LABOR-MARKET CONCENTRATION AND LABOR COMPENSATION

YUE QIU AND AARON SOJOURNER*

This article estimates the effect of labor-market concentration on labor compensation across the US private sector since 2000. The authors distinguish between concentration in local labor markets and local product markets while guarding against bias from confounded product-market concentration. The analysis extends beyond wages to rates of employment-based health insurance coverage. Reported results suggest negative effects of labor-market concentration on labor compensation. These effects are exacerbated when product-market concentration is higher or when workers are older.

We study the effects of local labor-market concentration on wages in the US economy. Workers may be subject to employer market power through a lack of competition between employers, a form of monopsonistic competition that is one of many varieties of employer wage-setting power derived from upward-sloping labor supply curves facing firms (Manning 2011; Naidu, Posner, and Weyl 2018). Employers in more-concentrated labor markets, proxied here by a high Herfindahl-Hirschman Index (HHI) on employment shares in a market, may have the wage-setting power to mark down workers' wages below their marginal product, analogously to how sellers in monopolistically concentrated product markets may have power to mark up consumer prices above their marginal cost. Given that US workers' wages, below the top end, have stagnated for decades (Shambaugh, Nunn, Liu, and Nantz 2017), investigating potential avenues for increasing them is of first-order economic importance (Shambaugh and Nunn 2018).

*Yue Qiu is an Assistant Professor in the Fox School of Business, Temple University. Aaron Sojourner (https://orcid.org/0000-0001-6839-2512) is a Senior Researcher at the W.E. Upjohn Institute for Employment Research.

Thanks to José Azar, Sandy Black, Kyle Herkenhoff, Ioana Marinescu, Suresh Naidu, and Kevin Rinz, participants in LERA@ASSA for constructive discussions about this research, and Claire Black for editing assistance. An Online Appendix is available at http://journals.sagepub.com/doi/suppl/10.1177/00197939221138759. Data replication files are available at https://research.upjohn.org/jrnlarticles/224/. For information, please address correspondence to sojourner@upjohn.org.

Keywords: labor-market concentration, wages, health insurance, unions

ILR Review, 76(3), May 2023, pp. 475–503
DOI: 10.1177/00197939221138759.
Journal website: journals.sagepub.com/home/ilr

We define labor markets as the combination of an occupation and a commuting zone (CZ) and focus on the relationship between changes in concentration and changes in wages within labor markets using market fixed effects and market-specific time trends. We use pooled cross-sections of worker-level wage data from the 2000 Decennial Census and the American Community Survey (ACS) in each year from 2005 to 2017. Distinct from the prior literature, our contributions include controlling for important potential confounding factors, such as the product-market concentration of the worker's local industry and the worker's individual human capital characteristics. This approach is possible because, unlike the prior literature, we analyze worker-level data with each individual's wage, industry, occupation, and human-capital characteristics. We also expand the analysis of effects beyond wages to include employment-based health insurance.

We make three main contributions relative to the prior literature. First, we distinguish local labor-market concentration in an occupation-year from local product-market concentration in an industry-year to build more-credible estimates. Industries describe how firms face consumers; occupations describe how they face workers. We measure both and control for the latter when trying to measure the effect of the former. This distinction is conceptually critical because the two are easily confounded. For example, if a town has only two nursing homes and they are the only local employers of registered nurses, they will have power in both the product and the labor markets. Industry concentration may generate economic rents for firms from consumers, which might provide a basis for rent-sharing with employees and higher wages. Occupational concentration may generate economic rents for firms from workers by suppressing wages. Because product-market and labor-market concentration may move together, the absence of product-market concentration from the prior literature's analysis creates a risk of omitted-variable bias. By distinguishing labor-market concentration from product-market concentration, our approach guards against this risk. To achieve this, we use a novel method to measure firm-level occupational employment in a local labor market.

Second, beyond examining the effect on workers' wages as the prior literature has, we also look at the effect of labor-market concentration on workers' probability of employment-based health insurance coverage, a substantial component of labor compensation. In recent data on US private-sector workers, the cost of employee health insurance to employers represents approximately 11% of wage and salary costs.[1]

Third, we study how the relationship between labor-market concentration and wage differs depending on the degree of product-market concentration, worker unionization, workers' ages, and occupational offshorability. Our novel ability to measure both labor- and product-market concentration enables this first test. Absent product-market concentration, the firm

[1]The statistic is from https://www.bls.gov/news.release/ecec.nr0.htm.

would have less economic rents to share with workers, whether the labor market is concentrated or not. With product-market concentration, the question becomes, does labor-market concentration prevent workers from receiving a share of it. We also test whether the presence of strong worker organizations countervails the negative effects of labor-market concentration on labor compensation. Benmelech, Bergman, and Kim (2022) used national-level unionization rates in the handful of industries they studied and found that concentration had more-negative effects of pay in less-unionized industries. We generalize this finding across the entire labor market and exploit variation in occupational unionization rates across state-year. Next, we examine whether the estimated effects are stronger for older workers as they have lower job mobility. Finally, we test the hypothesis that predicts weaker negative effects of concentration on wages in occupations that are more likely to be shifted offshore (i.e., where localness matters less).

Data and Methods

Several recent papers have studied trends in local labor-market concentration and its effects on wages, using various labor-market definitions and identification strategies. Such work finds evidence of labor-market concentration based on firms' shares of vacancy postings in an occupation-locale during 2010 to 2013 (Azar, Marinescu, and Steinbaum 2022) and in 2016 (Azar, Marinescu, Steinbaum, and Taska 2020). The former, studying vacancies that included a posted wage, presented evidence that greater concentration caused lower wages. The latter focused on how to define local labor markets and offered a cross-sectional description of concentration across markets but did not analyze its relation to wages.

Four papers leveraged the US Census Bureau's Longitudinal Business Database (LBD) to measure labor-market concentration at the industry-locale-year level using employment shares rather than vacancy shares. Benmelech et al. (2022) focused on a handful of industries within manufacturing and defined a labor market at the industry-county or industry-CZ level. They focused on manufacturing because Census provides establishment-level measures of labor productivity for that segment. This variable is important in explaining establishment-year average wage. Benmelech and colleagues found evidence that greater concentration causes slightly lower wages. Also, this effect is weaker in industries with higher national unionization rates. Rinz (2022) expanded this approach to the entire US economy from 1976 to 2015, again focusing on concentration within an industry-locale and using CZ as the measure of locale. He produced the first evidence of trends in local labor-market concentration broadly and found evidence of effects on individual wages. Rinz observed a positive effect, contrary to expectations, in two-way fixed-effect models without instruments. The sign reverses to a negative effect, consistent with increasing concentration lowering wages, when he used an instrumental variable based on contemporaneous changes in the structure of the same kind of labor market across other locales. Lipsius (2018) also analyzed the LBD, defining labor markets along local industry lines but using LBD's establishment average wage aggregated up to the local firm level rather than linking to individual worker-level wages. He interpreted the evidence in the context of a more well-developed theoretical model, which highlights the importance of controlling for labor-market size and labor productivity. Berger, Herkenhoff, and Mongey (2022) used a similar empirical strategy and added value by interpreting estimates in the context of a general-equilibrium model to assess the welfare implications of changes in labor-market concentration and of minimum-wage policy changes. Because product-market and labor-market concentration may move together, the absence of product-market concentration from the prior literature's analysis creates a risk of omitted-variable bias. In this regard, the article that is most similar to ours is Prager and Schmitt (2021). They focused on the US hospital industry specifically and used changes in average wages within a few occupational groups among hospitals that experienced a merger. The authors leveraged hospital-year-occupational group data on employment and wages along with data on mergers and found negative effects on wage growth. Relative to Prager and Schmitt (2021), we generalize across all industries and occupations and use different variation in concentration.

Our primary sample comes from the ACS between the years 2005 and 2017 supplemented with the 5% public use subsample of the 2000 Decennial Census, drawn from IPUMS-USA (Ruggles et al. 2018). Let t index years. The virtue of this sample relative to other data used in this literature is its measures of worker wage, industry, occupation, and locale at the individual level. This feature enables us to separately measure the degree of labor-market concentration in the workers' occupation-locale and the degree of consumer-market concentration in their industry-locale. Other articles in this line of research have not included both industry and occupation. Instead, they have focused on one or the other, interpreted the focal variable as the labor-market boundary, and left the concentration of the other in the residual (Azar et al. 2020, 2022; Benmelech et al. 2022; Berger et al. 2022; Rinz 2022). The one exception is Prager and Schmitt (2021), who focused on a select set of occupations within one industry. Additionally, our microdata include measures of each worker's human-capital characteristics, which provide additional insight into mechanisms. Let i index workers in the sample.

Labor Market

We define a labor market as the combination of an occupation and a CZ. Markets are indexed by m, with $o(m)$ and $l(m)$ denoting a market's occupation and locale, respectively. Conceptually, occupation is superior to industry as the basis for defining a labor market. Occupation is an aspect of a job and is rooted in the knowledge, skills, and abilities that workers and firms trade in the labor market. Industry is an aspect of a product and is rooted in the characteristics that consumers and firms trade in goods and service markets. We use the 3-digit 1990 Census occupational codes and examine 334 occupations.[2]

To define locales, we use the 1990 definition of CZs and use all years of data for which it is available in IPUMS. IPUMS does not have information on CZ but has information on the public use microdata area (PUMA) of an individual's residence. To map each PUMA to a CZ, we use the crosswalk from Autor and Dorn (2013).[3] This step yields 224,942 distinct labor markets observed in 14 years, for a total of 1,955,643 possible market-years observed.

Labor Compensation

The primary outcome measure (*Y*) is an employee's log hourly wage measured in 1999 dollars. We include individuals between age 16 and 64 who work in for-profit firms in the private sector. We drop those associated with institution group quarters and those with missing wage, 1990 Census industry, 1990 CZ, or 3-digit 1990 Census occupational codes. This leaves 23,978,672 observations of workers. Hourly wages average $15.30, with a standard deviation of $21.11 and a median of $11.15 (see Table 1).[4]

As a supplemental measure of labor compensation, we also study whether a worker reports having employment-based health insurance and how this relates to labor-market concentration. This factor is observed in the ACS from 2008 forward. In our sample, approximately 90% of workers have employment-based insurance.[5] In addition to viewing health insurance as an alternative form of compensation, economists have considered it as a source of job lock, which diminishes the value of outside options and mobility for workers with employment-based insurance (Gruber and Madrian 1994). To the extent that this condition motivates employers to offer it,

[2]Occupational categories provide imperfect proxies for market boundaries, and the proper level of aggregation is not clear. Our analysis focuses on the Census's preferred occupational boundaries. We explore an alternative aggregation, however, replicating the analysis using the Census's more-aggregated occupational groups to define boundaries. The results are sensitive to this choice and estimated labor-market concentration effects become smaller, mixed in signs between wage and health insurance coverage outcomes, and not statistically significant at the 10% level.

[3]The crosswalk is available at https://www.ddorn.net/data.htm.

[4]To measure hourly wage, we divide annual earnings by 52 times the reported usual hours per week. Weeks of work are not reported in all years so we do not use it. Results using direct data on annual, rather than constructed hourly, earnings are very similar. In the prior literature, only Rinz used individual-level wage data and he used annual earnings, lacking any measures of time worked. Others used posted wages on vacancies or establishment average wage or earnings. Note: 2017 $1.471 = 1999 $1.00.

[5]If a person is covered by their own or another family member's current employer, former employer, or union, then this person is coded as covered by employment-based health insurance. See https://usa.ipums.org/usa-action/variables/HINSEMP#description_section.

Table 1. Summary Statistics

	Mean	*Std. dev.*	*P10*	*Median*	*P90*	R^2
Hourly wage	15.301	21.111	3.565	11.153	29.385	0.189
Employment HHI	0.066	0.113	0.005	0.027	0.161	0.803
Sales HHI	0.315	0.313	0.029	0.190	1.000	0.817
Sales HHI missing	0.198	0.398	0.000	0.000	1.000	0.819
Labor productivity ($ 000)	201.861	169.397	127.451	184.164	289.426	0.462
Fraction of missing estab labor productivity	0.202	0.223	0.013	0.122	0.496	0.927
Employment in Data Axle (000)	1.000	4.598	0.007	0.126	1.823	0.986
Age	38.557	12.693	22.000	38.000	56.000	0.125
Male	0.551	0.497	0.000	1.000	1.000	0.344
Black	0.109	0.312	0.000	0.000	1.000	0.157
Other race	0.140	0.347	0.000	0.000	1.000	0.148
Married	0.502	0.500	0.000	1.000	1.000	0.095
Hispanic	0.171	0.376	0.000	0.000	1.000	0.300
US-born	0.799	0.401	0.000	1.000	1.000	0.226
Full-time job	0.738	0.440	0.000	1.000	1.000	0.231
Bachelor degree	0.257	0.437	0.000	0.000	1.000	0.359
Instrumental variable	−7.480	0.814	−8.208	−7.712	−6.544	0.881
Health insurance through employers/unions	0.703	0.457	0.000	1.000	1.000	0.171
Unionization rate	0.072	0.072	0.006	0.049	0.187	
Offshorability	0.009	1.305	−1.745	0.011	1.680	

Notes: This table reports the summary statistics of variables used in the estimations. A labor market is defined as the interaction between a 3-digit 1990 Census occupation (OCC) and a 1990 commuting zone (CZ). A product market is defined as the interaction between an industry, which is derived from 3-digit 1990 Census industry, and a CZ. For employment HHI, the instrumental variable, and total employment, they are defined at the labor-market-year level. For sales HHI and variables related to labor productivity, they are defined at the product-market-year level. For each CZ-OCC-year cell, the *Instrumental variable* for employment HHI is the average of the natural logarithm of 1 over the number of firms in the same occupation but in other CZs in that year. The sample includes 23,978,672 individuals, 1,955,643 CZ-OCC-year, and 1,459,297 CZ-industry-year observations. The bachelor degree dummy and US-born dummy are missing for 0.7% and 41% of individuals, respectively. Labor productivity is missing for 0.5% of CZ-occupation-year observations. *Health insurance through employers/unions* is available since 2008 in the American Community Survey and we have data for 14,403,816 individuals. *Unionization rate* is the 5-year average unionization rate in a major occupation group-state cell centered around a year in the Current Population Survey. It is available for 4,284 OCC-state-year observations. *Offshorability* measures the extent to which the tasks performed by occupations are offshorable; the data are from David Dorn's webpage (https://www.ddorn.net/data.htm) and available for 322 occupations derived from the 3-digit 1990 Census occupation codes. The last column reports the R^2 from regressing each variable on labor-/product-market and year fixed effects at the individual level. The hourly wage and labor productivity are in year 1999 dollars. HHI, Herfindahl-Hirschman Index.

employment concentration may serve as a strategic substitute for employer-based health insurance. Greater concentration would make its offer less strategically compelling, thereby reducing insurance coverage.

Labor-market Concentration

For each labor market m each year t, we measure concentration by combining data on each establishment's employment level, industry, parent firm, and location. Our source is Data Axle (DA), which provides information on the joint distribution of industry and employment locally for the entire country over time.

DA's data on each establishment's location, industry, firm, ultimate parent firm, and total employment level (but not workers' occupations) have the same structural features as the Census LBD used by Lipsius (2018), Benmelech et al. (2022), Berger et al. (2022), and Rinz (2022) to measure concentration. The DA data are not produced by required official reporting, therefore employment and revenue measures are sometimes imputed by DA or are missing. However, the database has been produced and sold for decades to support business-to-business marketing and analysis, and many researchers have used these data. For each establishment, DA provides information on establishment name, firm name, establishment ID, location (street address, city, county, and state), 4-digit 1987 version Standard Industrial Classification (SIC) industry code (which we index by d), DA ultimate parent ID, employment, and sales.[6] To match each SIC code to a 3-digit detailed Census industry code, we use the crosswalk from the US Census Bureau.[7]

To go from establishment's industry and employment to an estimate of establishment's employment by occupation, we multiply each establishment's employment level by the national occupational distribution of employment for the establishment's industry. We estimate a distribution of occupational employment by industry each year ($Pr(o)_{dt}$) from the Census microdata's joint distribution of occupation among US workers in industry d in year t. For an establishment e in industry d in year t employing E_{dte} workers, its number of employees in occupation o is measured as $E_{ote} \equiv E_{dte} Pr(o)_{dt}$.

Multiple establishments within the same parent firm by locale combination are considered as a single employer. Each firm's annual employment in a market is the sum of its establishments' employment levels: $E_{mtf} \equiv \sum_{e \in f} E_{mte}$. Letting N_{mt} be the number of firms employing workers in occupation $o(m)$ in year t, each firm's employment share is $s_{mtf} \equiv E_{mtf} / E_{mt}$ where E_{mt} is total market employment, the sum of E_{mtf} across firms. A positive employment level is observed in 1,955,643 market-years.

---

[6] The establishment-level data in Data Axle are verified by the company's phone verification process every year, and the data team takes great efforts to create longitudinal links across establishments over time. The establishment-level sales data are estimated based on statistical models performed by Data Axle.

[7] The crosswalk is available at "CPS Industry Classifications (1992–2002)" at http://unionstats.com/. For each detailed Census industry code, this crosswalk provides the equivalent SIC codes, mostly at the 3-digit level. For a few cases, we have to construct the crosswalk such that a group of the Census industry codes is uniquely mapped to a group of SIC codes and vice versa. Specifically, we aggregate the Census codes 272 and 280 as code "272,280"; 371 and 372 as code "371,372"; 771 and 790 as code "771,790"; and 862 and 863 as code "862,863."

Labor-market concentration is measured by an employment Herfindahl-Hirschman Index (EHHI) based on firms' employment shares:

$$EHHI_{mt} = \sum_{f=1}^{N_{mt}} s_{mtf}^2.$$

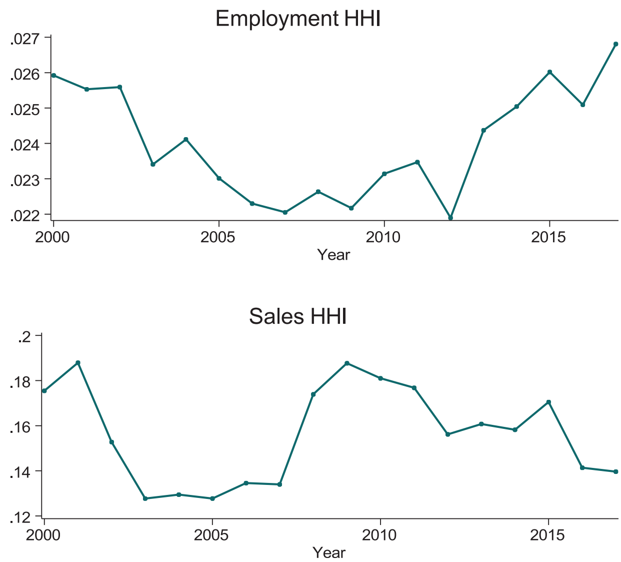
Firm mergers and acquisitions will tend to lift concentration in labor markets in which both firms operate; death and exit of small employers and growth of large employers also tend to lift concentration. Our measure of concentration very likely underestimates true concentration. If a firm in a metro has only 1 employee but it is in an industry that nationally has a positive probability of employing people in 100 occupations, the firm is measured as having a fraction of an employee in each of those 100 occupational labor-markets in that metro. This calculation mechanically forces a high number of employers into each labor market but most will have very small shares, thereby pushing down EHHI. For this reason, we recommend interpreting neither our estimate of EHHI levels nor the number of employers literally. In our analysis of the effects of concentration, we use market fixed effects and focus on how changes in log(*EHHI*) predict changes in log(*Wage*). The essential question is whether changes in our EHHI measure capture changes in true labor-market concentration. Conceptually, it should.

Our study provides the first economy-wide estimate of labor-market concentration based on employment shares in occupationally defined labor markets in the recent literature.[8] Across market-years, the measured average EHHI is 0.066 (or 660 of 10,000) with a standard deviation of 0.113 and a median of 0.027, which is consistent with a skew toward higher concentration (see Table 1).

It is useful to compare our measures to others derived from the LBD. Figure 1, top panel, shows the trend in average of EHHI between years 2000 and 2017. Rinz (2022) computed a measure of EHHI (in his figure 2) but used a different measure of labor market, industry by CZ instead of occupation by CZ. Despite these differences and our levels being consistently lower by a factor of approximately four-fifths, changes in our EHHI measures follow a similar path. Both Rinz's and our estimates fall steadily from 2000 until the start of the Great Recession in 2007 and rise abruptly during the Recession, though some divergence occurs in the last couple of

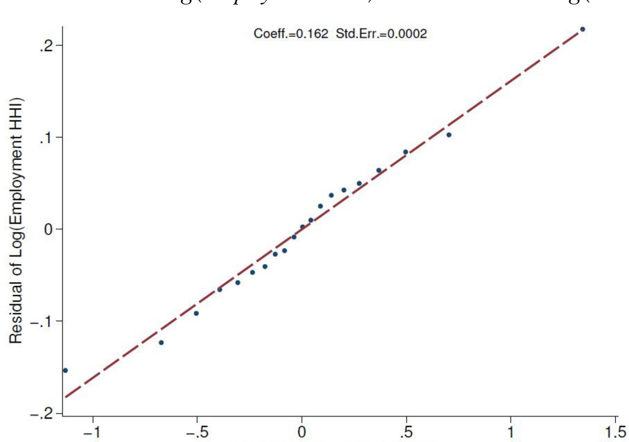
[8]Other studies have been based on vacancy shares, not employment shares (Azar et al. 2020, 2022), or labor markets defined along industrial lines (Lipsius 2018; Berger et al. 2022; Rinz 2022), not occupational lines. A potential downside of defining labor-market boundaries according to self-reported occupation rather than self-reported industry is evidence that occupation exhibits greater classification error than industry (Mellow and Sider 1983; Bound, Brown, and Mathiowetz 2001). This classification error of occupation would tend to introduce error into our measure of labor-market concentration and, if classical, to create attenuation bias in ordinary least squares (OLS) analysis. Other articles in this literature have used only industry or occupation. By using both of them in theoretically motivated and distinct ways, we aim to avoid confounding them.

Figure 1. Employment and Sales HHI Trends



Notes: This figure reports the employment and sales HHI trends between years 2000 and 2017. The statistics are calculated using all firms in the Data Axle database. We calculate average employment HHI across labor markets using employment-weights and average sales HHI across product markets each year using sales-weights. A labor market is defined as the interaction between an occupation (1990 Census definition at the 3-digit level) and a commuting zone (1990 definition). A product market is defined as the interaction between an industry (based on 3-digit 1990 Census definition) and a commuting zone (1990 definition). HHI, Herfindahl-Hirschman Index.

Figure 2. Residual of Log(*Employment HHI*) on Residual of Log(*Sales HHI*)



Notes: This figure reports the relation between the residualized log(*Employment HHI*) and the residualized log(*Sales HHI*). The residualized log(*Employment HHI*) is computed at the individual level and is conditional on the labor-market and year fixed effects. The residualized log(*Sales HHI*) is computed at the individual level and is conditional on the product-market and year fixed effects. HHI, Herfindahl-Hirschman Index.

years. As shown in the bottom panel of Figure 1, sales HHI (SHHI) follows a different trend, presenting evidence of independent variation.

As shown in the last column of Table 1, 80% of variation in labor-market concentration across individual workers is absorbed by a set of market fixed effects and year fixed effects. We exploit the 20% of variation that remains, which represents changes in labor-market concentration within labor market over years driven by local establishment entry, exit, employment growth or decline, firm merger or spin off, and changes in occupational shares within industry over time.

We use other important, time-varying influences on average wages that may be correlated with labor-market concentration as control variables.

Product-market Concentration

A key contribution of our article beyond the prior literature is that we distinguish product-market concentration from labor-market concentration to separately measure both and to estimate the relationship of each to wages conditional on the other. Each worker in our Census data is observed in an occupation and an industry. Within a CZ, workers in the same labor market (occupation-CZ) can be in different product markets (industry-CZ).

We measure each industry-CZ-year's product-market concentration with HHI based on firms' sales shares in the industry-CZ-year constructed from the DA establishment data on location, industry, and annual sales, aggregated in a similar way to that described above for labor-market concentration.

Both our labor-market and product-market HHI measures derive from the CZ-industry-year firm shares. The product-market HHI uses firms' local shares of sales (SHHI). The labor-market HHI uses firms' shares of occupational employment after projecting establishment employment into occupational employment. To illustrate how the EHHI and SHHI measures work, we present a simple example in the Online Appendix.

Workers are in local industries with an average SHHI of 0.315 (3,150 out of 10,000). The median SHHI is 0.190, consistent with a skew toward high concentration, and the standard deviation is 0.313.[9]

On one hand, EHHI and SHHI are positively correlated, creating a risk of omitted variable bias when EHHI is analyzed without controlling for SHHI. On the other hand, are they too highly correlated to be sensibly separated? To understand the covariation in these variables that will be relevant in the regression analysis, we regress logs of each at the worker level on labor-market or product-market fixed effects and year fixed effects, create residuals of each, and study their relationship. The bin scatter in Figure 2 shows that, as we expected, a positive association exists. Both EHHI and SHHI depend on the number and sizes of local firms. However, because they focus on distinct markets (labor versus product or, equivalently,

[9]In our sample, SHHI is missing for approximately 20% of observations at the product-market–year level. We impute these as equal to 0 and include an indicator for missing.

occupation versus industry) a lot of independent variation remains. Conditional on labor or product market and year, residuals of log(*EHHI*) and log(*SHHI*) have a correlation of only 0.16 across workers.

Other Market Controls

Changes in wages may also be related to changes in employment levels. The sign depends on whether the changes are attributable to supply or demand shocks. We are not focused on this factor but construct and control for a time-varying measure of log total employment by market-year using our occupational employment projections. The average market has one thousand employees with a median of 126 (Table 1, row 7). Lipsius (2018) also conditioned on employment levels, arguing that it is a proxy for labor productivity.

Other studies in this literature, with the exception of Benmelech et al. (2022), do not include any direct measure of labor productivity. Benmelech and colleagues focused on manufacturing, primarily because a measure of establishment-year labor productivity is available in this sector, which they linked to establishment-year average wage. This approach is a very nice feature of their study, as labor productivity may lift wages. Though we cannot do as well as this because we focus outside of manufacturing, we can do better than most who have looked broadly. We begin by measuring labor productivity as the ratio of sales to employees at the establishment-year level. We do not know which establishments employ the workers whose wages we observe, however, so we aggregate establishment-year labor productivity to the product-market–year level and use this as a control in some specifications tied to each worker's local industry, that is, product market. Labor productivity averages $200,861 per employee with a standard deviation of $169,397.[10]

Including market-year labor productivity in the analysis has benefits and costs. Omitting productivity may make OLS estimates misleading, reflecting the influence of time-varying unobservable factors that drive both changes in wages and changes in labor-market concentration.

For instance, a positive productivity shock to a local firm could cause quick employment growth and increase concentration, product-market rents, and wages if workers obtain a constant share of rents, yielding a positive bias. Further, compositional changes in the set of employing firms across the business cycle could create confounds. For example, when lower-productivity, lower-wage firms exit during an economic contraction, concentration would increase and average wages of those still employed might rise, representing a similar kind of positive bias. Controlling for local labor-market productivity can help reduce these kinds of omitted-variable bias but it has a cost. If rising concentration itself affects labor productivity, then

[10]We winsorize labor productivity to the 1st and 99th percentiles across all establishments in each year before aggregating to the product-market–year level.

including labor productivity may over-control by including an intermediate outcome or mechanism as a predictor, causing attenuation of the estimated effect of concentration on wages. For instance, if higher concentration itself raises productivity and wages by allowing greater economies of scale in production and technology use, as is commonly asserted in merger proposals, then part of concentration's effect on wage will be absorbed into the estimated productivity effect. This context leads us to conduct and discuss robustness analysis related to the issue.

Specifications

The logic of a Cournot model of oligopsonistic competition between employers motivates use of a common, straightforward relationship between wages and labor-market concentration (Azar et al. 2022; Rinz 2022). We estimate a regression of log wage on log employment HHI at the individual worker level:

$$Y_{mti} = \beta \log(EHHI_{mt}) + \alpha X_{mti} + \gamma_t + (\gamma_{0m} + \gamma_{1mt}) + \varepsilon_{mti} \tag{1}$$

where Y_{mti} is the natural log of the individual worker's real hourly wage and salary income or a dummy variable indicating whether a worker is covered by employer-sponsored health insurance, $\log(EHHI_{mt})$ is the natural logarithm of the labor market's employment concentration in that worker's occupation-CZ-year, and X_{mti} contains various observable characteristics of the worker and the market-year. All models also include year fixed effects, labor-market-specific fixed effects, and market-specific linear time trends. The term ε_{mti} is the idiosyncratic residual. Various additional fixed effects that further partition ε will be introduced as results are discussed. In particular, specifications add CZ by year fixed effects to control for unobserved common shocks to the local economy affecting wages in all local occupations similarly. Adding industry by year fixed effects controls for unobserved wage determinants that affect all workers in an industry nationally each year.

Assuming that changes in labor-market concentration are mean independent of changes in average unobserved influences on wages (ε) conditional on X identifies the parameters in OLS models. All estimations are weighted by the worker's personal weights (variable "perwt" in the IPUMS data). Standard errors are clustered at the labor-market level.

Making comparisons across occupations controlling for industry helps insulate against the influence of demand and supply shocks that operate at the product level. For example, if demand for oil falls, this will reduce demand for operating engineers in CZs with oil wells and refineries and might spur firm exit and increase labor-market concentration in these locales, but not in locales without wells or refineries. Our design helps answer whether wages fall more in locale-years where operating engineer concentration increases more and controls for if the operating engineer works in the oil industry rather than construction. In terms of internal validity, this improves over designs that also include observations across multiple industries and multiple occupations but do not measure both and, therefore, confound the two. In terms of external validity, it improves over designs that focus only on a single industry.

In our study and others, idiosyncratic local demand or supply shocks by occupation could create systematic confounds in OLS analysis. For instance, negative local labor demand shocks might reduce employment and the number of employers, increasing concentration, and also reduce wages. In OLS analysis, this outcome will present as a negative "effect" of concentration on wages but it would be a spurious association through the demand shock, $\varepsilon_{mti} < 0$ with $Corr(EHHI_{mti}, \varepsilon_{mti}) < 0$ violating the standard exogeneity condition for unbiased OLS. This possibility of spurious associations between local market changes in EHHI and wages through shocks to local demand is a primary threat to unbiased estimation. Unobserved shocks to local labor supply might tend to generate countervailing bias. For instance, unobserved changes in supply would tend to move both concentration and wages in the opposite direction, generating a spurious positive association between them. These concerns motivate the use of an instrumental variable (IV) identification design. The goal is to move away from a focus on contemporaneous covariation in local concentration and wages, which is especially prone to these kinds of biases, and shift to a focus on variation in local labor-market concentration driven by national-level changes in the occupation, thereby insulating our observations from idiosyncratic local supply and demand shocks to the local wage outcome.

We instrument local labor-market concentration in each year and market (occupation by CZ) with a function of the number of firms employing workers in that occupation in all other CZs in that year. The instrument is designed to focus on variation in local concentration deriving from changes in the fundamentals of the occupation, apart from idiosyncratic changes in local market concentration that could be confounded with idiosyncratic unobserved influences on local wages such as shocks to local supply or demand. Azar et al. (2022) used this instrument. For a robustness check, we use a different but similar instrument, the average local labor-market concentration in the same year in other CZs for the same occupation, as in Rinz (2022). Our instrument also focuses on local variation in concentration driven by national shocks to occupational concentration, rather than using all local variation in concentration, some of which might be driven by local demand or supply shocks.

We focus on evidence from IV identification, following strategies in the prior literature (Azar et al. 2022; Rinz 2022). Our primary instrument for labor-market concentration in each market-year $\log(EHHI_{mt})$ is the average of the natural log of the reciprocal of the number of firms in that same occupation-year in all other CZs (Azar et al. 2022). The instrument averages

-7.48 with a median of -7.71 and standard deviation of 0.81.[11] In a robustness check, instead of the leave-this-market-out mean reciprocal of the number of employing firms, we use an alternative instrument from Rinz (2022), the leave-this-market-out mean labor-market concentration. That is, we instrument focal market concentration with average concentration of the same occupation in other CZs that year. Azar et al. (2022) pointed out that the former instrument is less likely to be endogenous, as the number of employing firms does not include information on firms' specific market shares.

With both instruments, the first stage will focus on local variation in EHHI correlated with these national changes and exclude local variation tied to idiosyncratic local supply or demand shocks that could create omitted-variable bias as discussed above. Identification based on changes in local labor-market concentration driven by occupational changes that are not specific to the local market should be more insulated from this threat. Studies commonly use these kinds of leave-this-market-out instrument to deal with endogeneity of local prices (Nevo 2001).

When will this kind of instrumental variables design provide unbiased estimates? The maintained exclusion restriction is that changes in the instrument (leave-this-market-out national mean reciprocal employer counts or HHI) affect the outcome mean (local labor-market wages) only through shifting the treatment variable (local labor-market concentration) conditional on other observable variables, rather than through any correlated unobservable. Exclusion restrictions are difficult to justify for many potential instruments that drive changes in local labor-market concentration because they are plausibly correlated with other determinants of local supply or demand shocks. Moving the focus away from the local market builds insulation to local wage determinants.

Shocks for which we would expect our design to deliver unbiased estimates include those to industry demand, supply, or productivity as discussed above. Our ability to control for industry apart from occupation and for local product-market concentration adds credibility beyond the prior IV literature, reducing the risk of some exclusion-restriction violations. Further, suppose adoption of a new technology raises the relative productivity of a particular occupation. Then, relative to other occupations and to prior years in the same occupation, the number of firms employing workers in that occupation will tend to increase in many industries and in many local labor markets. The instrument's value will tend to fall in every market. Changes in the instrument express these kinds of broad occupation-specific shifts in the propensity of employers using that occupation to enter, exit, grow, shrink, merge, or split, driving changes in concentration. Our ability

[11]This level implies high numbers of employers per market, an artifact of our strategy for measuring occupational employment. The essential element is that it captures meaningful variation across time and markets.

to control for changes in local labor productivity helps pull out any direct effect of productivity on wages, allowing the estimate to focus on residual covariation in wages and concentration.

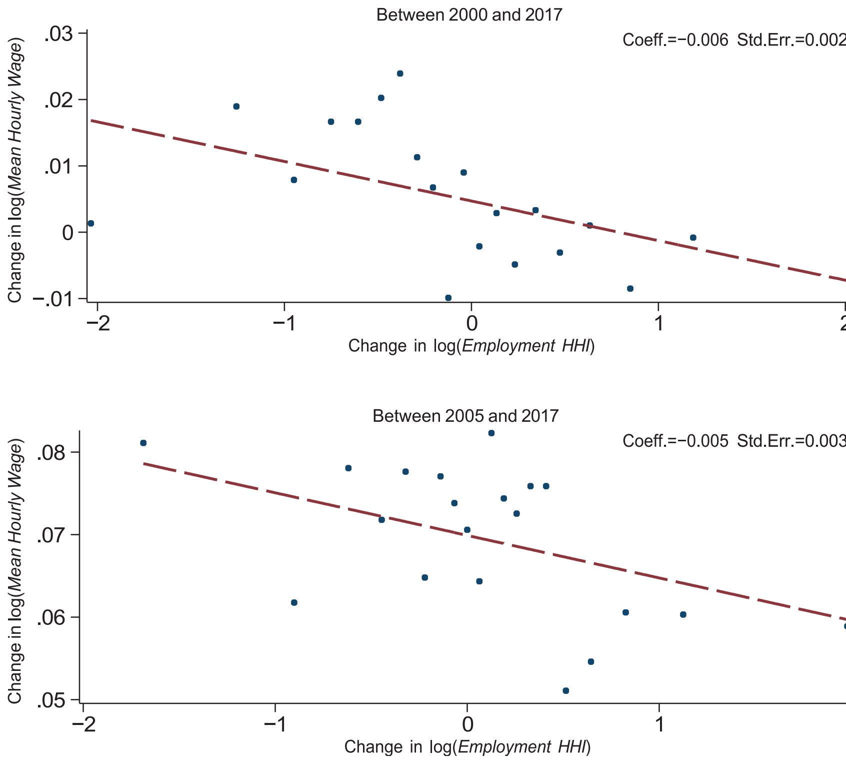
For transparency, we also highlight real possibilities that would violate the maintained exclusion restriction and create bias in the IV estimates. In particular, consider an unobservable, positive, national demand shock to consumer tastes for the services of one occupation (not for an industry's product) in one year. Wages and employment would increase for that occupation in most locales. If this led to smaller employers growing employment faster than larger ones, new employers entering, and existing employers exiting less, labor-market concentration would fall for that occupation in most locales. The first-stage of the instrument would work—the leave-this-market-out mean reciprocal for employing firms decreases along with each local market's concentration. However, the exclusion restriction would be violated, creating bias. The unobservable positive demand shock for the occupation drives up wage independently of concentration and is confounded for the effect of falling concentration. A national, occupation-year–specific supply shock would generate countervailing bias, which would look like a positive effect of concentration on wages.

Results

We begin by presenting simple, bivariate evidence using long first-differences to understand the relationship between changes in labor-market concentration and wages within labor markets over time. For each labor market, we compute the change in average log(*Wage*) between the first year in our data, 2000, and the last year, 2017, compared to the same kind of change in log(*EHHI*). The top panel of Figure 3 displays a bin scatter of the result along with an estimated best-fit line. Labor markets in which concentration increases more tend to experience *smaller decreases* in real wages, with a negative estimated relationship. The bottom panel of our figure displays results from a parallel exercise with the base year 2005 and produces similar results. These results are similar to those presented in Rinz's figure 21, panel (d), and contrary to the expected sign. Increasing concentration could, theoretically, lead to higher wages if expectations of lower turnover led to increased investments by the workers and firms in the relationship and, for some reason, the firm shared its value with the workers despite the reduction in competition between employers (Acemoglu and Pischke 1998; Benson 2013).

The next section introduces the basic specifications we use and estimation results for wages under OLS. The following section discusses the analogous IV results. As in Rinz, the results differ substantially between OLS and IV. For subsequent outcomes, we focus on IV results and present OLS results in the Online Appendix.

Figure 3. Change in Log(*Mean Hourly Wage*) on Change in Log(*Employment HHI*)



Notes: This figure reports the relationship between changes in concentration and wages within labor markets over time using long first-differences. In the top panel, for each labor market, we compute the change in average log(*Hourly wage*) and in log(*Employment HHI*) between 2000 and 2017. In the bottom panel, we use the year 2005 as the base year and perform a parallel exercise as in the top panel. HHI, Herfindahl-Hirschman Index.

Ordinary Least Squares

To begin, we look only at changes within market over time without controlling for other observables. Point estimates suggest that increases in labor-market concentration have no effect on wage changes, as expressed in the −0.001 point estimate with a 95% confidence interval (CI) of (−0.003, 0.001). This specification includes only year fixed effects, labor-market (occupation-CZ) fixed effects, and labor-market-specific linear time trends (column (1), panel A of Table 2). This finding is robust to allowing for very flexible sets of fixed effects. Allowing CZ-specific annual wage shocks by replacing the year fixed effects with CZ-year fixed effects does not change the estimate (column (2), panel A of Table 2). Adding the possibility of national occupation-year specific wage shocks shifts the estimate very slightly to 0.001 with a CI of (−0.001, 0.003) (column (3), panel A of Table 2). This specification is more flexible than the prior literature. Allowing the possibility of national industry-year specific wage shocks in addition to

Table 2. Effect of Labor-market Concentration on Hourly Wage Using OLS

	Panel A: Only fixed effects			
	(1)	*(2)*	*(3)*	*(4)*
Log(*Employment HHI*)	−0.001	−0.001	0.001	0.001
	[0.001]	[0.001]	[0.001]	[0.001]
CZ × OCC FE	Yes	Yes	Yes	Yes
Year FE	Yes			
CZ × Year FE		Yes	Yes	Yes
OCC × Year FE			Yes	Yes
Ind × Year FE				Yes
Labor-market trends	Yes	Yes	Yes	Yes
Adj. R^2	0.337	0.337	0.338	0.368
	Panel B: Add market-level controls			
	(1)	*(2)*	*(3)*	*(4)*
Log(*Employment HHI*)	−0.001	0.001	−0.000	−0.000
	[0.001]	[0.001]	[0.001]	[0.001]
Log(*Labor productivity*)	0.003	0.006**	0.011***	0.011***
	[0.003]	[0.003]	[0.004]	[0.004]
Missing labor productivity	0.012	0.027*	0.018	0.018
	[0.015]	[0.016]	[0.019]	[0.019]
Fraction of missing estab labor productivity	−0.026***	−0.022***	−0.004	−0.004
	[0.007]	[0.008]	[0.009]	[0.009]
Log(*Total employment,* Data Axle)	0.001	−0.015***	0.012***	0.010***
	[0.003]	[0.003]	[0.004]	[0.004]
CZ × OCC FE	Yes	Yes	Yes	Yes
Year FE	Yes			
CZ × Year FE		Yes	Yes	Yes
OCC × Year FE			Yes	Yes
Ind × Year FE				Yes
Labor-market trends	Yes	Yes	Yes	Yes
Adj. R^2	0.337	0.337	0.338	0.368
	Panel C: Add sales-based HHI			
	(1)	*(2)*	*(3)*	*(4)*
Log(*Employment HHI*)	−0.009***	−0.006***	−0.007***	−0.003***
	[0.001]	[0.001]	[0.001]	[0.001]
Log(*Sales HHI*)	0.026***	0.026***	0.026***	0.010***
	[0.001]	[0.001]	[0.001]	[0.001]
Sales HHI missing	−0.075***	−0.076***	−0.076***	0.004
	[0.003]	[0.003]	[0.003]	[0.003]
Log(*Labor productivity*)	0.003	0.006**	0.009**	0.010**
	[0.003]	[0.003]	[0.004]	[0.004]
Missing labor productivity	0.045***	0.062***	0.082***	0.019
	[0.015]	[0.016]	[0.019]	[0.019]
Fraction of missing estab labor productivity	−0.019**	−0.017**	0.037***	0.011
	[0.008]	[0.008]	[0.009]	[0.009]
Log(*Total employment,* Data Axle)	0.004	−0.012***	0.009**	0.010***
	[0.003]	[0.003]	[0.004]	[0.004]
CZ × OCC FE	Yes	Yes	Yes	Yes
Year FE	Yes			
CZ × Year FE		Yes	Yes	Yes

(continued)

Table 2. Continued

	Panel C: Add salesased HHI			
	(1)	(2)	(3)	(4)
OCC × Year FE			Yes	Yes
Ind × Year FE				Yes
Labor-market trends	Yes	Yes	Yes	Yes
Adj. R^2	0.338	0.338	0.339	0.368

Notes: The dependent variable in all estimations is worker's log(*Hourly Wage*). All estimations are weighted by the worker's personal weight. The sample includes 23,978,672 individuals. Standard errors in brackets allow clustered errors at the CZ-occupation level. CZ, commuting zone; FE, fixed effects; HHI, Herfindahl-Hirschman Index; OCC, 3-digit 1990 Census occupation; OLS, ordinary least squares.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

CZ-year and occupation-year shocks does not affect the estimate (column (4), panel A of Table 2). One justification for including industry × year fixed effects is that local product-market concentration better captures product-market power within an industry than across industries. For example, technology differences across industries would make local product-market concentration less informative about the scope of competition. For this reason, we prefer specification 4, which makes comparisons between workers in the same industry. Each additional panel in Table 2 adds a set of observable control variables while maintaining the same structure of fixed effects across specifications.

Table 2, panel B, presents estimates after adding controls for time-varying labor-market observables beyond EHHI, specifically log(*average labor productivity*) in the labor market, an indicator for market-years in which all establishments are missing productivity, a market-specific measure of the share of establishments missing the productivity measure, and log(*total employment*) in the labor market to capture changes in workforce size. Adding these controls barely changes the estimated labor-market concentration results. The basic pattern of null or small positive effects is stable. Any attenuation it causes to the concentration coefficient from over-controlling appears to be minimal. Productivity itself has the positive sign we expect.

Table 2, panel C, adds each worker's local industry's product-market concentration measure, log(*SHHI*), as well as an indicator of when this is missing. The prior literature has omitted this factor but, given the positive correlation between labor- and product-market concentration (see Figure 2), a positive product-market concentration effect on wages might mask a negative labor-market concentration effect. Including product-market concentration in the model changes the labor-market concentration estimates in a predictable way. Compared to panel B, estimated coefficients on labor-market concentration become negative and statistically significant but remain substantially small in every column. Coefficients on product-market concentration are positive, consistent with firms sharing some product-market rents with their workers.[12] Estimated coefficients on employment concentration range from -0.009 to -0.003 across specifications, implying that a one standard deviation increase in employment concentration from the mean reduces wages by -0.89% to -0.30%, respectively.[13]

Two-Stage Least Squares

We reproduce the same structure as the OLS results described above except instrumenting for log(*EHHI*). Because the instrumental variables are almost at the occupation-year level, we do not include those corresponding fixed effects in these estimations. The instrument is strong, as can be seen in Online Appendix Table A.1. (Hereafter, numbering for Online Appendix material is prefaced with an "A.")

Table 3 presents the two-stage least squares (2SLS) estimates, which suggest that higher labor-market concentration substantially reduces wages. Panel A presents results based on just the various sets of fixed effects and labor-market specific trends, without observable controls. Specification 1 looks at changes across years within labor market beyond a linear trend in unobservables and finds a -0.061 effect with a CI of (-0.104, -0.018). That point estimate implies that labor-market concentration one standard deviation above the mean is associated with a 5.9% lower wage than at mean concentration.[14] The estimated effect varies but remains substantially similar using more-flexible sets of fixed effects across the columns of panel A, increasing in magnitude to -0.091 in the richest specification, implying an 8.7% lower wage given a one standard deviation increase in labor-market concentration from the mean. In panel B, adding market-level controls changes the results little. Increased concentration does appear to raise labor productivity substantially.[15] However, workers do not appear to benefit from that productivity increase through higher wages. Across specifications, the coefficients on productivity are small and the coefficients on concentration are similar when productivity is included (panel B) or excluded (panel A) from the wage model. In panel C, adding product-market concentration

[12]Our product-market concentration measure focuses on local product markets. Industries vary in the extent to which their markets are truly localized, however, rather than being more nationalized or globalized. In highly localized industries (e.g., nursing homes), the measure should perform as intended. At the other extreme, all locales in an industry with a global market actually share the same (global) concentration and our local product-market concentration measure is simply noise that is likely positively correlated with labor-market concentration. Whereas most studies allow product- and labor-market concentration to be fully confounded by omitting product-market concentration altogether, we offer an (admittedly imperfect) measure of each by which we hope to reduce confounding, but we cannot credibly eliminate it. In later sensitivity analysis, we allow for heterogeneous effects of labor-market concentration depending on the tradability of each industry's products.

[13]Estimated Percent Wage Effect $= \exp(\hat{\beta}\,[\ln(E(E\hat{H}HI) + SD(E\hat{H}HI)) - \ln(E(E\hat{H}HI))]) - 1$.

[14]Calculated as $\exp(-0.061 \times [\ln(0.066 + 0.113) - \ln(0.066)]) - 1 = -5.9\%$.

[15]Table A.2 reports estimates of a positive effect of employment concentration on labor productivity using an analogous 2SLS model. Higher labor-market concentration causes a large increase in labor productivity (revenue per worker).

Table 3. Effect of Labor-market Concentration on Hourly Wage using 2SLS

	Panel A: Only fixed effects		
	(1)	(2)	(3)
Log(*Employment HHI*)	−0.061***	−0.054**	−0.091***
	[0.022]	[0.022]	[0.023]
CZ × OCC FE	Yes	Yes	Yes
Year FE	Yes		
CZ × Year FE		Yes	Yes
Ind × Year FE			Yes
Labor-market trends	Yes	Yes	Yes
	Panel B: Add market-level controls		
	(1)	(2)	(3)
Log(*Employment HHI*)	−0.059***	−0.055***	−0.089***
	[0.022]	[0.021]	[0.022]
Log(*Labor productivity*)	0.005**	0.009***	0.009***
	[0.003]	[0.003]	[0.003]
Missing labor productivity	0.016	0.036**	0.026
	[0.015]	[0.016]	[0.017]
Fraction of missing estab labor productivity	0.039	0.034	0.093***
	[0.025]	[0.023]	[0.024]
Log(*Total employment*, Data Axle)	0.089***	0.072**	0.140***
	[0.032]	[0.033]	[0.037]
CZ × OCC FE	Yes	Yes	Yes
Year FE	Yes		
CZ × Year FE		Yes	Yes
Ind × Year FE			Yes
Labor-market trends	Yes	Yes	Yes
	Panel C: Add sales-based HHI		
	(1)	(2)	(3)
Log(*Employment HHI*)	−0.094***	−0.089***	−0.089***
	[0.024]	[0.023]	[0.022]
Log(*Sales HHI*)	0.027***	0.027***	0.014***
	[0.001]	[0.001]	[0.001]
Sales HHI missing	−0.078***	−0.079***	0.009***
	[0.003]	[0.003]	[0.003]
Log(*Labor productivity*)	0.006*	0.010***	0.008***
	[0.003]	[0.003]	[0.003]
Missing labor productivity	0.052***	0.076***	0.028
	[0.016]	[0.017]	[0.017]
Fraction of missing estab labor productivity	0.076***	0.066***	0.104***
	[0.028]	[0.025]	[0.025]
Log(*Total employment*, Data Axle)	0.132***	0.114***	0.134***
	[0.036]	[0.035]	[0.037]
CZ × OCC FE	Yes	Yes	Yes
Year FE	Yes		
CZ × Year FE		Yes	Yes
Ind × Year FE			Yes
Labor-market trends	Yes	Yes	Yes

Notes: The dependent variable in all estimations is worker's log(*Hourly Wage*). All estimations are weighted by the worker's personal weight. The sample includes 23,978,672 individuals. Standard errors in brackets allow clustered errors at the CZ-occupation level. CZ, commuting zone; FE, fixed effects; HHI, Herfindahl-Hirschman Index; OCC, 3-digit 1990 Census occupation; 2SLS, two-stage least squares. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

makes estimates from all three specifications very similar. Our preferred estimate from the richest model is -0.089 with a CI of $(-0.132, -0.046)$. It implies an 8.5% effect on wage from a one standard deviation higher concentration. In Table A.15, we further control for individual human capital characteristics and find our baseline results to be robust. The estimated coefficient on log(*EHHI*) is -0.062 and remains statistically significant at the 1% level.

Robustness

We assess robustness to use of alternatives for the instrument, the concentration measure, the order of introducing controls, the functional form of the product-market concentration variable, and the labor compensation measure.

The negative effect of increased labor-market concentration on wages is also apparent when we use an alternative variable as an instrument. Our primary instrument is the leave-me-out (same-occupation, same-year, other-CZ) mean reciprocal of the number of employing firms, as in Azar et al. (2022). When we instead use the leave-me-out mean HHI as the instrument, as in Rinz (2022), it produces similar results (Table A.3).

A concern with this type of IV that was not addressed in prior work is that a positive national-level consumer demand shock that raises labor demand for an occupation could directly drive up labor compensation and drive down labor-market concentration in every CZ for that occupation. As a result, the exclusion restriction could be violated. To mitigate this concern, we further control for the natural logarithm of total sales at the occupation-year level in the estimations.[16] We present these results in Table A.5, which is analogous to Table 3 but for the addition of this control. The estimated labor-market concentration effects are similar, adding credibility to the IV results.

Another concern with the IV estimate in Table A.3 is that insufficient attention is paid to conditioning on productivity. While the relationship between own-market wage and own-market HHI are analyzed conditional on own-market productivity, the instrument is average other-market HHI, but not average other-market HHI conditional on other-market productivity. To estimate this, we first regress log(*EHHI*) on the natural logarithm of labor productivity in each CZ-occupation-year, obtain the residual, and then define an instrument for own-market residualized HHI as the average of residuals in other CZs for the same occupation-year. We report the 2SLS estimates in Table A.4. Compared to the results in panel C of Table 3, the estimated effect of HHI on wages falls substantially in panel C of Table A.4,

[16]To measure total sales at the occupation-year level, we use the establishment-level sales information from Data Axle and the annual joint distribution of occupation and industry employment nationally from the Census samples, apportioning sales value across occupations within establishment based on employment share and then totaling across establishments within occupation-year.

but all the estimated coefficients are statistically significant. This finding pushes the analysis beyond what others have done in the literature and shows some degree of fragility of the IV estimate above and in the literature.

We also assess robustness to use of an alternative measure of employer concentration. Given that Amazon employs workers both in warehouses and in Whole Foods, this alternative measure allows for the possibility that, within the same CZ, concentration across Amazon's multiple Whole Foods stores matters, across its multiple warehouses matters, but that concentration across warehouses and Whole Foods does not. Rather than our primary way of defining concentration within parent firm, occupation, and CZ, this alternative measure defines it within parent firm, occupation, CZ, and industry. We present estimates analogous to panel C of Table 3 but using this alternative EHHI measure in Table A.6. Point estimates are similar, though a bit larger in magnitude.

In our analysis, all models that include local product-market concentration also include productivity measures. As discussed earlier, this method could bias downward the estimated effect of concentration on wages. Increased firm concentration in product-markets and/or employment may increase labor productivity (the claimed rationale for most merger and acquisition activity), and higher labor productivity may be passed through into higher wages, including labor productivity as a control while modeling effects of labor-market concentration on wages may over-control for a variable on the causal pathway. This variable could remove a channel by which concentration might positively affect wages. To assess this concern empirically, we estimate versions of the main model in Table 3 that add our measure of product-market concentration, the log of sales HHI, as a control variable without including labor productivity (Table A.7). In the most stringent specification (column (3)), the estimated coefficients on employment concentration conditional on product-market concentration are substantially comparable: -0.089 when including labor productivity (Table 3) and -0.095 when excluding it (Table A.7).

One of the innovations of our article is the ability to control for product-market concentration along with labor-market concentration. However, the main model controls for it only linearly. To assess robustness to this functional form restriction, we also estimated models that include log(*SHHI*) more flexibly, using a fifth-order polynomial. Results are robust, and the estimated coefficient on log(*EHHI*) becomes slightly larger. Adding the higher-order polynomials of product-market concentration to the specification in column (3), panel C of Table 3 increases the log(*EHHI*) coefficient slightly from -0.089 to -0.094 and it remains statistically significant. We report the results in Table A.8.

As an alternative measure of labor compensation beyond wage, we use an indicator, available in the Census ACS data since 2008, that reflects whether each worker is covered under employment-based health insurance. The

2SLS point estimate based on the richest specification (column (3), panel (C) of Table 4) implies that a one standard deviation increase in labor-market concentration at the sample mean reduces the probability of employment-based health insurance coverage by 2.9 percentage points, representing a 4.1% reduction relative to the sample mean of 70.3 percentage points. This result is statistically significant at the 10% level and is the first evidence on the effect of labor-market concentration on non-wage compensation.[17] We do not see evidence of reduced wages being offset by increases in non-wage compensation. The direction of change appearing similar for wage and the largest non-wage form of compensation provides some evidence that studies that focus only on wage get the basic compensation story correct.

Heterogeneous Effects

We examine if the effects of labor-market concentration differ depending on the levels of each of three factors: local product-market concentration, worker unionization rate, and occupational offshorability. We use only the richest specification (as shown in column (3), panel C of Table 3) and instrument for labor-market concentration and its interaction with our measures of each of these factors in turn. We report the results in Table 5.

For both wages (panel A) and employment-based health insurance probability (panel B), labor-market concentration has a more negative effect on labor compensation in the context of more-concentrated product markets compared to less-concentrated product markets (column (1) of Table 5). Any greater product-market rents within the firms driven by increased product-market concentration do not seem to translate into higher labor compensation and especially not in the context of greater labor-market concentration.

In manufacturing, Benmelech et al. (2022) found evidence that stronger unions reduce the negative effect of labor-market concentration on wages. This evidence is consistent with unions increasing worker bargaining power and protecting against negative compensation effects attributable to market concentration limiting workers' outside options. We broaden the analysis to the entire private sector and leverage variation in unionization rates and find similar results. Because the Census does not measure worker's union status, we rely on the data from the Current Population Survey (CPS). To obtain enough observations per cell, we aggregate occupations into six broad groups and estimate unionization rates within each group-state.[18] For

[17]Although we do not discuss them in the text, for this table and all 2SLS estimates in the rest of the article, Online Appendix tables report first-stage estimates, estimates of all control-variable coefficient estimates, and OLS analogue estimates.

[18]We follow broad groups as presented in Census 1990 data by IPUMS: "Managerial and professional specialty occupations," "Technical, sales, and administrative support occupations," "Service occupations," "Farming, forestry, and fishing occupations," "Precision production, craft, and repair occupations," and "Operators, fabricators, and laborers."

Table 4. Effect of Labor-market Concentration on Employer-sponsored Health Insurance Coverage using 2SLS

	Panel A: Only fixed effects		
	(1)	*(2)*	*(3)*
Log(*Employment HHI*)	0.032*	0.032*	−0.028
	[0.017]	[0.017]	[0.022]
CZ × OCC FE	Yes	Yes	Yes
Year FE	Yes		
CZ × Year FE		Yes	Yes
Ind × Year FE			Yes
Labor-market trends	Yes	Yes	Yes
	Panel B: Add market-level controls		
	(1)	*(2)*	*(3)*
Log(*Employment HHI*)	0.015	0.015	−0.029*
	[0.013]	[0.013]	[0.016]
CZ × OCC FE	Yes	Yes	Yes
Year FE	Yes		
CZ × Year FE		Yes	Yes
Ind × Year FE			Yes
Labor-market trends	Yes	Yes	Yes
	Panel C: Add sales-based HHI		
	(1)	*(2)*	*(3)*
Log(*Employment HHI*)	−0.002	−0.003	−0.029*
	[0.014]	[0.014]	[0.016]
CZ × OCC FE	Yes	Yes	Yes
Year FE	Yes		
CZ × Year FE		Yes	Yes
Ind × Year FE			Yes
Labor-market trends	Yes	Yes	Yes

Notes: The dependent variable in all estimations is a dummy variable indicating whether an individual has health insurance through a current or former employer or union. All estimations are weighted by the personal weight. The sample includes 14,403,816 individuals. Standard errors in brackets allow clustered errors at the CZ-occupation level. The full estimation results are available in Table A.10. CZ, commuting zone; FE, fixed effects; HHI, Herfindahl-Hirschman Index; OCC, 3-digit 1990 Census occupation; 2SLS, two-stage least squares.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

each year, we pool CPS observations in that group-state across a five-year window centered on the study year and use CPS earner weights to estimate a unionization rate.[19] We match this back to Census data for each worker whose wages we are explaining based on occupation-state-year. Our estimated coefficients on the interaction between labor-market concentration

[19]Benmelech et al. (2022) used national-level, industry-specific unionization rates. We use state-level, occupation-group-specific unionization rates.

Table 5. Heterogeneous Effects of Labor-Market Concentration using IV

	Panel A: Log(Hourly Wage)			
	(1)	*(2)*	*(3)*	*(4)*
Log(*Employment HHI*)	−0.115***	0.011	−0.089***	−0.083***
	[0.022]	[0.023]	[0.022]	[0.031]
Log(*Employment HHI*) × Log(*Sales HHI*)	−0.008***			
	[0.001]			
Log(*Employment HHI*) × Unionization rate			0.048	
			[0.060]	
Unionization rate			0.309	
			[0.286]	
Log(*Employment HHI*) × Age between 25 and 44		−0.080***		
		[0.006]		
Log(*Employment HHI*) × Age between 45 and 64		−0.089***		
		[0.006]		
Age between 25 and 44		0.205***		
		[0.029]		
Age between 45 and 64		0.314***		
		[0.028]		
Log(*Employment HHI*) × Offshorability				−0.006
				[0.013]
CZ × OCC FE	Yes	Yes	Yes	Yes
CZ × Year FE	Yes	Yes	Yes	Yes
Ind × Year FE	Yes	Yes	Yes	Yes
Labor-market trends	Yes	Yes	Yes	Yes
	Panel B: Employer-sponsored health insurance			
	(1)	*(2)*	*(3)*	*(4)*
Log(*Employment HHI*)	−0.030*	−0.028*	−0.031*	−0.029
	[0.016]	[0.016]	[0.017]	[0.020]
Log(*Employment HHI*) × Log(*Sales HHI*)	−0.001*			
	[0.000]			
Log(*Employment HHI*) × Unionization Rate			0.059	
			[0.044]	
Unionization rate			0.315	
			[0.214]	
Log(*Employment HHI*) × Age between 25 and 44		0.004		
		[0.004]		
Log(*Employment HHI*) × Age between 45 and 64		0.003		
		[0.003]		
Age between 25 and 44		0.006		
		[0.018]		
Age between 45 and 64		0.072***		
		[0.017]		
Log(*Employment HHI*) × Offshorability				0.001
				[0.013]
CZ × OCC FE	Yes	Yes	Yes	Yes
CZ × Year FE	Yes	Yes	Yes	Yes
Ind × Year FE	Yes	Yes	Yes	Yes
Labor-market trends	Yes	Yes	Yes	Yes

Notes: This table reports the heterogeneous effects of labor-market concentration on labor compensation using 2SLS estimations. We use two instrumental variables: 1) the average of the natural logarithm of 1 over the number of firms in the same occupation but in other CZs in a year and 2) the interaction between the first instrument and log(*Sales HHI*) or *Unionization rate* or *Offshorability*. The

dependent variables in panels A and B are the natural logarithm of real hourly wage and a dummy variable indicating whether an individual has health insurance through a current or former employer or union, respectively. *Unionization rate* is the 5-year average unionization rate in a major occupation group-state cell centered around a year in the Current Population Survey (CPS). Occupation represents the major group in CPS. *Offshorability* measures the extent to which the tasks performed by occupations are offshorable, and the data are from David Dorn's webpage (https://www.ddorn.net/data.htm). In column (1) in both panels, we drop observations in which *Sales HHI* is missing. The control variables are the same as those in column (3), panel C of Table 3. All estimations are weighted by the personal weight. Standard errors in brackets allow clustered errors at the CZ-occupation level. CZ, commuting zone; FE, fixed effects; HHI, Herfindahl-Hirschman Index; IV, instrumental variables; OCC, 3-digit 1990 Census occupation; 2SLS, two-stage least squares.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

and unionization rate for wage and employment-based health insurance coverage are both positive but not statistically significant.[20]

Furthermore, we test whether the estimated effects of labor-market concentration vary with workers' ages. We divide workers into three groups: 1) ages 16 to 24, 2) ages 25 to 44, and 3) ages 45 to 64. Older workers have lower job mobility.[21] Therefore, we expect the estimated effect of labor-market concentration to be stronger for older workers. For hourly wages, the estimated effects are consistent with our expectations. But the estimated effect on health insurance is not statistically significant by age group of workers.

Finally, we test whether occupational offshorability changes the effect of labor-market concentration, measuring offshorability following Autor and Dorn (2013). For more-offshorable occupations, locale is a less-meaningful labor-market boundary and $\log(EHHI)_{mt}$ is a noisier measure of concentration. Therefore, we expect a negative main effect of concentration and a positive interaction term. Instead, we observe a negative main effect of concentration but a null interaction term with offshorability for both wages and health insurance.

Conclusion

We develop new evidence that suggests recent results that link higher labor-market concentration to lower wages are robust to potential confounders such as product-market concentration, labor productivity, and labor force composition. OLS estimates imply null or small, negative effects but, as in Azar et al. (2022) and Rinz (2022), analysis using a market-type's average

[20]The estimated effects are small for both wages and insurance. A one standard deviation increase in unionization (14.8%) above the sample mean (7.4%) moves the marginal effect of log(*EHHI*) on log(*hourly wages*) from −0.085 to −0.078, and moves the marginal effect of log(*EHHI*) on employment-based health insurance coverage from −0.027 to −0.018.

[21]For example, see https://www.bls.gov/opub/ted/2006/jan/wk1/art02.htm?view_full.

market structure in other locales as an instrument for each local market's concentration yields estimates that are substantially larger, significant, and negative.[22] For a rough sense of magnitude, increasing concentration of a labor market by one standard deviation (0.113) at the mean level (0.066) would imply an 8.5% decrease in wages based on the estimate in column (3), panel C of Table 3.

In terms of changes in labor-market concentration as an explanation for recent *changes* in US wages, the potential importance is limited by the fact that average concentration has not changed much. As a back-of-the-envelope calculation, the move from the average annual concentration level in 2000 (0.026) to the minimum (0.022 in 2012) would imply a predicted 0.65% increase in wages and, then, a 0.79% wage decrease as concentration moved to 0.027 in 2017. Our finding aligns with results in Rinz (2022) and Lipsius (2018), which suggests that changes in labor-market concentration do little to explain changes in labor share. However, findings of negative concentration effects on wages suggest that concentration has been a factor consistently depressing wages across the period.

Comparing our estimates to those of Rinz (2022), the most-similarly specified, is particularly illuminating. We both define locale as CZ and have similar underlying data and analytic structures. Rinz used the LBD and defined labor markets along local industrial lines. We use the similarly structured Data Axle data combined with national occupational distributions for each industry to measure labor-market concentration along local occupational lines while controlling for local product-market concentration. The most comparable estimates are between his richest model in the 2005 to 2015 period (column (5) of his table 5) and our specification in column (2), panel A of Table 3, which excludes market-level controls and product-market concentration. Our estimate is −0.054 (SE = 0.022) while his is −0.134 (SE = 0.028). When we add market-level and product-market concentration controls (column (2), panel C of Table 3), our estimate becomes −0.089 (SE = 0.023), suggesting that omitting these factors may lead to underestimation of the labor-market concentration effect. Further adding industry-year fixed effects does not change our estimate.

We add novel evidence that the negative effect of labor-market concentration on wage is robust to conditioning on local product-market concentration, a theoretically important, potential confounder. Increases in workers' local product-market concentration predicts increased wages, consistent with rent-sharing within the firm. The effect of labor-market concentration is strengthened when product-market concentration is added as a control. The estimated effect on the probability of employer-provided

[22]Rinz (2022) did not present OLS results, instead focusing only on the IV. His figure 21(d), however, is a visual analogue to OLS with worker-level wage data and also describes a weak positive association between concentration and wages.

health insurance is in the same direction and is marginally statistically significant.

This evidence suggests reductions in labor-market concentration could lift labor compensation levels toward competitive levels. Marinescu and Hovenkamp (2019) and Naidu et al. (2018) fleshed out applications of traditional legal and economic anti-trust analysis into the labor market. In labor markets with employer market power derived from concentration and other sources (information frictions, mobility costs, legal barriers, and so on), labor-market power in the form of stronger workers' organizations can countervail and shift the balance of bargaining power (Lee and Mas 2012; Sojourner et al. 2015). In theory, this can increase labor-market efficiency in some cases. Blunting employer market power is possible through tools beyond traditional anti-trust enforcement as well: reducing covenants not to compete and no-poach agreements (Krueger and Posner 2018; Starr 2019), addressing workers' information problems with respect to unobserved employer heterogeneity (Benson, Sojourner, and Umyarov 2020), and setting labor-market standards through regulation (Shierholz 2018). For markets in which concentration suppresses labor compensation, these kinds of reforms are likely to have larger benefits and smaller costs.

References

Acemoglu, Daron, and Jörn-Steffen Pischke. 1998. Why do firms train? Theory and evidence. *Quarterly Journal of Economics* 113(1): 79–119.

Autor, David, and David Dorn. 2013. The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review* 103(5): 1553–97.

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

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

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

Benson, Alan. 2013. Firm-sponsored general education and mobility frictions: Evidence from hospital sponsorship of nursing schools and faculty. *Journal of Health Economics* 32(1): 149–59.

Benson, Alan, Aaron Sojourner, and Akhmed Umyarov. 2020. Can reputation discipline the gig economy? Experimental evidence from an online labor market. *Management Science* 66(5): 1802–25.

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2022. Labor market power. *American Economic Review* 112(4): 1147–93.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2001. Measurement error in survey data. In James J. Heckman and Edward Learner (Eds.), *Handbook of Econometrics*, Vol. 5, pp. 3705–843. Amsterdam: North Holland, Elsevier.

Gruber, Jonathan, and Brigitte C. Madrian. 1994. Health insurance and job mobility: The effects of public policy on job-lock. *Industrial and Labor Relations Review* 48(1): 86–102.

Krueger, Alan B., and Eric A. Posner. 2018. A proposal for protecting low-income workers from monopsony and collusion. Policy Proposal 2018-05. Washington, DC: Hamilton Project.

Lee, David S., and Alexandre Mas. 2012. Long-run impacts of unions on firms: New evidence from financial markets, 1961–1999. *Quarterly Journal of Economics* 127(1): 333–78.

Lipsius, Ben. 2018. Labor market concentration does not explain the falling labor share. Technical Report, University of Michigan Department of Economics.

Manning, Alan. 2011. Imperfect competition in the labor market. In Orley Ashenfelter and David Card (Eds.), *Handbook of Labor Economics*, Vol. 4B, pp. 973–1041. Amsterdam: North Holland, Elsevier.

Marinescu, Ioana, and Herbert Hovenkamp. 2019. Anticompetitive mergers in labor markets. *Indiana Law Journal* 94(3): Article 5.

Mellow, Wesley, and Hal Sider. 1983. Accuracy of response in labor market surveys: Evidence and implications. *Journal of Labor Economics* 1(4): 331–44.

Naidu, Suresh, Eric A. Posner, and Glen Weyl. 2018. Antitrust remedies for labor market power. *Harvard Law Review* 132(2): 536–601.

Nevo, Aviv. 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica* 69(2): 307–42.

Prager, Elena, and Matt Schmitt. 2021. Employer consolidation and wages: Evidence from hospitals. *American Economic Review* 111(2): 397–427.

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

Ruggles, Steven, Josiah Grover, Erin Meyer, Jose Pacas, Sarah Flood, Ronald Goeken, and Matthew Sobek. 2018. IPUMS USA: Version 8.0 [dataset]. Minneapolis, MN.

Shambaugh, Jay, and Ryan Nunn (Eds.). 2018. *Revitalizing Wage Growth: Policies to Get American Workers a Raise*. Washington, DC: Hamilton Project.

Shambaugh, Jay, Ryan Nunn, Patrick Liu, and Greg Nantz. 2017. Thirteen facts about wage growth. Technical Report. Washington, DC: Hamilton Project.

Shierholz, Heidi. 2018. Strengthening labor standards and institutions to promote wage growth. Policy Proposal 2018-07. Washington, DC: Hamilton Project.

Sojourner, Aaron J., Brigham R. Frandsen, Robert J. Town, David C. Grabowski, and Min M. Chen. 2015. Impacts of unionization on quality and productivity: Regression discontinuity evidence from nursing homes. *ILR Review* 68(4): 771–806.

Starr, Evan. 2019. The use, abuse, and enforceability of non-compete and no-poach agreements. Technical Report. Washington, DC: Economic Innovations Group.