Establishing stylized facts of concentration in US regional labor markets

Ansel Schiavone, Swayamsiddha Sarangi

January 20, 2026

Abstract

This paper establishes stylized facts on U.S. labor market concentration from 1998–2016. Using worker-transition-pattern clustering, we identify six aggregate labor markets across commuting zones. We find a decline in average concentration prior to 2008 via reduced firm-size variance, followed by an increase due to rising firm-size variance falling firm count.

Keywords: labor market concentration; monopsony; regional labor markets; unsupervised learning

1 Introduction

The primary purpose of this paper is to establish a set of stylized facts on labor market concentration in the United States. Doing so requires addressing two fundamental questions: (1) what constitutes a labor market? and (2) what mechanism(s) drive labor market concentration dynamics.

To answer the first question, we propose an unsupervised learning method to detect latent structure within employment data. Specifically, we employ a clustering algorithm on the transition matrix of workers moving between industries to identify industry clusters in which certain workers tend to concentrate. We find that national employment data organize into six broad labor market clusters (LMs). Merging these industry cluster definitions with commuting zones (CZs) provides delineation of approximately 3,500 regional labor markets (CZ-LMs). We propose this approach as an agnostic, empirical method for identifying labor markets in both national and regional industry-based employment data.

To address the second question, we adapt a concentration decomposition method originally developed by de Gioia (2017) to the context of labor markets. This method decomposes labor market concentration into two components – the *scale effect* and *distribution effect* – which sum to total concentration as measured by the Herfindahl–Hirschman Index (HHI).

To our knowledge, these approaches have not previously been used – independently or in combination – to analyze labor market concentration. We argue that, taken together, they provide a coherent and flexible framework for empirical analysis. Using this framework, we establish several empirical facts about U.S. regional labor markets. First, we find that monopsony is more prevalent in non-urban regions. Second, we find that labor markets based in Broader Services exhibit the lowest average levels of concentration, while Information, Finance & Advanced Services and Nondurable Manufacturing & Wholesale trade exhibit the highest. Interestingly, there is significant heterogeneity in demographic characteristics between these two highly concentrated labor markets; the former is highly educated and disproportionately female and non-White, while the latter is less educated and disproportionately male and White.

Finally, we find that employment-weighted average labor market concentration declined modestly between 1998-2016. Reductions in firm-size variance (distribution effect) have tended to decrease concentration, while reductions in the number of firms (scale effect) have tended to increase it. On balance, the distribution effect has dominated, producing moderate declines in average concentration within most labor markets.

The remainder of the paper proceeds as follows. Section 2 provides a review of the existing literature. Section 3 develops the unsupervised learning approach for delineating labor markets using worker transition data. Section 4 presents the concentration decomposition method. Section 5 integrates these two methods to establish a set of stylized facts regarding concentration dynamics in U.S. labor markets. Section 6 concludes.

2 Literature Review

Imperfect competition in labor markets is now widely regarded as pervasive in the United States, with strong empirical evidence of adverse effects on wages, employment, and the labor share of income (Card, 2022). The literature, however, diverges along three dimensions: the sources of monopsony power, its empirical measurement, and its observed characteristics.

Regarding sources, the dominant explanation emphasizes search frictions arising from imperfect information and workers' heterogeneous preferences over non-wage attributes such as commuting distance (Manning, 2013; Bhaskar & To, 1999). Industrial organization research instead highlights institutional sources of monopsony, including non-compete and no-poaching agreements, which have been shown to suppress wages (Ashenfelter *et al.*, 2022; Callaci *et al.*, 2024). Macroeconomic explanations are less developed, though existing work suggests that monopsony power increases during recessions (Depew & Sørensen, 2013) and that trade exposure may have ambiguous effects, as firm exit among domestic producers can be offset by foreign entry (Amiti & Heise, 2025).

Empirical measurement of monopsony has followed three main approaches. The first directly measures labor market concentration using employment data, drawing on sources such as the Quarterly Census of Employment and Wages (QCEW) (Thompson, 2024), Longitudinal Business Database (LBD) (Rinz, 2022; Benmelech *et al.*, 2022), County Business Patterns (CBP) (Schiavone, 2023), and online vacancy data (Azar *et al.*, 2020, 2022). These studies generally find that many workers face high concentration and that concentration is associated with lower wages or labor shares.

A second approach estimates firm-level labor supply elasticities to infer monopsony power (Manning, 2013; Bassier *et al.*, 2022). While grounded in standard market-power theory, this approach typically requires linked employer–employee data and relies on narrowly defined labor markets, limiting generalizability.

A third approach estimates wage markdowns from firms' production functions, with evidence documented in manufacturing (Yeh *et al.*, 2022a; Ribeiro, 2023), retail (Bonanno & Lopez, 2012a), and healthcare (Matsudaira, 2014). Like elasticity-based methods, this approach is inherently micro-level and does not directly characterize economy-wide monopsony.

Given these varied methodologies, it is unsurprising that no consensus has emerged regarding aggregate trends in U.S. labor market concentration. Some common patterns nevertheless appear. Concentration tends to be higher in manufacturing (Yeh *et al.*, 2022a; Schiavone, 2023), with more adverse wage effects in rural labor markets (Bonanno & Lopez, 2012a). Labor market frictions may also exacerbate exploitation of women and Black and Latino workers (Stelzner & Bahn, 2022), though these findings sit uneasily with evidence on the demographic composition of highly concentrated industries (Schiavone, 2023). Importantly, several studies find little evidence of a secular rise in labor market concentration comparable to that observed in product markets (Rinz, 2022; Azar *et al.*, 2022; Schiavone, 2023).

Against this backdrop, our paper makes three contributions. First, we provide a broad analysis of monopsony at the regional-industry level by identifying six national labor market clusters and approximately three thousand regional labor markets. Second, we adapt the decomposition framework of de Gioia (2017) to labor market concentration, allowing us to distinguish the structural forces driving concentration dynamics. Third, by combining the decomposition method with empirical delineation of labor markets, we establish several emirical facts on U.S. labor market monopsony.

3 Delineating Labor Markets

A fundamental step in studying concentration is defining the relevant "market." In labor markets, this definition spans physical space, skills, technology, and worker preferences. Questions such as the geographic scope of job search, the role of remote work, and workers' ability to reskill or relocate admit no universal answers. As a result, empirical analysis necessarily relies on empirically motivated definitions.

3.1 Existing delineations

U.S. regional economic data are highly granular. Datasets such as County Business Patterns (CBP) and the Longitudinal Business Database (LBD) provide county- and firm-level observations with associated NAICS codes. The central challenge is selecting appropriate dimensions and levels of aggregation.

Spatial aggregation is largely standardized through USDA commuting zones (CZs), which group counties linked by commuting flows. There are 706 CZs in the United States, and they serve as the primary spatial unit in much of the literature on regional labor markets (Rinz *et al.*, 2018; Berger *et al.*, 2022; Schiavone, 2023; Azar *et al.*, 2022).

Aggregation along the skill dimension is less settled. Labor markets are defined not only by geography but also by the types of workers and skills involved. Many studies rely on industry-based definitions using NAICS codes, though results are sensitive to the chosen level of aggregation (Schiavone, 2023). An alternative approach uses occupational classifications such as Standard Occupation Classification (SOC) or Census occupation codes to group jobs with similar skill requirements (Azar *et al.*, 2022; Macaluso, 2025). While highly granular, this approach yields thousands of distinct labor markets and limits comparability with industry-level regional data.

A common criticism of industry-based delineations is "occupational spillover," whereby industries employ heterogeneous occupations, weakening the link between employer concentration and labor market power. While valid, industry-based definitions offer the practical advantage of compatibility with regional economic data. The challenge, therefore, is to construct industry-based labor markets that better reflect underlying occupational structure.

3.2 Non-negative Matrix Factorization

We address this challenge by identifying labor markets as clusters of 3-digit NAICS industries linked by worker mobility using non-negative matrix factorization (NMF), an unsupervised method first developed by Lee & Seung (2000)[1]

[1] We refer the reader to Lee & Seung (2000) as well as Appendix B for complete discussion of the approach and its application.

In brief, using individual-level longitudinal data from the Current Population Survey (CPS) for 1998–2016, we construct an industry transition matrix capturing worker movements across industries. Row-normalizing these transitions yields a stochastic transition probability matrix P. NMF decomposes the transition matrix into the non-negative matrix product:

$$P \approx WH, \tag{1}$$

where W $(n \times k)$ captures membership weights of n industries across k (latent) labor markets, and H $(k \times n)$ captures how industries relate within each labor market. Each industry is assigned to the LM with the highest membership weight in W.

Industry clustering may be informed by observable characteristics that affect worker mobility. We incorporate such information by weighting transitions between industries with similar attributes. Let $\Pi_{i,j}$ denote a similarity measure between industries i and j, constructed from observable characteristics such as education, age, and gender composition. The transition matrix is adjusted via element-wise multiplication:

$$P' = \Pi \odot P \tag{2}$$

and renormalized prior to factorization. This procedure increases the influence of transitions between industries with similar workforce characteristics. In our estimations, we consider priors based on average years of education, worker age, and sex composition.

[TABLE 1 ABOUT HERE]

Because no ground truth exists for latent labor markets, model evaluation focuses on minimizing occupational spillover. We assess classification performance using two metrics: Cramer's V, which measures correspondence between industry clusters and occupation codes, and normalized occupational entropy, which measures occupational concentration within clusters.

For each prior specification, we search over weights $\lambda_\rho \in [0, 1]$ select the number of clusters using BIC, and retain the configuration that minimizes occupational entropy. We compare NMF-based classifications with standard NAICS 2- and 3-digit groupings. Table 1 shows that all NMF specifications outperform NAICS classifications. A model with a single educational prior achieves the lowest entropy and highest Cramer's V, while additional priors yield diminishing returns. Because all NMF variants produce highly similar clusterings, we select the educational-prior specification (NMF2) for subsequent analysis. Full algorithmic details are provided in Appendix C.

[**FIGURE 1 ABOUT HERE**]

Figure 1 presents the six labor markets identified by NMF2. Industry composition is reported in Table 2, and summary statistics appear in Table 3. Education levels are lowest in Distribution, Transportation, & Utilities (LM1), Durable Manufacturing & Construction (LM3), and Agriculture (LM6), and highest in Information, Finance, & Advanced Services (LM2) and Broader Services (LM5). Higher-education labor markets are younger and more female and Asian, while lower-education markets are older and more male and White. Broader Services (LM5) accounts for nearly half of private-sector employment.

[**TABLE 3 ABOUT HERE**]

4 Decomposition Framework

We turn now to a decomposition method intended to uncover underlying dynamics driving labor market concentration. This approach is developed by de Gioia (2017) to decompose concentration in product markets. To our knowledge, the approach has not been used to study labor market concentration. The method is intended to distinguish the effects of firm number and market share inequality in measuring concentration. Applied specifically to shares of employment, the decomposition allows for a deeper understanding of the driving forces of monopsony within labor markets.

Suppose a labor market has n firms with employment L_1, L_2, ..., L_n. Let:

$$\mathbf{L} = (L_1, L_2, ..., L_n) \tag{3}$$

Let $\bar{L}$ denote the vector of firm employment where all firms are equal, such that:

$$\bar{\mathbf{L}} = (\bar{L}, \bar{L}, ..., \bar{L}) \tag{4}$$

The Euclidean distance between market $\bar{\mathbf{L}}$ and $\mathbf{L}$ is given by:

$$d(\mathbf{L}, \bar{\mathbf{L}}) = \sqrt{\sum_{i=1}^{n} L_i^2 - n\bar{L}^2} \tag{5}$$

Consider a situation $\mathbf{L}^*$ where all workers are employed in a single firm, where total employment in the market is equal to that of $\bar{\mathbf{L}}$ (such that the employment of the single firm is $n\bar{L}$). This represents the most concentrated the market can be, holding total employment constant. This maximum distance is given as:

$$d(\mathbf{L}^*, \bar{\mathbf{L}}) = \sqrt{(n-1)n\bar{L}^2} \tag{6}$$

It follows that a market $\mathbf{L}$ will lie between 0 and $d(\mathbf{L}^*, \bar{\mathbf{L}})$, given the total employment in the market. The relative concentration index of the market, given total employment, is thus:

$$\tau = \frac{d(\mathbf{L}, \bar{\mathbf{L}})}{d(\mathbf{L}^*, \bar{\mathbf{L}})} = \sqrt{\frac{\sum_{i=1}^{n} L_i^2 - n\bar{L}^2}{(n-1)n\bar{L}^2}} \in [0, 1] \tag{7}$$

The Herfindahl–Hirschman Index (HHI) is expressed as:

$$HHI = \sum_{i=1}^{n} \Big(\frac{L_i}{n\bar{L}}\Big)^2 \tag{8}$$

It follows that the concentration index τ may be expressed as a function of HHI:

$$\tau = \sqrt{\frac{n}{n-1}(HHI - \frac{1}{n})} = \sqrt{\frac{n \cdot HHI - 1}{n-1}} \tag{9}$$

Note that as $n \to \infty$, $\tau \to \sqrt{HHI}$. It follows, therefore, that the difference between the limit case and the actual concentration index of the market is given by:

$$\sqrt{HHI} - \tau \tag{10}$$

This difference is purely a function of n, and the effect of n is an unambiguous convergence of τ towards $\sqrt{HHI}$. As a result, HHI can be written as:

$$HHI = E_n + E_i \tag{11}$$

where the first term:

$$E_n = HHI - \tau^2 \tag{12}$$

is the *scale effect*, that is, the effect that the number of firms has on HHI, holding overall employment in the market constant. The second term:

$$E_i = \tau^2 \tag{13}$$

is the *distribution effect*, which captures the impact of variance in firm employment size, holding total employment and number of firms constant.

[FIGURE 2 ABOUT HERE]

Figure 2 provides a sketch for how decomposition results are presented throughout the following analysis. Note that positive movements in either the x or y axes results in rising concentration, despite potentially different causal mechanisms.

5 Bringing it together

We now combine the labor market delineation developed in Section 3 with the decomposition framework in Section 4 to analyze monopsony in U.S. regional labor markets. This requires merging labor market classifications with regional establishment-level employment data from County Business Patterns (CBP). CBP reports the distribution of establishment employment sizes by county and NAICS industry, where establishments correspond to physical locations rather than firms; measured concentration therefore likely understates true monopsony.

We restrict attention to the period 1998–2016. Earlier years are excluded due to the SIC–NAICS transition, while later years are affected by disclosure suppression following the removal of observations with three or fewer firms. Agriculture is excluded because CBP underrepresents nonemployer farms and exhibits classification breaks, following standard practice (Schiavone, 2023; Mendieta-Muñoz *et al.*, 2021). CBP data are aggregated to commuting zones and merged with labor market clusters, yielding a panel of CZ–labor markets (CZ-LMs) over time.

5.1 Initial Observations

The scope and dimensionality of the data imply that many results must be omitted. We therefore highlight several basic patterns regarding the levels and dynamics of concentration across labor markets and regions.

[**FIGURE 3 ABOUT HERE**]

Figure 3 shows the distribution of 2016 HHI values among CZs, grouped by LMs. LM5 exhibits the lowest average concentration, consistent with the relatively large number of employers in much of the service economy. LM1–LM4, which span the goods-producing economy as well as technology and information services, tend to be more concentrated. In the spirit of Baumol (1967), these activities are more systematically affected by technical change and scale economies. Importantly, the distributions overlap substantially, indicating meaningful regional heterogeneity in concentration within each labor market.

We next consider concentration for the "average" worker at the CZ level. For each CZ, we compute employment-weighted average concentration across CZ-LMs, where weights reflect each LM's employment share within the CZ. Figure 4 maps this measure for 2016, overlaid with cities with population greater than 100,000.

[**FIGURE 4 ABOUT HERE**]

Average concentration is highest in CZs *without* large population centers. This pattern reflects both fewer employers in smaller markets and the heavier employment weight of inherently concentrated industries in rural regions (Bonanno & Lopez, 2012b; Schiavone, 2023). This urban–rural gradient is also correlated with broader political and institutional differences (see Figures 5 and 6).

What about the demographic characteristics of workers most exposed to monopsony? Directly measuring CZ-LM demographics would require linked worker data with county of residence and industry of employment, which are not widely available. We therefore approximate demographic exposure by combining (i) national average concentration within each LM (employment-weighted across CZ-LMs) with (ii) LM demographic characteristics from Section 3. This approach relies on homogeneity of industry demographics across regions, an assumption that is correct to question.

[**TABLE 4 ABOUT HERE**]

Table 4 highlights two highly concentrated labor markets: Information, Finance, Advanced Services (LM2) and Nondurable Manufacturing & Wholesale Trade (LM4). Despite similar concentration levels, their demographic profiles differ sharply: LM2 is highly educated and disproportionately young, female, and non-White, while LM4 is below-average education and disproportionately male and White. Broader Services (LM5) exhibits substantially lower average concentration and is disproportionately educated, young, female, and non-White. Overall, high concentration is not confined to a single segment of the economy.

[**FIGURE 7 ABOUT HERE**]

Figure 7 displays employment-weighted HHI for each LM over time. Concentration declines from 1998 through the mid-2000s, then reverses during or immediately after the 2008 financial crisis for several labor markets (especially LM2, LM3, and LM4). By 2016, average concentration in LM3 and LM4 exceeds its 1998 level, while LM2 remains below its initial value but increases post-crisis. LM1 and LM5 exhibit steadier declines throughout. At the aggregate level (weighting LMs by employment size), overall concentration declines modestly, with most of the decline occurring before 2008. These patterns suggest distinct mechanisms across labor markets and an important macro inflection point around 2008, motivating the decomposition analysis below.

5.2 Decomposing labor market concentration

Recall from Section 4 that the decomposition separates concentration into a *scale effect* (number of firms) and a *distribution effect* (variance in firm size). Applying this decomposition to the regional panel (weighted by CZ-LM employment) yields the evolution of concentration faced by the average worker within each LM.

[FIGURE 8 ABOUT HERE]

We conclude the section with a discussion as to how these results fit into to the broader literature on labor market concentration and other macroeconomic phenomena that have shaped the US economy throughout the period considered.

Beginning with existing investigations of labor market concentration dynamics, the findings of this paper are largely consistent. Some authors find an increase in concentration in certain areas of the economy, including Yeh *et al.* (2022b) who find that manufacturing concentration decreased from 1970s until early 2000s but then began to rise again. Similarly, Schiavone (2023) finds that average unweighted labor market concentration at the commuting-zone level has increased for certain industries such as mining and finance, but has remained constant for others. Thompson (2024) finds that on average the US labor market has become 7% less concentrated when looking at metropolitan statistical area (MSA)-industry labor markets over the period 2002-2023. The results of this paper are largely in line with existing findings, with certain labor markets experiencing share increases in concentration post-2008, despite constant or negative trends in other labor markets as well as at the aggregate level.

We turn now to the how the findings fit within the broader scope of macroeconomic phenomena during the period in question. We consider briefly the following: trade shocks (specifically the "China shock" beginning in the early 2000s), the 2008 financial crisis, the rise of superstar firms, and finally structural change.

China shock

Beginning with the trade shocks of the early 21st century, Autor *et al.* (2016) serves as the seminal analysis of the impact on US labor markets from China's entrance into the World Trade Organization (WTO) and penetration of domestic product markets. Using the same geographic unit of analysis as this paper, the authors find import penetration explains approximately one quarter of the decline in manufacturing between 1990-2007. Importantly, the analysis treats CZs as separate, individual labor markets (without industry delineation). Unsurprisingly, therefore, they find that the impact of foreign imports on local labor markets were most pronounced in CZ with high levels of manufacturing. However, they also find that trade shocks impact other non-manufacturing sectors that are linked to manufacturing as well.

Based on these findings, it seems almost certain that trade shocks have impacted labor market concentration. However, the direction of this impact requires a deeper analysis of the impact of import exposure along the distribution of firm employment size. Specifically, does import exposure explain the decline in firm size variance observed within manufacturing and adjacent labor markets (LM1, LM3, LM4)? If larger firms tend to offshore (Bernard *et al.*, 2003), this will result in decline in firm size variance without significant decline in overall number of firms, as is observed in Figure 8 for 1998-2007. In other words, the idea that more productive firms (which tend to be larger due to increasing returns to scale) are more likely to offshore would mean that globalization is likely to make domestic labor markets *more* competitive.

At the same time, one would expect trade shocks to result in a decline in the overall number of firms within manufacturing and adjacent labor markets, thus working to increase concentration via the scale effect. Interestingly, there is little evidence of this in the pre-2008 years. Some possible explanations for this are: 1) a delayed effect – while China joined the WTO in 2002, it likely took some time for the effects of this integration to impact domestic firm counts, and 2) a generally strong economy leading up to the financial crisis, whereby robust economic conditions before 2008 meant that firms were able to remain open even under increasingly fierce foreign competition.

2008 Financial Crisis

The financial crisis is clearly a point of inflection within the trajectory of the US economy. This is observed in the results of the decomposition, with a pronounced decline in the number of firms contributing to a rise in labor market concentration in the post-crisis period via the scale effect. At the national level, all labor markets have experienced a decline in establishment count per worker since 2008, with all but LM5 reaching their zenith in 2010 followed by a sustained decline for the remainder of the period (see Figure 9). Similarly, Federal Reserve Bank (2025) data show a steady decline in the number of publicly listed companies per million people throughout the period, from 16.9 in 2007 to 13.4 in 2016.

It is notable that contributions to rising concentration from a declining number of firms is persistent throughout the remainder of the period, despite recovery in aggregate employment statistics. Hershbein & Stuart (2024) find that local labor markets that experience greater decline in employment during recessions also suffer from persistent declines in the employment-population ratio. This may point to hysteresis when it comes to number of firms, thereby explaining the persistent pro-concentration impact of number of firms component of the decomposition on concentration post-2008.

The Rise of Superstar Firms

A phenomenon that has garnered significant attention particularly in the literature surrounding income distribution is the rise of the "superstar firm", a term first coined by Autor *et al.* (2020). In short, the forces of globalization and technical change have resulted in increased market shares of the most productive firms within industries, leading to eventual domination of a handful of firms across broad swathes of the economy. While the focus of the Autor *et al.* (2020) is specifically on rising share of these superstar firms in product markets, others have pointed out that a likely corollary is increased market power within labor markets of these same firms (Schiavone, 2023). Put differently, firms that dominate entire industries are likely to exhibit characteristics of monopsonistic employers.

So does this explain the observed dynamics of contributions from firm size variation, specifically the large positive contributions in later years? Perhaps, but there is a timing issue with this possibility. Recall that firm employment size variance is declining until 2008 for all labor markets. However, Autor *et al.* (2020) put the rise of superstar firms as beginning in the early 1990s in many industries. If this is the case, it seems puzzling that labor markets simultaneous became more competitive during the period of rising concentration within product markets.

One possible explanation for this is that superstar firms may have been quicker to dominate product markets due to increased productivity over competitors, but remained relatively smaller in employment shares due their greater productivity or alternative production modes. It may be the case that the 2008 financial crisis served as the final straw for non-superstar firms to exit both product *and* labor markets, leading to rapid concentration in the latter.

Deindustrializing Structural Change

Lastly, we consider how these results fit in with the literature on structural change. Put simply, this literature examines the implications on growth and distribution associated with the secular decline in manufacturing activities in levels as well as a share of the overall economy. Here, we focus particularly on the two manufacturing labor markets – LM3 and LM4. Within the context of the decomposition, there are two avenues whereby structural change impacts concentration: 1) the number of firms within manufacturing labor markets, and 2) the distribution of employment by manufacturing firms. In addition, structural change impacts concentration within the "aggregate" US labor market; leaving aside the direct impacts on firm count and variance within manufacturing, secular shifts of labor between industries (e.g. from manufacturing to services) will lead to an aggregate reweighing of the various labor markets in the overall economy.

Beginning with the impact of number of firms, throughout the period the gross number of establishments has declined for both LM3 and LM4 over the period (-5.6% and -10.9%, respectively). Thus, firms either shutting down or going overseas contributes to rising concentration via a decline in firm count (this is discussed briefly in Schiavone 2023).

The impact of structural change on the distribution of firm size is less clear. As discussed previously, there is evidence that more larger, productive firms are more likely to offshore activities (Bernard *et al.*, 2003). In this case, the impact would contribute negatively towards concentration. However, if import exposure results in the shutdown of smaller firms, this could increase variance of firm size, resulting in rising concentration. If we were to explain the movements along the x-axis of the decomposition purely through the lens of structural change, we would say that the former mechanism dominated in the pre-2008 period, while the latter dominated in the post-2008 period.

Finally, what about the effect of structural change on concentration on the "aggregate" US labor market? Structural change means that the shares of employment in each labor market are evolving over time, in particular with shifts towards LM5 (see Figure 10). If workers decide to transition to service-sector work (LM5) because of inability to find sufficient employment in increasingly concentrated and shrinking manufacturing labor markets (LM3 or LM4), the net effect may be a decline in the exposure to concentration of the average American worker.[2] In other words, workers are transitioning to less concentrated labor markets, because of factors contributing to higher concentration *within* the labor markets they are exiting. For example, LM4, one of the most concentrated labor markets in our results, includes Beverage & Tobacco, Chemical, and Petroleum & Coal Products Manufacturing—low labor share sectors that have consistently led manufacturing labor share decline (Sarangi, 2026). The impacts on distribution and worker bargaining power are therefore unclear, and in our opinion worthy of further investigation. This echoes the point raised by Mendieta-Muñoz *et al.* (2021); Rada *et al.* (2021) that structural shifts that are nominally "pro-worker" in distribution are not always so in worker welfare.

A Note on Future Work

As is often the case, these results reveal as many questions as they do answers. As the primary concern of monopsony is distributional (i.e. wage suppression), the question of mechanisms comes to the forefront. The decomposition demonstrates two different ways in which labor market concentration can change – rising variance of firm sizes and changes in the number of firms. In addition, shifting weights between labor markets also change the overall degree of labor market concentration. Do these underlying dynamics carry the same importance when it comes to wage determination?

This question fits within a growing set of complications that exist at the intersection of labor and macroeconomics. For example, consider the observation that larger firms

[2] While the method in this paper does opt for an industry-based delineation of labor markets, it does take into consideration transitions of workers between industries. However, because the delineation is created using a panel of worker transition patterns, it may not be robust to evolution in these linkages. This is something that should be addressed in future work.

tend to be higher paying in many industries (see for example Berlingieri *et al.* (2018) and Cobb & Lin (2017)), despite workers receiving a lower share of total income (Autor *et al.*, 2020). Similarly, one might ask whether a worker is really better off in a labor market that is becoming less concentrated due to the exit of large productive firms. To answer this question requires careful examination of the mechanism(s) linking monopsony to distribution, and how these are impacted by macroeconomic forces.

6 Conclusion

This paper consists of three exercises. First, we introduce an unsupervised learning approach to detect labor markets within employment transition data. This method is more sophisticated in its delineation of labor markets than traditional industry classifications by minimizing the risk of occupational spillover between labor markets, but still preserves the ability to merge labor market classifications onto regional industry data. The optimized classification model produces six labor markets.

Second, we adapt the Herfindahl–Hirschman index decomposition of de Gioia (2017) to labor markets, providing a framework for separating the effects of scale and distribution underlying dynamics of labor market concentration.

Finally, we conduct an empirical investigation into regional labor market concentration using commuting-zone level data from the County Business Patterns dataset merged with Current Population Survey worker data. Descriptive statistics reveals significant heterogeneity in terms of industry, education, and demographics among workers most exposed to concentrated labor markets. The subsequent decomposition analysis uncovers a variety of important dynamics. Most notably, while labor market concentration largely declined in the period 1998-2007 due to declining firm-size variance, 2008-2016 saw a reversal in these dynamics for several labor markets, driven both by rising firm-size variance and falling firm count.

In terms of contributions, the paper establishes a method for decomposing the drivers labor market concentration dynamics, which is useful in informing development theoretical mechanisms linking concentration to distribution. Additionally, we develop a novel method for delineating labor markets that is both sensitive to empirical patterns and occupational spillover, while still operable with regional industry data. Lastly, the paper uncovers important dynamics for the period 1998-2016. For most labor markets, we observe a fall in concentration in the early years driven by falling variance in firm size, followed by a rise in concentration in the later years driven by both rising firm-size variance and a falling number of firms. This leaves open important questions as to how these patterns impact the distribution of income between labor and capital.

References

Amiti, Mary, & Heise, Sebastian. 2025. US market concentration and import competition. *Review of Economic Studies*, **92**(2), 737–771.

Ashenfelter, Orley, Card, David, Farber, Henry, & Ransom, Michael R. 2022. Monopsony in the labor market: new empirical results and new public policies. *Journal of Human Resources*, **57**(3), S1–S10.

Autor, David, Dorn, David, Katz, Lawrence F, Patterson, Christina, & Van Reenen, John. 2020. The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, **135**(2), 645–709.

Autor, David H, Dorn, David, & Hanson, Gordon H. 2016. The China shock: Learning from labor-market adjustment to large changes in trade. *Annual review of economics*, **8**(1), 205–240.

Azar, José, Marinescu, Ioana, Steinbaum, Marshall, & Taska, Bledi. 2020. Concentration in US labor markets: Evidence from online vacancy data. *Labour Economics*, **66**, 101886.

Azar, José, Marinescu, Ioana, & Steinbaum, Marshall. 2022. Labor market concentration. *Journal of Human Resources*, **57**(S), S167–S199.

Bassier, Ihsaan, Dube, Arindrajit, & Naidu, Suresh. 2022. Monopsony in movers: The elasticity of labor supply to firm wage policies. *Journal of Human Resources*, **57**(S), S50–s86.

Baumol, William J. 1967. Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis. *The American Economic Review*, **57**(3), 415–426.

Benmelech, Efraim, Bergman, Nittai K, & Kim, Hyunseob. 2022. Strong employers and weak employees: How does employer concentration affect wages? *Journal of Human Resources*, **57**(S), S200–S250.

Berger, David, Herkenhoff, Kyle, & Mongey, Simon. 2022. Labor market power. *American Economic Review*, **112**(4), 1147–1193.

Berlingieri, Giuseppe, Calligaris, Sara, & Criscuolo, Chiara. 2018. The productivity-wage premium: Does size still matter in a service economy? *Pages 328–333 of: AEA Papers and Proceedings*, vol. 108. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Bernard, Andrew B, Eaton, Jonathan, Jensen, J Bradford, & Kortum, Samuel. 2003. Plants and productivity in international trade. *American economic review*, **93**(4), 1268–1290.

Bhaskar, Venkataraman, & To, Ted. 1999. Minimum wages for Ronald McDonald monopsonies: A theory of monopsonistic competition. *The Economic Journal*, **109**(455), 190–203.

Bonanno, Alessandro, & Lopez, Rigoberto A. 2012a. Wal-Mart's monopsony power in metro and non-metro labor markets. *Regional Science and Urban Economics*, **42**(4), 569–579.

Bonanno, Alessandro, & Lopez, Rigoberto A. 2012b. Wal-Mart's monopsony power in metro and non-metro labor markets. *Regional Science and Urban Economics*, **42**(4), 569–579.

Callaci, Brian, Gibson, Matthew, Pinto, Sérgio, Steinbaum, Marshall, & Walsh, Matt. 2024. The effect of franchise no-poaching restrictions on worker earnings. *Review of Economics and Statistics*, 1–35.

Card, David. 2022. Who set your wage? *American Economic Review*, **112**(4), 1075–1090.

Cobb, J Adam, & Lin, Ken-Hou. 2017. Growing apart: The changing firm-size wage premium and its inequality consequences. *Organization Science*, **28**(3), 429–446.

de Gioia, Giacomo. 2017. A Decomposition of the Herfindahl Index of Concentration.

Depew, Briggs, & Sørensen, Todd A. 2013. The elasticity of labor supply to the firm over the business cycle. *Labour Economics*, **24**, 196–204.

Federal Reserve Bank, St. Louis. 2025. *Number of Listed Companies for United States (DDOM01USA644NWDB)*. https://fred.stlouisfed.org/series/DDOM01USA644NWDB. Accessed: 2025-05-06.

Hershbein, Brad, & Stuart, Bryan A. 2024. The evolution of local labor markets after recessions. *American Economic Journal: Applied Economics*, **16**(3), 399–435.

Lee, Daniel, & Seung, H Sebastian. 2000. Algorithms for non-negative matrix factorization. *Advances in neural information processing systems*, **13**.

Lee, Daniel D, & Seung, H Sebastian. 1999. Learning the parts of objects by non-negative matrix factorization. *nature*, **401**(6755), 788–791.

Macaluso, Claudia. 2025. Skill remoteness and post-layoff labor market outcomes. *American Economic Journal: Macroeconomics*, **17**(2), 134–176.

Manning, Alan. 2013. *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press.

Matsudaira, Jordan D. 2014. Monopsony in the low-wage labor market? Evidence from minimum nurse staffing regulations. *Review of Economics and Statistics*, **96**(1), 92–102.

Mendieta-Muñoz, Ivan, Rada, Codrina, & von Arnim, Rudi. 2021. The decline of the US labor share across sectors. *Review of Income and Wealth*, **67**(3), 732–758.

Rada, Codrina, Schiavone, Ansel, & von Arnim, Rudi. 2021. The Labor Share and Strucutral Change: Insights From Baumol and Lewis. *Working Paper*.

Ribeiro, Eduardo Pontual. 2023. Monopsony Power and Coordination in the Broiler Industry. *The Antitrust Bulletin*, **68**(1), 24–46.

Rinz, Kevin. 2022. Labor market concentration, earnings, and inequality. *Journal of Human Resources*, **57**(S), S251–S283.

Rinz, Kevin, *et al.* 2018. Labor market concentration, earnings inequality, and earnings mobility. *Center for Administrative Records Research and Applications Working Paper*, **10**.

Sarangi, Swayamsiddha. 2026. Labor share decline across US manufacturing sectors: 1979–2019. *International Review of Applied Economics*, **40**(1), 63–89.

Schiavone, Ansel. 2023. Labor market concentration and labor share dynamics for US regional industries. *Economic Modelling*, **125**, 106342.

Stelzner, Mark, & Bahn, Kate. 2022. Discrimination and monopsony power. *The Review of Black Political Economy*, **49**(2), 152–174.

Thompson, Trent L. 2024. Measuring labor market concentration using the QCEW. *MONTHLY LABOR REVIEW*.

Yeh, Chen, Macaluso, Claudia, & Hershbein, Brad. 2022a. Monopsony in the US labor market. *American Economic Review*, **112**(7), 2099–2138.

Yeh, Chen, Macaluso, Claudia, & Hershbein, Brad. 2022b. Monopsony in the US labor market. *American Economic Review*, **112**(7), 2099–2138.

A Figures & Tables

Method	**Clusters (k)**	**Priors (Optimal weights)**	**Cramer's V**	**Entropy**
NAICS 3-digit	77	–	0.461	140.375
NAICS 2-digit	19	–	0.501	95.001
NMF1	6	–	0.4807	66.238
NMF2	6	Ed. (0.65)	0.4809	66.044
NMF3	6	Ed. (0.65), Sex (0.50)	0.4807	66.228
NMF4	6	Ed. (0.65), Sex (0.50), Age (0.15)	0.4807	66.228

Table 1: Table with clustering model characteristics

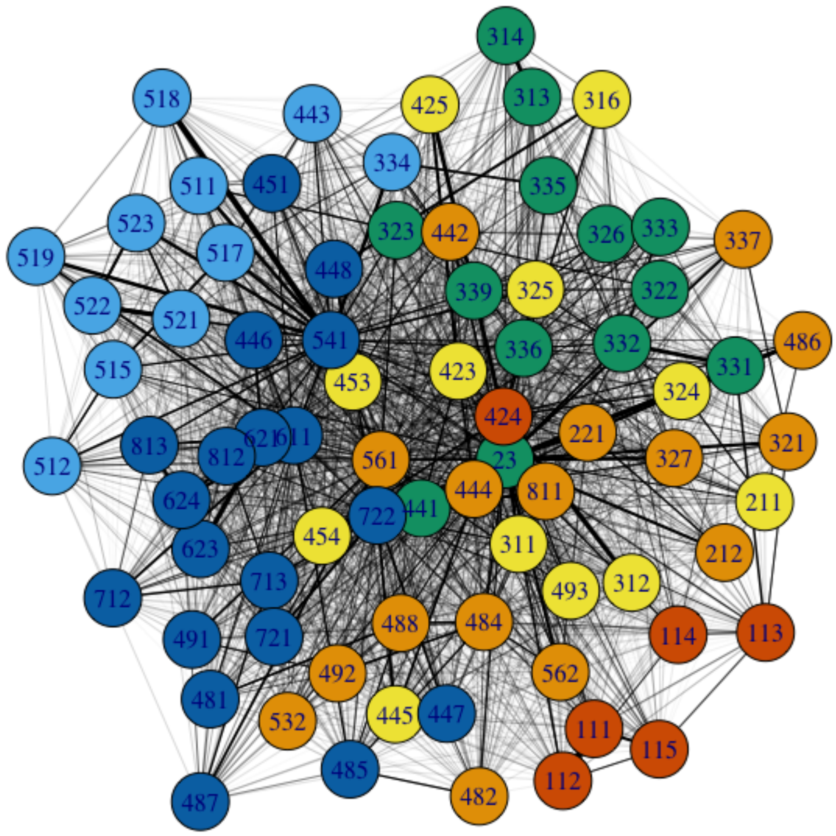

Figure 1: Non-negative matrix factorization NAICS 3-digit industry clusters, educational weighting ($\lambda = 0.65$).

Table 2: NMF Cluster Assignments by NAICS 3-digit Industry

Cluster	**NAICS**	**Industry Description**
1	212	Mining (except Oil & Gas)
1	221	Utilities
1	321	Wood Product Manufacturing
1	327	Nonmetallic Mineral Product Manufacturing
1	337	Furniture & Related Product Manufacturing
1	442	Furniture & Home Furnishings Stores
1	444	Building Material & Garden Equipment & Supplies Dealers

Continued on next page

Continued from previous page

Cluster	NAICS	Industry Description
1	482	Rail Transportation
1	484	Truck Transportation
1	486	Pipeline Transportation
1	488	Support Activities for Transportation
1	492	Couriers & Messengers
1	532	Rental & Leasing Services
1	561	Administrative & Support Services
1	562	Waste Management & Remediation Services
1	811	Repair & Maintenance
2	334	Computer & Electronic Product Manufacturing
2	443	Electronics & Appliance Stores
2	511	Publishing Industries (except Internet)
2	512	Motion Picture & Sound Recording Industries
2	515	Broadcasting (except Internet)
2	517	Telecommunications
2	518	Data Processing, Hosting, & Related Services
2	519	Other Information Services
2	521	Monetary Authorities - Central Bank
2	522	Credit Intermediation & Related Activities
2	523	Securities, Commodities, & Other Financial Investments
3	23	Construction
3	313	Textile Mills
3	314	Textile Product Mills
3	322	Paper Manufacturing
3	323	Printing & Related Support Activities
3	326	Plastics & Rubber Products Manufacturing
3	331	Primary Metal Manufacturing
3	332	Fabricated Metal Product Manufacturing
3	333	Machinery Manufacturing
3	335	Electrical Equipment & Component Manufacturing
3	336	Transportation Equipment Manufacturing
3	339	Miscellaneous Manufacturing
3	441	Motor Vehicle & Parts Dealers
4	211	Oil & Gas Extraction
4	311	Food Manufacturing

Continued on next page

Continued from previous page

Cluster	NAICS	Industry Description
4	312	Beverage & Tobacco Product Manufacturing
4	316	Leather & Allied Product Manufacturing
4	324	Petroleum & Coal Products Manufacturing
4	325	Chemical Manufacturing
4	423	Merchant Wholesalers, Durable Goods
4	425	Wholesale Electronic Markets & Agents & Brokers
4	445	Food & Beverage Stores
4	453	Miscellaneous Store Retailers
4	454	Nonstore Retailers
4	493	Warehousing & Storage
5	446	Health & Personal Care Stores
5	447	Gasoline Stations
5	448	Clothing & Clothing Accessories Stores
5	451	Sporting Goods, Hobby, Musical, & Book Stores
5	481	Air Transportation
5	485	Transit & Ground Passenger Transportation
5	487	Scenic & Sightseeing Transportation
5	491	Postal Service
5	541	Professional, Scientific, & Technical Services
5	611	Educational Services
5	621	Ambulatory Health Care Services
5	623	Nursing & Residential Care Facilities
5	624	Social Assistance
5	712	Museums, Historical Sites, & Similar Institutions
5	713	Amusement, Gambling, & Recreation Industries
5	721	Accommodation
5	722	Food Services & Drinking Places
5	812	Personal & Laundry Services
5	813	Religious, Grantmaking, Civic, Professional, & Similar
6	111	Crop Production
6	112	Animal Production & Aquaculture
6	113	Forestry & Logging
6	114	Fishing, Hunting & Trapping
6	115	Support Activities for Agriculture & Forestry
6	424	Merchant Wholesalers, Nondurable Goods

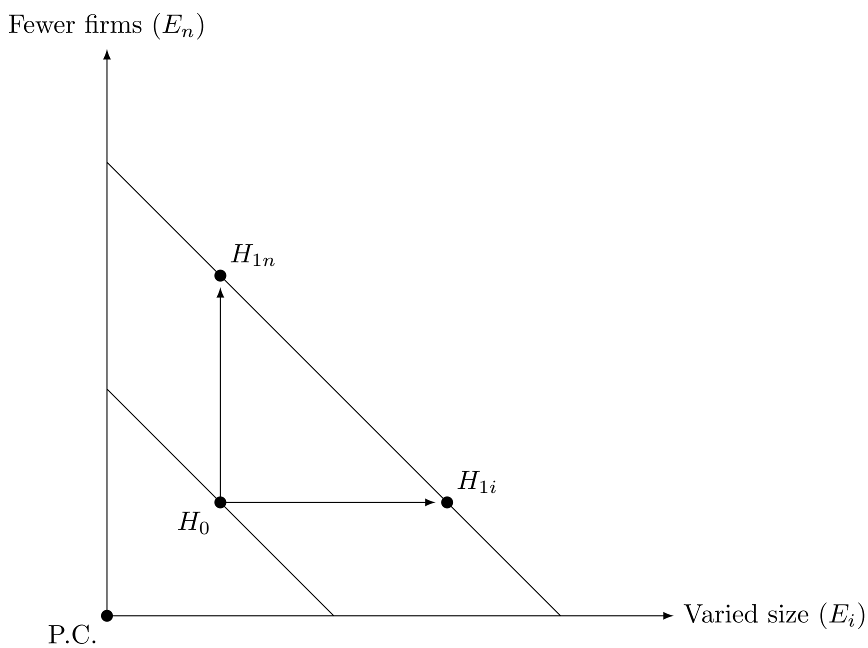

Figure 2: Decomposition component graph. Higher concentration may be observed from either an increase in the variance in firm size (increase in E_i) or a decrease in number of firms (increase in E_n).

Labor Market (Cluster)	Years Ed.	Age	Female %	White %	Black %	Asian %	% of Total
Distribution, Transportation, & Utilities (LM1)	12.80	43.87	26.43	86.41	8.14	2.64	13.89
Information & Financial Services (LM2)	14.53	43.28	48.55	84.99	7.01	5.92	7.96
Durable Manufacturing & Construction (LM3)	12.72	43.85	17.80	89.92	5.05	2.56	18.26
Nondurable Manufacturing & Wholesale Trade (LM4)	13.22	44.24	38.43	86.59	6.81	4.39	6.64
Broader Services (LM5)	14.43	43.23	62.68	83.70	8.43	5.06	48.82
Agriculture (LM6)	12.65	45.77	25.13	90.94	3.93	2.65	4.43

Table 3: Summary statistics for labor markets, as defined by Table 2 1998-2016.

Labor Market (Cluster)	Years Ed.	Age	Female %	White %	Black %	Asian %	HHI
Logistics, Utilities, & Transportation (1)	12.80***	43.87***	26.43***	86.41***	8.14***	2.64***	56.88
Information, Finance, Advanced Services (2)	14.53***	43.28***	48.55***	84.99***	7.01***	5.92***	95.13
Durable Manufacturing & Construction (3)	12.72***	43.85	17.80***	89.92***	5.05***	2.56	73.93
Nondurable Man. & Wholesale Trade (4)	13.22***	44.24**	38.43***	86.59	6.81***	4.39***	87.73
Broader Services (5)	14.43***	43.23***	62.68***	83.70***	8.43**	5.06***	16.93
National Average	13.74	43.62	45.05	85.83	7.35	4.18	52.49

Table 4: Demographics and concentration statistics for labor markets, as defined by Table 2 1998-2016. Asterisks denote cluster mean differs significantly from the national average based on regressions of centered outcomes on clusters; $(* = p < 0.1), (** = p < 0.05), (*** = p < 0.01)$.

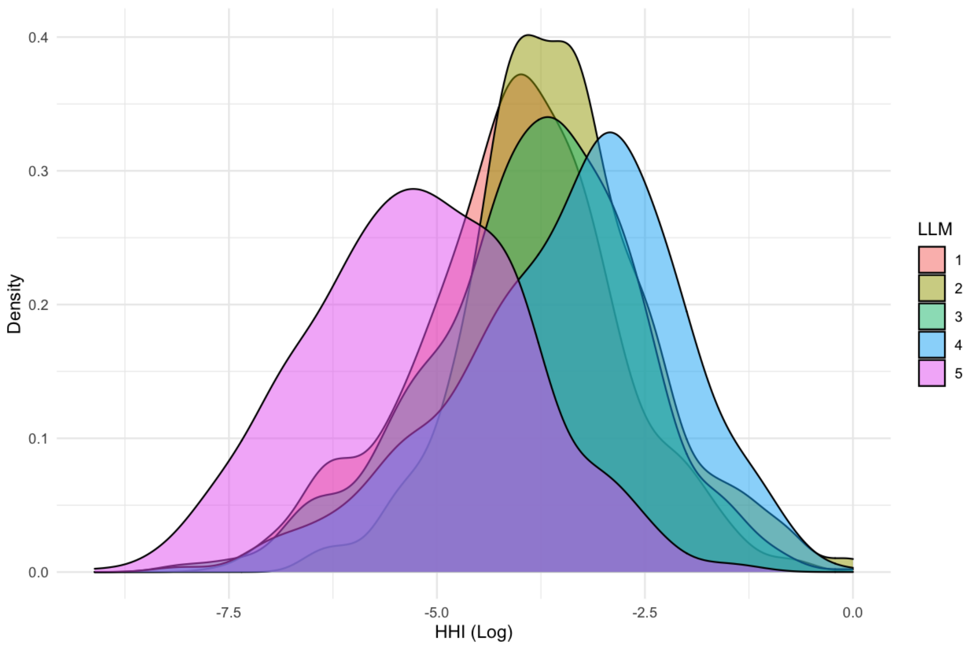

Figure 3: Density plot of HHI at CZ-level, grouped by latent labor markets. 2016.

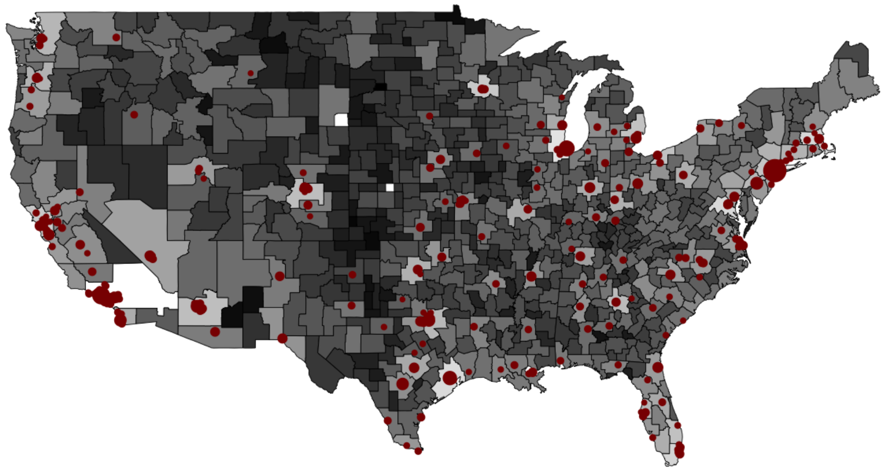

Figure 4: Average labor market concentration, 2016. Darker shades correspond to higher levels of concentration. Point size corresponds to city population.

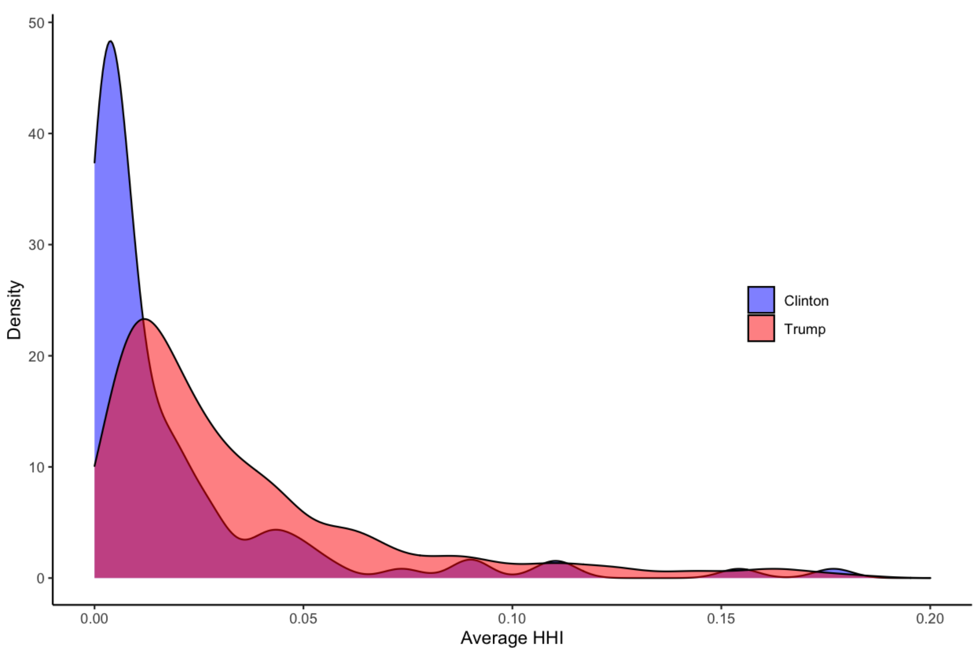

Figure 5: Density plot of HHI at CZ-level, grouped by 2016 election outcomes.

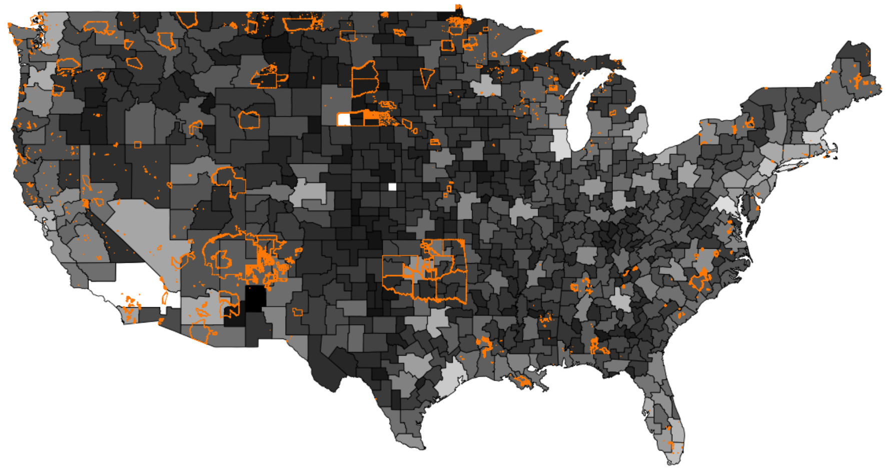

Figure 6: Average labor market concentration, 2016. Darker shades correspond to higher levels of concentration. Overlays display Native American tribal lands.

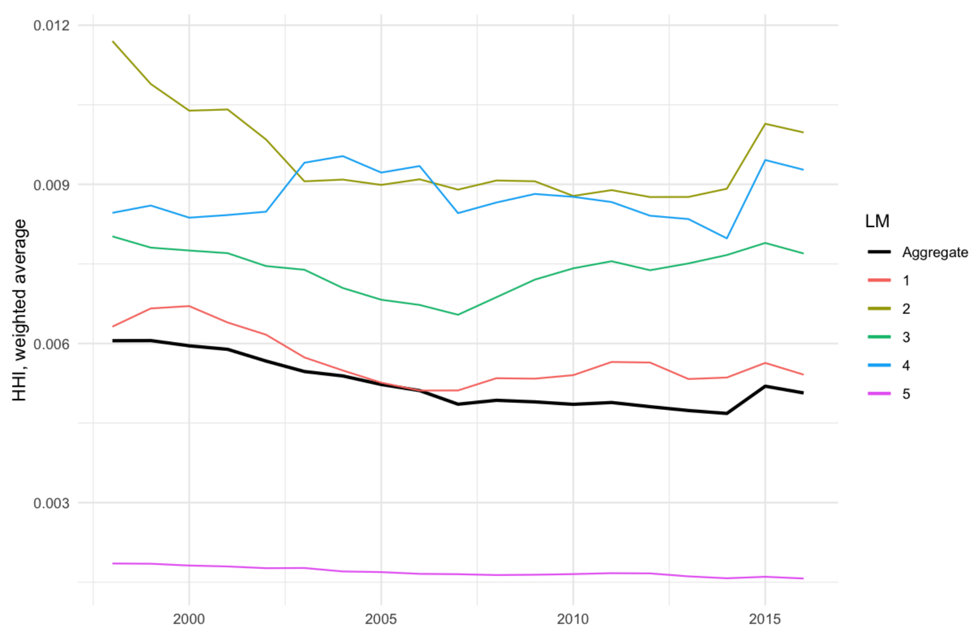

Figure 7: Weighted-average HHI across labor markets, 1998-2016.

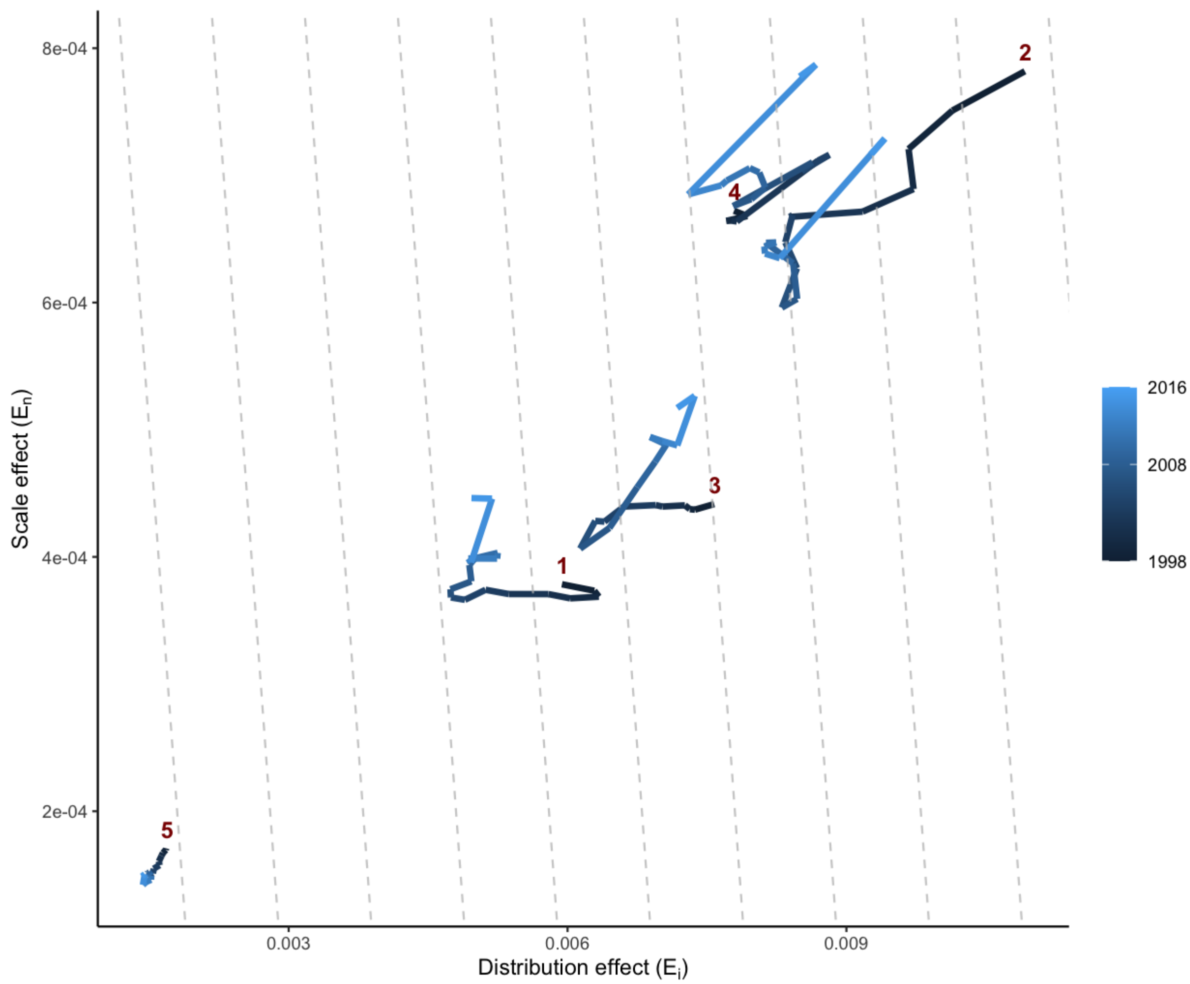

Figure 8: Decomposition paths, NMF2 labor market clusters. 1998-2016. Dashed lines depict HHI contour lines.

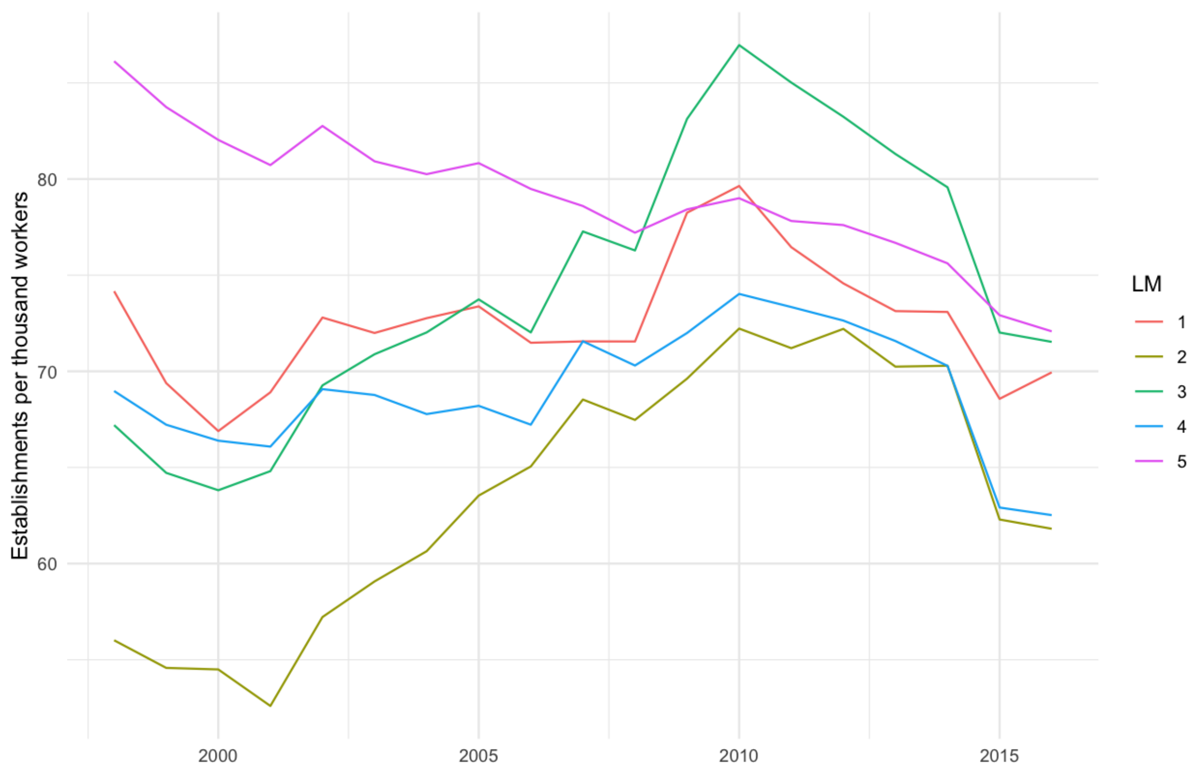

Figure 9: Number of establishments per one thousand workers, 1998-2016

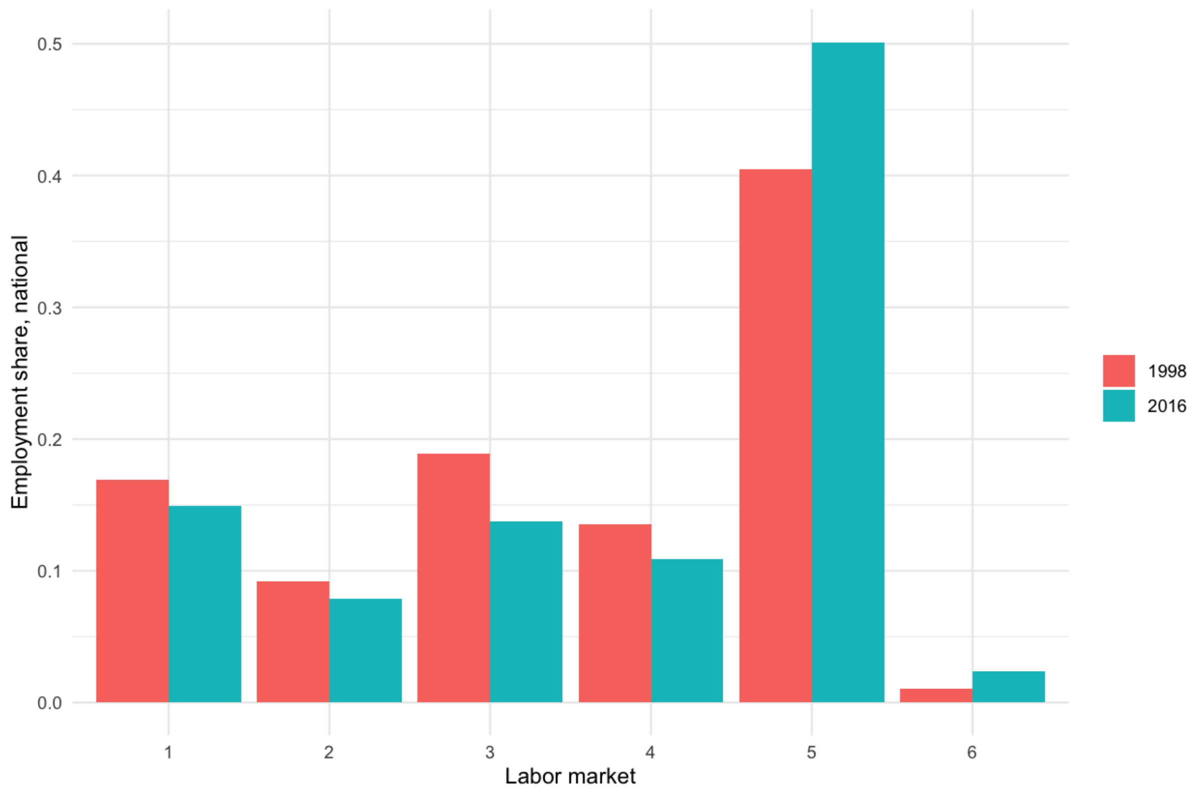

Figure 10: Labor market employment shares; national, 1998 vs. 2016