Revisit labor market concentration under Work-From-Home: An Integration of Labor Market Boundaries

Chengmou Lei*

December 12, 2024

Abstract

Post-COVID work-from-home (WFH) policies have reduced geographic constraints, enabling workers to access remote job opportunities across regions. This shift has expanded local labor market boundaries and altered labor market concentration levels. In this study, I adjust a traditional concentration measurement to include remote job postings and analyze the impact of WFH on labor market concentration using the difference-in-differences (DID) approach. I find that excluding remote jobs significantly biases the concentration estimate. Additionally, I re-estimate the effect of concentration on wages using the two-stage least squares (2SLS) method, finding that concentration exerts downward pressure on wages for occupations with lower education requirements. In contrast, wages for occupations with higher education requirements increase with concentration. These results suggest supply-side shifts in the labor market.

Keywords: remote work, work-from-home, WFH, labor market concentration, pandemic, COVID-19

JEL: J2, J3, J4, J6, D22, O33, R23

*Finance Department, Smith Business School, University of Maryland (College Park)

Contents

1 Introduction

Labor market concentration and its impact on wages is not new in labor economics. However, the remote job offerings and the WFH trend sparked by COVID-19 suggest a revisit of this topic. The critical shift under WFH is the integration of local labor markets into national or global ones. Pre-COVID, remote jobs were scarce, with workers typically required to work on-site. This restriction limits workers' opportunities at local firms. WFH reduces the friction of mobility in labor markets and dramatically expands workers' outside options, allowing them to consider jobs from other states or countries. This integration of local labor market boundary into a global one will change the concentration pattern.

In this paper, I ask two research questions: 1) How does WFH affect local labor market concentration? 2) How does concentration affect wages?

The answer to the first question is ex-ante unclear. On the one hand, remote jobs offered by firms not previously present in the local labor market are equivalent to new firm entries, potentially decreasing local concentration. On the other hand, as Berger et al. (2022) points out, despite many firms and workers in the US labor market, employment is concentrated mainly in a few "superstar" firms. Local concentration may increase if remote jobs primarily come from these already-present and dominant firms. For this question, I cannot use the traditional Herfindahl-Hirschman Index (HHI) developed by Azar et al. (2020) as a measure of labor market concentration. This traditional measurement considers only jobs posted in the local labor market delineated by occupation and commuting zone (CZ) and does not account for remote job offerings. It will produce biased concentration estimates when there is plenty remote jobs. I develop a new concentration measurement based on the HHI index proposed by Azar et al. (2020) incorporating nationwide remote jobs. The remote-adjusted concentration measurement, on average, is far smaller than the traditional one. I present summary statistics of concentration for prevalent labor markets in the US.

Accurately measuring labor market concentration is crucial, as it has been linked to inequality, low wages, and stagnant pay growth. Labor market concentration is also a critical factor in antitrust authorities' reviews of mergers and acquisitions (M&A) (Schubert et al. (2024)). If mismeasured, M&As that could enhance efficiency may be blocked due to unfounded concerns over its labor market effects. My calculation using two different HHI shows significantly different results. According to antitrust law, markets with HHI above 2500 are highly concentrated, while those between 1500 and 2500 HHI are moderately concentrated (Department of Justice / Federal Trade Commission 2010 horizontal merger guidelines). My traditional HHI calculation shows that the average market has an HHI of 3291.52 (simple average), which is considered a highly concentrated market and is equivalent to only three firms with equal shares of the total vacancy pool hiring. Across the year, there are 56.59% of markets have high concentration, 3.27% of markets have medium concentration, and 28.91% of markets have low concentration. Using 2019 employment data for each market (occupation-CZ), 92% of national-wide total employment works in low-concentrated markets, 5.6% in medium-concentrated markets, and only 2.3% in highly-concentrated markets. This explains why later when weighted by 2019 employment, the nationwide average traditional HHI decreased significantly and converged to remote-adjusted HHI. My modified HHI calculation shows that the average market has an HHI of 609.08 (simple average), considered a low-concentrated market, is equivalent to 16.5 recruiting firms with equal shares of the total vacancy pool. Under the new HHI measurement, only 1.31% of markets have a concentration greater than 2500. Another 5.92% of markets have a concentration between 1500 and 2500. 92.77% of markets have a concentration below 1500. Many highly concentrated labor markets now have low or medium concentrations. This suggests remote job offerings significantly alter the labor market concentration pattern.

Then I use a difference-in-difference (DID) method to analyze the effect of WFH on labor market concentration by comparing occupations that tend to post more remote jobs post-COVID (teleworkable) with those that tend to post fewer (non-teleworkable).

To complete the analysis, I examine the relationship between labor market concentration and wages using the new concentration measurement. I use the difference-in-difference

(DID) and two-stage least squares (2SLS) methods for the analysis. I find different patterns for occupations with varying levels of skill requirements. For low-skilled occupations, higher concentration puts downward pressure on wages, as pointed out by prior literature (Azar et al. (2020), Schubert et al. (2024), Qiu and Sojourner (2023), etc.). Meanwhile, higher concentrations are accompanied by higher wages for occupations with high education requirements at entry. The aggregate positive relationship is mainly driven by four occupations: Management Occupations (11-0000), Business and Financial Operations Occupations (13-0000), Healthcare Practitioners and Technical Occupations (29-0000), Healthcare Support Occupations (31-0000). Since I control for demand-side factors, the positive pattern suggests that supply-side factors may play a role. It is also consistent with Macaluso et al. (2019), which finds that labor market power not only manifests itself through downward pressure on wages but also through higher requirements for skill.

My research intersects with three streams of literature. The mainstream papers measure labor market concentration. Benmelech et al. (2022), Rinz (2018), and Qiu and Sojourner (2023) use employment share as a measure of labor market concentration and argue it is a good measurement for labor market power. Yeh et al. (2022), and Berger et al. (2022) estimate markdown, which is the difference between the marginal revenue productivity of labor and wage, and argue it is a proxy for firms' labor market power. Azar et al. (2022) is the first to propose the HHI measurement, which is widely used later by other papers in the field (Posner (2021), Prager and Schmitt (2021), etc). Schubert et al. (2024) adopt a "probabilistic" approach to redraw the delineation of the labor market. They identify other options for workers in the focal occupation to switch. However, none of the existing research considers the expansion of local labor market boundary due to WFH when calculating HHI.

The second stream of literature utilizes the concentration measurement from Azar et al. (2020) to quantify its downward pressure on wages (Berger et al. (2019), Benmelech et al. (2018), Jarosch (2021), Rinz (2018), etc). My study contributes to this literature by re-quantifying it while considering the changing boundaries of local labor markets. I found great heterogeneity of the impact across occupations and levels of concentration.

The third stream of literature examines changing labor market dynamics post-COVID. Gallant et al. (2020) distinguish between temporary and permanent unemployment in post-pandemic recovery. Coibion et al. (2020) quantify job loss and labor force participation post-pandemic. Bartik et al. (2020) find that low-wage services drive major employment declines for hourly workers in small and large businesses. My study contributes to this literature by providing evidence suggesting a change in the labor supply structure.

Section 2 outlines the construction of the new concentration measurement. Section 3 describes the Burning Glass data. Section 4 provides labor market concentration estimates. Section 5 compares the concentration between teleworkable and non-teleworkable jobs. Section 6 analyzes the effect of WFH on wages. Section 7 concludes.

2 Construction of Remote-Adjusted HHI

Azar et al. (2020) construct the labor market Herfindahl-Hirschman Index (HHI) to measure labor market concentration using formula (1).

$$HHI_{m,t} = \sum_{j=1}^{J} s_{j,m,t}^2 \tag{1}$$

where m represents the labor market defined at ONET-SOC6 - commuting zone (CZ) level, j is the firm, and t is the year. $s_{j,m,t}$ is the job posting share of firm j at time t in a given labor market m. The numerator is the number of job postings by firm j at time t in this labor market, and the denominator is the total number of job postings in the same labor market at the same time regardless of firms.

With WFH, remote jobs from other commuting zones are also part of workers' choice sets and should be factored into the numerator and denominator. Consider the following example for computer and information systems managers to give a quick and clear illustration of how my new measurement is constructed (6-digit ONET code: 11-3020). The numbers and calculations are summarized in table 1. In this scenario, Amazon, and IBM have offices in Ohio, but Google does not. In Oregon, only Amazon and Google have offices. The job posting numbers used are hypothetical.

1. Without remote jobs: Each firm posts 50 jobs in different states. According to the traditional labor market HHI formula by Azar et al. (2020), the HHI of Ohio will be $\left(\frac{50}{100} \times 100\right)^2 + \left(\frac{50}{100} \times 100\right)^2 = 5000$. It's the same for Oregon.

2. Now, imagine Google's headquarters in California posts 30 remote jobs. Since workers in both Ohio and Oregon can access these jobs, the total number of job postings in each market increases to 130. In Ohio, Google's remote jobs can be seen as a new firm entry, diluting the market share of both Amazon and IBM to $50/130 = 38\%$, while Google's market share rises to $30/130 = 23\%$. As a result, Ohio's HHI drops to $(38\%^2 + 38\%^2 + 23\%^2) * 10000 = 3491$. In Oregon, where Google already had a presence, its market share further rises to 61%, while Amazon's decreased to 38%, causing the HHI to increase to 5266.

Ohio			Oregon	
No Remote Jobs				
Amazon		IBM	Amazon	Google
50		50	50	50
$HHI = \left(\frac{50}{100} \times 100\right)^2 + \left(\frac{50}{100} \times 100\right)^2 = 5000$			$HHI = \left(\frac{50}{100} \times 100\right)^2 + \left(\frac{50}{100} \times 100\right)^2 = 5000$	
Google in California post 30 remote jobs				
Amazon	IBM	Google	Amazon	Google
50	50	30	50	80
$HHI = \left(\frac{50}{130} \times 100\right)^2 + \left(\frac{50}{130} \times 100\right)^2 + \left(\frac{30}{130} \times 100\right)^2 = 3491$			$HHI = \left(\frac{50}{130} \times 100\right)^2 + \left(\frac{80}{130} \times 100\right)^2 = 5266$	

Table 1: Example

Formally, my remote-adjusted HHI formula is as follows:

$$HHI_{\text{remote-adjusted},zkt} = \sum_{j \in \mathrm{M}_{\mathrm{zkt}}} \left(\frac{n_{\mathrm{zjkt,nrm}} + n_{cjkt,rm,c!=z}}{\sum_{i \in \mathrm{M}_{\mathrm{zkt}}} (n_{zikt,nrm} + n_{cikt,rm,c!=z})} \times 100 \right)^2 \tag{2}$$

Where $n_{\mathrm{zjkt,nrm}}$ is the number of non-remote jobs posted by firm j at quarter t for occupation k and commuting zone z. $n_{jkt,rm}$ is the same but for remote jobs. M_{zkt} is the set of firms in the labor market defined by commuting zone z - occupation k - quarter t. The inverse of the HHI multiplied by 10000 (10000/HHI), gives the number of firms that would result in such an HHI if each had the same share of the market.

One caveat of my construction is that it could indicate reduced concentration and downward wage pressure in local labor markets where residents cannot realistically compete for remote job offerings. For instance, in a rural village where residents are unfamiliar with computers, remote programming jobs from a company like Google are unlikely to affect their labor market outcomes. My concentration measurement would misleadingly suggest lower concentration in such areas without any adjustment, implying less wage pressure. Whereas in reality, labor market conditions remain unchanged. To address this concern, I adjust my concentration measurement in Appendix B by incorporating the educational attainment levels of the population within each CZ and the educational requirements of remote job postings. The adjustments do not significantly change my main labor market concentration.

3 Data and Sample

I use data from Burning Glass Technologies (BGT) covering 2015 to 2023. This dataset captures a near-complete record of online job vacancies in the US and is widely utilized in labor market research (Azar et al. (2020), Schubert et al. (2024), Braxton and Taska (2023), Goldfarb et al. (2023), Acemoglu et al. (2022), Yeh et al. (2022), etc). It is important to note that BGT records only the posting and expiration dates of job listings in the month they are posted, without indicating when vacancies are filled or the number of hires per vacancy. For the analysis, I assume that each vacancy corresponds to only one hire. For a detailed discussion of other limitations related to the BGT data, see Azar et al. (2020).

Of interest to my work is the identifier of the employer, the occupation code (ONET-

SOC6), the remote type (remote, non-remote, hybrid), the location of the vacancy (county), and salary. The employer identifier allows me to calculate HHI for each firm separately. Some firms are staffing companies, as identified by BGT, and they hire for others, which are not directly observed. Following Azar et al. (2020) and Schubert et al. (2024), I treat the staffing companies as one firm.

I exclude internships, starting with a total of 312,072,080 job postings. I use the crosswalk provided by ONET to convert the 2019 ONET code in BGT to the 2010 ONET code and classify the occupation at the 6-digit level (ONET-SOC6). This conversion is necessary because I intend to utilize the occupation teleworkability data from Dingel and Neiman (2020), based on the 2010 ONET-SOC6 codes. I retain only those occupations with a corresponding match in 2010, resulting in a final sample of 301,855,847 observations across 836 occupations.

For the benchmark analysis, I follow Azar et al. (2020) to trim away the narrowly defined labor market. Within each year, I rank occupations by the numbers of job postings from highest to lowest and retain the top 90% of occupations. I apply this filter across all years, keeping only those occupations that appear every quarter. This process leaves me with 245 occupations, which account for 92% of the vacancy postings in the BGT dataset. Next, I identify the commuting zones corresponding to different counties based on delineation records. The total number of markets (ONET-SOC6 - CZ) considered in my primary analysis is 171,197.

I also calculate the traditional HHI weighted by population to mitigate concerns of the narrowly defined labor market and to compare my sample with Azar et al. (2020) and Schubert et al. (2024). My employment data comes from the Occupational Employment and Wage Statistics (OEWS) database. I used the "all data" file from BLS, which contains employment figures by year and the ONET-SOC6 code. The data is at the CBSA level, so I follow the methodology of Azar et al. (2020) to get commuting zone-level employment. This involves using estimated county population shares from the Census and multiplying these shares by the BLS employment data at the CBSA level for each ONET-SOC6 code. I then aggregate the ONET-SOC6 employment numbers across the counties that comprise each commuting zone. Finally, I merge the teleworkability scores from Dingel and Neiman (2020), retaining only those occupations for which this score is available. This process results in 245 ONET-SOC6 occupations across 171,197 labor markets.

I need BGT observations with non-missing annual salary information for the wage-related results. The dataset has limited coverage in this regard, with only 23.41% of observations providing salary information[1]. After applying this filter, I have 229 ONET-SOC6 occupations and 160,300 labor markets. Figure 1 shows the distribution of log real wages across markets and years. For comparison, I also plot the distribution of occupational wages for the same period using data from the OEWS. The two have a very similar distribution. Posted wages have less mass in the right tail of the distribution and more mass in the left tail, consistent with lower starting wages.

[1]For comparison, in Azar et al. (2020), which includes only 2016 data, only 16% of postings contained wage information.

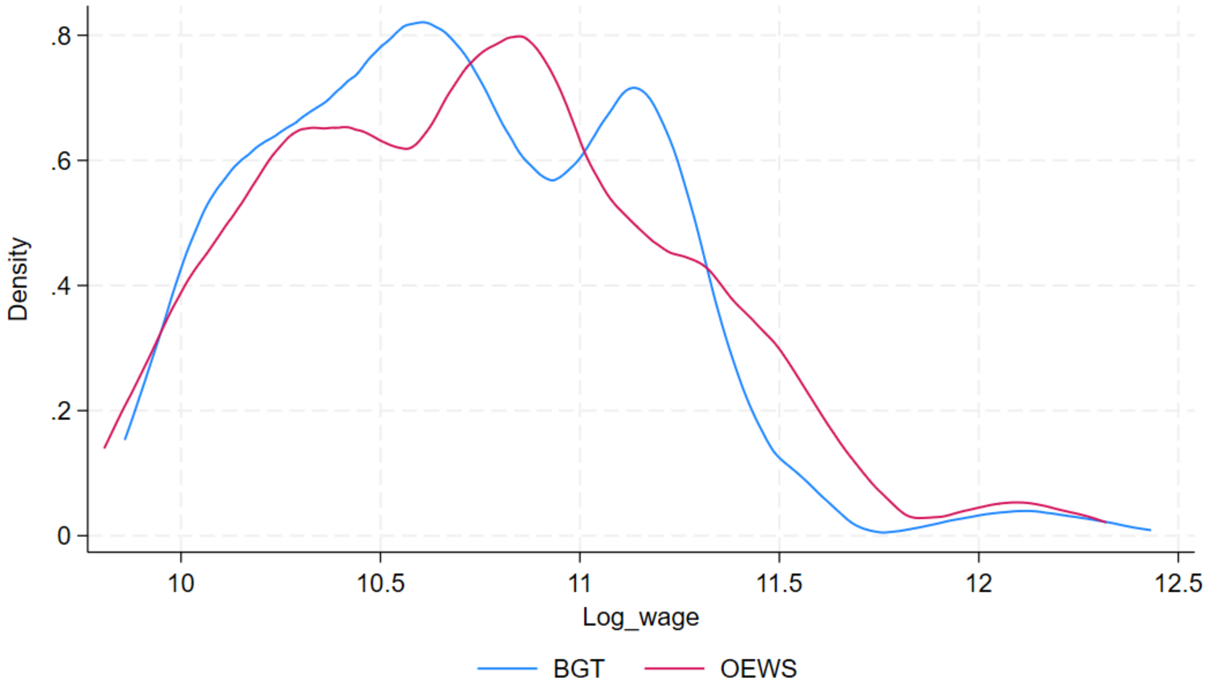


Figure 1: Comparison of wage data from BGT and OEWS

I also use the number of establishment and employment data from the BLS Quarterly Census of Employment and Wages (QCEW) to reduce the concern of omitted variable bias. My geographic area delineation file comes from David Dorn's website. My ONET-SOC structure data (used to aggregate 6-digit occupations to 2-digit) is from ONET.

4 Labor Market Concentration Estimates and Analysis

4.1 Remote Job Offerings

Figure 2 illustrates the trend in the weighted average share of remote job offerings over the years. As shown, less than 2% of jobs were remote before COVID-19. With the onset of the pandemic, the share of remote jobs increased significantly.

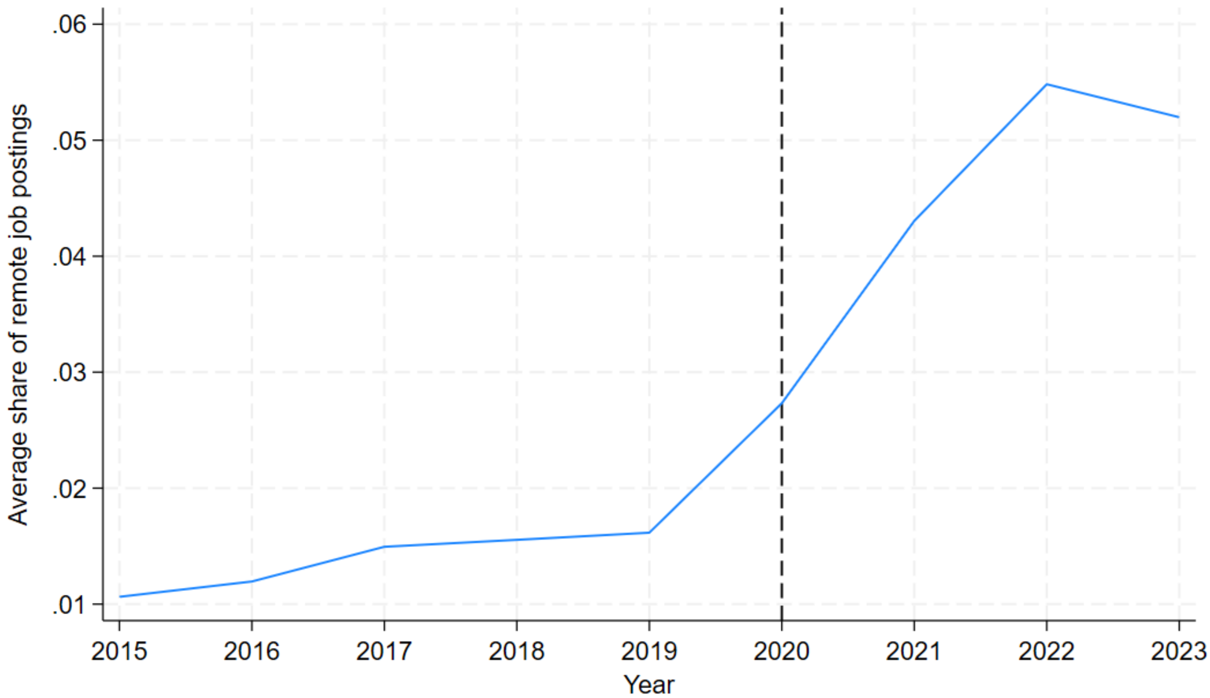


Figure 2: Average share of remote job postings across the year (weight: 2019 employment)

I also plot the average share of remote jobs for each commuting zone over the years on a U.S. map. Figures 3 and 4 display the distribution of remote jobs across commuting zones before and after the pandemic, respectively. Figures 5 and 6 show the population distribution across counties during the same periods. Comparing the distribution of remote jobs and population, we can observe that, before the pandemic, remote jobs were concentrated in commuting zones with relatively higher populations. After the pandemic, remote jobs became more evenly distributed across the country.

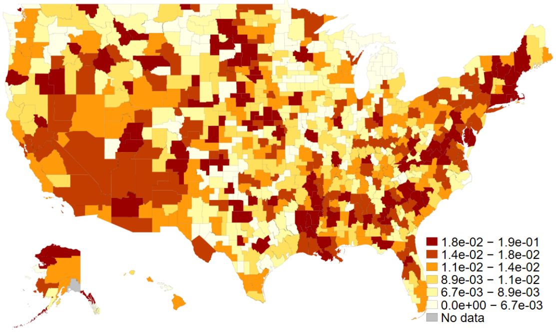


Figure 3: Average share of remote job postings (2015-2019, weighted using 2019 employment)

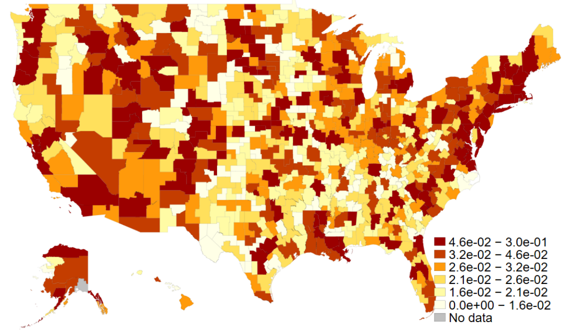


Figure 4: Average share of remote job postings (2020-2023, weighted using 2019 employment)

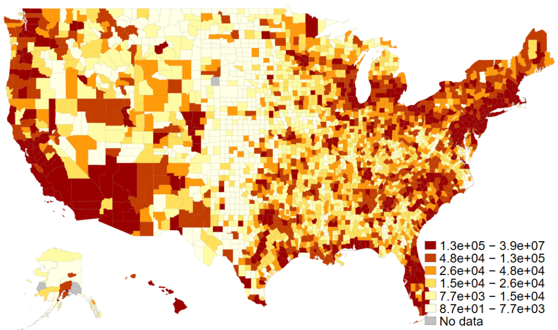


Figure 5: Average population (2015-2019)

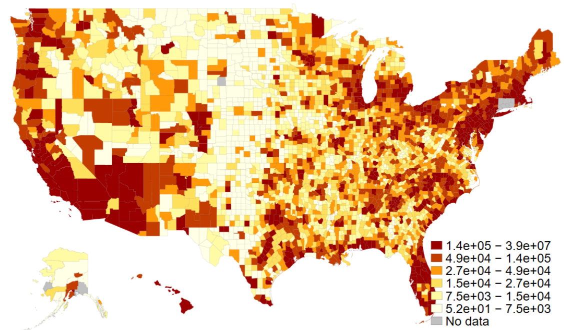


Figure 6: Average population (2020-2023)

4.2 Comparison of Traditional and Remote-Adjusted HHI

My remote-adjusted HHI differs significantly in magnitude from the traditional HHI developed by Azar et al. (2020). I provide the following summary statistics to highlight this difference and present a simple calculation to illustrate why the discrepancy is so substantial even before the pandemic (when the share of remote jobs was small).

Table 2 presents summary statistics of the HHI calculated using both the traditional and remote-adjusted methods, based on data from 2015 to 2023[2]. Figure 7 displays the corresponding distributions.

[2] 1,230,349 represents the number of market-years

	count	mean	min	max	p25	p50	p75	p90
Traditional HHI	1230349	3291.521	158.2979	10000	962.9761	2160.494	5000	10000
Remote-adjusted HHI	1230349	609.0799	.0001882	5656.316	31.65407	212.5643	672.9913	1748.98
Observations	1230349							

Table 2: Summary Statistics (unweighted, 2015-2023)

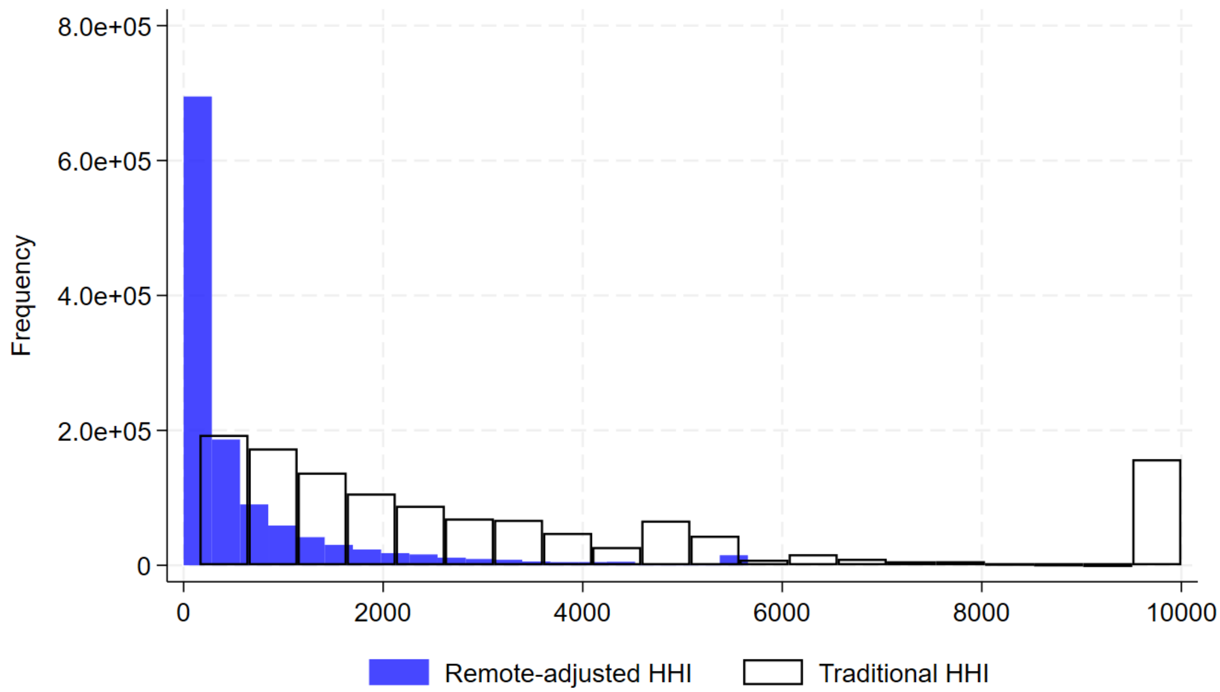


Figure 7: Distribution of Traditional HHI and remote-adjusted HHI

The difference may be driven by the increase in remote jobs post-COVID, as there were very few remote jobs before the pandemic. Consequently, the two measurements should be pretty similar in the pre-COVID period. However, as shown in table 3, even when using only pre-COVID data, the unweighted remote-adjusted HHI is still significantly smaller than the traditional HHI. The difference observed before COVID arises from the fact that, in the remote-adjusted HHI, the denominator accounts for all remote jobs posted by all firms in other CZs nationwide. But at the same time, the numerator considers only the remote jobs posted by the same firm in those other CZs. This can lead to substantial differences in magnitude. A detailed quantitative analysis is provided in Appendix A.

	count	mean	min	max	p25	p50	p75	p90
Traditional HHI	675763	3465.246	158.2979	10000	1136	2364.569	5000	10000
Remote-adjusted HHI	675763	897.111	.0001882	5656.316	100.9508	432.4993	1178.134	2451.62
Observations	675763							

Table 3: Summary Statistics (unweighted, 2015-2019)

Table 4 presents the summary statistics of the weighted HHI, using 2019 employment as the weight. The employment data is at the ONET-SOC6 - CBSA level. To obtain employment figures at the ONET-SOC6 - CZ level, I follow the methodology of Azar et al. (2020). First, I use county population estimation data from the Census to calculate the population share of each county within the same commuting zone. I then assume that total employment by occupation is distributed to each county based on population share. This assumption is credible if counties within the same commuting zone exhibit similar economic development and education levels. Afterward, I aggregate the county employment data to derive commuting zone-level employment for each ONET-SOC6 occupation. The two measurements yield similar results across the years (2015-2023). The increase in remote-adjusted HHI and the decrease in traditional HHI suggest that before COVID, markets with high employment levels had low traditional HHI but high remote-adjusted HHI.

	count	mean	min	max	p25	p50	p75	p90
Traditional HHI	1114296	669.453	36.22005	10000	240.4133	426.5789	780.7496	1411.133
Remote-adjusted HHI	1114296	685.6787	6.90e-08	9781.635	211.9146	409.4407	788.9755	1512.524
Observations	1114296							

Table 4: Summary Statistics (weighted using 2019 employment data, 2015-2023)

In Figure 8, I visualize the weighted average of remote-adjusted HHI across CZs. Figure 9 shows the weighted average of traditional HHI across CZs, while figure 10 presents the 2019 population estimates by county. As expected, before the pandemic, CZs with low populations exhibited high traditional HHI, consistent with the findings from Azar et al. (2022). In contrast, the remote-adjusted HHI displays a slightly more dispersed pattern.

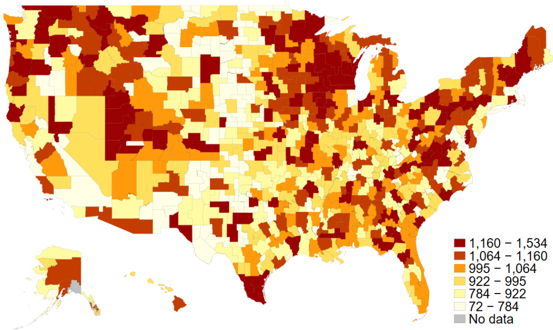


Figure 8: Average remote-adjusted HHI across CZ before COVID (2015-2019)

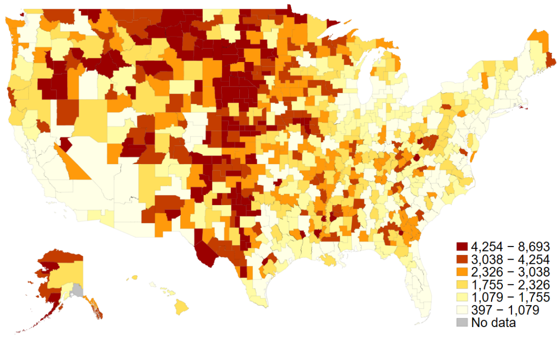


Figure 9: Average traditional HHI across CZ before COVID (2015-2019)

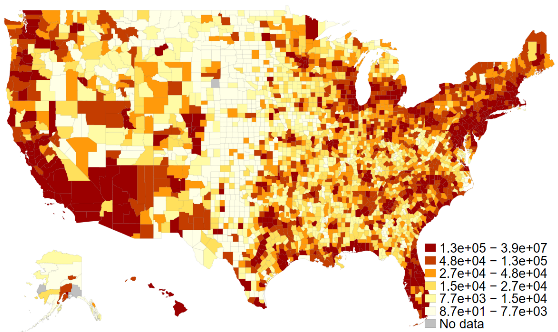


Figure 10: Estimated population by counties (2019)

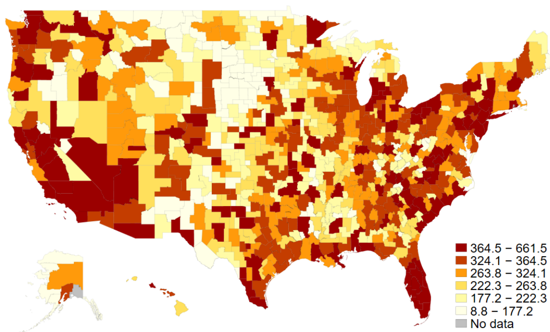


Figure 11: Average remote-adjusted HHI across CZ after COVID (2020-2023)

The contrast becomes more striking when examining the average remote-adjusted HHI across CZs after the pandemic, as shown in figure 11. The remote-adjusted HHI is positively correlated with population density. Suggesting these high-populated areas are also homes to firms that offer most remote jobs nationwide. Figure 12 compares my remote-adjusted HHI with the traditional HHI measurement across years, using 2019 market employment as the weight. Figure 13 provides evidence for this guess, showing that in larger markets, firms with a high share of remote job postings tend to be larger than those with a low share of remote job postings. The fact that the remote-adjusted HHI is lower than the traditional HHI after COVID-19 indicates that more firms began offering remote jobs post-pandemic, reducing labor market concentration.

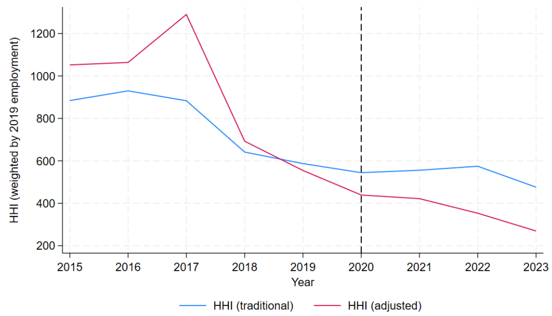


Figure 12: Comparison of traditional HHI vs. remote-adjusted HHI (weighted using 2019 employment)

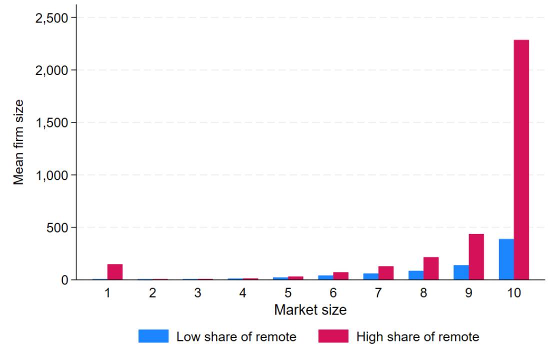


Figure 13: Remote jobs, firm size, and market size (2015-2019) (Markets are classified into ten deciles based on market (occupation-CZ) employment, firm size is based on number of remote job postings)

4.3 Analysis using Remote-Adjusted HHI

Table 5 presents the 2019 employment-weighted HHI across years using these two methods. Compared to Azar et al. (2022)'s estimate of the 2016 HHI at the occupation-CZ-quarter level, my weighted and unweighted values for 2016 are slightly lower. This discrepancy may be attributed to three differences in the filters we applied: 1) I excluded internship job postings, whereas Azar et al. (2022) does not explicitly mention this filter; 2) I applied a stricter filter on narrowly defined labor markets than Azar et al. (2022). Specifically, I used data from 2015 to 2023, ranking job postings from high to low for each year and flagging the top 90% of occupations. I then retained only those occupations flagged in all years, while Azar et al. (2022) applied this process using only 2016 data. This would lead to a lower average since the excluded narrowly defined labor markets will have extremely high concentration; 3) We used different years of employment data for weighting.

The weighted average remote-adjusted HHI is 685.6787, which falls below the low concentration threshold of 1,500 HHI indicated by the Department of Justice (DOJ) and the Federal Trade Commission's 2010 horizontal merger guidelines. This suggests that, on average, approximately 15 firms are recruiting in each market. Considering the average across years, 28% (44,805)[3] of markets exceed the 2,500 HHI high concentration threshold. The 90th percentile HHI is 1,512.524 across years and markets, corresponding to an average of 6.61 firms recruiting.

	count	mean	min	max	p25	p50	p75	p90
2015								
Traditional HHI	121045	883.4133	92.2568	10000	391.0671	615.567	1030.177	1727.778
Remote-adjusted HHI	121045	1053.094	.0000103	8226.068	479.5839	771.8877	1303.164	2130.748
2016								
Traditional HHI	121956	929.3015	95.47257	10000	413.5151	661.5792	1125.854	1828.35
Remote-adjusted HHI	121956	1069.869	5.47e-06	8748.178	453.2195	730.5641	1258.724	2157.611
2017								
Traditional HHI	122444	882.7476	80.58363	10000	364.6443	608.983	1063.99	1783.265
Remote-adjusted HHI	122444	1308.726	6.18e-06	9241.989	483.6157	858.2512	1550.379	2941.496
2018								
Traditional HHI	123158	636.7479	59.90124	10000	227.6827	381.1742	696.2963	1336.686
Remote-adjusted HHI	123158	694.4241	2.96e-06	8203.235	247.5026	398.468	699.3038	1541.217
2019								
Traditional HHI	131201	581.8831	52.10181	10000	210.3102	358.0329	651.1248	1184.867
Remote-adjusted HHI	131201	554.6352	2.53e-06	8336.484	257.3113	402.9262	629.4113	1092.627
2020								
Traditional HHI	122366	535.0484	43.20059	10000	187.0373	315.9155	560.147	1096.514
Remote-adjusted HHI	122366	443.8135	8.34e-07	8850.502	134.1187	258.5554	473.0795	889.9492
2021								
Traditional HHI	124196	542.636	36.22005	10000	164.1763	320.6789	620.0373	1180.556
Remote-adjusted HHI	124196	424.1701	1.78e-07	9781.635	114.1926	245.4099	506.2228	912.2512
2022								
Traditional HHI	124110	568.066	55.48668	10000	201.5097	370.9443	653.8942	1191.908
Remote-adjusted HHI	124110	353.5556	6.90e-08	8510.094	135.9237	226.6713	417.673	717.4464
2023								
Traditional HHI	123820	465.4693	48.10834	10000	169.268	298.1908	522.2402	928.68
Remote-adjusted HHI	123820	269.2896	3.27e-07	5126.758	107.0589	194.0696	313.3138	553.5803
Total								
Traditional HHI	1114296	669.453	36.22005	10000	240.4133	426.5789	780.7496	1411.133
Remote-adjusted HHI	1114296	685.6787	6.90e-08	9781.635	211.9146	409.4407	788.9755	1512.524
Observations	1114296							

Table 5: Summary stats by year (weighted)

The significant drop in my remote-adjusted HHI from 2017 to 2018 is surprising. To investigate which occupations drive this change, I plotted the remote-adjusted HHI across years by different teleworkable groups[4] in figure 14. The third teleworkable group is responsible for the substantial decline in concentration from 2017 to 2018.

[3]Here, I consider only occupation-commuting zone pairs; if I include years, there will be 72,836 markets, representing 6% of the total number of markets.

[4]Instead of using quartiles, I split the teleworkable groups based on specific cutoff points.

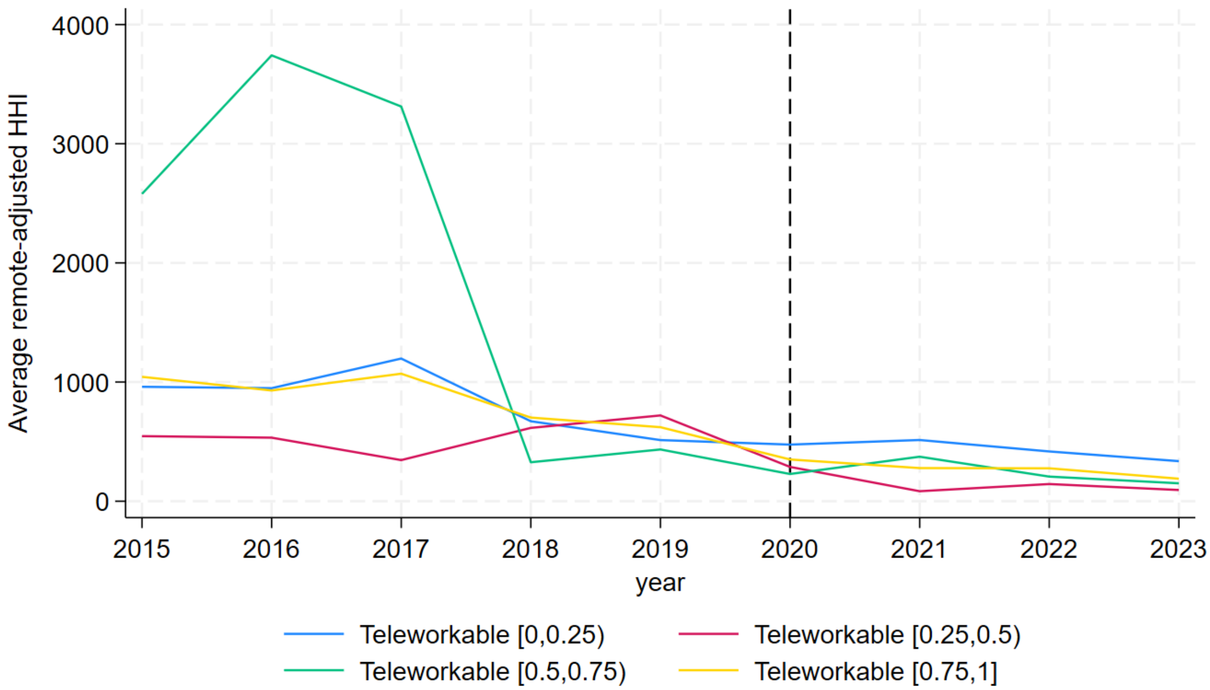


Figure 14: Weighted average remote-adjusted HHI across year and teleworkable groups

The occupations in the third teleworkable group include Transportation, Storage, and Distribution Managers (11-3071); Architects, Except Landscape and Naval (11-9041); Life, Physical, and Social Science Technicians, All Other (19-4099); First-Line Supervisors of Personal Service Workers (39-1021); Billing and Posting Clerks (43-3021); and Customer Service Representatives (43-4051). I plotted the remote-adjusted HHI evolution for these occupations in Figure 15. It is evident that Customer Service Representatives (43-4051) are driving the aggregate trend for the third teleworkable group, while First-Line Supervisors of Personal Service Workers (39-1021) exhibit a different pattern compared to other occupations in the group, the average HHIs of the rest do not change significantly after the pandemic. These results highlight the considerable heterogeneity across occupations.

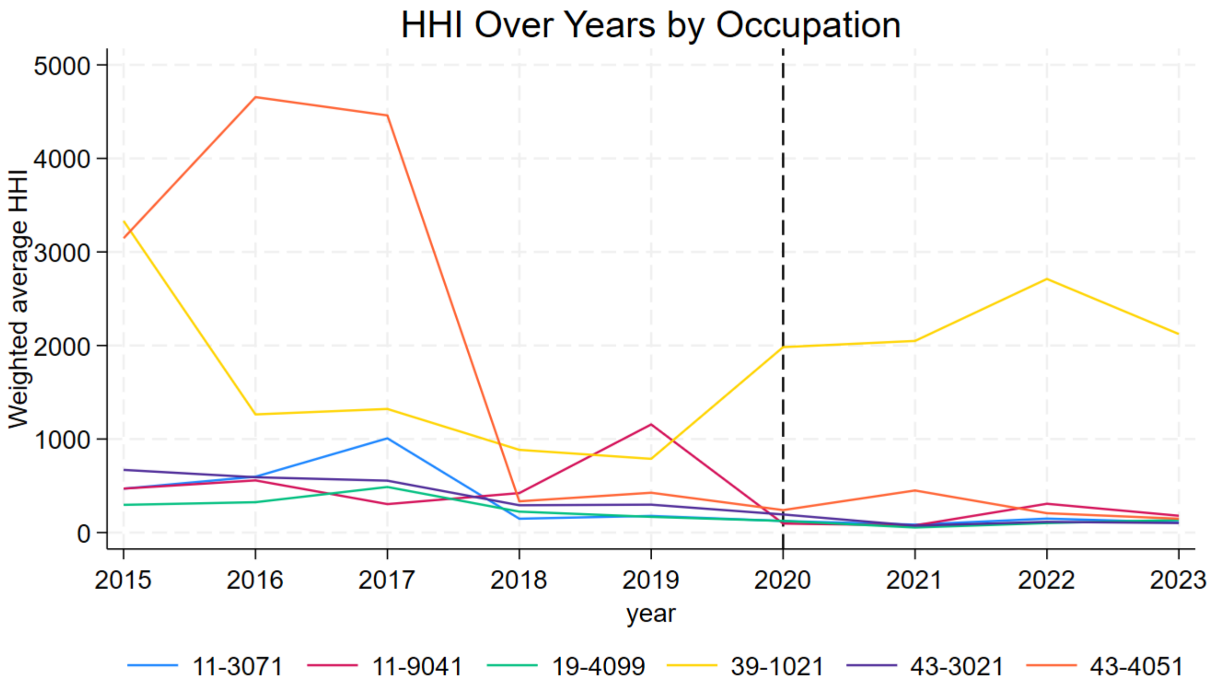


Figure 15: Weighted average remote-adjusted HHI across year and occupation (teleworkable: [0.5,0.75))

5 Teleworkable and Concentration

My intended analysis is to examine how adopting work-from-home (WFH) affects labor market concentration. There are two potential outcomes, as stated in the introduction. Ideally, I would like to compare the labor market concentration of two occupations: one with higher exposure to WFH and the other with lower exposure. I use COVID-19 as a natural experiment introducing variation in occupational WFH exposure. Specifically, I assign different teleworkability indices to occupations based on the work of Dingel and Neiman (2020). The teleworkability index is calculated from responses to two questionnaires conducted long before COVID-19: the Work Context Questionnaire and the Generalized Work Activities Questionnaire. These questionnaires assess physical working condition requirements, communication frequency, and other factors related to the feasibility of WFH for each occupation. While all occupations have the potential to offer remote jobs, those with a higher teleworkability index are more suitable for WFH. Therefore, when COVID hit, occupations with a higher teleworkability index were more likely to post remote jobs, creating the desired variation. To verify that this indeed occurred after COVID-19, I split the sample into two groups based on the occupational teleworkability index. The index itself is a continuous variable ranging from 0 to 1. If an occupation's teleworkability index is below or equal to the median, it belongs to the control group, and the treatment dummy is set to 0; if the index is above the median, it belongs to the treated group, and the treatment dummy is set to 1. In my data, the median value is 0. As a result, any occupations with a positive teleworkability index are classified into the treated group, while those with a null value are classified into the remote group. Figure 16 shows the dynamics of remote job posting shares across years for the two groups, and COVID indeed triggered the desired variation in the share of remote job postings across the treated and control groups.

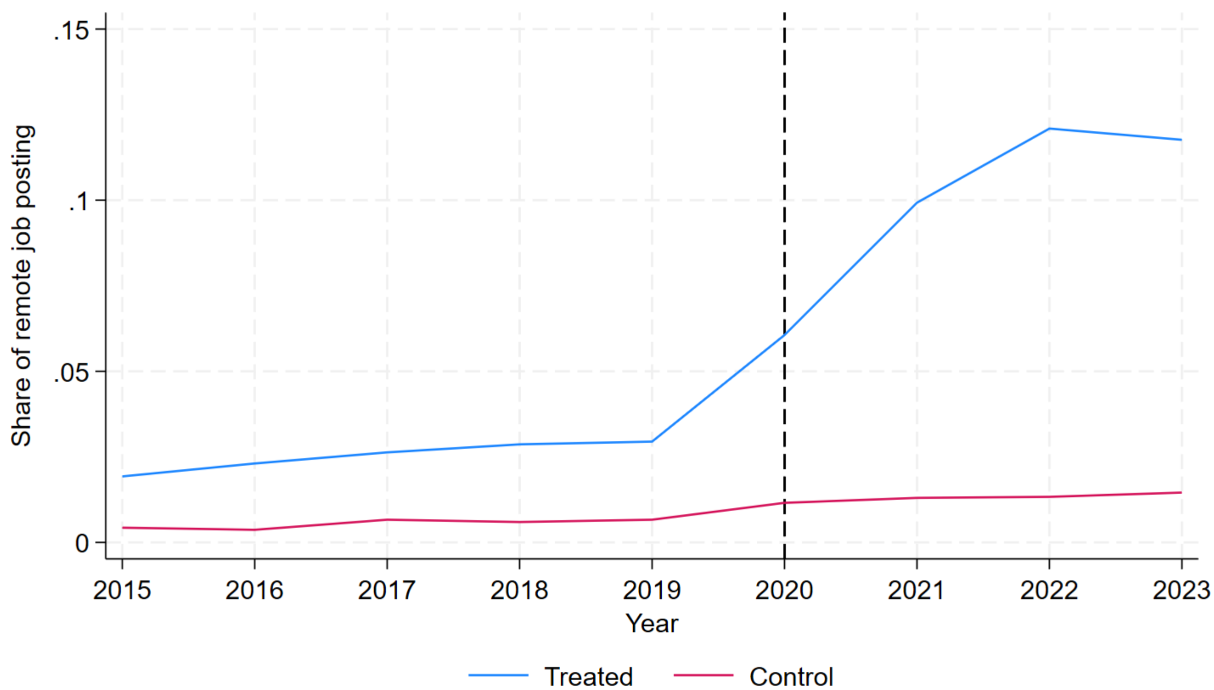


Figure 16: Average share of remote job postings by treated and control groups (weight: 2019 employment)

I run the following baseline regression using 2019 market employment as weight:

$$HHI_{zkt} = \alpha_{zk} + \lambda_{zt} + \sum_{j=-6}^{+6} \beta_j D_{kt}(t = s + j) + \sum_{j=-6}^{+6} X_{zkt} I(t = s + j)\delta_j + \epsilon_{zkt} \tag{3}$$

D_{kt} represents the interaction between the treatment dummy and the year dummies. α_{zk} denotes the occupation-CZ fixed effect (FE), while λ_{zt} indicates the CZ-year FE. X_{zkt} comprises control variables that influence market HHI and also correlate with D_{kt}. I define the treatment period as starting in 2020, following the World Health Organization (WHO)'s formal declaration of COVID-19 as a public health emergency in March. To mitigate the confounding effect of employment changes across years on my results, I use 2019 employment as the fixed weight. Figure 17 suggests the two groups have parallel trends in remote-adjusted HHI before the treatment. Figure 18 does the same using traditional HHI. It can be seen that the trends of the two groups are similar when using traditional HHI. However, the magnitude of changes varies across years, and the regression results using traditional HHI in Table **??** indicate a pre-trend. While the similar results between the remote-adjusted HHI and the traditional HHI suggest some common factors beyond WFH contribute to the observed effect observed in the DID analysis, WFH still plays a role.

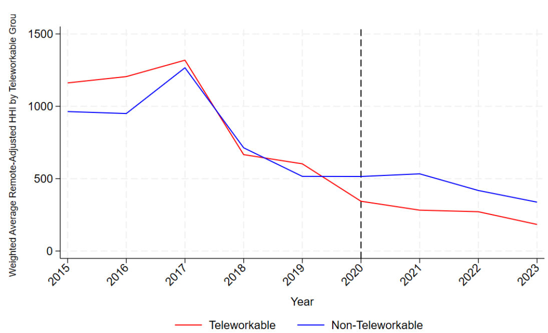


Figure 17: Remote-adjusted HHI across years for treated and control groups (weighted using 2019 employment)

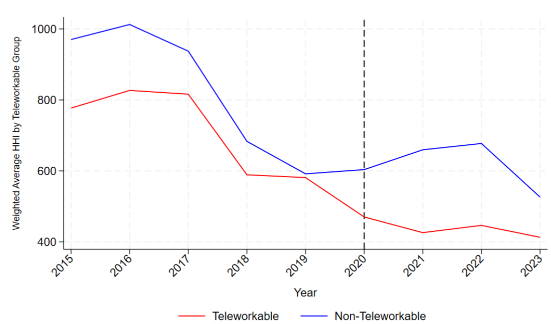


Figure 18: Traditional HHI across years for treated and control groups (weighted using 2019 employment)

My baseline regression results are presented in Table 6. Column (1) displays the result without controls using remote-adjusted HHI, while column (2) includes controls for the number of job postings by market-year and uses remote-adjusted HHI. I normalize the 2019 (t-1) coefficient to 0. The dependent variable is the level of remote-adjusted HHI. The event study plots for the two regressions are illustrated in Figures 19 and 20. None of the coefficients in the pre-period is statistically significant. Although there is a slight downward pre-trend according to the point estimates, this trend reverts starting in 2021, and the effect remains significantly negative after 2021. These results suggest that, compared to non-teleworkable occupations, teleworkable occupations, on average, exhibit 250-300 lower remote-adjusted HHI. The economic magnitude is very large, given the weighted average values of both the traditional and the remote-adjusted HHI are 600-800. Including controls for the number of job postings does not substantially alter the coefficients.

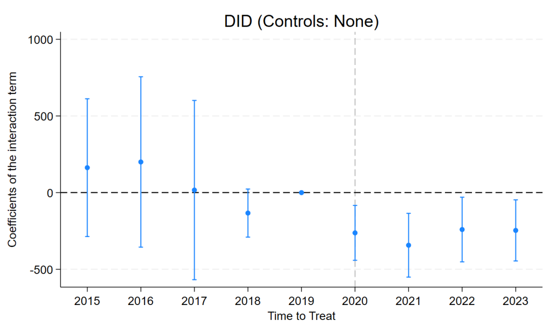


Figure 19: Event study plot (remote-adjusted HHI, no control)

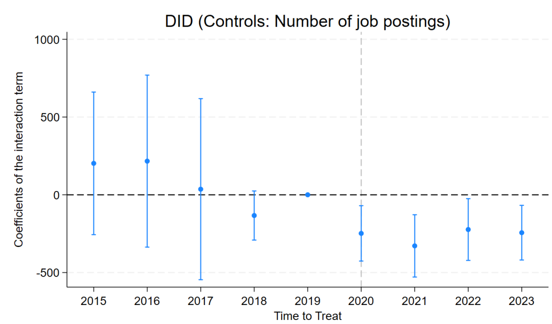


Figure 20: Event study plot (remote-adjusted HHI, control)

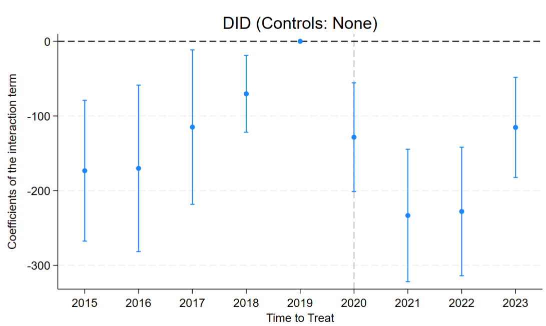


Figure 21: Event study plot (traditional HHI, no control)

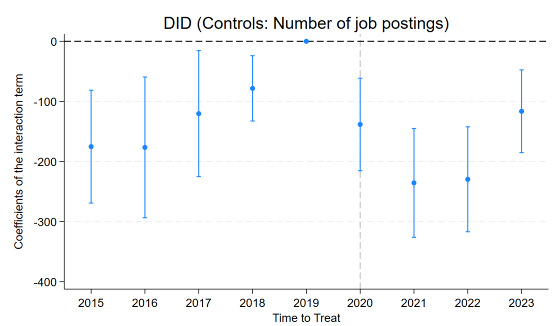


Figure 22: Event study plot (traditional HHI, control)

	Remote-adjusted HHI		Traditional HHI	
	(1)	(2)	(3)	(4)
	Without	With	Without	With
2015	162.6503	202.2257	-1.7e+02***	-1.8e+02***
	(228.038)	(232.678)	(47.8419)	(47.656)
2016	199.8837	216.4929	-1.7e+02***	-1.8e+02***
	(282.017)	(280.691)	(56.5528)	(59.4496)
2017	16.4367	36.1019	-1.1e+02***	-1.2e+02***
	(296.714)	(295.445)	(52.4912)	(53.2288)
2018	-1.3e+02***	-1.3e+02***	-70.3093***	-78.3424***
	(79.6226)	(80.1176)	(26.1117)	(27.5791)
constant	0	0	0	0
	(.)	(.)	(.)	(.)
2020	-2.6e+02***	-2.5e+02***	-1.3e+02***	-1.4e+02***
	(90.8088)	(90.3777)	(36.9376)	(39.1183)
2021	-3.4e+02***	-3.3e+02***	-2.3e+02***	-2.4e+02***
	(105.573)	(101.688)	(44.9903)	(45.9775)
2022	-2.4e+02***	-2.2e+02***	-2.3e+02***	-2.3e+02***
	(106.996)	(100.805)	(43.6647)	(44.3046)
2023	-2.5e+02***	-2.4e+02***	-1.2e+02***	-1.2e+02***
	(101.187)	(89.1937)	(34.0137)	(34.9158)
Constant	723.7086***	762.3409***	736.1232***	724.8509***
	(49.103)	(40.0304)	(13.3687)	(19.859)
ControlVar	NO	YES	NO	YES
CZ_Occupation_FE	YES	YES	YES	YES
CZ_Year_FE	YES	YES	YES	YES
N	1126996	1126996	1126996	1126996
F	3.4781***	10.9243***	4.2364***	6.2367***
r2_a	0.4478	0.4512	0.6909	0.6913
r2_within	0.0237	0.0299	0.0068	0.0081

Standard errors in parentheses
* $p < 0.1$, ** p<0.05, *** p<0.01

Table 6: Baseline regression

6 Concentration and Wage

To complete the analysis, I estimate the effect of concentration on wages using the remote-adjusted HHI. Given the observed decrease in concentration, the literature that finds a negative relationship between the two suggests a smaller downward pressure on wages (Barkai (2020), Qiu and Sojourner (2023), Azar et al. (2020), Schubert et al. (2024), etc.). To test this hypothesis, I follow Azar et al. (2022) by regressing wages on concentration. Additionally, I run a baseline regression using concentration as the dependent variable. The rationale for this approach is that if teleworkable occupations exhibit lower average HHI than non-teleworkable ones while also having higher wages, it could indicate that WFH reduces firms' labor market power, ultimately benefiting workers by increasing their wages.

Figure 23 is the binscatter plot of wages and concentration (both in level). The nonlinear pattern observed suggests that a nonlinear model (in variables) may be more appropriate for capturing the dynamics of this relationship. Figure 24 plots the same relationship for both remote-adjusted HHI and traditional HHI but using the logarithm versions of the variables.

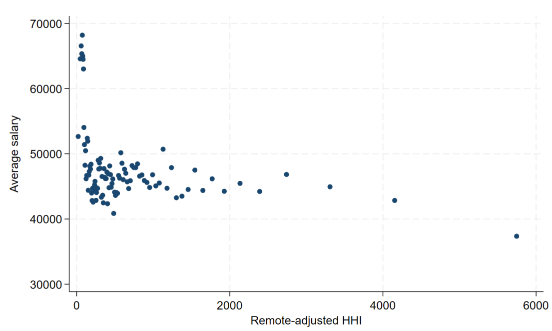


Figure 23: Wage and salary (weighted using 2019 employment, binscatter)

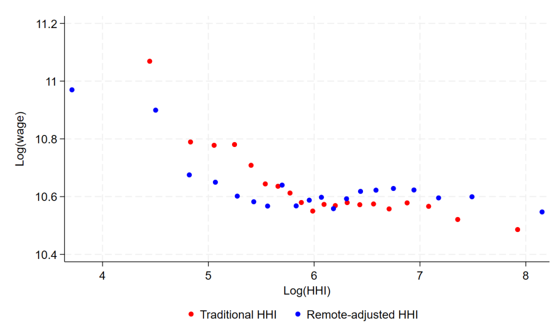


Figure 24: Wage and salary (weighted using 2019 employment, log, binscatter)

Azar et al. (2020) runs regression (5). However, my data does not have clear cutoffs of concentration such that the dependent and independent variables between adjacent cutoffs show a clear functional form.

$$\text{Log(wage)}_{zkt} = \alpha + \beta \text{Log(HHI)}_{zkt} + \delta\theta_{zkt} + \epsilon_{zkt} \tag{4}$$

To avoid model misspecification, I split HHI into three groups based on the DOJ classification of low, medium, and high concentration and ran regression (6). The remote-adjusted HHI result is shown in table 7 column (1). The dependent variable is the logarithm of wage. Surprisingly, I found that the higher the concentration, the higher the wage. Column (2) shows the same regression but using traditional HHI. It gives consistent results as prior literature.

$$\text{Log(wage)}_{zkt} = \beta_1 \mathbb{I}\left\{\text{Medium HHI}\right\}_{zkt} + \beta_2 \mathbb{I}\left\{\text{High HHI}\right\}_{zkt} + \delta\theta_{zkt} + \epsilon_{zkt} \tag{5}$$

	(1) Remote-adjusted HHI	(2) Traditional HHI
Medium HHI	0.0176***	-0.0017
	(0.0021)	(0.002)
High HHI	0.0058**	-0.0119***
	(0.0023)	(0.0029)
Constant	10.6349***	10.6366***
	(0.0002)	(0.0002)
Control Var	NO	NO
CZ_Occupation_FE	YES	YES
N	8.10E+05	8.10E+05
F	34.5187***	8.4128***
r2_a	0.9429	0.9428
r2_within	0.0012	0.0002

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Wage and concentration (HHI 3 groups)

The results may be biased by endogeneity. One potential omitted variable affecting both wage and HHI is productivity. Higher productivity may enable firms to expand, leading to increased HHI; simultaneously, higher productivity can raise the marginal productivity of labor, resulting in higher wages. While the literature suggests that higher concentration should correlate with lower wages, this negative effect could be offset by the upward bias from productivity. Thus, I observe positive results using remote-adjusted HHI. Regression using traditional HHI also suffers from this endogeneity problem. However, the fact that my result is negative despite the upward bias suggests that the unbiased result is also negative, aligning with prior literature.

I follow Azar et al. (2022) and address the issue with a 2SLS regression using the average of the logarithm of the inverse number of employers posting job ads in all other CZs for the same occupation in a given year (leave-one-out IV). This instrument captures variation in market concentration driven by national-level changes in occupational hiring over time rather than endogenous shifts in local productivity. One caveat is that there might be a nationwide/regional productivity shock. In this case, the increase in posting firms in other geographic markets due to productivity shock is also positively correlated with wage and concentration increases in the focal market. The proposed IV cannot take care of such endogeneity. With regression (5), I need two instruments for two endogenous variables (medium HHI, high HHI). However, I cannot find another instrument. I thus implement regression (5) with just one HHI dummy (if a market has above median HHI, it is classified into the high HHI group, otherwise it is classified into the low HHI group), and instrument it with the leave-one-out IV. I follow Azar et al. (2020) and add occupation-CZ FE. The first and second stage results are in table 8 column (1)(2). We can see from the first stage result that IV has the expected sign. The higher the number of employers in other markets, the higher the national demand and the lower the focal concentration. However, I found a positive effect in the second stage. The 2nd stage coefficient suggests that compared to the low concentration group, the high concentration group has, on average, 2.03% higher wage.

In columns (3)-(6), I split occupations based on their typical entry education requirement into two groups: occupations that require a bachelor's degree or higher (above-bachelor occupations) and occupations that require a degree less than bachelor's (below-bachelor occupations). Then, I run the same regression separately. Columns (3)-(6) show that higher concentration has a downward pressure on wages for below-bachelor occupations, as found in prior literature. Still, a higher concentration is accompanied by higher wages for the above-bachelor occupations. Since I already control demand-side factors (productivity), the latter suggests that supply-side factors also play a role. For example, while the concentration of above-bachelor occupations may increase, the supply of high-skilled labor might decline. I lack direct supply-side data to test this hypothesis, but other research shows that many developed countries faced labor shortages during the pandemic, primarily due to senior workers exiting the labor force Causa et al. (2022), Shibata (Shibata), Tavares (Tavares). According to European Employment Services (EURES), 60% of the top twenty shortage occupations were in skilled trades, 15% were healthcare professionals, and 10% were software professionals, with consistent shortages over time. In fast-growing IT and healthcare sectors, demand frequently outpaces supply. The reasons for the post-COVID high-skilled labor shortage, whether due to shifts in workers' working preferences or others, are worth further investigation.

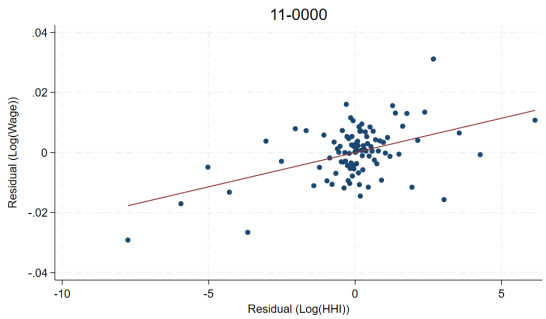
	(1)	(2)	(3)	(4)	(5)	(6)
	All Sample		Above Bachelor Degree		Below Bachelor Degree	
	D(Above Median HHI)	Log(Wage)	D(Above Median HHI)	Log(Wage)	D(Above Median HHI)	Log(Wage)
IV	0.428***		0.572***		0.363***	
	-0.00324		-0.00595		-0.00379	
D(Above Median HHI)		0.0198**		0.163***		-0.127***
		-0.00826		-0.00718		-0.0139
Constant	2.017***		2.317***		1.892***	
	-0.0109		-0.0189		-0.0131	
Observations	745,992	745,992	285,962	285,962	460,030	460,030
CZ_Occupation_FE	YES	YES	YES	YES	YES	YES
R-squared	0.366	0.016	0.319	-0.126	0.381	-0.291
Ftest IV	1438.103		1153.167		835.984	

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

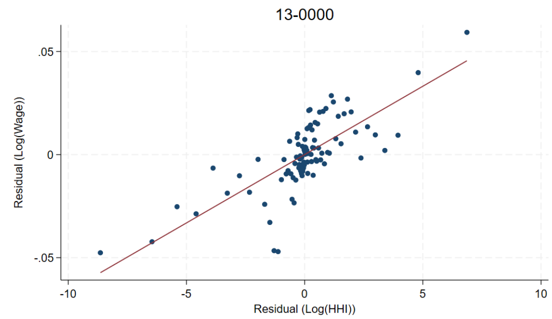
Table 8: 2SLS

To investigate whether specific occupations drive the pattern, Figure 25[5] shows the residual plot for each occupation with a positive correlation. As illustrated, these occupations are either high-skilled or healthcare-related positions with increased demand due to the pandemic.

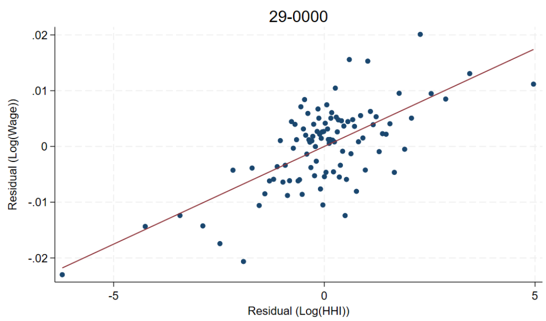
[5]Note here I did not use IV for the residual plots because the residuals from the second stage of 2SLS are not the residual of regressing the dependent variable on the original independent variable, but the residual of regressing dependent variable on the projection of independent variable onto the space of the instrument.



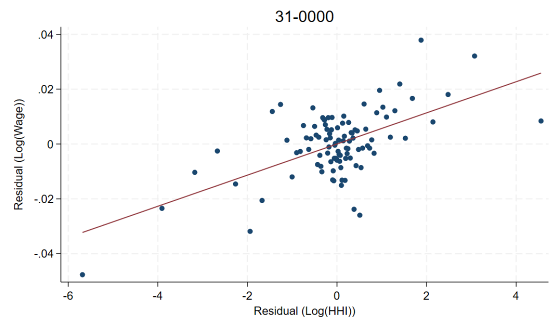
(a) Management Occupations



(b) Business and Financial Operations Occupations



(c) Healthcare Practitioners and Technical Occupations



(d) Healthcare Support Occupations

Figure 25: Overall caption for all subplots

I also tested the wage effect using my baseline specification, replacing the dependent variable with the logarithm of wage. Figures 26 and 27 present the results with and without controls respectively, detailed coefficients are shown in table 9. Teleworkable occupations exhibit lower wages than non-teleworkable occupations after COVID-19, which is consistent with the counter-intuitive findings obtained using 2SLS and the whole sample.

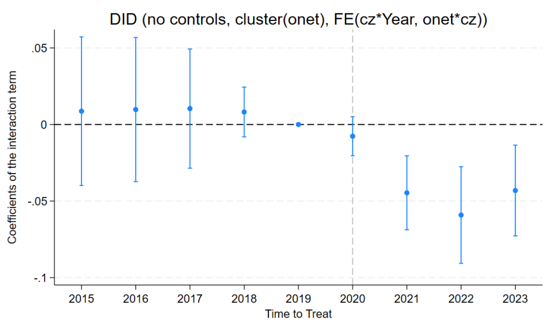


Figure 26: Wage difference-in-difference result (no control)

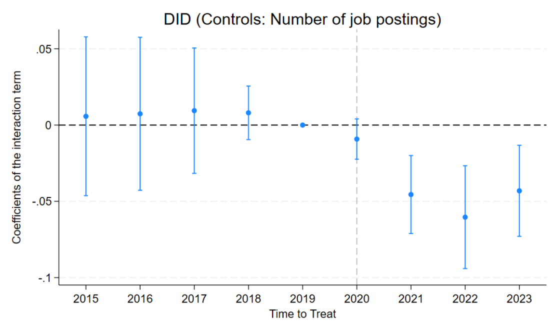


Figure 27: Wage difference-in-difference (control)

	(1) Without	(2) With
2015	0.0087	0.0057
	(0.0246)	(0.0264)
2016	0.0097	0.0074
	(0.0239)	(0.0254)
2017	0.0104	0.0094
	(0.0198)	(0.0209)
2018	0.0082	0.0080
	(0.0082)	(0.0089)
constant	0.0000	0.0000
	(.)	(.)
2020	-0.0077	-0.0092
	(0.0065)	(0.0067)
2021	-0.0446***	-0.0455***
	(0.0123)	(0.0130)
2022	-0.0591***	-0.0604***
	(0.0160)	(0.0171)
2023	-0.0431***	-0.0431***
	(0.0150)	(0.0151)
Constant	10.6421***	10.6396***
	(0.0024)	(0.0031)
ControlVar	NO	YES
CZ_Occupation_FE	YES	YES
CZ_Year_FE		YES
N	809594	809594
F	3.9705***	5.4463***
r2_a	0.9471	0.9472
r2_within	0.0149	0.0164

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Wage difference-in-difference

To determine whether specific occupations drive the trend, I replicated the difference-in-difference regressions for each 2-digit ONET-SOC occupation. The results, shown in table 10, include only the occupations with observations in both the treated and control groups and insignificant coefficients. As evident, there is significant heterogeneity across broad occupational categories. There is no clear pattern regarding education requirements or skill levels when using the DID method. Additional data on the supply side is required to further clarify the driving forces behind the observed effects.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	11-0000	13-0000	17-0000	19-0000	25-0000	27-0000	29-0000	41-0000	43-0000
2015	0.0433	0.1410**	-0.0723***	-0.0083	0.0126	0.1699**	0.0301	-0.1810***	0.0160
	(0.0406)	(0.0627)	(0.0198)	(0.0162)	(0.0260)	(0.0607)	(0.0334)	(0.0552)	(0.0271)
2016	0.0375	0.1430**	-0.0525**	0.0144	0.0417*	0.1309***	0.0451***	-0.1322**	0.0137
	(0.0374)	(0.0652)	(0.0184)	(0.0233)	(0.0194)	(0.0317)	(0.0146)	(0.0478)	(0.0251)
2017	0.0419	0.0810	-0.0379**	0.0224	0.0525	0.0661*	0.0211	-0.0657	0.0068
	(0.0444)	(0.0525)	(0.0147)	(0.0178)	(0.0322)	(0.0324)	(0.0193)	(0.0759)	(0.0232)
2018	0.0084	0.0188	-0.0336**	-0.0287*	-0.0203	0.0144	0.0031	0.0463	-0.0036
	(0.0166)	(0.0182)	(0.0104)	(0.0098)	(0.0190)	(0.0111)	(0.0147)	(0.0432)	(0.0097)
constant	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
2020	0.0134	-0.0486***	0.0311**	-0.0214	0.0258**	0.0228	0.0136	-0.0049	-0.0065
	(0.0105)	(0.0162)	(0.0116)	(0.0162)	(0.0088)	(0.0175)	(0.0119)	(0.0275)	(0.0139)
2021	0.0366**	-0.1463**	-0.0055	0.0990	0.0107	-0.0189	-0.0282	-0.0604	-0.0327**
	(0.0147)	(0.0667)	(0.0171)	(0.0526)	(0.0101)	(0.0334)	(0.0421)	(0.0367)	(0.0141)
2022	0.0064	-0.1521**	-0.0465***	0.1667**	0.0527*	-0.0132	-0.0815	-0.1466**	-0.0175
	(0.0123)	(0.0635)	(0.0094)	(0.0489)	(0.0230)	(0.0327)	(0.0543)	(0.0590)	(0.0102)
2023	0.0337*	-0.0635	-0.0592***	0.0732	-0.0545***	-0.0103	-0.0564	-0.1961*	-0.0045
	(0.0170)	(0.0389)	(0.0069)	(0.0384)	(0.0162)	(0.0448)	(0.0477)	(0.0921)	(0.0116)
Constant	11.2302***	10.9921***	11.1056***	10.9855***	10.6650***	10.6072***	11.1385***	10.5125***	10.3515***
	(0.0088)	(0.0143)	(0.0039)	(0.0077)	(0.0096)	(0.0166)	(0.0024)	(0.0036)	(0.0054)
ControlVar	YES	YES	YES	YES	YES	YES	YES	YES	YES
CZ_Occupation_FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
CZ_Quarter_FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
N	83279	60160	35031	9295	25972	23441	104619	52774	83972
F	9.2333***	133.9629***	.**	.**	.***	.*	63.8464***	.	24.4957***
r2_a	0.8454	0.7782	0.9191	0.7112	0.8121	0.8574	0.9372	0.9097	0.8952
r2_within	0.0055	0.0589	0.0420	0.0475	0.0045	0.0562	0.0850	0.1063	0.0078

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Wage difference-in-difference (by ONET-SOC2)

7 Conclusion

Labor market concentration has been a key topic in literature due to its strong connection with firms' labor market power. The latter affects the gap between workers' marginal revenue productivity and their wages. Given the importance of intangible human capital, labor market concentration can also affect firms' innovation capacity and broader operational efficiency. It is thus important to accurately measure labor market concentration, especially with the rise of remote work, which has changed the boundary of the local labor market.

In this paper, I adopt the traditional HHI index by incorporating remote job postings, building on the framework developed by Azar et al. (2020). Using BGT job posting data, I demonstrate that when adjusted for remote jobs, labor market concentration reduces significantly compared to the traditional measurement. I also re-estimate the impact of labor market concentration on wages. Using the 2SLS method, I find occupations with above-bachelor education requirements at entry have higher wages when the concentration is higher. The results are driven by Management Occupations (11-0000), Business and Financial Operations Occupations (13-0000), Healthcare Practitioners and Technical Occupations (29-0000), and Healthcare Support Occupations (31-0000). Using the DID method, I find substantial heterogeneity in this relationship across occupations. Given that I already control demand-side factors, my results highlight the need for supply-side data for further investigation. My results suggest some occupations have labor demand outpace labor supply post-COVID, whether this is due to changes in industry structure or changes in workers' preferences are interesting question for future research.

8 Appendix

8.1 Appendix A: Explanation of the Difference between Traditional HHI and Remote-Adjusted HHI

The difference observed before COVID arises from the fact that, in the remote-adjusted HHI, the denominator accounts for all remote jobs posted by all firms in other commuting zones nationwide, while the numerator considers only the remote jobs posted by the same firm in those other commuting zones. This can lead to substantial differences in magnitude. To illustrate this, we can express the remote-adjusted share of firm j in commuting zone z for occupation k in year t (s_{zkjt}) as follows:

$$s'_{zkjt} = \frac{a+x}{b+y} \tag{6}$$

Where the firm market share calculated using the traditional method is expressed as $s_{zkjt} = \frac{a}{b}$. Let x represent the number of remote jobs posted by firm j in other commuting zones at time t, and y denote the total number of remote jobs posted by all firms in those commuting zones nationwide at time t. Consequently, we have the relationship $\frac{s'_{zkjt}}{s_{zkjt}} = \frac{b(a+x)}{a(b+y)}$. Therefore, if $\frac{a}{b} < \frac{x}{y}$, then $s' > s$; if $\frac{a}{b} > \frac{x}{y}$, then $s' < s$. Since I aim to illustrate how the increase in the denominator reduces the remote-adjusted share relative to the traditional one, I will focus on observations where the remote-adjusted share is lower. Notably, these observations account for 95% of all data points from 2015 to 2019. Among them, 90% have $\frac{x}{y} = 0$ because $x = 0$. However, the value of y can be quite large, as shown in table 11. I emphasize these observations because they represent the majority of the sample.

	count	mean	min	max	p25	p50	p75	p90
y	2.54e+07	7757.355	1	62869	480	1605	8151	27609
Observations	25394113							

Table 11: Summary Statistics of y ($x = 0$, $s'_{zkjt} < s_{zktj}$, 2015-2019)

To further illustrate how an increase in y would decrease s'_{zkjt} relative to s_{zkjt}. I split observations into five groups based on the percentiles of s_{zkjt}. The distribution of y for each group is presented in table 12.

	count	mean	min	max	p25	p50	p75	p90
1	5379957	18006.43	5	62813	2945	10387	28851	52704
2	5287703	7732.757	2	62849	810	2210	8745	24534
3	5189633	5345.585	2	62867	446	1214	4328	15832
4	5032236	3918.672	1	62869	297	885	2755	10401
5	4504584	2612.344	1	62869	150	544	1602	5202
Total	2.54e+07	7757.355	1	62869	480	1605	8151	27609
Observations	25394113							

Table 12: Summary Statistics of y ($x = 0$, $s'_{zkjt} < s_{zktj}, 2015 - 2019$)

I then plot $\frac{s'_{zkjt}}{s_{zkjt}}$ against y for each group, as shown in figure 28. The legend indicates the range of traditional shares for each group. It is evident that for firms with a small traditional s_{zkjt} (0 - 0.026), if there are no remote jobs from the same firm elsewhere but the same occupation has 20,000 remote jobs elsewhere, the new remote-adjusted share s'_{zkjt} would only be 40% of the traditional share. Since HHI is the sum of the squares of the shares, this gap will be further amplified when we take the square. When s_{zkjt} is large (e.g., close to 1, as indicated by the purple line), the diluting effect of a large y is even more pronounced.

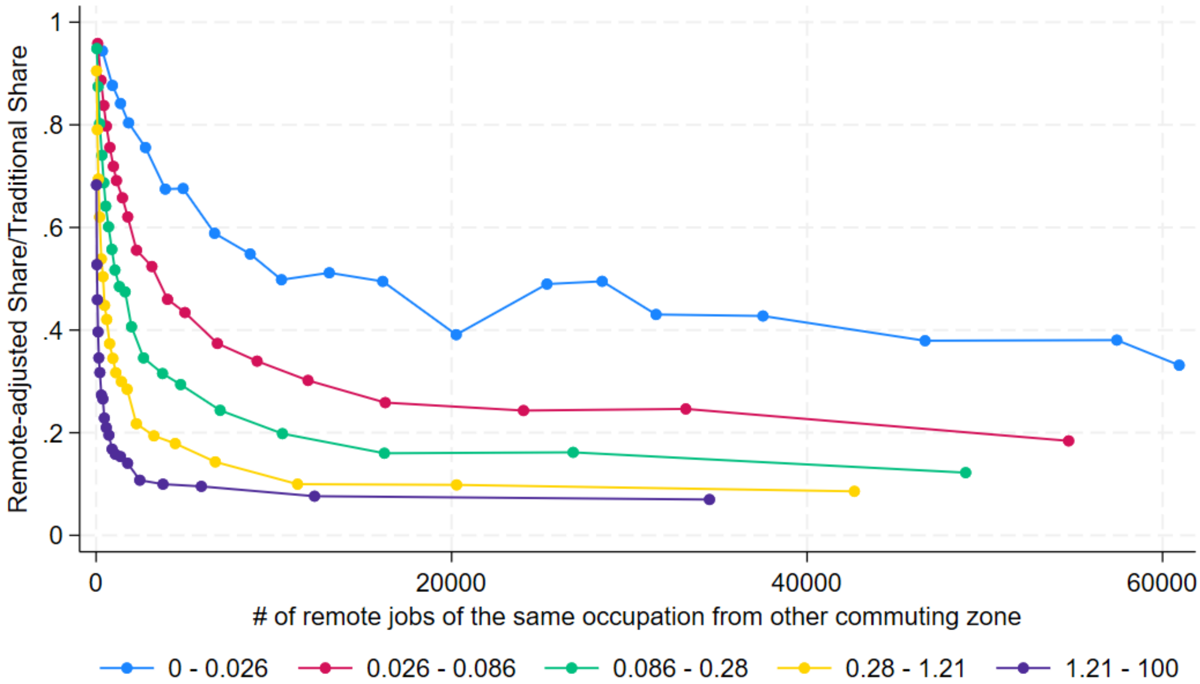


Figure 28: Remote-adjusted share relative to traditional share ($\frac{s'_{zkjt}}{s_{zktj}}$) across different numbers of remote jobs in other commuting zones of the same occupation by different levels of traditional share (y) ($x = 0$, $s'_{zkjt} < s_{zktj}$, 2015-2019)

Figure 29 focuses on the segment where $y < 1600$, centering around the whole sample median value of y. Figures 28 and 29 collectively illustrate why my remote-adjusted HHI is significantly smaller than the traditional HHI. They show the differences between traditional and remote-adjusted firm shares based on varying numbers of remote job postings for the same occupation in different geographic markets and the level of traditional share. Table **??** reveals that for firms with a small traditional share (group 1), the median y is 10,387, corresponding to a share ratio of approximately 50%. Consequently, when squared, the remote-adjusted share represents only 25% of the traditional share. In contrast, for firms with a large traditional share (group 5), the median y is 544, corresponding to a share ratio of about 20%, resulting in the remote-adjusted share being only 4% of the traditional share when squared.

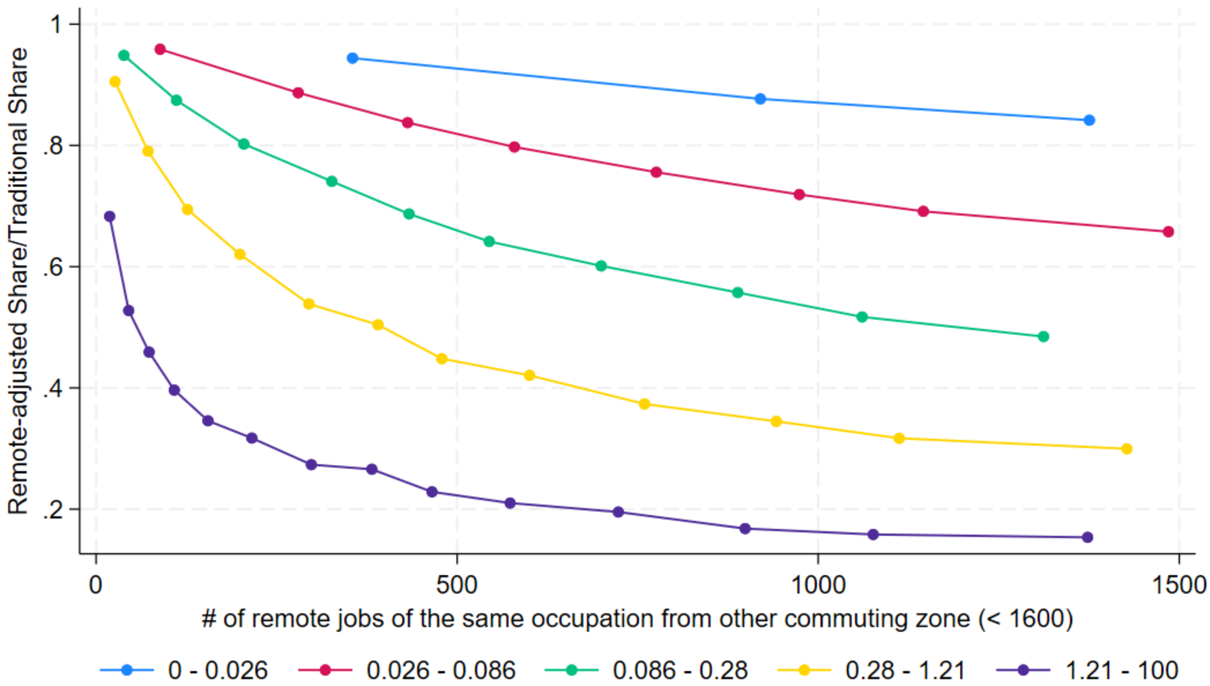


Figure 29: Remote-adjusted share relative to traditional share across a different number of remote jobs in other commuting zones of the same occupation by different levels of traditional share ($x = 0$, $s'_{zkjt} < s_{zktj}$, $y < 1600$)

This argument helps explain why my remote-adjusted HHI differs significantly from the traditional HHI even before the COVID-19 pandemic.

8.2 Appendix B: Education Adjustment

In this appendix, I present the results of refining my remote-adjusted HHI measurement by weighting remote job postings according to the probability that workers in a specific market can meet the job's typical entry education requirements. Specifically, if an occupation requires a minimum of a bachelor's degree, only residents with a bachelor's degree or higher are considered potential matches. For occupations with no (low) educational requirement, all residents (all residents with degrees above the requirement) are considered eligible. However, my primary analysis assumes segmented labor markets, where workers with advanced degrees tend to avoid jobs typically requiring lower qualifications (later referred to as "highly segmented"). I also consider a scenario where all higher-educated individuals are willing to accept jobs with lower entry requirements (later referred to as "weakly segmented"), but this alternative adjustment has minimal impact on HHI concentration.

Specifically, my education-adjusted concentration formula is as follows. Where η_{ez} is the share of the population in the focal CZ that matches the focal occupation's education requirement.

$$HHI_{\text{remote-adjusted},zkt} = \sum_{j \in \mathrm{M_{zkt}}} \left(\frac{n_{\mathrm{zjkt,nrm}} + n_{cjkt,rm,c!=z} \times \eta_{ezk}}{\sum_{i \in \mathrm{M_{zkt}}} (n_{zikt,nrm} + n_{cikt,rm,c!=z} \times \eta_{ezk})} \times 100 \right)^2 \tag{7}$$

Figures 30 and 31 show the distribution comparison between education-adjusted and unadjusted HHI. It can be seen that education adjustment does not alter much of my concentration measurement. Therefore, this adjustment does not change the results of my further analysis above.

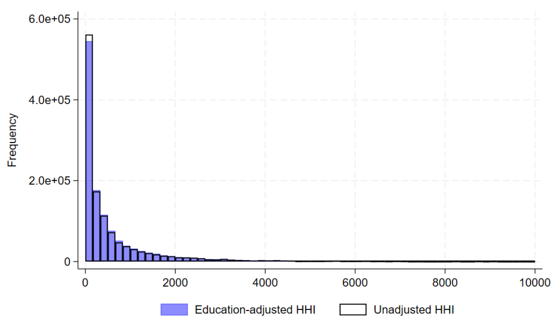


Figure 30: Distribution of education-adjusted HHI (weakly segmented) and unadjusted HHI

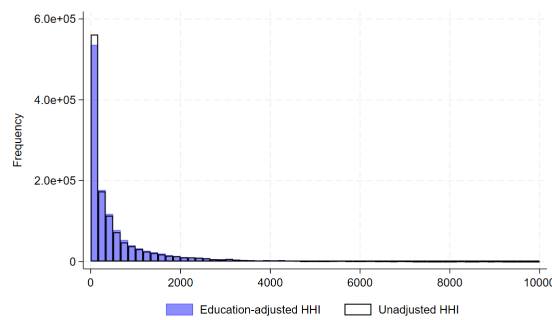


Figure 31: Distribution of education-adjusted HHI (highly segmented) and unadjusted HHI

References

Acemoglu, D., D. Autor, J. Hazell, and P. Restrepo (2022, April). Artificial Intelligence and Jobs: Evidence from Online Vacancies. *Journal of Labor Economics 40*(S1), S293–S340. Publisher: The University of Chicago Press.

Azar, J., I. Marinescu, and M. Steinbaum (2022, April). Labor Market Concentration. *Journal of Human Resources 57*(S), S167–S199. Publisher: University of Wisconsin Press Section: Articles.

Azar, J., I. Marinescu, M. Steinbaum, and B. Taska (2020, October). Concentration in US labor markets: Evidence from online vacancy data. *Labour Economics 66*, 101886.

Barkai, S. (2020). Declining Labor and Capital Shares. *The Journal of Finance 75*(5), 2421–2463. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12909.

Bartik, A. W., M. Bertrand, F. Lin, J. Rothstein, and M. Unrath (2020, July). Measuring the labor market at the onset of the COVID-19 crisis.

Benmelech, E., N. Bergman, and H. Kim (2018, February). Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages?

Benmelech, E., N. K. Bergman, and H. Kim (2022, April). Strong Employers and Weak Employees How Does Employer Concentration Affect Wages? *Journal of Human Resources 57*(SpecialIssue 1), S201–S250.

Berger, D., K. Herkenhoff, and S. Mongey (2019, March). Labor Market Power. pp. w25719. Place: Cambridge, MA Publisher: National Bureau of Economic Research Report Number: w25719.

Berger, D., K. Herkenhoff, and S. Mongey (2022, April). Labor Market Power. *American Economic Review 112*(4), 1147–1193.

Braxton, J. C. and B. Taska (2023, February). Technological Change and the Consequences of Job Loss. *American Economic Review 113*(2), 279–316.

Causa, O., M. Abendschein, N. Luu, E. Soldani, and C. Soriolo (2022, July). The post-COVID-19 rise in labour shortages.

Coibion, O., Y. Gorodnichenko, and M. Weber (2020, April). Labor Markets During the COVID-19 Crisis: A Preliminary View.

Dingel, J. I. and B. Neiman (2020, April). How Many Jobs Can be Done at Home?

Gallant, J., K. Kroft, F. Lange, and M. J. Notowidigdo (2020, October). Temporary Unemployment and Labor Market Dynamics During the COVID-19 Recession.

Goldfarb, A., B. Taska, and F. Teodoridis (2023, January). Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings. *Research Policy 52*(1), 104653.

Jarosch, G. (2021, February). Searching for Job Security and the Consequences of Job Loss.

Macaluso, C., B. Hershbein, and C. Yeh (2019). Concentration in U.S. local labor markets: evidence from vacancy and employment data. *2019 Meeting Papers*. Number: 1336 Publisher: Society for Economic Dynamics.

Posner, E. A. (2021, September). Labor Monopsony in the United States. In E. A. Posner (Ed.), *How Antitrust Failed Workers*, pp. 0. Oxford University Press.

Prager, E. and M. Schmitt (2021, February). Employer Consolidation and Wages: Evidence from Hospitals. *American Economic Review 111*(2), 397–427.

Qiu, Y. and A. Sojourner (2023). Labor-market concentration and labor compensation. *ILR Review 76*(3), 475–503. ISBN: 0019-7939 Publisher: Sage Publications Sage CA: Los Angeles, CA.

Rinz, K. (2018). Labor Market Concentration, Earnings Inequality, and Earnings Mobility.

Schubert, G., A. Stansburry, and B. Taska (2024). Employer Concentration and Outside Options.

Shibata, Ippei, C. P. Has COVID-19 Induced Labor Market Mismatch? Evidence from the US and the UK.

Tavares, Yi Ji, L. L. M. O. C. P. I. S. A. S. K. S. M. M. R. A. D. Labor Market Tightness in Advanced Economies.

Yeh, C., C. Macaluso, and B. Hershbein (2022, July). Monopsony in the US Labor Market. *American Economic Review 112*(7), 2099–2138.