Product Market Concentration, Wage Inequality and Worker Sorting*

Guido Matias Cortes[†] — York University

Leyla Gilgen[‡] — University of Bern

Jeanne Tschopp[§] — University of Bern

This version: September 2023

Abstract

Product markets have become increasingly dominated by a smaller number of firms with high market shares. At the same time, wage dispersion between firms has been increasing. In this paper, we show that product market concentration is associated with higher wage dispersion between firms within industries. Using rich administrative data from France covering the near-universe of workers and firms over the period 2009-2019, we find a positive and statistically significant correlation between sectoral concentration and different measures of between-firm wage inequality. The relationship is driven by (i) increased sorting of workers in high-paying occupations towards more productive firms within industries, and (ii) higher wage differentials between more and less productive firms in more concentrated industries, even conditional on their workers' occupations. In a model that features wage heterogeneity between firms, a shock to consumer price sensitivity – which has been posited by Autor et al. (2020) as a driver of the rise in product market concentration – generates predictions that are consistent with the empirical patterns that we document.

Keywords: Concentration, Wage Inequality, Sorting, Firm Heterogeneity

JEL Codes: J31, J23, L11, L13

*We thank workshop participants at the University of Bern and participants at the Young Swiss Economists Meeting 2023 and the 2023 Annual Congress of the Swiss Society of Economics and Statistics for valuable comments and suggestions. We are grateful to CASD for providing us access to the data.

[†]E-mail: gmcortes@yorku.ca

[‡]E-mail: leyla.gilgen@unibe.ch

[§]E-mail: jeanne.tschopp@unibe.ch

1 Introduction

The overall increase in wage inequality observed in many countries over recent decades has drawn a lot of attention from researchers and policymakers. Recent literature has shown that a large share of the increase in inequality can be attributed to increasing wage dispersion between firms (e.g. Card et al. 2013; Barth et al. 2016; Card et al. 2018; Song et al. 2018). At the same time, product markets have become increasingly dominated by a smaller number of firms (Autor et al. 2017; Bajgar et al. 2019; Grullon et al. 2019; Autor et al. 2020; Bighelli et al. 2021). Recently, Cortes and Tschopp (2023) have argued that the rise in product market concentration and the rise in between-firm wage inequality are linked. Using European data at the industry-country-year level, they show that markets that feature higher levels of concentration also exhibit higher levels of between-firm wage inequality.

This paper deepens the analysis of Cortes and Tschopp (2023) by exploring the micro-level mechanisms underlying the link between product market concentration and between-firm wage inequality. We use rich administrative micro-data from France covering the near-universe of French workers and firms over the period 2009-2019. The dataset includes information on the characteristics of workers, including their earnings, hours, age, gender, and occupation, as well as firm-level outcomes such as sales and value added. Our final sample contains nearly 13 million firm-year observations. We focus on the patterns observed across firms within detailed 4-digit industries.

Using an econometric strategy that exploits differential changes in concentration over time within industries for identification, and consistent with Cortes and Tschopp (2023), we find that higher product market concentration – as measured by the Herfindahl-Hirschman Index (HHI) or the share of sales of the top 10 firms within an industry – is associated with higher wage inequality across firms within industries. We show that wages are higher throughout the entire firm-wage distribution when concentration is higher, but particularly so at higher percentiles – thus accounting for the higher levels of inequality.

We then explore the mechanisms that account for the observed link between concentration and inequality. Given the well-established link between productivity and wages, we first explore whether the increased levels of wage inequality in more concentrated markets can be attributed to higher levels of productivity dispersion.

We find no evidence of a statistically significant relationship between concentration and the variance of productivity within industries; if anything, our estimated coefficients suggest that higher concentration is associated with lower levels of productivity dispersion. Thus, we can rule out the possibility that the higher level of inequality observed in more concentrated markets is due to higher levels of heterogeneity in firm-level productivity. In other words, more concentrated industries feature higher levels of inequality in spite of having similar (or even slightly lower) levels of dispersion in the underlying firm productivity distribution.

In order to further understand this pattern, we then explore whether the relationship between productivity, sales, employment, and wages varies with the degree of industry concentration. Using only variation across firms within industry-year cells, we find that, indeed, the relationship between productivity and other firm-level outcomes is stronger in more concentrated markets. In particular, while more productive firms have higher sales than less productive firms within all industries, this relationship is stronger in more concentrated industries. The same is true for employment: more productive firms are disproportionately larger than less productive firms when they are in more concentrated industry-year cells. The positive relationship between productivity and wages is also stronger when concentration is higher. These results imply that higher concentration is associated with higher levels of dispersion across firms in terms of sales, employment and wages, *conditional on productivity*, and that between-firm wage inequality is higher in more concentrated industries because firms that are relatively more productive within those industries tend to be disproportionately larger and pay disproportionately higher wages.

In order to explore the reasons why more productive firms pay disproportionately higher wages in more concentrated industries, we analyze the patterns of occupational sorting.

We find that roughly half of the stronger productivity-wage nexus in more concentrated industries is driven by the stronger sorting of individuals in high-wage occupations towards high-productivity firms in these industries. High productivity firms in more concentrated industries employ a disproportionately higher share of workers in abstract (high-paying) occupations, and a disproportionately lower share of workers in routine and manual (lower-paying) occupations.

We conclude by showing that the empirical patterns that we have documented can be rationalized by embedding a shock to consumer price sensitivity – a shock that has been posited as a driver of concentration by Autor et al. (2020) – within the framework of Helpman et al. (2010), which features wage inequality between firms and allows for worker sorting along occupational dimensions. An increase in consumer price sensitivity in the model leads to higher inequality in average firm wages due to (i) increased sorting of workers in high-wage occupations to high-productivity firms, and (ii) increased dispersion in firm wages conditional on occupation.

Our key contribution to the literature is to analyze the firm-level patterns underlying the link between product market concentration and between-firm wage inequality. While the link between these two outcomes was established by Cortes and Tschopp (2023) using data at a much higher level of aggregation, that paper was unable to pin down the mechanisms that drove this correlation. Our analysis allows us to provide a rich exploration of the underlying micro-level patterns, highlighting the importance of sorting along occupational dimensions as a major driver of the higher levels of between-firm wage inequality observed in more concentrated markets.

Our results are consistent with Akerman (2021) who, using detailed employer-employee data from Sweden, finds evidence of a strong correlation between a firm's size and its share of workers that are college educated. He shows that an increase in a sector's market concentration is associated with an increase in relative demand for skilled workers, due primarily to the reallocation of production from less skill-intensive small firms to more skill-intensive large firms. Our results imply that rising concentration has broader implications for inequality that go beyond changes in relative skill demand, as they are also associated with rising wage differences between firms within occupations.

Our paper also contributes to the broader literature that studies the effect of concentration on labor market outcomes (e.g. Azar et al. (2020); Rinz (2020); Benmelech et al. (2020)). Several papers, including Azkarate-Askasua and Zerecero (2021), Bassanini et al. (2021), and Marinescu et al. (2021), have focused particularly on the case of France.

Webber (2015) and Rinz (2020) study the effect of labor market concentration on wage inequality and find a positive relationship between local labor market concentration and higher wage inequality. Our results are consistent with their findings even if our study differs with respect to the country (France versus US) and the type of concentration (product market concentration rather than local labor market concentration). While a large part of the literature focuses on (local) labor market concentration, we focus on the product market. Our analysis is therefore also related to Prager and Schmitt (2021) and Arnold (2021) who find negative labor market effects as a result of an increase in product market concentration. Labor and product markets can, but do not necessarily overlap. Deb et al. (2022) derive a theoretical model to quantify the macroeconomic implications of firm heterogeneity, labor and product market power on wage inequality. They find that a less competitive environment substantially contributes to higher between-firm wage inequality.

We also contribute to the literature on firm heterogeneity, especially to the strand of the literature which highlights the importance of higher between-firm wage differentials contributing to the rise in wage inequality (e.g. Card et al. 2013; Barth et al. 2016; Card et al. 2018; Song et al. 2018). Through our analysis of firm-level mechanisms, we show that higher product market concentration is not only related to wage and employment outcomes, but also to between-firm wage inequality. This is also consistent with Bighelli et al. (2021) who show that higher concentration in Europe is positively associated with changes in productivity through a reallocation of economic activity towards large and concentrated industries.

The remainder of the paper is structured as follows. Section 2 describes the data and the construction of our measures of concentration and wage inequality. Section 3 presents the main results and analyzes the underlying firm-level mechanisms. Section 4 discusses a theoretical framework that can rationalize our results, while Section 5 concludes.

2 Data

2.1 Data Sources and Sample Restrictions

Our analysis uses two administrative datasets from France: the job-level BTS-Postes (formerly known as DADS-Postes), and the firm-level FARE.[1] BTS-Postes stems from mandatory annual employment reports provided by French firms, and contains job-level records for all employees and firms in the private and public sector in France. For each job at each firm, we observe the worker's age and gender, their gross and net annual wages, the number of hours worked during the year, and their detailed occupation code. We also observe the detailed industry code of the firm.[2] Meanwhile, FARE is based on fiscal data from the Public Finances Directorate General, and covers all non-agricultural and non-financial firms in France, including micro-firms.[3] It provides firm-level balance sheet data such as total sales, value added, capital stock and entry and exit dates. Firms can be linked across the two datasets via the unique firm identifier ('*Siren*'). The data is available for the period 2009-2019. All nominal values are converted to real 2015 euros using the Consumer Price Index (CPI).[4]

[1] The data is compiled and prepared by the French statistical office (Insée) and can be accessed by researchers via the platform CASD (*Centre d'accès sécurisé aux données*).

[2] Industry codes are based on the 5-digit APE nomenclature, which corresponds to NACE level 2. The codes are assigned based on the principal activity of the firm (as determined by the industry from which they derive the largest share of turnover), and thus can change over time for a given firm.

[3] Micro-firms are firms with less than 10 employees and an annual turnover or total balance sheet of less than 2 million euros. They represent about 80% of the firms in our sample, but only account for approximately 17% of total employment and 11% of total sales.

[4] We use the annual CPI series for all items in France (not seasonally adjusted), as obtained from the Federal Reserve Economic Database (FRED).

We restrict the analysis to individuals aged between 16-65 in the DADS-Postes data. We exclude individuals working in farming occupations, as well as apprentices and interns. Since much of our analysis will focus on patterns within 4-digit industries, we exclude firms that are in 4-digit industries that have less than 10 firms in total in any given year. We also exclude firms in the following industries: agriculture, public administration & defense, arts & recreation, museums, gambling, amusement, unions, membership organisations, extraterritorial organizations, other personal services, and activities of households.

We compute job-level hourly wages by dividing annual gross wages by the number of hours worked during the year. We drop a small number of observations with negative or missing annual wages or hours of work, as well as observations with real hourly wages below the French minimum wage level.[5] We winsorize top wages at 0.01%.

We estimate a residual job-level wage (conditional on the worker's observable characteristics) by running a set of Mincer-type wage regressions where we control for age, age squared, gender, birthplace (i.e. an indicator variable for whether the individual was born in France), as well as interaction terms of age and age squared with gender and birthplace. We run this regression separately for each year and obtain the residual wage for each job-year observation.[6] We also estimate a residual wage that controls for occupation fixed effects at the 3-digit level, in addition to the worker characteristics listed above.[7]

We aggregate the BTS-Postes data to the firm level in order to obtain firm-level measures of employment and wages in each year, and we match this aggregated data to the balance sheet data from FARE. Total firm-level employment corresponds to the sum of all jobs in a firm, expressed in full-time equivalents (FTE).[8] Average wages (in levels or in logs) are

[5]The data for the French minimum wage ('*Salaire minimum de croissance*') is obtained from *Insée*. Around 3.7% of the observations in 2019 are dropped due to the minimum wage restriction.

[6]It is not possible to track workers over time in our dataset; therefore, we are unable to estimate an AKM-style (Abowd et al. 1999) fixed effects regression. Although there is a separate dataset in which employees can be followed over time, it includes only 1/12th of the workers in each year (rather than the universe of workers and firms), and thus would not allow us to construct reasonable measures of product market concentration due to its limited coverage of firms within industries.

[7]Occupation codes are from the French nomenclature of occupations PCS-ESE 2017.

[8]FTE employment is available in the data for private sector workers, computed based on individuals' an-

computed by weighting each job at the firm by its FTE value. Average firm-level residual wages can be interpreted as a firm wage premium conditional on workers' observable characteristics (and also conditional on the occupational composition of the firm, in the case of the residual that controls for occupation fixed effects).

We can match 86% of the firm-level observations from BTS-Postes to FARE. These firms represent 77% of total employment. Using this merged dataset, we construct a measure of firm-level labor productivity, measured as the firm's annual real value added (from FARE) divided by its annual full-time equivalent employment (from BTS-Postes). To adjust for extreme values, we right-winsorize the top 1% of the labor productivity distribution. We also construct a (year-specific) measure of the market share of each firm by dividing their total sales by the total sales across all firms in their 4-digit industry.

Our final dataset at the firm-level contains 12,782,118 observations with around 1.1 million firms per year across 477 4-digit industries.

2.2 Industry-Level Concentration and Wage Inequality Measures

We use two different measures of product market concentration, both of which are based on the market share of each firm f in industry i and year t, ms_{fit}, computed as described above.

Our first measure is the Herfindahl-Hirschman Index (HHI), which is defined as the sum of squared market shares across firms in an industry and year:

$$HHI_{it} = \sum_{f} ms_{fit}^2 \tag{1}$$

The HHI index takes into account the entire distribution of firms' market shares, and ranges from close to 0 (perfect competition) to 1 (monopoly).

Our second measure of concentration focuses on the market share of top firms, namely

nual hours of work. For workers in the public sector, we compute an FTE measure based on the recommended annual work hours in the public sector, which corresponds to 1820 hours per year.

the total market share of the 10 largest firms in industry i in year t:

$$CR10_{it} = \sum_{f \in Q} ms_{fit} \tag{2}$$

where Q is the set of the 10 largest firms in industry i in year t.

To measure wage inequality across firms within industries, we use the variance of average firm-level log wages within industry i in year t, as well as the 90-10, 90-50 and 50-10 percentile differences, based on percentiles of average firm-level log wages within an industry and year. All wage inequality measures are weighted using the total full-time equivalent employment of each firm.

2.3 Summary Statistics

Panel A of Table 1 presents summary statistics at the firm level. All statistics are employment-weighted, except the first row, which shows that the median firm has less than two FTE employees and the average firm size is around 10 FTE. Average (employment-weighted) firm-level wages are around €19. The average worker workers at a firm that has sales of around €1 billion per year and value added of around €344 million per year. There is large cross-firm heterogeneity in value added and labor productivity.

Panel B presents summary statistics for the industry-level concentration and wage inequality measures. Means, standard deviations, and percentiles are computed across industry-years, where each industry-year is weighted by its total FTE employment. On average, the top 10 firms in an industry account for 35% of total sales. There is large heterogeneity across industry-years, with the 10th percentile of the CR10 measure being 6%, and the 90th being 74%. The average 90-10 firm wage difference is 0.46, which means that, within the average industry, firms at the 90th percentile of the (employment-weighted) firm-wage distribution pay average wages that are 46 log points, or around 58% higher, than firms at the 10th percentile. There is also quite a bit of heterogeneity in inequality in firm-level wages across industry-year cells.

Table 2 provides further information about the composition of firms in our sample. Firms are classified according to their size based on the official groupings used in the French system: Micro-firms (MICRO) are those with less than 10 employees and an annual turnover or total balance sheet of less than €2 million; small and medium sized firms (PME) are those with less than 250 employees and an annual turnover of less than €50 million or total balance sheet of less than €43 million; intermediate sized firms (ETI) are all firms that cannot be classified as PME and have less than 5,000 employees and an annual turnover of less than €1.5 billion or a total balance sheet of less than €2 billion; and big firms (GE) are the remaining firms. The first column shows that the vast majority of firms in our dataset are micro-firms. However, as the second and third columns show, these firms account for a small share of overall sales and employment. Large firms (GE) account for a small share of all firms (less than 1%), but for 36% of total sales and 28% of total employment.

3 Link between Concentration and Inequality

3.1 Industry-Level Relationship

We begin by analyzing the industry-level correlation between product market concentration and between-firm wage inequality. Specifically, we estimate the following regression, analogous to Cortes and Tschopp (2023):

$$INEQ_{it} = \beta CONC_{it} + \gamma_i + \delta_t + u_{it} \tag{3}$$

The dependent variable, $INEQ_{it}$, represents different measures of between-firm wage inequality in industry i in year t. The independent variable, $CONC_{it}$, is a measure of concentration (either log HHI or log CR10) in industry i and year t. γ_i and δ_t represent industry and year fixed effects, respectively. The inclusion of these fixed effects ensures that identification is achieved from differential changes in concentration and inequality within industries over time, rather than overall time patterns or persistent differences between industries. In order to abstract from the effect of changes in the industry structure over time, each industry-year observation is weighted by the employment share of that industry in 2009.[9]

The results of the estimation of Equation (3) are presented in Table 3. Column (1) uses the variance of average firm log wages as the measure of inequality, while Columns (2)-(4) focus on different percentile differences. The top panel uses log HHI as the measure of concentration, while the bottom panel focuses on the log CR10.

The results in Table 3 show a robust positive correlation between concentration and between-firm wage inequality, in line with the findings of Cortes and Tschopp (2023). Higher concentration in an industry is associated with an increase in both top-half (p90-50) and bottom-half (p50-10) inequality.

In order to provide a richer description of the relationship between concentration and the entire distribution of firm-level wages in an industry, we run an additional set of regressions in which we regress a particular percentile of the firm-wage distribution in industry i and year t, w_{itp}, on the same set of controls as above, i.e.:

$$w_{itp} = \beta_p CONC_{it} + \gamma_{ip} + \delta_{tp} + u_{itp} \tag{4}$$

The coefficient β_p reflects how a particular percentile p of the firm-wage distribution varies across industry-year cells with different levels of concentration, after accounting for permanent percentile-specific differences in wages across industries (through the fixed effect γ_{ip}), and time-varying changes in wages that are common across industries at particular percentiles (through the fixed effect δ_{tp}).

The results of the estimation of Equation (4) are displayed in Figure 1. The figure plots

[9] The results are not sensitive to using using average industry employment as a weight, or using time-varying weights (i.e. current industry employment shares).

the estimated coefficient $\widehat{\beta}_p$ along with its 95% confidence interval at each ventile of the firm-wage distribution. Panel (a) uses the log HHI as the measure of concentration, while Panel (b) uses the log CR10.

In both panels, we see a near-monotonic increase in wages as we move to higher percentiles in more concentrated industry-year cells (after accounting for industry and time effects). This indicates that higher concentration is associated with a widening of the entire firm-wage distribution, with the only exception being the top 5%, where wages are higher in more concentrated industry-year cells by a smaller margin than in the next-highest ventile. Overall, this indicates that the results in Table 3 (in terms of the positive correlation between concentration and wage inequality) are not particularly driven by specific parts of the firm-wage distribution; rather, they reflect a fanning out of the entire distribution.

An additional interesting conclusion that can be obtained from Figure 1 is that all of the coefficients across all of the percentiles are positive. In other words, higher concentration is associated with higher average wages throughout the entire firm-wage distribution. This is in line with Cortes and Tschopp (2023), who find a positive correlation between average firm wages and concentration, and consistent with Bighelli et al. (2021), who show that changes in concentration are positively associated with changes in productivity and allocative efficiency.

3.2 Mechanisms

We next exploit our detailed firm-level data in order to provide a novel exploration of the mechanisms that can account for the positive industry-level link between concentration and between-firm wage inequality.

Differences in the Productivity Distribution It is well established that firm sales and firm wages are positively correlated with firm productivity. Given this link, it is possible that differences in the productivity distribution among firms operating in different industry-year cells could account for the positive link between concentration and wage inequality. In other words, holding the relationship between productivity, sales and wages constant, higher levels of dispersion in productivity among operating firms would naturally lead to a more dispersed sales distribution (i.e. more concentration) and a more dispersed wage distribution (i.e. more wage inequality).

To test whether this is the case, we run a set of regressions analogous to those in Equation (3), but with the industry-level variance of log firm *productivity* as the dependent variable.[10] Table 4 displays the results. We find that there is no statistically significant relationship between concentration and the variance of productivity. This means that higher concentration is *not* associated with a more dispersed distribution of productivity among operating firms, as measured by the variance; if anything, the signs of the estimated coefficients point in the opposite direction.

In order to further explore how the entire shape of the productivity distribution varies with concentration, we also run a set of regressions analogous to those in Equation (4), with particular percentiles of the productivity distribution in industry i and year t as the dependent variable. The results are shown in Figure 2. We can see that both the bottom percentiles and the top percentiles of productivity tend to be higher in more concentrated industry-year cells (after accounting for industry and year fixed effects). Moreover, this variation is generally statistically insignificant at conventional levels. Overall, the figure clearly shows that there is no systematic evidence that higher concentration is associated with a more dispersed distribution of productivity among operating firms.

Differences in the Relationship between Productivity and Other Firm-Level Outcomes The results above indicate that the positive relationship between concentration and between-firm wage inequality is not driven by differences in the composition of operating firms along the productivity margin. In other words, more concentrated industries feature

[10] As we are interested in dispersion across firms, we compute this variance assigning equal weights to all firms. Below we further analyze the implications of employment reallocation across firms.

higher levels of concentration and inequality in spite of having similar levels of dispersion in the underlying firm productivity distribution.

In order to further understand this pattern, we now explore whether the relationship between productivity, sales, employment, and wages varies with the degree of industry concentration. We perform this analysis using the firm-level (rather than the industry-level) data. Specifically, we estimate the following regression:

$$y_{fit} = \beta_1 PROD_{fit} + \beta_2 PROD_{fit} \times CONC_{it} + \gamma_{it} + u_{fit} \tag{5}$$

where y_{fit} represents a firm outcome for firm f in industry i in year t (which may be either log sales, log employment or log wages). $PROD_{fit}$ represents the firm's labor productivity (in logs). As above, $CONC_{it}$ is our measure of concentration (either log HHI or log CR10) in industry i and year t. γ_{it} represent fully interacted industry-year fixed effects, so that all of our identifying variation is derived from differences between firms within industry-year cells, and u_{fit} is a standard error term. Observations are weighted using the firm's full-time equivalent employment.

Our coefficient of interest is β_2. While β_1 captures the baseline relationship between productivity and other firm-level outcomes, β_2 indicates whether this relationship is stronger or weaker within more concentrated industry-year cells.

Table 5 presents the results. Column (1) shows that, on average, more productive firms have higher sales than less productive firms within their industry, and this is particularly so in more concentrated industries. Column (2) shows a similar pattern for employment: more productive firms are larger than less productive firms, and disproportionately so when they are in more concentrated industry-year cells. Column (3) shows that more productive firms pay higher wages relative to less productive firms on average, and that this relationship is stronger when concentration is higher. These results imply that higher concentration is associated with higher levels of dispersion across firms in terms of sales, employment and wages, *conditional on productivity*. Between-firm wage inequality is higher in more concentrated industries because firms that are relatively more productive within those industries tend to be disproportionately larger and pay disproportionately higher wages.

Exploring the Role of Sorting We can shed further light on the reasons why more productive firms pay higher wages in more concentrated industries by determining whether these firms disproportionately employ workers with certain observable characteristics, or in certain occupations.

First, we re-run Equation (5) using, as the dependent variable, the average residual wage in each firm, based on the Mincer-type job-level regression that controls for workers' observable characteristics (age, gender and birthplace). This allows us to verify whether more productive firms pay higher wages conditional on worker characteristics, and whether this is disproportionately so in more concentrated industries. If more productive firms in more concentrated industries tend to pay higher wages solely because they tend to employ workers with characteristics that are associated with higher wages, then we would expect the estimate of $\widehat{\beta}_2$ in this regression to be close to zero.

The result is presented in Column (1) of Table 6. The estimated coefficients are nearly identical to those obtained when using average wages in Column (3) of Table 5. This means that worker sorting along age, gender and birthplace dimensions cannot account for the disproportionately higher mean wages observed in high productivity firms in more concentrated industries.

In Column (2) we instead use as the dependent variable the average residual wage in each firm when we additionally control for occupation fixed effects in the job-level Mincer-type regression. This allows us to verify whether the wage differentials are partly driven by differences in the occupational composition of more productive firms (within industries). The results show that the estimated coefficients $\widehat{\beta}_1$ and $\widehat{\beta}_2$ each fall by about half. This implies that roughly half of the reason why more productive firms pay higher wages than less productive firms in their same industry is because their workforce is disproportionately made up of workers in higher-paying occupations. Higher concentration is associated with a stronger degree of sorting of high-paying occupations towards high-productivity firms.

As a way to further visualize this stronger occupational sorting in more concentrated industries, in Column (3) we use as the dependent variable an occupational composition index for each firm. This is obtained by first assigning each occupation the value of the estimated coefficient on its fixed effect in each year from the Mincer-type regression, and then computing a weighted average of these values for each firm in each year. Firms that employ a larger share of workers in higher-paying occupations will have higher values of this index. As the results in Column (3) show, more productive firms in more concentrated industries disproportionately employ workers in high-paying occupations.

To understand the types of occupations that differ between more and less productive firms in more and less concentrated industries, we classify occupations as either abstract, routine, or manual, following the type of classification used by Acemoglu and Autor (2011) and the literature on job polarization.[11] As the literature has discussed, abstract occupations tend to be high-paying; routine occupations tend to be middle-paying; and manual occupations tend to be low-paying.

The results in Columns (4)-(6) show that more productive firms within industries tend to employ a higher share of abstract workers, and particularly so in more concentrated industries. These firms also tend to employ a lower share of routine and manual workers, and again particularly so in industry-year cells that exhibit higher levels of concentration.

Figure 3 provides more details on the heterogeneity in occupational composition by estimating Equation (5) using the employment share in each 2-digit occupation as the dependent variable. The figure plots the estimated $\widehat{\beta}_2$ for each occupation, along with the 95% confidence interval. Occupations are ordered along the x-axis from lowest-paying to highest-paying and are color-coded. Green dots represent manual, black dots routine and blue dots

[11] Appendix Table A.1 shows the mapping of each 2-digit occupation code in our data to the three task groups.

abstract occupations. The figure shows that the coefficients are negative for most of the bottom-half occupations, and positive for most of the top-half occupations. In particular, more productive firms in more concentrated industries employ a particularly high share of workers in engineering and technical occupations (38) and executive positions (37), and a particularly low share of workers in unqualified industrial occupations (67) and administrative positions (54).

Summary To summarize, we find that much of the variation in concentration across industry-year cells can be attributed to differences in the level of dominance (market share) of the most productive firms within the industry. High-productivity firms in more concentrated industries are not only disproportionately large in terms of terms of their market share, but also in terms of their total employment. They also pay disproportionately higher wages compared to less productive firms in the same (detailed) industry. This rationalizes the industry-level link between product market concentration and between-firm inequality: the increased dispersion in market shares, employment, and wages conditional on productivity implies that more concentrated industries also feature higher levels of between-firm wage inequality (as we found in Table 3). The finding that more productive firms in more concentrated industries pay higher wages is partly due to the fact that these firms disproportionately employ workers in high-paying occupations (increased sorting of high-paying occupations to high-paying firms), and partly due to higher firm wage premia conditional on occupational composition.

4 Conceptualizing the Link

Autor et al. (2020) argue that increased consumer price sensitivity is an important driver of the rise in concentration. This section shows how this shock would also account for a rise in between-firm inequality, with implications that are consistent with our empirical results.

To consider the impact of an increase in consumer price sensitivity, we follow Cortes and Tschopp (2023) and embed this shock within a framework that allows for wage inequality between firms, namely the search and bargaining framework of Helpman et al. (2010).

As in Melitz (2003), each sector features a continuum of horizontally differentiated varieties, with total consumption Q being given by a constant elasticity of substitution (CES) aggregate:

$$Q = \left[\int_{j \in J} q(j)^{\beta} dj \right]^{1/\beta},$$

where j indexes varieties, J is the set of varieties within the sector, $q(j)$ denotes consumption of variety j and $\beta \in (0, 1)$ is a function of the elasticity of substitution between varieties, σ, namely $\beta \equiv (\sigma - 1)/\sigma$. An increase in consumer price sensitivity is modeled as an increase in σ.

Firms are heterogeneous in terms of their productivity, θ, and produce according to the following production function:

$$y = \theta \left(\overline{a}_H h_H^{\gamma_H} \right)^{\lambda_H} \left(\overline{a}_L h_L^{\gamma_L} \right)^{\lambda_L}, \qquad \lambda_H + \lambda_L = 1 \tag{6}$$

where H denotes skilled (high average wage) occupations and L denotes unskilled (low average wage) occupations; $\overline{a}_H$ and $\overline{a}_L$ are the average match-specific abilities of workers in skilled and unskilled occupations, respectively, and h_H and h_L are the number of workers in skilled and unskilled occupations employed by the firm.

The match-specific ability of workers in a given occupation has a Pareto distribution with shape parameter k_o and lower bound $a_{min,o}$, $o = H, L$. Helpman et al. (2010) show that, assuming that $k_H < k_L$, i.e. that the match-specific ability distribution is more dispersed among workers in skilled vs unskilled occupations, the share of employment in skilled occupations is increasing in firm productivity.

They also show that average firm wages will be higher in more productive firms because they: (i) employ a larger proportion of workers in high-paying occupations, and (ii) pay higher wages conditional on occupation. Hence, wages differ across firms both because of the occupational composition/sorting of workers, and because of firm premia conditional on occupations – in line with the empirical patterns documented above.

The concentration of employment in the top $\mu\%$ of firms for occupation o is given by:

$$C_{h,o} = \mu^{1-\frac{\beta}{\Gamma z}(1-k_o/\delta)} \tag{7}$$

where:

$$\Gamma = 1 - \beta(\lambda_H\gamma_H + \lambda_L\gamma_L) - \frac{\beta}{\delta}[1 - (\lambda_H\gamma_H k_H + \lambda_L\gamma_L k_L)]$$

and δ is a parameter of the firm's screening cost function.

Given that $k_H < k_L$, we have that $C_{H,h} > C_{L,h}$.

Prediction: An increase in the elasticity of substitution, σ, increases concentration of employment in the most productive firms, particularly so for skilled occupations:

$$\frac{\partial C_{h,H}}{\partial \sigma} > \frac{\partial C_{h,L}}{\partial \sigma} > 0$$

Corollary: The disproportionate increase in employment concentration for skilled occupations implies stronger sorting of workers in high-paying occupations to high productivity firms. This increased sorting and the implied changes in the occupational composition of employment across firm types will increase between-firm inequality in average firm-level wages.

The distribution of wages across firms for workers in occupation o is given by:

$$G_f(w_o) = 1 - \left(\frac{w_{d,o}}{w_o}\right)^{\frac{\delta\Gamma z}{\beta k_o}}$$

This is a Pareto distribution with scale parameter $w_{d,o}$ (a function of various parameters of the model), and shape parameter $\frac{\delta \Gamma z}{\beta k_o}$. Inequality, as measured by any scale-invariant measure, will be a function of the shape parameter only.

Prediction: An increase in the elasticity of substitution, σ, increases within-occupation, between-firm wage inequality for both occupations.

Corollary: An increase in the elasticity of substitution, σ, increases inequality in average firm wages both because of (i) increased worker sorting along occupational dimensions and (ii) increased dispersion in firm premia conditional on occupation.

Our empirical results show that more concentrated industry-year cells indeed feature stronger occupational sorting, as well as larger between-firm wage differences in residual wages (i.e. conditional on occupation). Hence, the empirical patterns that we have documented are consistent with the implications of a shock to consumer price sensitivity, as posited by Autor et al. (2020), when embedded within a framework that allows for heterogeneous wages across firms.

5 Conclusion

This paper shows that higher levels of product market concentration are associated with higher levels of between-firm wage dispersion within industries. Using rich administrative data for France covering the near-universe of workers and firms over the period 2009-2019, we find a positive and statistically significant correlation between sectoral concentration and different measures of wage inequality between firms. Importantly, we are able to shed light on the firm-level patterns that underlie this relationship. We find that markets that feature higher levels of concentration are not characterized by a more dispersed distribution of firm productivity; instead, they are characterized by higher levels of wage dispersion *conditional* *on productivity*. We show that this is partly due to the increased sorting of workers in high-paying occupations towards high-productivity firms, and partly due to larger firm pay differences within occupations. The empirical patterns that we document can be rationalized by embedding a shock to consumer price sensitivity within a model that features wage differences between firms.

References

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, 67(2):251–333.

Acemoglu, D. and Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics*, 4:1043–1171.

Akerman, A. (2021). Market Concentration and the Relative Demand for College-Educated Labor. *Working Paper*.

Arnold, D. (2021). Mergers and Acquisitions, Local Labor Market Concentration, and Worker Outcomes. *Working Paper*.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Reenen, J. V. (2017). Concentrating on the Fall of the Labor Share. *American Economic Review*, 107(5):180–185.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, 135(2):645–709.

Azar, J., Marinescu, I., and Steinbaum, M. (2020). Labor Market Concentration. *Journal of Human Resources*, 57(S):S167–S199.

Azkarate-Askasua, M. and Zerecero, M. (2021). The Aggregate Effects of Labor Market Concentration. *Working Paper*.

Bajgar, M., Berlingieri, G., Calligaris, S., Criscuolo, C., and Timmis, J. (2019). Industry Concentration in Europe and North America. *OECD Productivity Working Papers*, 18.

Barth, E., Bryson, A., Davis, J. C., and Freeman, R. (2016). It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States. *Journal of Labor Economics*, 34(S2):S67–S97.

Bassanini, A., Cyprien, B., and Caroli, E. (2021). Labor Market Concentration and Stayers' Wages: Evidence from France. *IZA Discussion Paper Series*, (14912).

Benmelech, E., Bergman, N. K., and Kim, H. (2020). Strong Employers and Weak Employees. *Journal of Human Resources*, 57(S):S200–S250.

Bighelli, T., di Mauro, F., Melitz, M., and Mertens, M. (2021). European Firm Concentration and Aggregate Productivity. *IWH-CompNet Discussion Papers No. 3.*

Card, D., Cardoso, A. R., Heining, J., and Kline, P. (2018). Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics*, 36(S1):S13–S70.

Card, D., Heining, J., and Kline, P. (2013). Workplace Heterogeneity and the Rise of West German Wage Inequality. *The Quarterly Journal of Economics*, 128(3):967–1015.

Cortes, G. M. and Tschopp, J. (2023). Rising concentration and wage inequality. *Scandinavian Journal of Economics, Forthcoming.*

Deb, S., Eeckhout, J., Patel, A., and Warren, L. (2022). Market Power and Wage Inequality. *BSE Working Paper 1360.*

Grullon, G., Larkin, Y., and Michaely, R. (2019). Are US Industries Becoming More Concentrated? *Review of Finance*, 23(4):697–743.

Helpman, E., Itskhoki, O., and Redding, S. (2010). Inequality and unemployment in a global economy. *Econometrica*, 78(4):1239–1283.

Marinescu, I., Ouss, I., and Pape, L.-D. (2021). Wages, hires, and labor market concentration. *Journal of Economic Behavior & Organization*, 184:506–605.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

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

Rinz, K. (2020). Labor Market Concentration, Earnings, and Inequality. *Journal of Human Resources*, 57(S):S251–S283.

Song, J., Price, D. J., Guvenen, F., Bloom, N., and von Wachter, T. (2018). Firming Up Inequality. *The Quarterly Journal of Economics*, 134(1):1–50.

Webber, D. A. (2015). Firm market power and the earnings distribution. *Labour Economics*, 35:123–134.

Table 1: Summary Statistics

	Mean	St. Dev.	p10	p50	p90
A: Firm-Level (N = 12,782,118)					
Employment (unweighted)	10.02	229.98	0.26	1.89	12.56
Employment	5,285.91	18,316.07	4.09	107.55	7,853.90
Average wages	19.14	8.03	12.16	16.93	28.88
Sales	1,123,177	4,552,05	518	18,391	1,907,473
Value added	344,047	1,321,373	226	5,976	538,113
Labor productivity	74.10	80.86	32.46	57.25	125.17
B: Industry-Level (N = 5,247)					
Employment	139,770	177,531	11,398	72,762	470,705
HHI	0.05	0.11	0.00	0.02	0.12
CR10	0.35	0.25	0.06	0.30	0.74
Variance of firm avg log wages	0.04	0.04	0.01	0.03	0.09
Wage inequality p90-10	0.46	0.23	0.16	0.41	0.78

Note: Panel A shows summary statistics at the firm-level. Wages are real hourly wages and employment is expressed in full-time equivalents. All nominal values are deflated using the 2015 CPI. Labor productivity, sales and value added are in thousands of euros. Labor productivity is measured as real value added per full-time equivalent employment. Summary statistics are employment-weighted. Panel B presents summary statistics aggregated to the industry level, where industries are weighted by their full-time equivalent employment. Observations are at the 4-digit industry-year level for the years 2009-2019. HHI and CR10 are based on firm sales; see text for details.

Table 2: Firm Composition

	Share of Firms	Share of Sales	Share of Employment
Firm Size			
GE	0.90	36.16	27.76
ETI	2.79	30.30	26.48
PME	15.80	22.11	28.40
MICRO	80.51	11.42	17.37
Total	100.00	100.00	100.00

Note: The table shows the composition of firm categories in our sample. Micro-firms (MICRO) are firms with less than 10 employees and an annual turnover or total balance sheet of less than €2 million; small and medium sized firms (PME) are firms with less than 250 employees and an annual turnover of less than €50 million or total balance sheet of less than €43 million; intermediate sized firms (ETI) are all firms that cannot be classified as PME and have less than 5,000 employees and an annual turnover of less than €1.5 billion or a total balance sheet of less than €2 billion; big firms (GE) are the remaining firms.

Table 3: Concentration & Between-Firm Wage Inequality

	(1) variance	(2) p90-10	(3) p90-50	(4) p50-10
log HHI	0.0029***	0.024***	0.014***	0.0092***
	(0.00061)	(0.0045)	(0.0042)	(0.0024)
Adjusted R-squared	0.96	0.95	0.89	0.93
N	5247	5247	5247	5247
	(1)	(2)	(3)	(4)
	variance	p90-10	p90-50	p50-10
log CR10	0.0062***	0.054***	0.031**	0.023***
	(0.0014)	(0.0100)	(0.0095)	(0.0047)
Adjusted R-squared	0.96	0.95	0.89	0.93
N	5247	5247	5247	5247
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Note: Observations are at the industry-year level. The dependent variables are measures of (employment-weighted) between-firm inequality in average log wages in the industry-year. In order to abstract from variation stemming from changes in the industry composition of employment over time, each observation is weighted by the corresponding industry's employment share in 2009. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Composition Effects

Variance of productivity	(1)	(2)
log HHI	-0.0052	
	(0.0085)	
log CR10		-0.018
		(0.023)
Adjusted R-squared	0.84	0.84
N	5247	5247
Industry FE	Yes	Yes
Year FE	Yes	Yes

Note: Observations are at the industry-year level. The dependent variable is the (unweighted) variance of log real labor productivity across firms in the industry-year. In order to abstract from variation stemming from changes in the industry composition of employment over time, each observation is weighted by the corresponding industry's employment share in 2009. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Differences in the Relationship between Productivity and Other Firm-Level Outcomes

	(1) Log Sales	(2) Log Employment	(3) Log Wages
Log productivity	1.55***	0.80***	0.21***
	(0.039)	(0.041)	(0.011)
Log productivity × log HHI	0.16***	0.17***	0.014***
	(0.0068)	(0.0073)	(0.0021)
Adjusted R-squared	0.50	0.48	0.63
N	12782118	12782118	12782118
	(1)	(2)	(3)
	Log Sales	Log Employment	Log Wages
Log productivity	1.27***	0.53***	0.20***
	(0.027)	(0.025)	(0.0067)
Log productivity × log CR10	0.29***	0.32***	0.037***
	(0.012)	(0.011)	(0.0034)
Adjusted R-squared	0.50	0.48	0.63
N	12782118	12782118	12782118
Industry-Year FE	Yes	Yes	Yes

Note: Observations are at the firm-year level. Log labor productivity is measured as real value added per full-time equivalent firm employment. Observations are weighted by current full-time equivalent firm employment. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

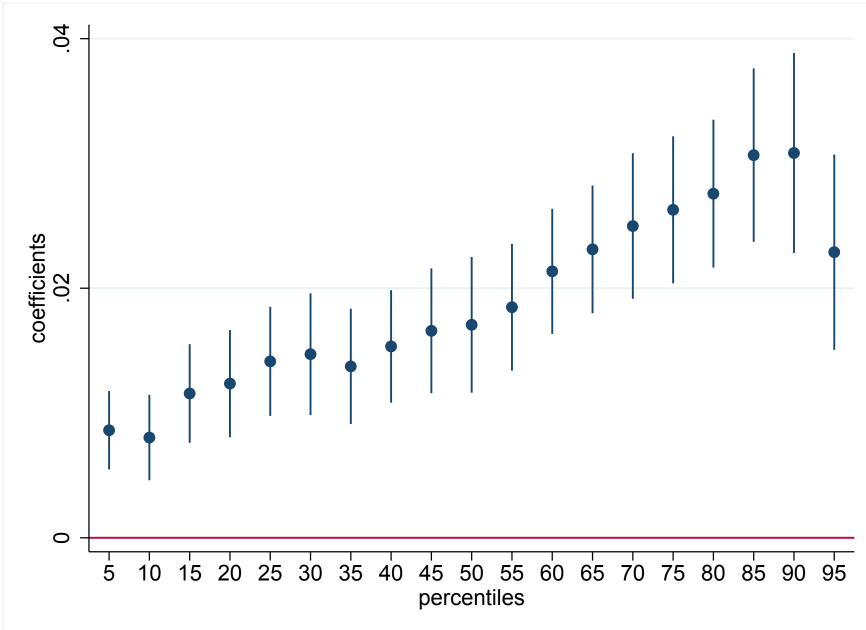
Table 6: Understanding the Variation in Firm-Level Wages

	(1) Residual Wages	(2) Residual Wages (occ.)	(3) Composition Index	(4) Share Abstract	(5) Share Routine	(6) Share Manual
Log productivity	0.20***	0.12***	0.090***	0.10***	-0.037***	-0.058***
	(0.010)	(0.0060)	(0.0046)	(0.0054)	(0.0027)	(0.0036)
Log productivity × log HHI	0.013***	0.0054***	0.0078***	0.0090***	-0.0043***	-0.0041***
	(0.0020)	(0.0011)	(0.00088)	(0.0010)	(0.00050)	(0.00071)
Adjusted R-squared	0.60	0.32	0.70	0.61	0.58	0.69
N	12782118	12782117	12782117	12782118	12782118	12782118
	(1) Residual Wages	(2) Residual Wages (occ.)	(3) Composition Index	(4) Share Abstract	(5) Share Routine	(6) Share Manual
Log productivity	0.19***	0.11***	0.083***	0.095***	-0.033***	-0.055***
	(0.0064)	(0.0038)	(0.0029)	(0.0035)	(0.0019)	(0.0022)
Log productivity × log CR10	0.034***	0.016***	0.020***	0.022***	-0.010***	-0.010***
	(0.0032)	(0.0019)	(0.0014)	(0.0017)	(0.00095)	(0.0012)
Adjusted R-squared	0.60	0.32	0.70	0.61	0.58	0.69
N	12782118	12782117	12782117	12782118	12782118	12782118
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes

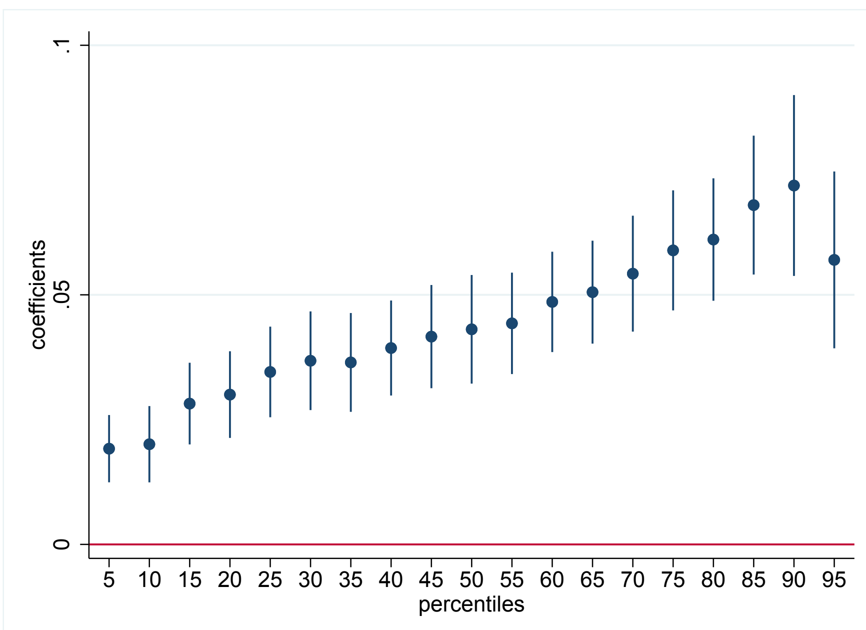
Note: Observations are at the firm-year level. The measures for the dependent variables in Columns (1) - (3) are obtained from job-level Mincer-type regressions. Column (1) controls for workers' observable characteristics (age, gender, birthplace). Column (2) adds 3-digit occupation fixed effects to control for the occupational composition within firms. In Column (3) we construct an occupational composition index by assigning each occupation the value of the estimated coefficient on its fixed effect in each year and then compute a weighted average of these values for each firm in each year. Higher values correspond to a larger share of workers in higher-paying occupations. Log labor productivity is measured as real value added per full-time equivalent firm employment. Observations are weighted by current full-time equivalent firm employment. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 1: Concentration and the Distribution of Firm-Level Wages

(a) HHI



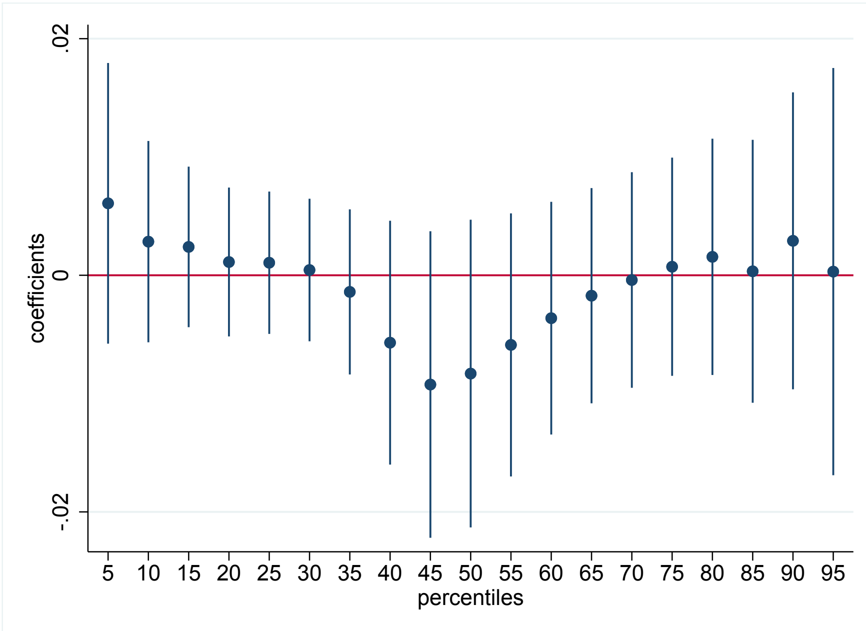
(b) CR10



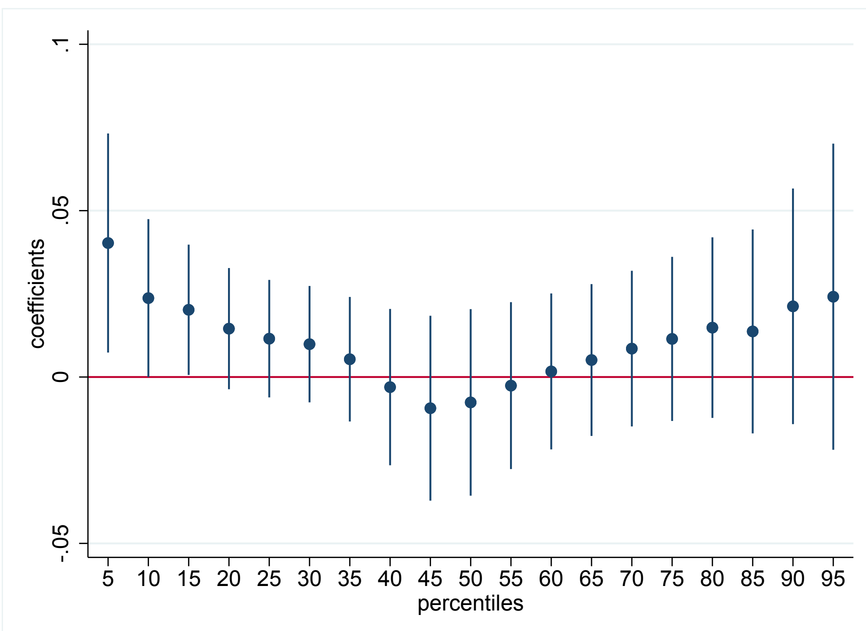
Note: The figure plots the estimated coefficients and the 95% confidence intervals from a set of regressions at the industry-year level, where the p-th percentile of the employment-weighted log-firm-wage distribution within an industry-year cell is regressed on concentration in that cell. The regressions control for industry and year fixed effects, and weight observations by the industry's employment share in 2009.

Figure 2: Concentration and the Distribution of Firm-Level Productivity

(a) HHI



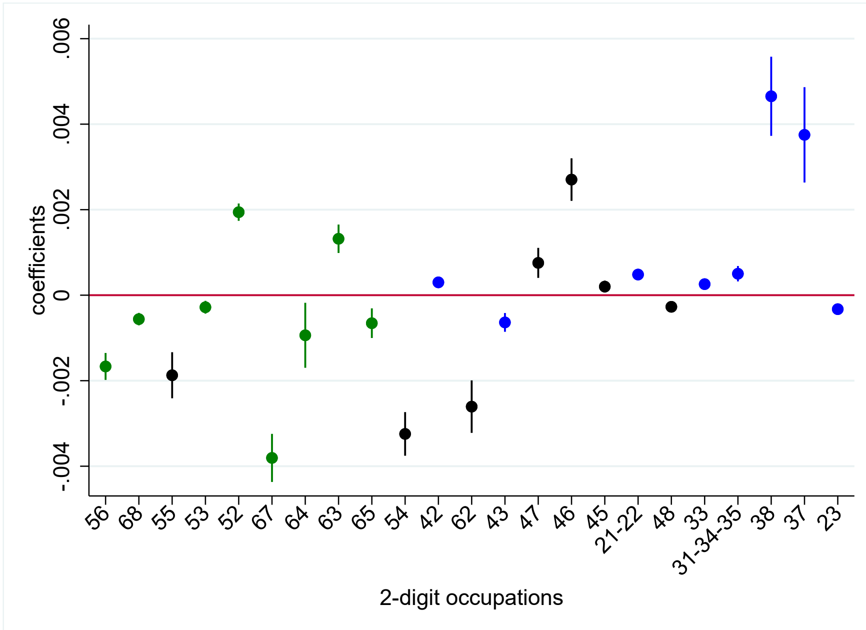
(b) CR10



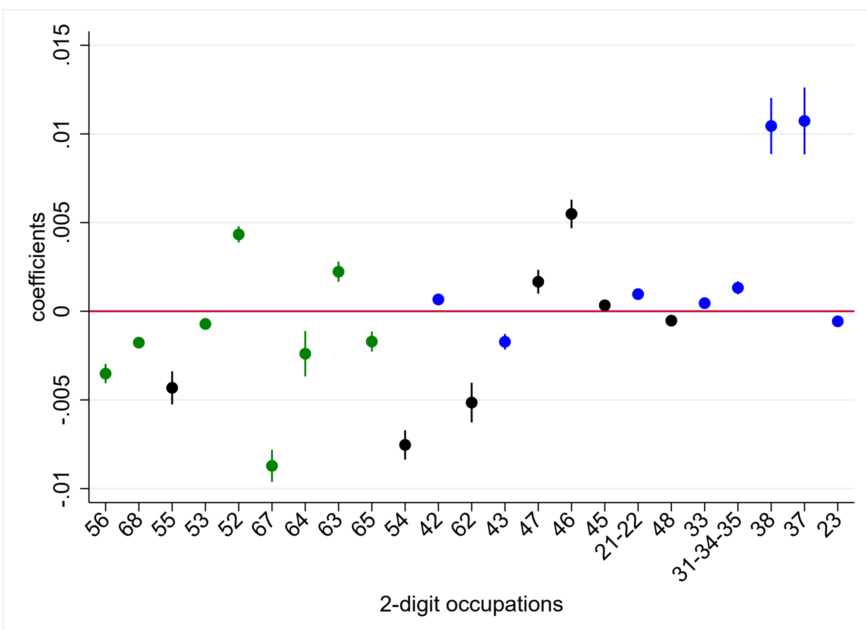
Note: The figure plots the estimated coefficients and the 95% confidence intervals from a set of regressions at the industry-year level, where the p-th percentile of the log-firm-productivity distribution within an industry-year cell is regressed on concentration in that cell. The regressions control for industry and year fixed effects, and weight observations by the industry's employment share in 2009. Labor productivity is measured as real value added per full-time equivalent employment.

Figure 3: Occupational Sorting

(a) log HHI



(b) log CR10



Note: The figure plots the estimated coefficients of the interaction term, and the 95% confidence intervals from the estimation of Equation (5) using the firm-level employment share of each 2-digit occupation as the dependent variable. Occupations are ordered from lowest- to highest-paying. Green dots represent manual, black dots routine and blue dots abstract occupations. See Appendix Table A.1 for a list of occupation titles. Observations are weighted by current full-time equivalent firm employment.

A Appendix

Table A.1: Occupation Classification in the French data

2-digit code	**Occupation Title**	**Type**	**Average Wage (log)**	**Employment Share** (%)
56	Consumer service occupations	manual	2.49	5.73
68	Unskilled craft workers	manual	2.50	4.02
55	Retail occupations	routine	2.53	8.79
53	Police, military and security workers	manual	2.57	1.24
52	Civil servants and public service agents	manual	2.57	1.82
67	Unqualified industrial workers	manual	2.60	5.96
64	Drivers	manual	2.62	4.96
63	Skilled craft workers	manual	2.65	8.74
65	Skilled storage and transportation workers	manual	2.69	2.98
44	Clergy, religious occupations	abstract	2.69	0.00
54	Administrative occupations	routine	2.72	10.22
42	Teachers and related professions	abstract	2.75	0.46
62	Skilled industrial workers	routine	2.77	9.17
43	Intermediate occupations in health and social work	abstract	2.83	1.72
47	Technicians	routine	2.94	5.94
46	Intermediate administrative and commercial professions	routine	2.95	6.90
45	Intermediate administrative civil service professions	routine	2.96	0.11
21	Craft workers (business heads)	abstract	2.96	0.12
22	Merchants (business heads)	abstract	3.01	0.31
48	Supervisors	routine	3.01	2.58
31	Self-employed professionals (with employment status)	abstract	3.26	0.10
33	Civil service executives	abstract	3.31	0.12
35	Information, arts and entertainment professions	abstract	3.34	0.48
34	Professors, scientific professions	abstract	3.39	0.43
38	Engineering and technical managers	abstract	3.40	8.19
37	Adminstrative and commercial executive positions	abstract	3.46	8.15
23	Top managers with 10 or more employees	abstract	3.69	0.77