delta_i t} + \mathrm{L}_0^{-1} \mathrm{l}_0 \right)$$

   (b) $\pi(\boldsymbol{\psi}|\boldsymbol{y}, \boldsymbol{f}, \boldsymbol{\delta}, \boldsymbol{\sigma}, \boldsymbol{\lambda}) = \prod_{i=1}^{N} N(q_i, Q_i) I_{\{Z(\psi_i)>1\}}$, with

$$Q_i = \left(\sigma_i^{-2} \varepsilon'_{i,-1} \varepsilon_{i,-1} + Q_0^{-1} \right)^{-1}, \quad q_i = Q_i \left(\sigma_i^{-2} \varepsilon'_{i,-1} \varepsilon_i + Q_0^{-1} q_0 \right) I_{\{Z(\psi_i)>1\}}$$

   and where ε_i and $\varepsilon_{i,-1}$ are, respectively, the appropriately designed left- and right-hand side matrices of the regression model:

$$\varepsilon_{it} = \psi_1 \varepsilon_{i,t-1} + \cdots + \psi_q \varepsilon_{i,t-q} + \epsilon_{it}, \ \epsilon_{it} \sim N(0, \sigma_i^2)$$

   (c) $\pi(\boldsymbol{\phi}, \boldsymbol{\mu}|\boldsymbol{f}, \mathbf{S}) \prod_{k=1}^{K} N(p_k, P_k) I_{\{Z(\phi_k)>1\}} I_{\{\mu_{k1}<\mu_{k2}, \mu_{k3}<\mu_{k4}\}}$, with

$$P_k = \left([f_{k,-1} \ D]' [f_{k,-1} \ D] + diag(P_0, M_0)^{-1} \right)^{-1} \tag{A.21}$$

$$p_k = P_k \left([f_{k,-1} \ D]' f_k + diag(P_0, M_0)^{-1} vec(p_0, m_0) \right) \tag{A.22}$$

   and where f_k, $f_{k,-1}$, D are respectively, the appropriate matrices of the regression

model ($D^{(s)} = 1$ if $S_t = s$ and 0 otherwise):

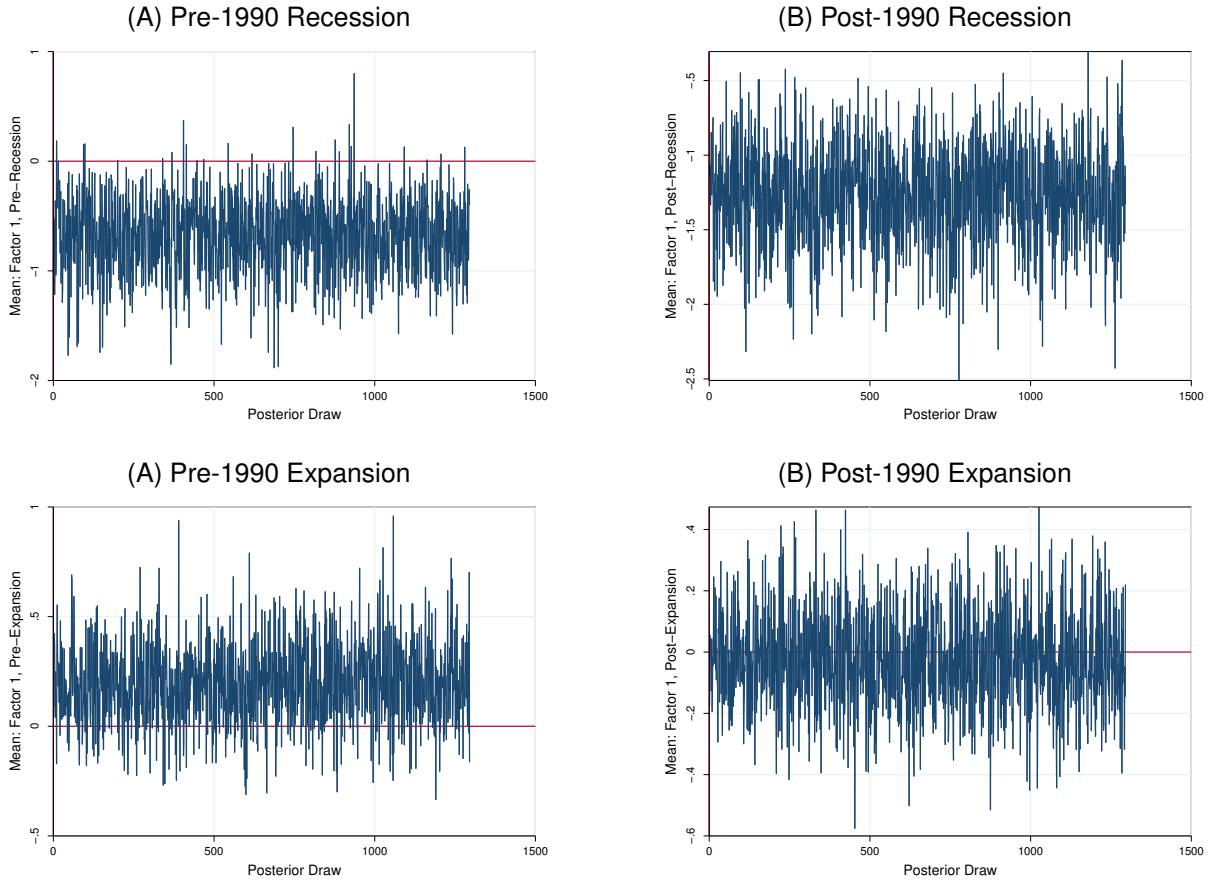
$$f_{kt} = \phi_1 f_{k,t-1} + \cdots + \phi_p f_{k,t-1} + \mu_1 D_t^{(1)} + \cdots + \mu_4 D_t^{(4)} + \nu_t, \nu_t \sim N(0,1)$$

(d) $\pi(\boldsymbol{\sigma}|\boldsymbol{y}, \boldsymbol{f}, \boldsymbol{\delta}, \boldsymbol{\psi}, \boldsymbol{\lambda}) = \prod_{i=1}^{N} IG(\mathrm{e}_i, \mathrm{E}_i)$ where

$$\mathrm{e}_i = \mathrm{e}_0 + 0.5(T-q), \quad \mathrm{E}_i = \mathrm{E}_0 + 0.5 \sum_{t=q+1}^{T} \epsilon_{it}^2$$

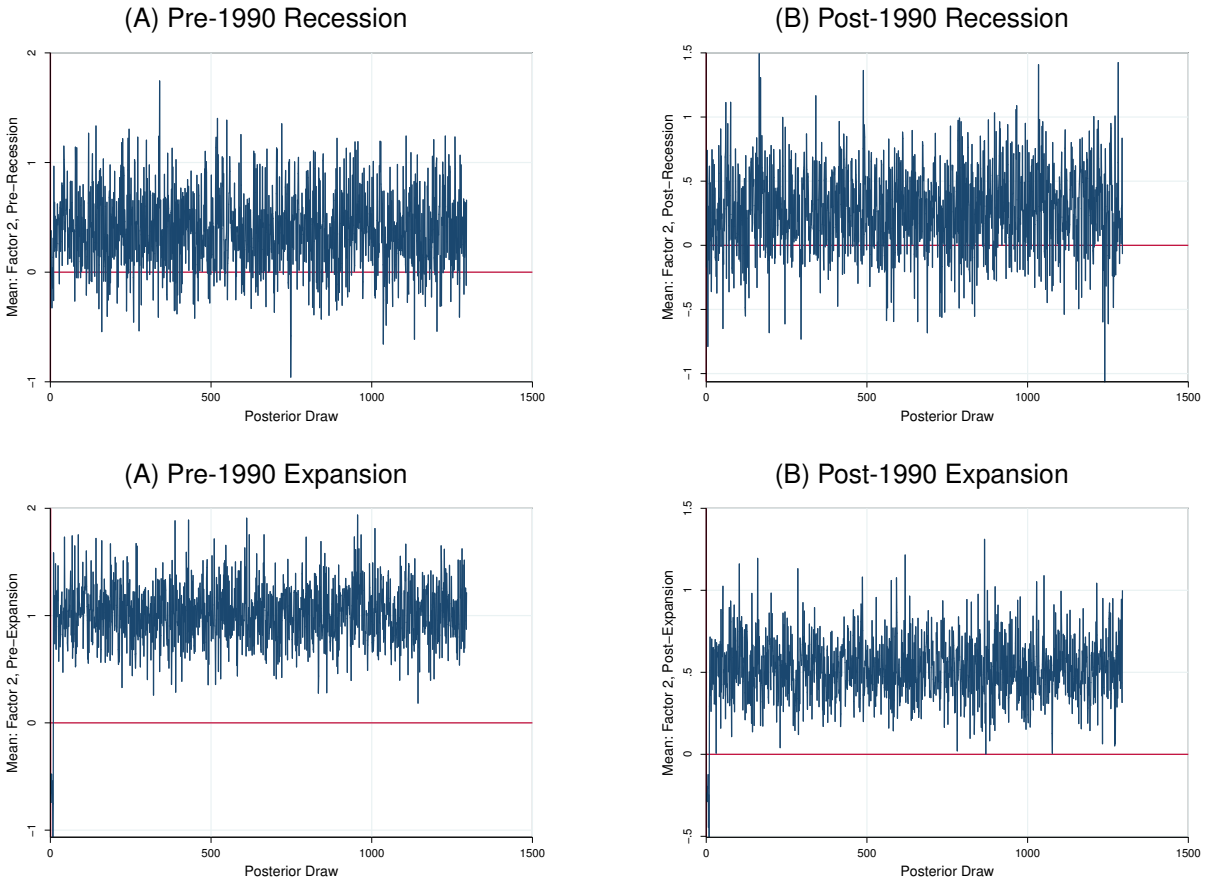
Appendix B. Sampler Convergence

Figure B.9: State Dependent Mean of Factor 1



Notes: The graphs illustrate the retained sample of posterior draws for μ_{kS_t} in which the median cluster assignemnt is sampled, i.e., when $\boldsymbol{\delta} = \boldsymbol{\delta_{50}}$.

Figure B.10: State Dependent Mean of Factor 2



Notes: The graphs illustrate the retained sample of posterior draws for μ_{kS_t} in which the median cluster assignemnt is sampled, i.e., when $\boldsymbol{\delta} = \boldsymbol{\delta_{50}}$.