Paper III

Wage effects from employer concentration and collective bargaining in Swedish labor markets

Carl Fredrik Söderqvist, Johan Eklund

Abstract

We explore how the wage bargain is affected by employer concentration and labor market institutions in Sweden, using a Herfindahl-Hirschman Index of concentration on hires and an outside occupation index to address concerns relating to market definition. The model attempts to separate the wage impact from exercising outside options (individual bargaining power) and from collective bargaining agreements (collective bargaining power). Holding value added constant, concentration has a modest negative impact on wage levels for white collar occupations, but a positive effect for blue collar occupations. The positive impact from concentration for blue collar workers is interesting and should generate more research on bargaining power dynamics in heterogeneous labor markets. The explanatory power of the model is largely driven by yearly occupational and regional fixed effects. Yearly changes of estimated fixed effects closely follow Sweden's centrally coordinated collective bargained yearly wage increases, validating the model's ability to separate individual from collective bargaining power. The strength of outside options thus approximates wage drift.

JEL-codes: J42, J30, J50, J60 **Key words:** Monopsony, Wages, Collective Bargaining, Labor Mobility

1 Introduction

A growing literature attempts to assess employer wage setting power in labor markets, finding that they often do, and thus labor markets are largely signified by monopsonistic competition (e.g. Card 2022 for an overview). Common sources of monopsony power[1] found in the literature include switching costs, search frictions, as well as worker and job heterogeneity (Robinson 1933, Manning 2003).

A conceptually interlinked source of wage setting power is employer concentration. In a labor market where workers have idiosyncratic tastes for jobs, the quantity and quality of available jobs (outside options) in the labor market can give an employer wage setting power. If there are few job postings concentrated to a single employer, employer competition for labor supply is reduced. With high employer concentration, individual bargaining power could suffer due to a reduced threat of *exit* to outside job options (Hirschman 1970). Estimates from a recent, growing, and geographically diverse literature suggests that concentration can have a significant negative impact on wages (e.g. Azar et al. 2019, Azar et al. 2022, Bassanini et al. 2023, Schubert et al. 2024, Thoresson 2024).

There is a salient distinction between having hypothetical wage setting power and exercising it; employers in labor markets signified by monopsonistic competition may be constrained by labor market institutions, legislation, or social conventions (Manning 2021). Manning (2003, p. 351) predicts that "the wages set by very powerful unions will be independent of the extent of monopsony power in the labor market". Then, what role does employer competition for labor play in labor markets where collective bargaining is (almost) omnipresent? And how do we assure that estimates of individual bargaining power are not tainted by the impact from collective bargaining power (*voice*)?

Here, we employ a direct method to assess effects to the wage bargain by assessing the strength of worker outside options, while controlling for plausibly strong collective bargaining effects in Swedish labor markets. We use employer-employee linked administrative data to estimate how wages and incomes are impacted by workers moving to outside options between 2005 and 2020. The model thus takes the perspective of labor supply to assess how the probabilities and payoffs in exercising outside options have impacted Swedish wage levels.

We adapt a Nash bargaining framework (Manning 2011) closely following Schubert et al. (2024, henceforth SST) to assess the wage impact from the strength of outside options, approximating labor market competition from the perspective of labor supply. The model is applied in labor markets defined by finely grained occupation codes and commuting zones. The dependent variable (the wage bargain) is represented by gross full-time equivalent wages (including variable payment compo-

[1]Monopsony signifies a market with a single buyer, which is the demand side equivalent of supply side monopoly; a single seller. Monopsony power does not necessarily imply a single buyer, but rather reflects a lack of competition resulting in below-equilibrium prices and quantities.

nents), and gross average monthly labor incomes from a worker's primary employer while correcting for working time (as a percentage of full time employment).

The strength of outside options is estimated by calculating a Herfindahl Hirschman Index (HHI) of employer concentration on hires in defined occupation-regional labor markets. However, defining concentration in occupation-regional labor markets can impose an overly strict market definition. High employer concentration should have a smaller negative impact on the wage bargain if a worker has strong outside options in nearby occupations, compared to if a worker has few outside options in other occupations. We address this concern by including a variant of SST's Outside Occupation Index (OOI), which considers the mean payoff of workers moving from the market-defining occupation to other occupations, multiplied by the probability of such occupation-changing moves occurring. Apart from addressing the market definition-problem, the OOI-variable also provides an indication of the quality of the job ladder available to workers in a defined labor market.

We address several plausible identification concerns, relating to concentration and firm size, as well as the model's ability to separate individual and collective bargaining power effects.

First, large firms producing large surpluses due to productivity effects can pay higher wages than less productive firms. Large growing firms may thus capture a greater market share of hires and pay higher wages, which may result in positive wage effects from concentration (Névo 2024). We address this concern by including a concentration and value added interaction in our regressions, where value added is the weighted mean of value added in the defined labor market.

Second, the payoff from changing occupations in the OOI-variable may capture wage increases that would have been realized in a counterfactual scenario where the worker did not change occupation; Sweden's extensive collective bargaining coverage and coordinated, centralized wage setting regime, implies that a certain percentage of wage increases likely would have been paid anyways. This raises identification concerns. We address this by including an instrumental variable where the yearly wage increases set by the Swedish industry norm – the Mark – is multiplied by the previous year's wage level of occupation movers to attain the instrument. The instrument is thus plausibly exogenous to the individual but with a likely impact on the wage.

Third, all models include yearly regional and occupational fixed effects. If the model is well specified, the outside option variables (e.g. HHI and OOI) should capture *only* individual bargaining power effects to the wage bargain, without being tainted by collective bargaining effects. Then, collective bargaining effects should be picked up by the yearly regional and occupational fixed effects. Thus, to assess if the model can separate outside option/individual bargaining power effects from collective bargaining effects, we study the R^2 and within-R^2 estimations resulting from our regressions, and how yearly changes to the fixed effects-estimations compare to national collective bargaining outcomes.

And fourth, we also include analyses on incomes and hours, which should approximate wages, allowing us to consider other motivations of worker mobility, such as increasing or reducing working hours.

We also explore salient heterogeneity in Swedish wage collectively bargained wage setting practices (e.g. Bhuller et al. 2022) by considering impacts to the wage bargain for large occupational white-collar (WC) and blue-collar (BC) subsets[2] of workers.

Our main results indicate that concentration has a relatively modest negative effect on Swedish wages. When holding value added constant, a one percent increase in employer concentration has a -0.8 percent negative impact on the wage. The effect is relatively small by international comparisons. For WC occupations, the negative effect on wages are -1.35 percent from a one percent increase in concentration, whereas the result indicates a 0.011 percent *increase* on wages for BC occupations when employer concentration is increased by one percent.

Recent Swedish employer concentration studies yield similar results for WC subsets of occupations. Thoresson (2024) finds that the de-monopolization of Swedish pharmacy markets in 2009 lowered employer concentration and had a positive impact on the wage bargain for pharmacists (a specialized WC occupation). Similarly, Zhao and Matti 2018 find a modest increases to midwife earnings (also a specialized WC occupation) when an additional maternity ward was opened in central Stockholm, and a modest decrease to earnings when the same ward was closed. To our knowledge, we provide a first (and perhaps counter-intuitive) estimation on BC effects from employer concentration in Sweden.

We argue that the positive concentration estimates for BC-occupations is an important contribution to the wage-concentration literature, suggesting that concentration estimates can be highly heterogeneous for groups with varying individual bargaining power. Such estimates can also be used to study a broader set of factors that may have adverse effects to the wage bargain. For example, if decreased concentration is a result of domestic outsourcing or extensive subcontracting, a positive concentration slope can possibly be seen as an indication of labor market fissurization (Weil 2014). These findings should motivate future research to explain the causes behind BC and WC heterogeneity in outcomes.

The results validate the ability of the SST model to separate individual from collective bargaining power effects to the wage bargain. For both the BC and WC-groups, the within-R^2 estimates for our main results show that approximately 7 percent of all variation in the data are explained by the outside option variables, with total R^2 explaining between 93.5 and 96.6 percent of all variation. Thus, the fixed effects are explaining most of the variation in the data. By extracting and summarizing the occupational and regional fixed effects estimations for each year, and calculating

[2] WC wage-setting practices tend to use individual negotiations to a higher extent than more collectivist practices regulated in BC agreements.

the yearly changes of these fixed effects summations, we find that they are nearly identical to the yearly central wage increases set in collective bargaining, as reported by the National Mediation Office (Medlingsinstitutet 2023b). In the Swedish context, the model appears to predict wage drift (e.g. Flanagan et al. 1976), which are the yearly wage increases above collective agreement levels.

The rest of the paper proceeds as follows. We begin with a brief summary of our theoretical model and the empirical approach. Next we present data and specify our models, followed by results, and ultimately a conclusion and reflection on our results.

2 Theory and empirical approach

The two main aims of this paper are to assess how labor market competition and labor market institutions affect the wage bargain. The latter is central if we believe Manning's (2003) prediction that powerful unions can mitigate monopsony effects in labor markets. If that is the case in Sweden, how does labor demand and the strength of worker outside options affect the wage bargain?

2.1 The Nash bargaining model

To assess labor market competition's effect on the wage bargain we employ an asymmetric Nash bargaining framework, as proposed by Manning (2011), to consider how the value of the wage bargain is affected by the probabilities and payoffs received by workers staying in their current job or switching to outside options. Conceptually, the strength of the outside option approximates how competitive the labor market is from the employer's perspective. In tight labor markets, labor is scarcer and competition among employers is greater than in slack labor markets.

To assess the strength of outside options we will primarily rely on two metrics: employer concentration, assessing the number of employers hiring workers in a defined labor market, and an index indicating the relative proximity of nearby labor markets and expected payoffs from making occupational changes (from SST), allowing us to assess market definition concerns and incorporate the wage impact from job ladders. The advantages of this model is that it allow us to incorporate aspects of a dynamic search-and-matching framework in labor markets with imperfect competition (Burdett and Mortensen 1998), define labor markets flexibly, and attain intuitive measures of employer concentration.

The Nash bargaining framework considers a worker and an employer at firm i bargaining over the worker's wage at some time-period, which can be represented as:

$$w_i = argmax_w (w_i - oo_i)^{\beta}(p_i - w_i)^{1-\beta}$$

where the worker's payoff is wage w_i minus a disagreement point in the form of taking some outside option (oo_i). The employer's payoff is the firm's generated surplus p_i minus the disagreement point w_i, which is the employer's value of the vacancy (i.e. the worker's wage). As the framework considers the distribution of a surplus between wages and profits, bargaining power can be treated as a relative and purely distributive term (Walton and McKersie 1991), where the worker's bargaining power is β and the employer's bargaining power is $1-\beta$, which is normalized to $\beta \in [0, 1]$. When maximizing the surplus from the match on wages, the Nash bargaining outcome yields:

$$w_i = \beta p_i + (1 - \beta) oo_i$$

where βp_i is interpreted as the product of the match, making w_i the difference between the matching product and outside options, subject to each party's relative bargaining power. Holding β constant, the wage bargain is a function of the firm's surplus and the value of available outside options.

The effect on the wage from labor market competitiveness can thus be interpreted as the relative value of the outside option compared to staying on and receiving the wage offered by the current employer. Assuming there is firm heterogeneity, two identical workers with identical outside options (neither of their respective firms have relevant vacancies) will find these outside options varyingly appetizing if the first worker is matched to a firm generating larger surpluses (paying higher wages) and the second worker is doing an identical job at a less productive firm.

Allowing for such heterogeneity is important, as worker information about outside options are seldom correct (Jäger et al. 2024) and temporal misalignment of quits and vacancies often produces "natural" frictions in labor markets (e.g. Boal and Ransom 1997, Burdett and Mortensen 1998, or Manning 2003).

2.2 Employer concentration

We proceed by expanding the outside options variable. First by considering how the wage bargain is affected by the number of vacant jobs offered by different employers in the defined labor market.

The value of the outside option is the probability (σ_j) of a worker changing to some other job j and receiving wage w_j (or becoming unemployed and receiving unemployment insurance b at probability σ_u). If the probability of staying with the current employer is σ_i, the probability of taking or receiving an outside option is $1 - \sigma_i = \sum_j \sigma_j + \sigma_u$, where $\sum_j \sigma_j$ is the total probability of taking any outside job option. The value of outside options thus be summarized as the probability weighted sum of expected payoffs from each outside option:

$$oo_i = \sum_i \sigma_j \times w_j + \sigma_u \times b$$

To extend the model to aggregate labor markets we consider mean wages ($\bar{w}$) and aggregate matching probabilities. The probability of matching with another employer ($1 - \sigma_i$) is proportional to the market share of her current employer. If the probabilities of receiving and accepting other offers is non-uniform, such as when employer i is relatively large, the probability of not being rehired by employer i is non-zero ($\sigma_i > 0$). Following SST, the mean wage can thus be expressed as a second-order approximation, $\bar{w} = \sum_i \sigma_i \cdot w_i$, allowing us to express the wage bargain as a function of the sum of square of employer shares (σ_j). If the market share of hires vary between firms, then the probability of taking job offers will vary over firms[3].

[3] See SST for proof.

Aggregate hiring probabilities can thus be conveniently expressed as a Herfindahl-Hirschman Index of employer concentration on hires ($HHI = \sum_i \sigma_i^2$), where an HHI-value of 1 indicates a market dominated by a single employer, and a value close to 0, a market populated by a large number of outside employer options.

We define aggregate labor markets by occupation o, in a regional commuting zone r, and year t. Employer concentration in the defined labor market is obtained by observing the total number of hires in firm j ($H^{x\to j}_{z\to o,r,t}$), where $x \to j$ defines all inflows to the defined labor market any previous state or activity (except $x \neq j$), and $z \to o$ defines an inflow to occupation o from any occupation or unknown occupation, except o if $x = j$, which we do not considered a new hire. Dividing each firm j's total number of hires (H) by the total number of observed hires in the defined labor market ($\sum_{j=1}^{N} H^{x\to j}_{z\to o,r,t}$), yields firm j's market share of all hires as:

$$s_{j,o,r,t} = \frac{H^{x\to j}_{z\to o,r,t}}{\sum_{j=1}^{N} H^{x\to j}_{z\to o,r,t}}$$

which is analogous to the probability of a worker being hired at firm j at time t in the defined labor market, conditional on the worker exercising some outside option. An aggregate measure of employer concentration on hires is obtained by taking the sum of squared of each firm's market share:

$$HHI_{o,r,t} = \sum_{j=1}^{N} s^2_{j,o,r,t}$$

We again note that market shares and concentration capture dynamic inflows of workers to firm j, marking an important conceptual distinction to the market share and concentration of jobs (stocks). Market shares and employer concentration of stocks can similarly be attained by considering the share of workers observed at firm j in the defined labor market[4].

2.3 Labor market definitions, nearby occupations, and the job ladder

A salient critique of price-on-concentration models relates to market definition (SST, Miller et al. 2022); measuring and assessing the wage effects from concentration on a labor market defined by a single occupation ignores the fact that workers may have non-uniform probabilities of taking jobs in other occupations. Some occupations may have a wide range of outside options in nearby occupations, while others are highly specialized or work in geographically isolated labor markets. If the former is true, negative wage effects from concentration should be mitigated by a greater

[4]The market share (and concentration) of hired worker (stocks) can be expressed as: $s^{stocks}_{j,o,r,t} = L^j_{o,r,t}/L_{o,r,t}$, where $L^j_{o,r,t}$ is the number of workers in firm j in the defined labor market.

abundance of outside options than the current market definition allows, whereas in the latter case, concentration might sufficiently define the relevant labor market.

SST addresses this issue by including a probabilistic flexible definition of outside occupation options to the outside option framework. If we assume that workers in occupation o also have relevant outside options in occupation p (at any employer) paying w_p, then the probability of taking an outside job option is $1 - \sigma_i = \sigma_{j,o\to o} + \sigma_{j,o\to p}$, where $\sigma_{j,o\to o}$ is the probability of taking a *new* job in the same occupation and $\sigma_{j,o\to p}$ is the probability of taking a new job in a different occupation. The value of the outside option can then be expressed as:

$$oo_{i,o} = \gamma_o \left(\sum_{j\neq i}^{N_o} \sigma_{j,o} w_{j,o} + \sigma_{u,o} b \right) + \quad (1-\gamma_o) \sum_{p\neq o}^{N_p} \sigma_{o\to p} w_p$$

where the aggregate probability of a worker in occupation o being matched in the same occupation or going into unemployment is γ_o from occupation o, while being matched into some other occupation p is $(1-\gamma_o)$. Thus, in an occupation with zero or very small probabilities of switching to outside occupations, the value of the outside option falls (almost) entirely on concentration – the within-occupation outside option.

Using probabilities of outward occupational mobility, we can construct an outside occupation index (OOI) similar to SST's, capturing the mean wage received by workers moving from o to p ($\bar{w}_{o\to p}$) in the observed period, multiplied by the probability of such occupational changes occurring ($\pi_{o\to p}$) .

If $N_{t-1}^{o\to p}$ are the number of workers observed in o at $t-1$ and p at time t, and $N_{o,r,t}$ are the number of workers observed in o at t, the probability (and share) of workers moving from o to p is:

$$\pi_{o\to p} = \frac{N_{t-1}^{o\to p}}{N_{o,r,t} + N_{t-1}^{o\to p}}$$

The mean value of the outside occupation option can then be expressed as the mean wage received by occupation changes moving from o to p multiplied by their corresponding probabilities of making such occupational moves, yielding the Outside Occupation Index:

$$OOI_{o,r,t}^{o\to p} = \sum_{p\neq o}^{occs} \pi_{o\to p} \times \bar{w}_{o\to p,r,t}$$

Here, our approach differs from SST in that we use the *unconditional* probability of observing occupational mobility in a given year, i.e. the probability of making an occupational switch during some year. SST uses a conditional probability of changing occupations, where the condition is the share of occupation movers among *all observed job changers*., i.e. the probability of changing occupation among all job changers.

Our different approach is based on the following intuition: in an occupation with little outward job mobility (most workers don't change jobs), the number of workers changing occupation might be close or identical to the number of job changers. If there is little outward job mobility, a promotion might be the only viable outside option. In such cases, $\pi_{o\to p}$ would be close to 1. If promotions are rare, and the promotion implies a wage premium, the OOI-variable will yield a relatively high value from both high probabilities and higher wages compared to taking the inside option. The relevant market will thus appear larger and the job ladder more favorable than it actually is[5].

2.4 Job mobility

Both the concentration and outside occupation index require us to attain probabilities of job changes which occur between periods $t-1$ and t. We separate probabilities as three distinct types of job flows moving along two dimensions: changing occupation and changing employer.

1. $\sigma_{\omega\to o}^{x\to i}$: are inflows to occupation o from any employer j which does not imply an occupation change. x denotes inflows from any (or no previous) other employer than i ($x \neq i$). ω are inflows from o or any unknown occupation ($\omega \neq p$, where p is a known occupation other than o).

2. $\sigma_{o\to p}^{j\to i}$: are occupation changes from occupation o to some occupation p, which implies moving to some outside employer employer j. o, p, and j must be known.

3. $\sigma_{o\to p}^{i\to i}$: are occupation changes from o to some occupation p, but staying with employer i. o, p, and i must be known.

4. $\sigma_{o}^{i\to u}$: is becoming unemployed[6].

5. $Stay$: is the probability of not changing jobs, i.e. not being identified by the above variables. It implies holding an identified occupation o and being employed at i in both $t-1$ and t. Workers with an unknown occupation ω in $t-1$ that are assigned an occupation o in t if there is no employer change between $t-1$ and t ($i \to i$). Thus, stayers who had an unknown occupations in $t-1$ are not considered as job changes.

The separate job mobility variables are summarized in table 1, below, along these two dimensions.

[5] We provide a graphical illustration of the difference between the measures later in the paper.

[6] Unemployment as an outside option is included for conceptual clarity, but will not be considered further in the analysis.

Table 1: Job-change matrix approximating probabilities for different types of labor market mobility

Firm \ Occupation	o or ω	p
i	$Stay$	$\sigma_{o\to p}^{i\to i}$
j	$\sigma_{\omega\to o}^{x\to i}$	$\sigma_{o\to p}^{i\to j}$
$Unemp.$	$\sigma_{o}^{i\to u}$	$-$

Although we only consider the sum of inter- and intra-firm occupation changes when calculating the outside occupation index, and exclude unemployment in our analysis of the data, this empirical approach allow us a wider net than we do here to capture the relative strength of relevant outside options. For example, the model allow us to observe job flows in both external and internal labor markets, and thus allowing us to assess impacts to the wage bargain from having strong or no internal labor markets (Osterman 2024).

2.5 Collective bargaining

The outside option-variables (HHI and OOI) are modelled to capture individual bargaining power as a function of labor demand in the defined labor market, where having more outside options implies a greater demand for labor. As labor demand improves, so does the threat of *exit* to outside options (Hirschman 1970). Rational employers should thus respond by adjusting wages to market rates, lest risk losing employees to the competition.

But if we suspect that collective bargaining has a large impact on the wage bargain in local labor markets, how do we separate improvements to the wage bargain which derive from individual bargaining power and collective bargaining power? Addressing this question is central to identification when estimating wage effects from concentration and outside occupation options (*exit*).

Collective bargaining power reflects a political *voice* mechanism resulting from unions monopolizing labor supply (Freeman and Medoff 1984), rather than the purely market-driven components of the outside option variables. Collective bargaining is a highly heterogeneous form of institutional intervention (Bhuller et al. 2022) which may vary between and within national settings. The Swedish National Mediation Office recognizes seven forms of wage setting practices found in the hundreds of national sectoral collective agreements, covering a wide variety of practices and wage setting procedures for different occupational groups, ranging from entirely individual to delegated wage setting procedures where union representatives negotiate for its entire membership group (Medlingsinstitutet 2022). Further, Sweden does not have legislated minimum wages and minimum wage regulations found in collective agreements are heterogeneous in multiple dimensions (Medlingsinstitutet 2023a) .

The complexity and heterogeneity of collective bargaining regimes thus poses serious modelling challenges, which are also central to our identification. In some labor markets, collective agreements may simply ”take wages out of competition”

(Freeman and Medoff 1984) resulting in wages or wage increases that are always ”on par” with the collectively bargained outcomes. In other labor markets, collective agreements may simply provide a ”floor” which still allow for considerable market-driven wage setting, resulting in positive wage drift (Flanagan et al. 1976), meaning aggregate wage increases above the collectively bargained levels. If collective bargaining coverage is low, and workers have little individual bargaining power due to poor market demand, wage drift may be negative resulting in wages below the collectively bargained levels.

Assessing individual and collective bargaining power simultaneously is important if we wish to understand the dynamics of individual wage and collective wage setting power in relation to employer wage setting (monpsony) power.

To address this concern we will assume that *individual* bargaining power components are captured by the outside option model specified above, whereas *collective* bargaining power effects are captured by the yearly regional ($\alpha_{r,t}$) and occupational ($\alpha_{o,t}$) fixed effects, and verify this assumption by comparing fixed effects estimates with collective bargaining outcomes[7]. Our intuition is that occupational and regional variation will largely capture heterogeneity in national collective agreements.

If this intuition holds, the fixed effects estimates of the model should yield yearly mean wage levels that approximates collective bargaining agreement outcomes. However, central/sectoral collective bargaining agreements tends to regulate yearly wage increases rather than levels. By calculating the yearly changes of the estimated fixed effects wage levels, these estimates should approximate observable wage increases found in collective agreements.

In sum, if the model is well-specified, the fixed effects estimates should capture collective bargaining wage increases, whereas the outside options components should capture wage drift. The fixed effects-approach provides a flexible form of identification without imposing implicit assumptions of collective bargaining causality to the model. However, it requires us empirically verify that the fixed effects estimates are consistent with collective bargaining outcomes.

[7]The usefulness and importance of studying how fixed effects estimates impact the wage bargain have recently been lifted by Card et al. (2024).

3 Data and model specifications

We now present the data and specify the models that will be used to estimate outside option and collective bargaining effects to the wage bargain. We begin by presenting market defining-variables, followed by wages, working hours, income, value added, and collective bargaining data, which require little to no processing. We follow by presenting variables that require more processing, including our definitions of labor market flows, concentration, and the Outside Occupation Index. The models that will be used in our analysis are sequentially presented as new data and instruments are added. We conclude the section by briefly discussion limitations to the study.

3.1 Main data sources and market defining variables

We make use of detailed employer-employee matched administrative data from Statistics Sweden's extensive databases using variables relating to job mobility, wages, working time, incomes, and firm characteristics. The dataset includes merges from the LISA[8], Business Registry, Wage Structure Survey, and Occupational Registry. Variables are merged to either individual or firm serialized identification numbers. Our data covers the years 2006 to 2020, which includes observed dynamics between 2005 and 2006. Individual and firm identifiers are largely based on information filed in November tax filings, whereas data on wages and working hours are collected between September and November each year. In total, the data set includes 51,520,795 individual observations.

Looking first to the labor market-defining variables, occupations (o) are drawn from the Occupational Registry using the finest available SSYK4 occupation codes, capturing a worker's primary occupation in the observed year. Between 2005 and 2013 occupations are reported according to the SSYK96-standard (355 occupations) and between 2014 and 2020 they are reported according to the newer ISCO08- compatible SSYK2012 standard (429 occupations).

Officially, the new SSYK2012-standard is incompatible with the older SSYK96 standard due to the inability to merge the coarser occupational definitions used in SSYK96 with the more finely grained occupations used in SSYK2012. To address this challenge, we use Yakymovych (2022) backward-compatible crosswalk, which also implies that we use the older SSYK96-standard of occupations throughout[9].

There are serious issues in the occupational data which relate to the switch of occupation standards in 2014, resulting in missing occupational data in this year as many employers continued to apply the old SSYK96-standard when they were sup-

[8]The longitudinal integrated database for health insurance and labor market studies, (*Longitudinell integrationsdatabas för Sjukförsäkrings- och Arbetsmarknadsstudier*).

[9]As the SSYK96-standard uses a coarser definition of occupations than the newer SSYK2012-standard, concentration measures will therefore be higher than if the SSYK2012 standard was used throughout.

posed to apply the new SSYK2012 standard[10]. We will address the issue of the 2014 gap-year in section 3.3 below, justifying an exclusion of 2014 in our analysis.

To explore heterogeneity and institutional variation in wage setting practices in the Swedish collective bargaining system (see, Bustos 2023 for example), we can further divide the SSYK96-codes into BC- and WC occupations. For this, we follow Statistics Sweden's recommended crosswalk. In Sweden, BC and WC occupational affiliation is determined by which central organization the occupation would belong to in terms of union membership and collective agreement coverage. BC occupations are organized by unions under the Swedish Trade Union Confederation's (LO-SE) umbrella, and WC occupations are most often organized in unions that are either members of the Swedish Confederation of Professional Employees (TCO) or The Swedish Confederation of Professional Associations (Saco).

We define regions as commuting zones (r) using Tillväxtanalys (2016) 60 functional labor market regions (FA-regions[11]) by matching FA-regions with workplace municipal codes. The FA-regions can be further categorized by regional types. To represent regional heterogeneity we will make use of the most aggregated categories, covering (UR1) metropolitan regions (Stockholm, Gothenburg, and Malmö), (UR2) non-metropolitan towns and cities, and (UR3) sparsely populated rural regions. In our data, region type UR1 corresponds to a share of 51.7 percent of the total population, UR2 a share of 41.7 percent, and UR3 a share of 6.6 percent.

Occupations and regions define the time-varying fixed effects variables $\alpha_{o,t}$ and $\alpha_{r,t}$ that will be used throughout, including in estimations of collective bargaining effects.

3.1.1 Wages, working time, and incomes

Gross wages and working hours are drawn from the Wage Structure Statistics, providing individual wages and working percentages in the months of September through November. Wages represent gross full time-equivalent wages, which includes base pay and variable pay (e.g. over-time and weekend compensation). The gross full-time wages for non-full time workers (working less than full time, 100 percent) are estimated to full time wage levels by Statistics Sweden using models that account for working hours.

Wage data quality varies by employer characteristics and worker types and differences in sampling methods. Generally, this implies that sampling frequency increases with firm size. Private sector employer association members report WC earnings and working time as monthly salaries, correcting for working percentages, and BC earnings as an hourly wage (both include variable components) to their respec-

[10]For example, a hypothetical healthcare employer in who mistakenly used the old SSSYK96 occupation codes in 2014 would have reported a large increase of bartenders in her employment (SSYK 2012 code: 5132) replacing assistant nurses (SSYK96-code 5132), which is also Sweden's largest occupation.

[11]Specifically, FA15-regions as defined by labor mobility observed in 2015.

tive central employer associations (mainly the Swedish Confederation of Enterprises). Non-employer association members provide data directly to Statistics Sweden via a sample survey, which aims to cover each worker at least every four years. Public sector employers, where collective bargaining coverage and employer association membership is 100 percent, report monthly wages and hours on a yearly basis based on observed wage payments in the sampled month for both BC and WC employees.

Gross labor incomes, which will be used intermittently with gross wages, are pretax incomes reported to the Swedish Tax Authority in November each year by a workers' primary employer[12]. As gross labor incomes are presented as yearly sums, we attain an average monthly income by dividing the reported income by 12, thus making the variable comparable to the full-time monthly wage variable.

We follow convention and remove workers with little or poor connection to the labor market by filtering out gross labor income earnings below a "price base amount" (*prisbasbelopp*) threshold, which are yearly adjusted legislated standard price units used to calculate social insurance payments (e.g. Bäckman and Nilsson 2016). We have chosen 2.5 price base amounts[13] as our threshold, as it is roughly equivalent to working 50 percent for the minimum wage stipulated in Sweden's largest collective agreement (Kommunal – SKL/Pacta) at the middle of our observed period (2016).

Figure 1, shows the weighted yearly mean regional wages, hours, and incomes of all workers with an observed occupation between 2006 and 2020, including regional and occupational sub-categories.

Notably, we observe a sudden increase of the hours-variable (as a percentage of full time employment) in the years 2014 to 2016. We can also observe small increases in both wages and incomes during these years, raising data concerns relating to the 2014 shift in occupational reporting standards.

3.1.2 Value added

Firm size and productivity may have a positive impact on wages and employer concentration if more productive firms hire a larger share of workers than less productive firms. We address this factor by including a measure of value added, collected from the Business Registry data, covering all organization numbers in the private sector. Value added numbers are merged to each individual's corresponding organization number. To attain aggregate value added estimates in defined labor markets, we take the mean value of each individual's corresponding employer value added, resulting in each defined labor market having a weighted value added value (see figure 2).

Public sector organizations are treated as empty values. Thus, a presence of large public sector employers will not be accounted for when we use value-added as a proxy for firm size.

[12]We use the variables *KU1Ink* (2005-2018) and *AGI1Ink* (2019-2020), and thus only consider incomes from the worker's primary employer (all incomes are found in the variable *LoneInk*).

[13]Bäckman and Nilsson (2016) use 3.5 price base amounts.

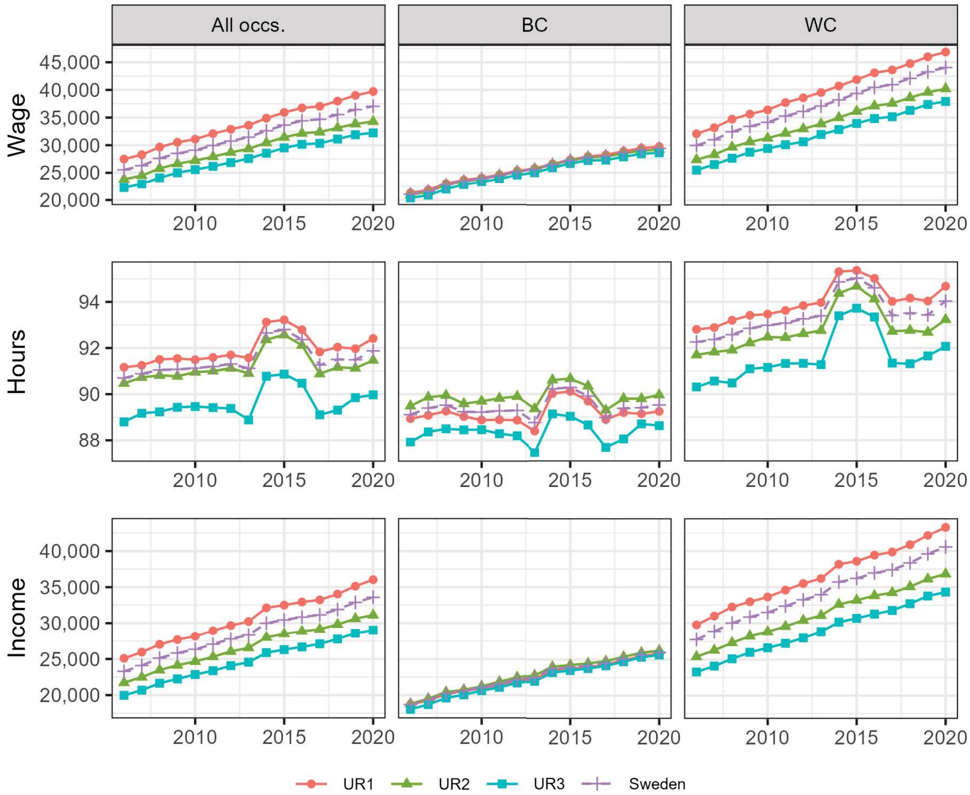

Figure 1: The figures show yearly mean wages, hours, and income-variables, which are the dependent variables used within. The figures are divided into all, blue-collar (BC), and white-collar (WC) occupation groups, and further subdivided into three regional types (UR1 - metropolitan, UR2 - non-metropolitan cities, UR3 rural regions), and the national average (Sweden). Monthly gross wages and working hours (percent of full time employment) are collected from the wage structure survey, while gross income are gross yearly labor incomes from an individual's primary employer reported to the tax authority, which we divide by 12 to get the mean monthly income.

3.1.3 *Collective bargaining*

As we also expect collective bargaining to have a significant impact to the wage bargain, our analysis also includes data for the yearly collectively bargained wage increases set in coordinated national wage negotiations collected from the Swedish National Mediation Office.

Swedish collective bargaining coverage is close to 90 percent throughout our observed time-period (Medlingsinstitutet 2023b) – without artificial extension[14] – and union density is close to 70 percent (Kjellberg 2023).

Since 1997, Swedish wage setting is based on a centrally coordinated collective

[14]Which is common in many other European settings, and implies that centrally negotiated collective agreements (on wages, for example) are extended to all firms in the sector which it regulates, regardless if the firms in the sector are organized or not.

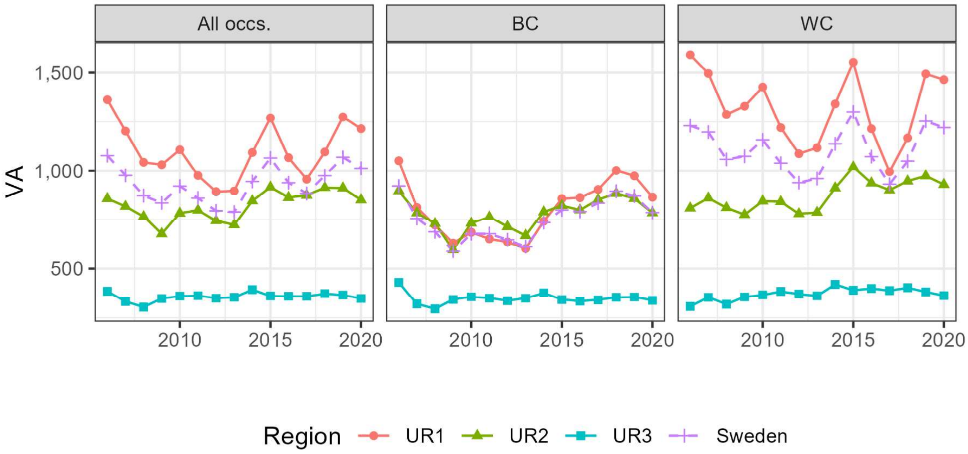

Figure 2: The figure shows weighted mean value-added in nominal terms, in millions of SEK, divided into the occupational groupings all, blue-collar (BC), and white-collar (WC) occupations, and further subdivided into three regional types (UR1 - metropolitan, UR2 - non-metropolitan cities, UR3 rural regions), and the national average (Sweden).

bargaining which sets percentage yearly wage increases, known as *Märket* or the Mark ($Mark_t$). The Mark's coordination is based on annual, biannual, or triennial national collective bargaining rounds beginning with negotiations between unions and employers associations in export-oriented industries. Once these agreements are concluded, wage agreements are negotiated in other sectors and industries, which almost always end up setting the Mark as the yearly wage increase. Thus, export-oriented industrial sectors set the norm for yearly wage increase (in percent) in subsequent collective bargaining rounds follow creating a national standard (e.g. Carlén and Hållö 2020).

The Mark is enforced in subsequent bargaining rounds by unions and employer organizations through member coordination. Unions coordinate across industries and occupations and can apply sympathy strike actions to assure that workers in all sectors receive wage increases set in the Mark, backed up by large strike funds. Employers similarly coordinate across industries to assure that subsequent contracts do not go above the Mark. Employers can enforce their demands by coordinating lockouts and offering strike- and lockout insurance, backed up by large lockout funds. Further, the National Mediation Office provides mandatory mediation, which implies that state appointed mediators are also tasked to assure that subsequent collective agreements fall within the Mark. (e.g. Bender 2024, Chapter 1)

The white bars in figure 3 shows the mean observed collectively bargained negotiated wage increases for each observed year ($Mark_t$). The shaded bars shows the remaining observed wage increases in the economy (with their percentage increases in brackets), indicating *wage drift*. The data is collected from the National Mediation

Office's "Wages and salaries in the private and public sector"-statistics[15] .

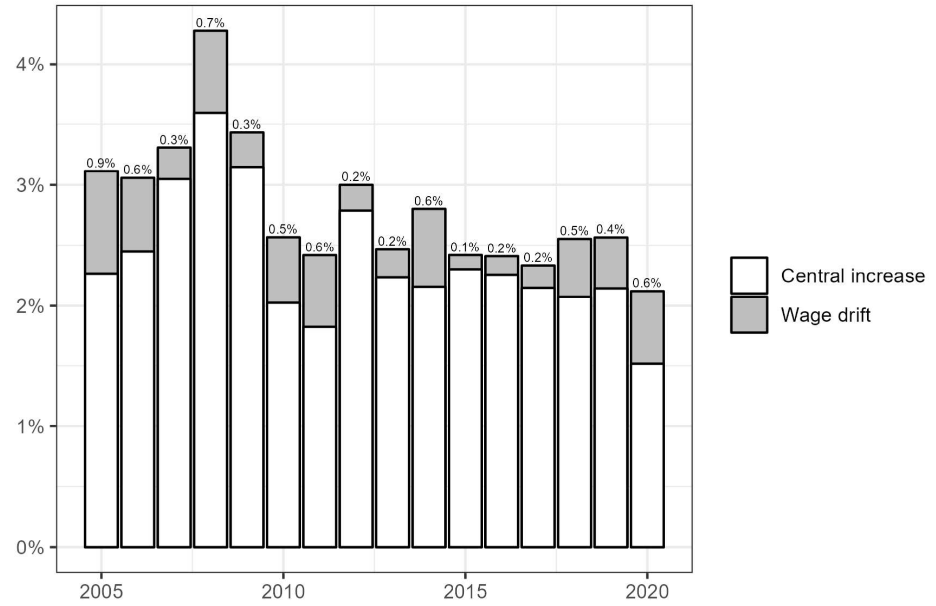

Figure 3: The figure shows the mean yearly collectively bargained wage increases in white, with the remaining observed wage increases in grey indicating wage drift in the Swedish economy between 2005 and 2020.

3.2 Estimating labor market flows

The concentration and outside occupation index-variables both require us to measure labor mobility (or flow) statistics, as presented in the job change matrix in table 1.

Starting with inter firm mobility ($i \to i$), we define the firm boundary by its "organization" or its corresponding "concern number", attained from the Business Registry Database. If a firm is a subsidiary to a concern (a mother corporation) the concern number identifies the firm. If not, the organization number defines the firm boundary.

Moves between firms is captured by a dummy variable indicating an individual being observed in different firms between in year $t-1$ and t. Further, we correct for "false" changes of employership, which may result from changed ownership structures that would typically not imply a new employment contract for a worker. We correct for such "false" changes by interacting the employer change dummy with a geographically consistent workplace number variable (*Cfarnummer*). If a worker is

[15] *Konjunkturlönestatistiken*, specifically the diagram *Faktiska och avtalade löner efter sektor per år och månad*)

observed changing firms, but the geographic workplace number is unchanged, we do not code the worker as changing firms.

Moves between occupations is similarly recorded as dummy variables by observing changes to the SSYK4-codes between years $t-1$ and t. Workers with missing occupational codes in year t are removed.

Within-occupation flows to o ($\sigma^{x \to i}_{\omega \to o}$) implies either observing a within-occupation employer change, observing a hire from an unobserved employer (e.g. unemployment or schooling), or inflows from an unknown occupation (s.t. $x \neq i$).

Outside-occupation flows ($o \to p$) for occupation o occurs when we observe a worker with occupation o in $t-1$ observed in some other occupation p in t. These flows are further divided into occupational changes within ($\sigma^{i \to i}_{o \to p}$) or outside ($\sigma^{i \to j}_{o \to p}$) the firm boundary, conditioned on the firm movement dummy described above.

All workers who are not recorded as changing occupation or employer are recorded as taking the *Inside option*. Figure 4 summarizes the size and share of each category of movers between 2006 and 2020. We also include the sum of individuals within the income threshold without occupation codes (dashed box).

Figure 4: The figure illustrates the total number of (and share in %) of workers taking outside options (three types of job changes) or inside options in a given year (i.e. being stationary), and the number of individuals in the data missing an occupation (in thousands, k) -- i.e. workers having more than 2.5 Price Base Amounts of labor income in a year.

3.3 Addressing data quality concerns

We note in figure 4 that the change of occupation standards which occurred in 2014 poses significant problems to our data[16] as correctly measuring occupational labor market flows are key to estimating the strength of outside options. Therefore, the years post 2014 (and 2014 in particular) risks tainting our analysis and results.

First, a significant problem relates to the large number and share (27.1 percent) of workers observed as changing occupations within the same employer in 2014 ($\sigma^{i \to i}_{o \to p}$). This comes despite using Yakymovych's (2022) crosswalk. Of consequence to our later analysis is a significantly smaller share of "stayers" taking the inside option (56.6 percent).

The sudden rise in within-employer occupational changes in 2014 relates to the practical exercise of employers having to change a large number of occupation codes due to the standard shift. This is likely a result of employers registering workers in occupational codes that were outside of the official crosswalk provided by Statistics Sweden. Although this may raise concerns about the quality of within employer occupational changes, it is plausible that employers will not make yearly active choices to describe a worker's occupation code. Rather, if employers used occupation codes outside of the provided crosswalk, and this definition is applied consistently on larger groups of workers, this should not be a problem, unless there occurs a significant change of the job description of individual workers between years which are not registered as occupation changes.

Second, the large number of missing occupations in 2014 and 2015, which then gradually reduces up to 2020, is a cause for concern. The most plausible explanation to the large number of missing variables relates to a lag in the collection of occupational codes after 2014. The steady decrease of missing workers over our income limit of 2.5 price base amounts indicates that missing workers are successively sampled for the occupational registry data. We should also note that the relatively large increase in observed individuals does not directly imply higher employment in Swedish labor markets, rather, a growing number of individuals with earnings above our income threshold [17].

We solve this problem by treating 2014 as a lost year and omitting all 2014 observations from the data. Consequent diagrams and variables that are presented all omit data from the year 2014. Uncensored diagrams (including 2014) can be found in the appendix.

[16] Using the unadjusted SSYK4 2012-variable results in a large omission of occupational codes; in 2014 there are over 962,000 individuals missing occupation codes and over 926,000 in 2015; a number which steadily reduces until 2020.

[17] Looking at employment statistics, there is no large bump in 2015. However, employment statistics use much lower income thresholds than 2.5 price base amounts.

3.4 Estimating employer concentration on hires

To attain measures of employer concentration on hires ($HHI_{o,r,t}$), we begin by calculating firm j's market share of new hires in our defined labor market ($s_{j,o,r,t}$), by dividing firm j observed yearly hires ($H^{x \to j}_{z \to o,r,t}$) by all hires in the defined labor market ($\sum_{j=1}^{N} H^{x \to j}_{z \to o,r,t}$). By taking the sum of squares of all market shares we attain a Herfindahl-Hirschman Index ($HHI_{o,r,t}$) of employer concentration on hires. Hires are attained by using all the labor mobility variables defined above, and targeting them towards occupation o.[18]

In figure 5 we can follow the weighted mean employer concentration in Sweden, in the metropolitan (UR1), non-metropolitan towns and cities (UR2) and in the sparsely populated rural regions (UR3) between 2006 to 2020, for all, blue- and white-collar occupations.

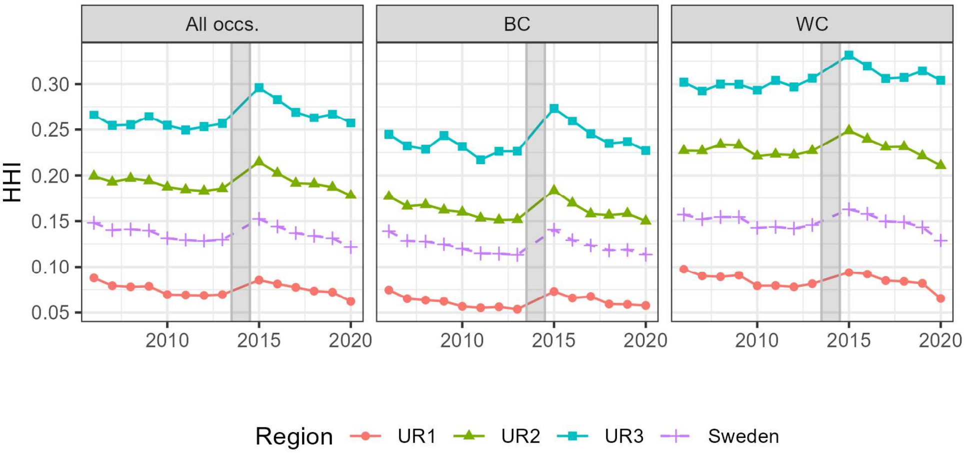

Figure 5: HHI of concentration for all occupations (All), and subdivided by blue-collar (BC), and white-collar (WC) occupations. Further divided by regional dimensions (UR1:3) and the national average (Sweden). A HHI-value of 1 indicates a single employer in the relevant labor market, whereas values closer to zero indicates a greater abundance of employers. Shaded areas indicate the censored year 2014.

Despite censoring values for 2014, we still note an uptick in concentration in 2015 which steadily drops in the following years. Comparing 2015 with 2013 (the year before the standard change), we record 141,929 occupational a inflows from unknown occupations in 2015, compared to 33,929 in 2013, but only 328,355 inflows from known occupations, compared to 442,511 in 2013.

Considering the regional aspects of concentration, the weighted mean of the HHI for all observed labor market regions for all years is illustrated as a map of Sweden in figure 6.

[18]Which implies setting $\sigma^{i \to i}_{p \to o}$ and $\sigma^{j \to i}_{p \to o}$ to attain the relevant inflows to the firm.

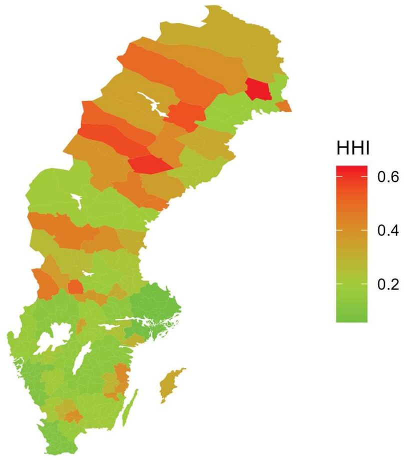

Figure 6: Employer concentration on hires, estimated as the weighted mean in Sweden's FA15-regions (commuting zones) between 2006 and 2020.

Similar to Azar et al. (2019) we can model the wage bargain as function of concentration using a log-log regression with time-varying occupational ($\alpha_{o,t}$) and regional ($\alpha_{r,t}$) fixed effects:

$$log(w_{o,r,t}) = \alpha_{o,t} + \alpha_{r,t} + \gamma_1 log(HHI_{o,r,t}) + \varepsilon_{o,r,t} \tag{III.1}$$

3.5 Estimating the outside occupation index

SST's outside occupation index is constructed from two components: the probabilistic measure of outward occupational mobility and the mean wage received by identified occupation changers.

Unconditional and conditional probabilities of occupational mobility

We begin by addressing the measure of outward occupational mobility ($\pi_{o\to p}$), which deviates from the SST model by using unconditional rather than conditional probabilities of changing occupations. The difference between these probability metrics are illustrated in figure 7, comparing the unconditional probability used within ($\pi^{Uncond.}_{o\to p}$, x-axis) to SST's conditional probability ($\pi^{Cond.}_{o\to p}$, y-axis), illustrated by the weighted mean values of all 355 occupations between 2006 and 2020 (excluding 2014).

Our unconditional weighted mean probability of changing occupations for all observed labor markets between 2006 and 2020 is 13.0 percent (13.0 percent for both BC and WC), whereas the conditional probability from the SST-method yields a mean of 53.2 percent (51.8 perecent BC and 54.5 percent WC). The (statistically

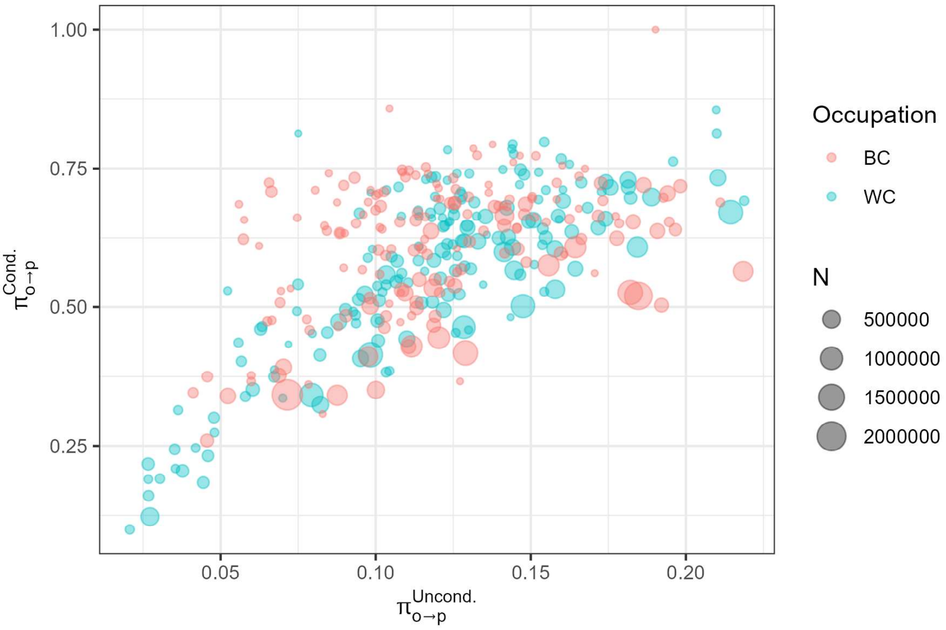

Figure 7: The diagram illustrates the differences between the unconditional ($\pi^{Uncond.}_{o \to p}$) and conditional ($\pi^{Cond.}_{o \to p}$) probability of changing occupation, summarized as the weighted mean between 2006 and 2020 among all blue- and white-collar occupations (total 355). Dot sizes are scaled to represent the relative size of the observed occupation. The unconditional probability is the "pure" probability of changing from occupation o to some other occupation p between $t-1$ and t (estimated as the share of workers in occupation o changing to some other occupation p), while the conditional probability is the probability of changing occupation among all workers observed changing jobs (as defined in table 1), similarly estimated as the share of all occupation changers among job changers.

significant) correlation between the unconditional and unconditional probabilities is 0.680. An ocular observation tell us that a best fitted slope would be non-linear, with considerable variation in conditional mobility in the mid range of unconditional occupational mobility (0.10 to 0.15). This confirms our concerns relating to plausible overestimation of outward occupational mobility from using SST's conditional probabilities, and validate our choice to use unconditional probabilities for the OOI-variable.

The OOI and its components

Using the same temporal, regional, and occupational divisions as the HHI-figure (5) above, the (unconditional) probability of changing occupations ($\pi_{o \to p}$), the wage levels of observed occupation changers ($w_{o \to p}$), and the outside option index on wages ($OOI_{o,r,t}$) are visualized in figure 8.

Apart from noting some cyclical patterns in the OOI-variable derived from $\pi_{o \to p}$, we also note a spike in the $\pi_{o \to p}$-variable in 2008, which appears to be driven by unusually large within-employer occupational shifts (see 2008 in figure 4).

Figure 8: The figures show the yearly mean development in the OOI, $\pi_{o\to p}$, and ΔOOI-variables. The figures are divided into all, blue- (BC), and white-collar (WC) occupations, and further subdivided into three regional types (UR1 - metropolitan, UR2 - non-metropolitan cities, UR3 rural regions), and the national average (Sweden). The shaded area indicates the omitted year 2014 due data quality issues discussed within, see figure 16, appendix for uncensored version.

The OOI-variable not only provides a means to address the market definition problem – it also tells us about the relative strength of the job ladder. Figure 8 supports this intuition to some extent. Outside occupation options are greater in metropolitan regions compared to smaller labor markets, and largest for WC occupations, where occupation-changer wages and probabilities of occupation changes are greatest.

Using this data we can regress the SST model where (log) wages are a function of (log) concentration and the (log) outside option index, including regional and occupational fixed effects as above:

$$log(w_{o,r,t}) = \alpha_{o,t} + \alpha_{r,t} + \gamma_1 log(HHI_{o,r,t}) + \gamma_2 log(OOI^{o\to p}_{o,r,t}) + \varepsilon_{o,r,t} \quad \text{(III.2)}$$

3.6 Identification concerns relating to value added and concentration effects

As noted by both SST and Névo (2024), employer concentration may also be a result of larger firms being more productive than smaller firms. If a firm is large due to higher productivity, this might have a positive impact on the wage bargain as it implies grater opportunities for rent sharing (e.g. Card et al. 2018). With improved productivity, large firms may end up hiring a larger share of workers than smaller less productive firms. Thus, higher concentration might have a positive impact on the wage bargain.

To correct for positive concentration effects on wages from firm size, we include the mean value-added variable $VA_{o,r,t}$ to our models. Figure (III.2) adds value added in a "naïve" regression, yielding:

$$\begin{aligned} log(w_{o,r,t}) = \alpha_{o,t} + \alpha_{r,t} + \gamma_1 log(HHI_{o,r,t}) + \gamma_2 log(OOI_{o,r,t}) \\ + \gamma_3 log(VA_{o,r,t}) + \varepsilon_{o,r,t} \end{aligned} \tag{III.3}$$

But as we expect concentration and value added to be correlated, we interact the concentration and value-added variables yielding:

$$\begin{aligned} log(w_{o,r,t}) = \alpha_{o,t} + \alpha_{r,t} + \gamma_1 log(HHI_{o,r,t}) + \gamma_2 log(OOI_{o,r,t}) \\ + \gamma_3 log(VA_{o,r,t}) + \gamma_4 log(HHI_{o,r,t}) \times log(VA_{o,r,t}) + \varepsilon_{o,r,t} \end{aligned} \tag{III.4}$$

where our analysis primary interest lies in studying concentration coefficients when holding value added constant[19].

3.7 Using incomes and hours to estimate wages

To extract more information about why workers may change jobs to outside options, we exploit the fact that the wage variable represent a full time equivalent monthly wage; by multiplying the full time equivalent wage by the working hours reported by employers, the result should approximate a worker's received average monthly income. Consequently, wages can be expressed as incomes divided by hours.

Increasing the number of hours worked, and thus incomes, are a probable motivation for workers changing jobs, and reflect important amenities relating to the nature of the job and its expected payoff.

Part-time work may be voluntary, reflecting preferences or private restrictions outside the control of the employer; or involuntary reflecting fewer hours worked due to actions by an employer. Regardless if part-time work is voluntary or involuntary, a voluntary part-time worker might switch to a job with full time employment if the occupation does not impede on private preferences or restrictions. A WC occupation

[19] We provide a guide on how to interpret interaction variables in the Results-section, below.

with ample opportunities to work flexible hours might make full-time employment feasible, compared to a BC occupation where being physically present in the workplace during set hours make full-time employment infeasible. Further, some workers may use a higher wage to reduce worked hours. Regardless of motives, including incomes and hours to our models is prudent as the variables may reflect important motivations driving workers towards outside options.

From our data we can extract yearly gross labor income from the primary employer ($Y_{o,r,t}$). Attaining the gross average monthly income is then $\bar{y}_{o,r,t} = Y_{o,r,t}/12$. Gross average monthly labor incomes can be expressed as $\bar{y}_{o,r,t} = w_{o,r,t} \times h_{o,r,t}$, where $w_{o,r,t}$ are monthly full-time wages and $h_{o,r,t}$, which are hours represented as a percentage of full-time employment[20].

Wages can thus be expressed as a function of hours and incomes. A model regressing (log) wages on some function of outside options ($f(oo_{o,r,t})$) can the be expressed as:

$$\begin{aligned} log(\overline{w}_{o,r,t} &= f(oo_{o,r,t}) \\ log(\bar{y}_{o,r,t}/h_{o,r,t}) &= f(oo_{o,r,t}) \\ log(\bar{y}_{o,r,t}) - log(h_{o,r,t}) &= f(oo_{o,r,t}) \\ log(\bar{y}_{o,r,t}) &= f(oo_{o,r,t}) + log(h_{o,r,t}) \end{aligned}$$

Figure 9 shows the mean reported working hours for workers who make occupation changes between $t-1$ and t, split into the BC and WC subsets of occupations. The x-axis ($h_{o\to p,t-1}$) shows the mean working time in year $t-1$, when the worker is employed in occupation o, and the y-axis ($h_{o\to p,t}$) shows the mean observed working hours for the workers in the year that they have switched to some occupation p. Circles on the dashed line indicate that there are no changes to mean working hours from changing occupation, whereas values above the dashed line indicate that workers, on average, increase their working hours when changing occupation. Circle sizes indicates the average number of occupation movers observed in each occupation group between 2006 and 2020.

We note that for for large BC groups, an occupational change often implies an increase in working hours, supporting the intuition that changes in hours might be an important parameter when workers change jobs or occupations.

The diagram reveals that incomes and hours may be a salient motivation for the BC group to change occupations, and thus we expect the Outside Occupation Index to be most affected by including incomes and hours in the analysis.

The approach requires wages and hours to approximate average monthly incomes relatively well, or more specifically, that hours estimates from a regression model should be able to adjust the dependent income variable to a full time equivalent wage. We explore this question in the appendix (How well do wages and hours predict

[20] $h_{o,r,t} \sim \beta[0,1]$

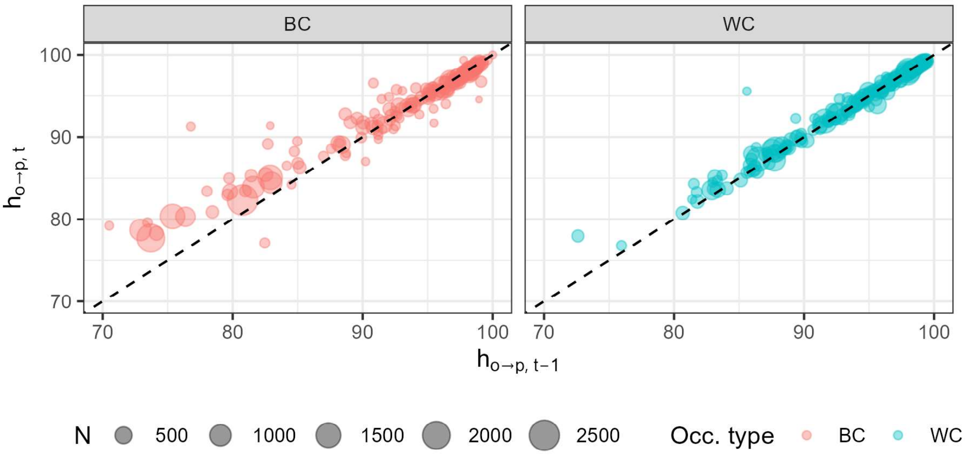

Figure 9: The figure shows weighted mean working hours of workers who change occupations in the year before the occupational move ($h_{o\to p,t-1}$, x-axis) and the working hours observed in the year when the worker is observed in new occupation p ($h_{o\to p,t}$, y-axis), for all BC and WC occupations. Each circle represents an SSYK4-level occupation for o, and the size of the circles indicate the mean number of individuals observed moving way from occupation o to p between 2006 and 2020. Values above (below) the black dashed line indicate that, on average, occupation changers increase (decrease) their working hours when changing occupations.

incomes?), finding mixed results; the method works well for blue collar occupations but less so for white collar occupations.

Assuming that average monthly gross labor incomes and hours roughly approximate wages by including hours in the right-hand side of the equation, a model of income and hours (estimating wages) yields:

$$\begin{aligned} log(y_{o,r,t}) = \alpha_{o,t} + \alpha_{r,t} + \gamma_1 log(HHI_{o,r,t}) + \gamma_2 log(OOI^{y}_{o,r,t}) \\ + \gamma_3 log(VA_{o,r,t}) + \gamma_4 log(HHI_{o,r,t} \times VA_{o,r,t}) \\ + \gamma_5 log(h_{o,r,t}) + \varepsilon_{o,r,t} \end{aligned} \tag{III.5}$$

where $log(OOI^{y}_{o,r,t})$, or the Outside Occupation Index for incomes, is a function of both incomes and hours received in occupation p:

$$log(OOI^{y}_{o,r,t}) = log(\sum_{p\neq o}^{occs} \pi_{o\to p} \times \frac{\bar{y}_{o\to p,r,t}}{h_{o\to p,t}})$$

where $\bar{y}_{o\to p,r,t}/h_{o\to p,t}$ is the estimated wage using the identity descdibed above. We present the components of the Outside Occupation Index from incomes in figure 15 in the appendix.

3.8 Addressing endogeneity concerns from collective bargaining in the OOI

As we expect job changes to (often) be motivated by opportunities provided in options, staying in a job can also have a positive impact on wages or incomes if there is collective bargaining, where wages are raised via the *voice* channel.

Figure 10 shows the average wage increases received by occupational movers for each occupation during the entire observed period ($\Delta w_{o \to p}$, y-axis), along with the probability of changing jobs ($\pi_{o \to p}$, x-axis). The circles are defined by BC and WC occupations, and dot size indicates the sum of all movers between 2006 and 2020. The average wage increases from The Mark is included in the black dashed line, while the blue and red dashed lines indicate the weighted means of $\Delta w_{o \to p}$ and $\pi_{o \to p}$ for both occupation groups.

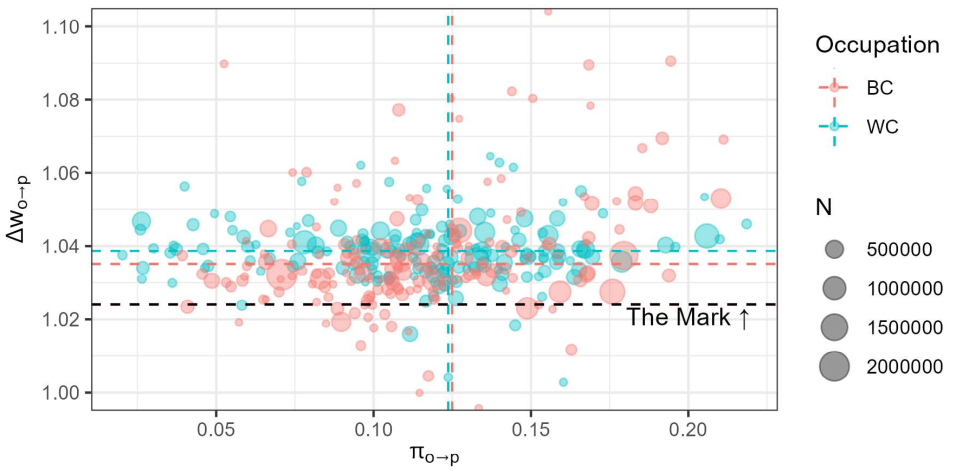

Figure 10: The figure shows the mean wage increases ($\delta w_{o \to p}$) and mean probabilities ($\pi_{o \to p}$) of all 355 occupations observed in our data, relating the wage increases to the mean increase of collectively bargained The Mark in the observed time period.

The mean wage increases observed by occupational movers are, on average, approximately one percentage unit higher than the average collective bargaining wage increases set by the Mark. The relative proximity and bunching just above the Mark indicates that collective bargaining might have an influence on the average worker's payoff when changing occupations[21].

This concern raises issues relating to identification in the Outside Option Index-variable, as it may capture wage increases that are a function of collective bargaining ($Mark_t +$ outside option markup) and not individual bargaining power.

[21]Why take a job which implies more responsibility or more effort if it does not imply a sufficiently large wage bump?

To address this concern we construct two instrumental variables (IV). For wages, the IV takes the wage of occupational movers in the year prior to the occupation change ($\bar{w}_{o\to p,t-1}$) and multiplies it with the yearly negotiated wage increases set in the Mark, ($Mark_t$), yielding:

$$Z^{w,OOI}_{o,r,t} = \sum_{p\neq o}^{occs} \pi_{o\to p} \times \bar{w}_{o\to p,r,t-1} \times Mark_t$$

where $Z^{w,OOI}_{o,r,t}$ is the instrument for wages. For income regressions, the instrument $Z^{y,OOI}_{o,r,t}$ is similarly attained by multiplying yesteryear's incomes divided by hours in model III.5, and multiplying this wage estimate by the Mark.

As the instrumental variable is based on wage increases set in national collective bargaining rounds by employers and unions ($Mark_t$), the instrument is arguably exogenous to decisions made by an individual occupation mover, but endogenous to the wage (or income) that the worker is likely to receive.

3.9 Limitations

As one of the motivations of this paper is to test, apply, and validate the SST model on registry data in a setting with a strong impact to wages and incomes from collective bargaining, our empirical strategy follows their approach of weighted regressions of mean values in occupation and commuting zone-defined labor market. This imposes several limitations to our study. For example an inability to use firm, industry, or individual fixed effects. However, as the SST model strives to assess aggregate developments from the perspective of labor supply, the choice of using defined labor market summations and weighted regression can still yield interesting and valuable results.

The study does not consider how the probability of unemployment or the size of unemployment insurance payments impact the wage bargain. Adding such variables in future studies could be interesting but is outside the scope of this study.

We do not separate the OOI as occupation changes as inter- or intra- firm mobility. This approach is plausible and should allow us to test wage effects from the presence of internal labor markets (e.g. Osterman 2024). Again, validating the SST model is a priority in this paper, but separating the wage impact from access to internal or external labor markets is an interesting concept for future research.

Another limitation in our study is that we use a single, aggregate definition of the Mark. In reality, different collective agreements expire at different intervals during the year, which should yield variation in the mean central collective agreement wage increases for workers under different collective agreements, as is done in Bustos (2023). Not taking such heterogeneity into account in our empirical strategy may yield deviations in our results.

3.10 Summary statistics

Summary statistics for the variables used in our regression models are presented in table 2.

Table 2: Summary Statistics for the BC and WC occupation subsets

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
			BC				
Year	63552	2013	4.5	2006	2009	2017	2020
N	63552	374	1307	1	35	261	34972
w (Wage)	63552	25813	4432	3360	22709	28377	95010
$\bar{y}$ (Income)	63552	23231	5245	6875	19538	26322	321350
Value Added	63552	601169	1281488	2.1	81244	600202	35155643
HHI	63552	0.29	0.26	0.0017	0.093	0.39	1
$\pi_{o \to p}$	63552	0.12	0.069	0.0013	0.073	0.16	0.63
$w_{o \to p,t}$	63552	25813	4432	3360	22709	28377	95010
$w_{o \to p,t-1}$	63552	25014	4769	3360	21830	27655	386910
OOI^w	63552	3167	1855	38	1855	4088	23142
$\bar{y}_{o \to p,t}$	63552	2788	629	825	2345	3159	38562
$\bar{y}_{o \to p,t-1}$	63552	2621	651	328	2152	3029	8545
OOI^y	63552	357	211	4	213	455	2886
			WC				
Year	72301	2013	4.5	2006	2009	2017	2020
N	72301	345	1302	2	27	221	46298
w (Wage)	72301	35084	10758	8610	27897	39284	290000
$\bar{y}$ (Income)	72301	31967	9778	7567	25154	36557	116698
Value Added	72301	791721	1562678	0.24	87970	840384	29802147
HHI	72301	0.34	0.29	0.00092	0.11	0.5	1
$\pi_{o \to p}$	72301	0.12	0.071	0.0017	0.071	0.15	0.8
$w_{o \to p,t}$	72301	35084	10758	8610	27897	39284	290000
$w_{o \to p,t-1}$	72301	33948	10715	5327	26962	38033	492088
OOI^w	72301	4184	2777	51	2257	5435	53247
$\bar{y}_{o \to p,t}$	72301	3836	1173	908	3018	4387	14004
$\bar{y}_{o \to p,t-1}$	72301	3676	1173	526	2859	4233	13514
OOI^y	72301	481	322	5.6	256	625	5185

Summary statistics for the data used in our analysis, separated by blue collar (BC) and white collar (WC) occupation subsets. HHI indicates employer concentration on hires (0,1]. $\pi_{o \to p}$ the probability of changing occupation between years. $w_{o \to p,t}$ are the wages received by workers who change occupations, $w_{o \to p,t-1}$ are the wages of occupation changes in the year prior to changing occupations. OOI^w is the Outside Occupation Index on wages. Similarly, $\bar{y}_{o \to p,t}$, $\bar{y}_{o \to p,t-1}$, and OOI^y are identical to the components of occupation movers on wages, but use average monthly incomes from the primary employer.

4 Results

All regressions are weighted log regressions, following SST, using natural logarithms. Weights are calculated from the number of observed individuals in each defined labor market.

4.1 Regressing wage levels

The result of the first set of regressions are presented in table 3 which report estimates for models III.1, III.2, III.3, and III.4, using time-varying fixed regional effects as $\mathrm{FA15}_t$ ($\hat{\alpha}_{r,t}$) and time-varying occupational effects as $\mathrm{Ssyk4}_t$ ($\hat{\alpha}_{o,t}$).

Table 3: Regression outputs from models III.1, III.2,III.3, and III.4

	logWage			
	(1)	(2)	(3)	(4)
logHHI	-0.0152***	-0.0146***	-0.0108***	0.1437***
	(0.0014)	(0.0014)	(0.0009)	(0.0094)
logOOI^w		0.0208***	0.0250***	0.0164***
		(0.0016)	(0.0009)	(0.0013)
logVA			0.0089***	1.26e-5
			(0.0003)	(0.0008)
logHHI × logVA				-0.0080***
				(0.0005)
Observations	135,968	135,968	135,968	135,968
R^2	0.96633	0.96673	0.89062	0.96883
Within R^2	0.01859	0.03029	0.03930	0.09152
FA15t fixed effects	✓	✓	✓	✓
Ssyk4t fixed effects	✓	✓	✓	✓

The results record weighted log-log regressions with time-varying occupational and regional fixed effects on 355-occupations (when available) in 60 labor market regions between 2006 and 2020, except 2014 which is omitted due to poor data quality. logHHI is log concentration on hires, logOOI^w is the outside occupation index (on wages), logVA is the mean value addedd in the observed temporal labor market, and logHHI × logVA is an interaction variable. The weights for the regional labor markets sum to 51,520,795 observations..

Regression (1) indicates a significant, but negative concentration slope of -0.0152. Adjusting for the outside option index in (2), yields a slight reduction in concentration. Including the ”naïve” value added-component in (3) further reduces the elasticity by half of a percentage unite, but reduces the model’s predictive power significantly, reducing the R^2 estimate by approximately 7 percent.

When interacting concentration with value added in (4), we further reduce the negative impact on concentration. To interpret the interaction result, we hold value added constant by taking VA’s partial derivative on the wage, yielding:

$$\frac{\partial w}{\partial VA} = \gamma_3 + \gamma_4 log(HHI)$$

Thus, holding value added constant in (4) yields: (insignificant) 0 - 0.0080 log(HHI), implying that a one percent change of concentration has a -0.8 percent negative effect on the wage bargain when controlling for firm size effects.

Concentration appears to have a relatively modest effect on wages in Sweden compared to other settings. In France, concentration elasticities to wages range from -0.0185 to -0.0230 for incumbents in Bassanini et al. (2023); in the US elasticities range from -0.01 to -0.015 according to SST, or -0.02 according to Azar et al. (2019) (who do not correct for market definition concerns).

The interaction effect also allow us to hold concentration constant to observe value added effects on the wage bargain. We note that workers, on average, receive a significant wage premium from increasing value added in their labor markets (.1437-0.008 logVA).

The Outside Occupation Index implies significant improvements to the wage bargain throughout, but is nearly halved in (4) when we apply the interacted concentration and value-added-component.

Although R^2 is high throughout, within-R^2 is dismally low in models (1), (2), and (3), implying that models (1)-(3) are primarily driven by occupational and regional fixed effects, however within-R^2 improves significantly in model (4).[22]

4.1.1 Regressing blue- and white-collar occupations separately

As indicated in figure 1, BC- and WC occupations differ significantly in terms of wage levels, wage dispersion, working hours (as percentages of full time employment), value added, and in outcomes from their outside occupation index. In terms of concentration (see figure 5) WC-groups outside the metropolitan regions seem to work in sectors with particularly high employer concentration. We are also aware that wage setting procedures in collective bargaining agreements can differ widely between these two groups, where WC workers often have more agency to bargain starting salaries and yearly wage increases individually compared to many BC occupations, that often use more objective wage setting criteria bargained for the entire collective by local or central union representatives.

In table 4 we repeat regressions (2), (3) and (4) from table 3, above, for the BC and WC occupation groups of occupations.

Comparing (1, BC) and (4, WC) the wage-concentration effect is close to zero but insignificant for the BC group. For the WC-group, concentration effects on wages are more in line with international estimates. The OOI-variable also shows relatively better payoffs from having good outside occupation options for the WC group compared to the BC group.

Adding the ”naïve” value added component to the regression (2 and 5) increases the negative concentration effects on the wage bargain somewhat for both groups. The interacted concentration-value-added effect in (3) and (6), yields a positive wage effect from concentration for the BC group (0.0135 - 0.0017 = 0.0121 from a one percent change) but a sizable negative effect for the WC group (-0.0135).

[22] We note that previous studies with high R^2-values do not report separate within-R^2 estimates, providing no opportunity to compare these outcomes to other settings.

Table 4: Regression outputs from models III.2,III.3, and III.4 by BC and WC occupation groups

	logWage					
	(1, BC)	(2, BC)	(3, BC)	(4, WC)	(5, WC)	(6, WC)
logHHI	-0.0011	-0.0033***	0.0300***	-0.0223***	-0.0230***	0.1258***
	(0.0009)	(0.0009)	(0.0068)	(0.0016)	(0.0015)	(0.0106)
logOOI^w	0.0084***	0.0072***	0.0068***	0.0176***	0.0165***	0.0154***
	(0.0010)	(0.0009)	(0.0009)	(0.0017)	(0.0017)	(0.0017)
logVA		0.0166***	0.0135***		0.0044***	-0.0063***
		(0.0007)	(0.0009)		(0.0006)	(0.0009)
logHHI × logVA			-0.0017***			-0.0073***
			(0.0003)			(0.0005)
Observations	63,594	63,594	63,594	72,374	72,374	72,374
R^2	0.93015	0.93494	0.93506	0.96484	0.96495	0.96600
Within R^2	0.00368	0.07199	0.07370	0.04213	0.04519	0.07370
FA15t fixed effects	✓	✓	✓	✓	✓	✓
Ssyk4t fixed effects	✓	✓	✓	✓	✓	✓

Holding concentration constant, we also note that the the rewards of working in labor markets with higher value added are larger for WC occupations (0.1258-0.0073 logVA) compared to the BC group (0.03 - 0.0017 logVA). The effect is likely regional, with WC workers in metropolitan regions driving the results (see figure 2). This lends some support to the notion of ”superstar” firms affecting the wage bargain for selected, higher skilled groups (e.g. Autor et al. 2017).

The separate BC and WC regressions retain high R^2 values, with somewhat higher within-R^2 (3, BC and 6, WC) compared to the estimates for the full population.

4.1.2 Adjusted wages from gross income and hours

Table 5 reports the outcomes from using gross average monthly incomes (y) and hours (h) from model III.5, comparing the results to model III.4's wage models for the entire population, as well as the BC and WC occupation groups.

Adding log hours implies small improvement to all R^2 and within-R^2 estimates, suggesting that incomes and hours provides slightly better fitting models than wages. The logHHI × logVA-interactions yield nearly identical effects from concentration when holding value added constant. The same is true for the Outside Occupation Index, save for BC occupations (4), where the previously small but positive coefficient is now small but negative. Although a surprising result, it likely reflects the greater discrepancy between full-time wages and actual incomes for BC workers.

As the log hours variable is necessary to estimate wages from income data, we can subtract the log hours coefficient by log Income to attain a wage estimate ($\hat{w}$).

For BC (4), incomes are approximately 11 percent lower than wages (0.1098), while the model suggests that WC incomes (6) are approximately 27 (0.2723) lower than wages. This WC esimate is some cause for concern.

To get an idea whether these income-to-wage adjustments are accurate we need to consider two factors. First, the percentage difference between mean incomes and

Table 5: Regression outputs from models III.4 and III.5 for all, and the BC and WC occupation groups

	logWage (1, All)	logKUInc (2, All)	logWage (3, BC)	logKUInc (4, BC)	logWage (5, WC)	logKUInc (6, WC)
logHHI × logVA	-0.0080***	-0.0083***	-0.0017***	-0.0009**	-0.0073***	-0.0077***
	(0.0005)	(0.0006)	(0.0003)	(0.0004)	(0.0005)	(0.0006)
logHHI	0.1437***	0.1518***	0.0300***	0.0177**	0.1258***	0.1337***
	(0.0094)	(0.0114)	(0.0068)	(0.0069)	(0.0106)	(0.0115)
logVA	1.26e-5	-0.0007	0.0135***	0.0126***	-0.0063***	-0.0060***
	(0.0008)	(0.0008)	(0.0009)	(0.0008)	(0.0009)	(0.0009)
$\log OOI^y$		0.0130***		-0.0053***		0.0179***
		(0.0017)		(0.0011)		(0.0019)
$\log OOI^y$	0.0164***		0.0068***		0.0154***	
	(0.0013)		(0.0009)		(0.0017)	
log h		0.1418***		0.1098***		0.2723***
		(0.0090)		(0.0075)		(0.0350)
Observations	135,968	135,968	63,594	63,594	72,374	72,374
R^2	0.96883	0.97207	0.93506	0.96023	0.96600	0.96521
Within R^2	0.09152	0.09115	0.07370	0.07585	0.07370	0.08959
FA15t fixed effects	✓	✓	✓	✓	✓	✓
Ssyk4t fixed effects	✓	✓	✓	✓	✓	✓

the sample wage ($\frac{y}{w_s}$); and second, the difference between estimated wages (derived from dividing incomes by hours) and the sample wage ($\frac{y \div h_s}{w_s}$). The first percentage difference tell us about the relative size of incomes and wages, which should be picked up by the hours coefficient in our regression outputs. The second estimate tell us how well the wage and hours data predict incomes, which is the identity.

As wages and hours are sampled in a single month, we can test how well these sample statistics compare to observed monthly incomes. We do this in figure 11, which shows monthly income data reported by the primary employer for all months in 2019 and 2020. The intuition behind this diagram is to show the percentage difference between incomes and the sample wage ($\frac{y}{w_s}$) – which, on average should be negative as the average worker does not work 100%; and the relative difference in percent between the estimated wages and the sample wage ($\frac{y \div h_s}{w_s}$). If the latter is greater than 0, estimated wages (from incomes and hours) *overestimate* the wage. The intuition is that if the difference between the estimated wage divided by the sample wage ($\frac{y \div h_s}{w_s}$) and incomes divided by the sample wage ($\frac{y}{w_s}$) should yield a percentage difference that is explained by the hours component from our regression estimates.

For the BC group, incomes are between 6.4 (2019) and 6.7 (2020) percent lower than wages, while the estimated wage levels (from income and hours) are between 4.1 and 4.4 percent higher than wages. For the WC group, the corresponding differences are between 3.8 and 4.0 for incomes/wages, and 3.2 and 1.7 higher for estimated wages and wages. The difference between the (over) estimated wages and the wage levels recorded in the wage structure statistics indicates that BC estimated wages from incomes need to be adjusted by between 11.0 and 11.1 percent to attain the wage levels from the wage structure statistics, while WC estimated wages are between 7.0

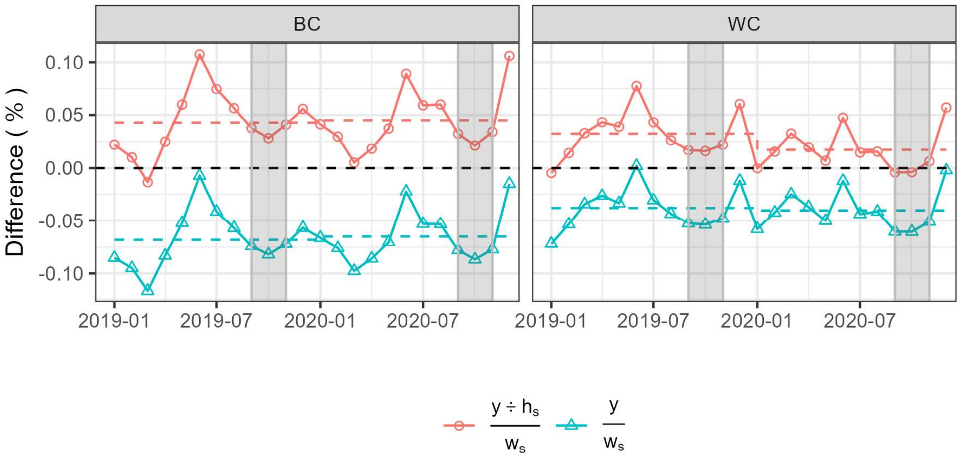

Figure 11: Monthly difference between incomes divided by hours (estimated wages) and wages; and incomes and wages (in percent)

and 5.8 percent higher than wages.

Assuming that the monthly incomes in 2019 and 2020 are representative for the entire observed period, the log hours variable in the BC model provides a good approximation of wages from incomes (11 percent vs 10.98 percent from the log hours coefficient), while the WC group (6 to 7 percent vs 27.23 percent from the log hours coefficient) does not, as the log hours variable for the WC-group appears to underestimate wages from incomes by about 20 percent.

A plausible issue here relates to having a β-distributed[23] hours-variable for the WC-group, where there is little observable variation around high working time percentages close to 100 %. Investigating these troubling (but interesting) results are outside the scope of this paper, so we conclude by suggesting that both the wage and income/hours models work well for the BC occupation group, but only the wage model for the WC occupation group[24].

4.1.3 *Identification in the job ladder*

To address issues relating to identification, where the wages of occupation movers, found in the independent OOI-variable, are endogenous to the dependent variable wages in occupation o via the collective bargaining channel, we use the IV-instruments for OOI on wage and income levels, with regression outputs in table 6 for both the wage (1-3) and income-hours models (4-6).

[23] I.e. some value with defined limits. In our case 0% to 100%, where the variable cannot go above or below these values.

[24] This follows the conclusion in appendix section "How well do wages and hours predict incomes?", where we show

Table 6: Regression outputs from models III.4 and III.5 using the OOI instrumental variable for all, and the BC and WC groups

	logWage			logKUInc		
	(1, All)	(2, BC)	(3, WC)	(4, All)	(5, BC)	(6, WC)
logHHI × logVA	-0.0081***	-0.0018***	-0.0073***	-0.0084***	-0.0009**	-0.0077***
	(0.0005)	(0.0003)	(0.0005)	(0.0006)	(0.0004)	(0.0006)
logHHI	0.1446***	0.0316***	0.1262***	0.1520***	0.0180***	0.1338***
	(0.0094)	(0.0068)	(0.0106)	(0.0114)	(0.0069)	(0.0115)
logVA	6.04e-5	0.0134***	-0.0062***	-0.0007	0.0126***	-0.0060***
	(0.0008)	(0.0008)	(0.0009)	(0.0008)	(0.0008)	(0.0009)
$\text{logOOI}^{w,IV}$	0.0121***	0.0021**	0.0116***			
	(0.0013)	(0.0009)	(0.0017)			
$\text{logOOI}^{y,IV}$				0.0118***	-0.0062***	0.0163***
				(0.0017)	(0.0011)	(0.0018)
logtjomf				0.1418***	0.1098***	0.2726***
				(0.0090)	(0.0075)	(0.0350)
Observations	135,968	63,594	72,374	135,968	63,594	72,374
R^2	0.96881	0.93499	0.96599	0.97207	0.96023	0.96521
Within Adjusted R^2	0.09100	0.07258	0.07330	0.09108	0.07574	0.08947
Wald (joint nullity), p-value	4.37e-162	1.24e-170	7.15e-103	1.49e-197	2.47e-221	9.6e-111
F-test (projected)	3,271.2	1,186.9	1,369.8	2,619.8	994.42	1,361.3
FA15t fixed effects	✓	✓	✓	✓	✓	✓
Ssyk4t fixed effects	✓	✓	✓	✓	✓	✓

The p-values from the Wald joint-nullity and projected F-test both indicate that our instrument is strong. However, the instrument has little effect on the coefficients in the regression outputs, yielding extremely small differences in both wage and income-hours model outputs (tenths of a percentage unit). Most notably, the negative outside occupation option index is still negative for the BC group in the income-hours regression. This suggests that the fixed effects estimates may already be addressing this identification concern.

4.2 Assessing collective bargaining from fixed effects estimates

To verify that the outside option variables-estimates capture *only* individual bargaining power effects, and that the fixed effects estimates capture *only* collective bargaining power effects, we use the BC and WC models from table 6. Looking to the R^2 regression outputs, we note that the wage and income models explain between 96 to 97 percent of the variation in our data, while the within-R^2 estimates show that the outside option variables only explain between 7.3 and 9.1 percent of all variation. Thus, all remaining variation in the model is picked up by the yearly regional ($\alpha_{r,t}$) and yearly occupational ($\alpha_{o,t}$) fixed effects.

If our intuition that collective bargaining has a high impact on Swedish wages is correct, the $\alpha_{r,t}$ and $\alpha_{o,t}$ estimates should summarize to wages that represent collective bargaining levels. Our best data source of collective bargaining impacts on wages are found in the Mark, which captures average yearly wage increases set in centrally coordinated national collective bargaining rounds. Thus, extracting and summarizing yearly fixed effects estimates, and calculating their yearly change rates yields:

$$\Delta\hat{\alpha}_t = \frac{\sum(\hat{\alpha}_{o,t} + \hat{\alpha}_{r,t})}{\sum(\hat{\alpha}_{o,t-1} + \hat{\alpha}_{r,t-1})} - 1$$

Figure 12 compares the yearly changes in summarized fixed effects estimates ($\Delta\hat{\alpha}_t$, red line) with the wage increases set by the Mark (black, dashed line).

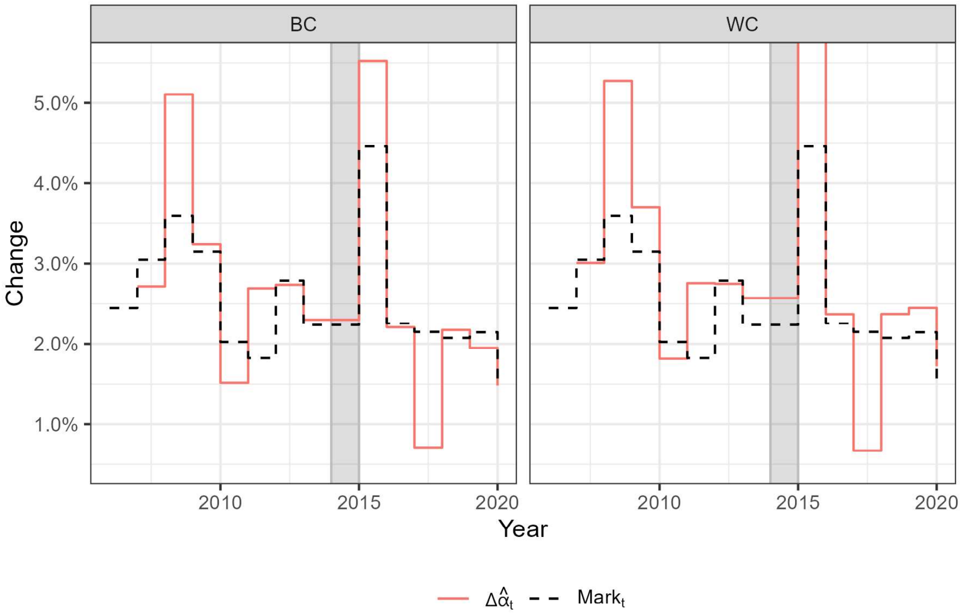

Figure 12: The figure compares the yearly changes to the estimated fixed effects summations ($\Delta\hat{\alpha}$) with the $Mark_t$ for observed years. 2015 includes 2014 and 2015's wage increases.

Although we see some notable upward and downward deviation throughout the years, the mean difference between yearly changes in the fixed effects estimates and the Mark are less than 0.3 percentage units, on average. For the WC group, the difference is 0.36 percentage units whereas for the BC group, the difference is only 0.08 percentage units, on average. As the fixed effects appear to pick up on dynamics from the collective bargaining regime, the model's ability to separate individual from collective bargaining power effects on wages are validated.

Using the models without instrumental variables (3, BC and 6, WC models from table 4), which does not explicitly introduce the Mark to the model, yields almost identical $\Delta\hat{\alpha}_t$ estimation (see figure 17 in the appendix). Thus, the low impact from the instrumental variable might be explained by the collective bargaining effects that the IV tries to correct for are already captured by the fixed effects.

The fixed effects summations perform worst at predicting the mark in 2008 (positive deviation) 2015 (positive) and 2017 (negative). The absolute deviation in percent are greatest for the WC-group of occupations. As noted in figure 8, there is a large spike in within-employer occupational change in in 2008 (see 2008 in 4). If this sudden change of within-employer occupation changes were the result of a purely administrative reclassification for a large group of workers, which did not imply changes to the wage bargain, we expect the outside option model to preform worse during this year. This would then spill over on the estimated fixed effects summations that compensate from an artificially low wage impact from the outside options.

5 Discussion and conclusion

The main motivation behind this study is to explore the role individual bargaining power in a labor market where collective bargaining has a high impact on the wage bargain, empirically assessing Manning's (2003) claim that monopsony effects will be mitigated in labor markets with a powerful union presence. We are also motivated to adapt the SST-framework on administrative data, and to verify if their framework is applicable in a labor market where collective bargaining has a large impact on wages.

We use a Nash Bargaining framework to study how the wage bargain in aggregate labor markets are affected by the value of outside options and collective bargaining outcomes. To gain insights into monopsony power, we study the role of concentration to the wage bargain, addressing concerns relating to wage-concentration measures being biased by faulty market definitions. We adapt SST's flexible definition of labor markets by constructing an Outside Occupation Index, which also reflects the quality of a worker's job ladder. We also include a model estimating wages from incomes and hours, which improves the explanatory power of our models, but produces mixed results in its ability to adjust incomes and hours to reported wage levels. Further, we address the role of collective bargaining power by considering its effect on our yearly regional and occupational fixed effects.

Overall, employer concentration has had a modest negative effect on the wage bargain, but the effects are much smaller compared to most settings. When considering occupational heterogeneity in the form of BC and WC occupations, which also reflects heterogeneous practices in wage setting regimes found in collective bargaining agreements, we find that concentration has a positive impact to the wage bargain for the BC workers, even when correcting for value added. For WC workers, we see similar negative effects from employer concentration as in international studies. WC effects are largely driven by higher concentration outside of Sweden's metropolitan regions. Further, we cannot exclude that WC worker wages are more sensitive to outside options, as WC groups (nearly) always have significant agency to bargain wages individually in their collective agreements. Concentration shows similar results when considering both effects to wage levels and yearly wage increases. In sum, collective bargaining appears to counterbalance monopsony power to a higher extent than other settings, but employers appear to be able to exercise wage setting power in many settings, as we observe significant variation on wage outcomes resulting from variation in the strength of worker outside options.

The positive concentration effects for the BC-group raises interesting questions about what a greater abundance of employers implies to groups with arguably lower levels of individual bargaining power. One possible explanation relates to firms with many employees being more likely to sign collective bargaining agreements than firms with fewer employees, regardless of their ”economic” size. A greater abundance of smaller firms might have a negative impact on the wage bargain if their abundance is related to subcontracting or domestic outsourcing, implying worse opportunities for rent-sharing (e.g. Weil 2014, Goldschmidt and Schmieder 2017 or Card et al. 2018). Another plausible explanation lies in a faulty or inflexible definition of firm size. Here, we assess firm size by its generated surpluses from production (as value added). Large public employers do not (formally) produce value added. Thus, the positive wage effect from concentration could also be an effect of public employers paying better wages, on average, but where firm size remains unaccounted for in the model due to a faulty definition of firm size. Another plausible explanation may be rooted in employment protection[25]. Arai and Heyman (2001), for example, find that in Swedish labor markets, observed worker flows are approximately 10 times higher for workers with temporary employment contracts compared to fixed contract workers. If a temporary contract give employers wage setting power, then observed job changes to temporary jobs may yield smaller wage increases than flows to permanent positions. And if permanent positions are rarer and concentrated to larger firms in specific labor markets, this could also imply that lower employer concentration has a negative impact on the wage bargain.

We verify the usefulness and ability of the SST framework to separate individual from collective bargaining power in a Nordic setting by studying how the fixed (wage) effects estimates compare between years, finding their yearly changes closely follow the collectively bargained wage increases set by the Mark.

We conclude that the Nash bargaining framework adapted from SST performs well at predicting the impact on wages from workers taking outside options in Swedish labor markets. Looking to the R^2 and within-R^2 outputs in table 6, we conclude that 88 percent (0.96023-.070574) of BC wage increases derive from the Mark, and 87 percent (0.96521-0.08947) of the WC wage bargain, indicating that the outside option variables (HHI and OOI) closely resemble the traditional notion of wage drift (Flanagan et al. 1976). Table 3 (showing central wage increases and wage drift), indicates that wage drift accounts for approximately 15 percent of all wage increases

[25] See for recent overview Belloc and D’Antoni 2020

between 2006 and 2020[26], leaving collective bargaining to explain the remaining 85 percent of wage increases. This is very close to the predictions yielded by our model, indicating that wage drift can be used as a simpler means to assess labor demand's impact on the wage bargain.

Our paper thus provides a good starting point to further explore the role of collective bargaining in supply-side models of labor markets, while validating the explanatory power of outside-option models in labor markets with strong labor market institutions. Future research may find that outside option models are useful when exploring other factors that may have an impact on the wage bargain, such as unemployment insurance, or the role of internal labor markets.

References

Arai, M. and F. Heyman (2001). "Permanent and temporary labour: job and worker flows in Sweden 1989:1998". In.

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. V. Reenen (2017). "Concentrating on the Fall of the Labor Share". In: *American Economic Review* 107.5, pp. 180–185.

Azar, J., E. Huet-Vaughn, I. Marinescu, B. Taska, and T. Von Wachter (2019). "Minimum wage employment effects and labor market concentration". In.

Azar, J., I. Marinescu, and M. Steinbaum (2022). "Labor market concentration". In: *Journal of Human Resources* 57.S, S167–S199.

Bäckman, O. and A. Nilsson (2016). "Long-term consequences of being not in employment, education or training as a young adult. Stability and change in three Swedish birth cohorts". In: *European Societies* 18.2, pp. 136–157.

Bassanini, A., C. Batut, and E. Caroli (2023). "Labor market concentration and wages: Incumbents versus new hires". In: *Labour Economics* 81, p. 102338.

Belloc, F. and M. D'Antoni (2020). "The elusive effect of employment protection on labor turnover". In: *Structural Change and Economic Dynamics* 54, pp. 11–25.

Bender, G. (2024). "Organizing the labor market: Power, ideas, and institutions in wage formation, digital automation, and migrations". PhD thesis. Stockholm School of Economics.

Bhuller, M., K. O. Moene, M. Mogstad, and O. L. Vestad (2022). "Facts and fantasies about wage setting and collective bargaining". In: *Journal of Economic Perspectives* 36.4, pp. 29–52.

Boal, W. M. and M. R. Ransom (1997). "Monopsony in the labor market". In: *Journal of economic literature* 35.1, pp. 86–112.

[26] I.e. the remaining wage increase levels above the collectively bargained wage increases set in the Mark.

Burdett, K. and D. T. Mortensen (1998). “Wage differentials, employer size, and unemployment”. In: *International Economic Review*, pp. 257–273.
Bustos, E. (2023). “The effect of centrally bargained wages on firm growth”. In: *IFN Working Paper*.
Card, D. (2022). “Who set your wage?” In: *American Economic Review* 112.4, pp. 1075–1090.
Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). “Firms and labor market inequality: Evidence and some theory”. In: *Journal of Labor Economics* 36.S1, S13–S70.
Card, D., J. Rothstein, and M. Yi (2024). “Location, location, location”. In: *American Economic Journal: Applied Economics*, forthcoming.
Carlén, T. and T. Hållö (2020). “Varför har lönerna utveck-lats svagt och hur bör det samhällsekonomiska löne-utrymmet beräknas?” In: *Ekonomisk Debatt*, pp. 45–50.
Flanagan, R. J., C. C. Holt, and B. Bosworth (1976). “Wage interdependence in unionized labor markets”. In: *Brookings Papers on Economic Activity* 1976.3, pp. 635–681.
Freeman, R. and J. Medoff (1984). *What Do Unions Do*. Basic Books.
Goldschmidt, D. and J. F. Schmieder (2017). “The rise of domestic outsourcing and the evolution of the German wage structure”. In: *The Quarterly Journal of Economics* 132.3, pp. 1165–1217.
Hirschman, A. O. (1970). *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*. Vol. 25. Harvard university press.
Jäger, S., C. Roth, N. Roussille, and B. Schoefer (2024). “Worker Beliefs About Outside Options”. In: *The Quarterly Journal of Economics*, forthcoming.
Kjellberg, A. (2023). *Kollektivavtalens täckningsgrad och vita fläckar*. Förlagets slutgiltiga version. OCLC: 1411259368. Stockholm: Arena Idé.
Manning, A. (2003). *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press.
— (2011). “Imperfect competition in the labor market”. In: *Handbook of labor economics*. Vol. 4. Elsevier, pp. 973–1041.
— (2021). “Monopsony in labor markets: A review”. In: *ILR Review* 74.1, pp. 3–26.
Medlingsinstitutet (2022). *Medlingsinstitutets årsrapport 2021 - Avtalsrörelsen och lönebildningen*. The Swedish National Mediation Office.
— (2023a). *”Så ser det ut med lägstalöner”*. URL: https://www.mi.se/nyheter/2023/sa-ser-det-ut-med-lagstaloner/.
— (2023b). *Kollektivavtal – Vilka tecknar avtalen och hur är löneavtalen konstruerade?* The Swedish National Mediation Office.
Miller, N., S. Berry, F. Scott Morton, J. Baker, T. Bresnahan, M. Gaynor, R. Gilbert, G. Hay, G. Jin, B. Kobayashi, et al. (2022). “On the misuse of regressions of price on the HHI in merger review”. In: *Journal of Antitrust Enforcement* 10.2, pp. 248–259.

Névo, G. (2024). “Large Firms, High Concentration, High Wages”. In: *Job Market Paper*.
Osterman, P. (2024). “Reconsidering Occupational Internal Labor Markets: Incidence and Consequences”. In: *ILR Review*, Forthcoming.
Robinson, J. (1933). *The economics of imperfect competition*. Springer.
Schubert, G., A. Stansbury, and B. Taska (2024). “Employer concentration and outside options”. In: *Available at SSRN 3599454*.
Thoresson, A. (2024). “Employer concentration and wages for specialized workers”. In: *American Economic Journal: Applied Economics* 16.1, pp. 447–479.
Tillväxtanalys (2016). *Sverige – ett attraktivt gruvland i världen?: En internationell jämförelse. Rapport 2016:06.* Tillväxtanalys.
Walton, R. E. and R. B. McKersie (1991). *A behavioral theory of labor negotiations: An analysis of a social interaction system*. Cornell University Press.
Weil, D. (2014). “The fissured workplace: Why work became so bad for so many and what can be done to improve it”. In: *The fissured workplace*. Harvard University Press.
Yakymovych, Y. (2022). “Consequences of job loss for routine workers”. In: *IFAU Working Paper Series 2022* 15.
Zhao, J. and T. Matti (2018). “Evaluation of Classical Monopsony in Nursing Labor Markets: A Natural Experiment”. In: *Master thesis*.

Appendix Paper III

How well do wages and hours predict incomes?

Including gross average monthly labor incomes and working hours in our analysis of wages require the quotient of incomes to hours to approximate the mean full-time equivalent wage relatively well. And thus by identity, the product of wages and hours should approximate monthly average incomes well.

We assess this by comparing by studying how well the sampled wage and hour variables predict observed monthly incomes between 2019 and 2020. First, observing the difference between monthly incomes and (sample) wages, and second, the difference between monthly incomes and (sample) wages times working hours

Observed monthly income vs. estimated wages and hours

Full time-equivalent wages (w_s) and hours (h_s) are both individually matched sample variables that aim to represent an average month in a working year[27]. The wage variable is probably the most used variable to represent unit wages in Swedish labor market studies. Incomes, on the other hand, is an administrative data variable, recording the yearly gross labor incomes paid to the Swedish Tax Authority by the primary employer, where we get the monthly average income by dividing yearly gross labor incomes by 12.

Thus, we need to test if the product of wages and hours are good at estimating average monthly incomes.

For workers who work full time (hours= 100 %), the income that they receive in the month when the Wage Structure Statistics are collected should reflect the reported monthly full-time equivalent wage. For workers who work less than 100 %, Statistics Sweden will employ estimations to make part-time incomes reflect a full time equivalent wage.

To check if wages and hours are good at approximating observed monthly incomes, we exploit the fact that since 2019 employers are obligated to pay employer contributions and report gross labor incomes to the Swedish Tax Authority on a monthly basis. This implies that we have access to monthly reported gross labor incomes for all months between 2019 and 2020.

To test how well the yearly wages and hours-variables predict observed monthly gross incomes, we compare the difference between observed monthly incomes and wages ($y - w_s$, in red), as well as monthly incomes and the product of wages and hours ($y - w_s \times h_s\%$, in blue) in figure 13 for four coarse SSYK occupation codes,

[27]Collected through the Wage Structure Statistics at the end of each year

where SSYK 1 through 4 are (mostly) WC occupations, and 5 through 9 are (mostly) BC occupations. Thus, 0 indicates no difference, whereas values greater than zero indicates that the variable overestimates incomes, whereas below 0 underestimates incomes. The shaded area indicates the months in the year when wages and hours are collected by the Wage Structure Statistics surveys. The percentage term in brackets indicate the share of workers that have jobs in the occupation groups in the total population. The occupation groups are presented in order of occupational of qualification, where SSYK 1 is management work and SSYK 9 is work that does not require specific vocational training (see figure 13's description).

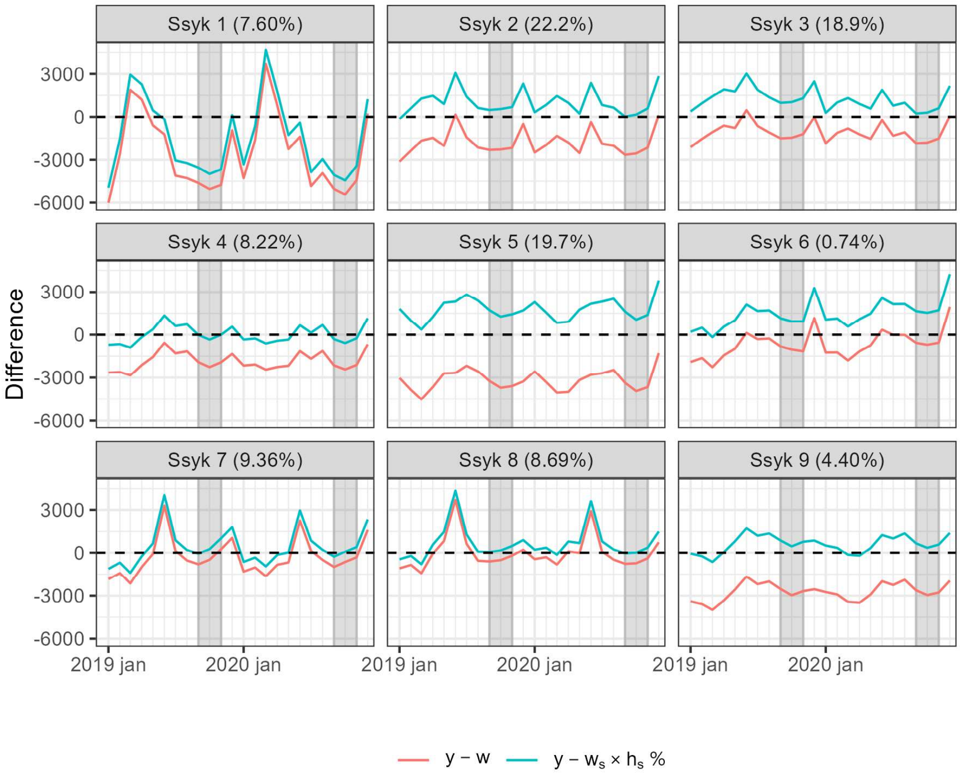

Figure 13: The figure shows the mean difference in SEK between observed monthly incomes y and the mean full time equivalent monthly wage w_s (in red); and the difference between observed monthly income y and the estimated income from wages and hours $w_s \times h_s\%$ from the wage structure survey (in blue). The shaded areas indicate periods when the wage structure statistics are gathered. SSYK 1 is management work; SSYK 2 is work that require theoretical specialist competences; SSYK 3 is work that requires shorter post secondary education or similar knowledge; SSYK 4 is office or customer service work; SSYK 5 is service, care, or sales work; SSYK 6 is work in agriculture, gardening, forestry, or fishing; SSYK 7 is craftsman's work in construction or manufacturing; SSYK 8 is process- or machine operator, and transportation work, etc; Ssyk 9 is work that does not require specific vocational training.

The difference between incomes and wages $(y - w)$ tell us how much lower incomes are than wages. Having incomes much below full time equivalent wages indicate that the occupation group has a large share of part time workers. We note for occupation groups 7 and 8 that the difference between income and wages are smallest, and largest for groups 5 and 9.

The difference between observed monthly incomes and the monthly income-approximation we get from multiplying wages by hours ($y - w_s \times h_s\%$) performs best for occupation groups 4, 7, 8, and 9, but worst for Ssyk 5 – service, care, or sales work – which is also the group with the highest share of part-time workers. The product of wages and hours appears to consistently overestimate actual observed incomes.

We note that the greater the distance between incomes and wages ($y - w_s$) the more do wages and hours overestimate received incomes.

This error likely comes from difficulties assigning correct values to part-time workers, as the actual number of working hours may vary over a year, with the mean not centering on the observed assigned percentage. Working part-time does not necessarily imply working exactly a percentage of a full time hours. Rather, part time work may imply a high variation in hours. For example, someone who works 50 % of full time at a comparatively low wage, getting extra hours may be valuable. If this is true, workers on lower working time percentages are more likely to work more than 50 %.

Regressing incomes, wages, and hours

If wages and hours from the Wage structure statistics can estimate average monthly incomes perfectly, regressing monthly observed incomes on wages $\times$ hours should yield a constant (intercept) at 0 and a slope coefficient of 1, while yielding an R^2-value at 1. The regression thus tell us how well the sample based wage (w_s) and working hours (h_s) variables predict the identity of monthly income.

A weighted regression of incomes on wages times hours, for blue- and white-collar occupations, are presented in Table 5 (2 and 4), as well as a regression of incomes on wages (1 and 3) for comparison.

Table 7: Predicting observed monthly incomes from wages and wages $\times$ hours

	y			
	(1, BC)	(2, BC)	(3, WC)	(4, WC)
Constant	-5,849.5***	6,560.5***	3,432.0***	6,912.9***
	(1,330.8)	(779.2)	(801.7)	(710.7)
w_s	1.117***		0.8476***	
	(0.0452)		(0.0178)	
$w_s \times h_s$		0.7662***		0.8132***
		(0.0289)		(0.0166)
Observations	128	128	153	153
R^2	0.82921	0.84770	0.93741	0.94079
Adjusted R^2	0.82785	0.84649	0.93700	0.94039

The weighted regressions test how well observed monthly incomes (y) are predicted by wages (w_s) and wages times hours ($w_s \times h_s$).

Outputs (2) and (4) indicates relatively low R^2 values for the BC group (0.84) but much higher for the WC group (0.94). The constant for the BC group is 6,560.5 and 6,912.9 for BC. As the WC group have higher incomes overall, the relatively high intercept should be considered less problematic than the BC intercept. The slope coefficients for both groups are well below 1 for both groups, but close for the WC occupation group.

Figure 14 illustrates the relative proximity of wages (w_s), and wages and hours ($w_s \times h_s\%$) (both y-axis), compared to observed monthly incomes (x-axis), summarized by their respective occupation. For wages and hours, dots on the dotted indicates a perfect approximation of incomes from wages and hours, whereas dots above the dotted line predict that wages times hours overestimate observed monthly incomes.

Thus, wages and hours tend to overestimate incomes for both groups, but more so for the BC groups (5 to 9) compared to the WC groups (1 to 4). We note that the estimations perform relatively well in the (shaded) sampling months, but that the average predictions are much higher than the observed incomes.

As a result, incomes divided by hours should also *overestimate* the wage. Using hours alone to approximate wages should thus yield estimated full-time equivalent wage levels above observed wages from Wage Structure Statistics sample, lest fixed effects for example, reduce the overestimation.

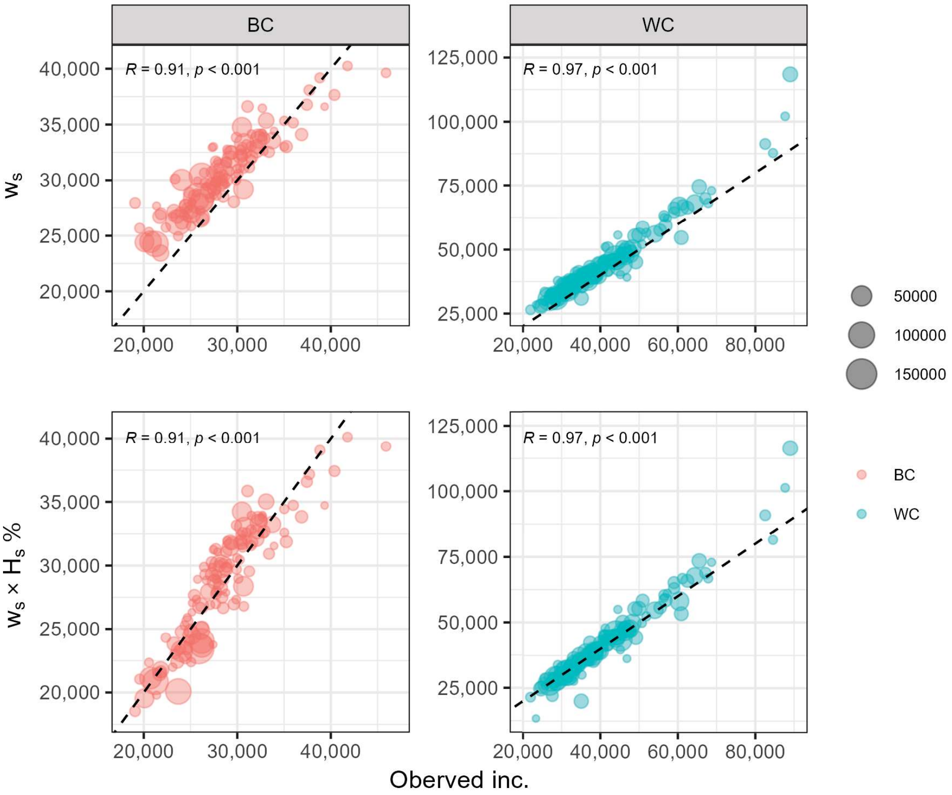

Figure 14: the relative proximity of wages, and wages and hours to incomes

Other diagrams referenced in Paper III

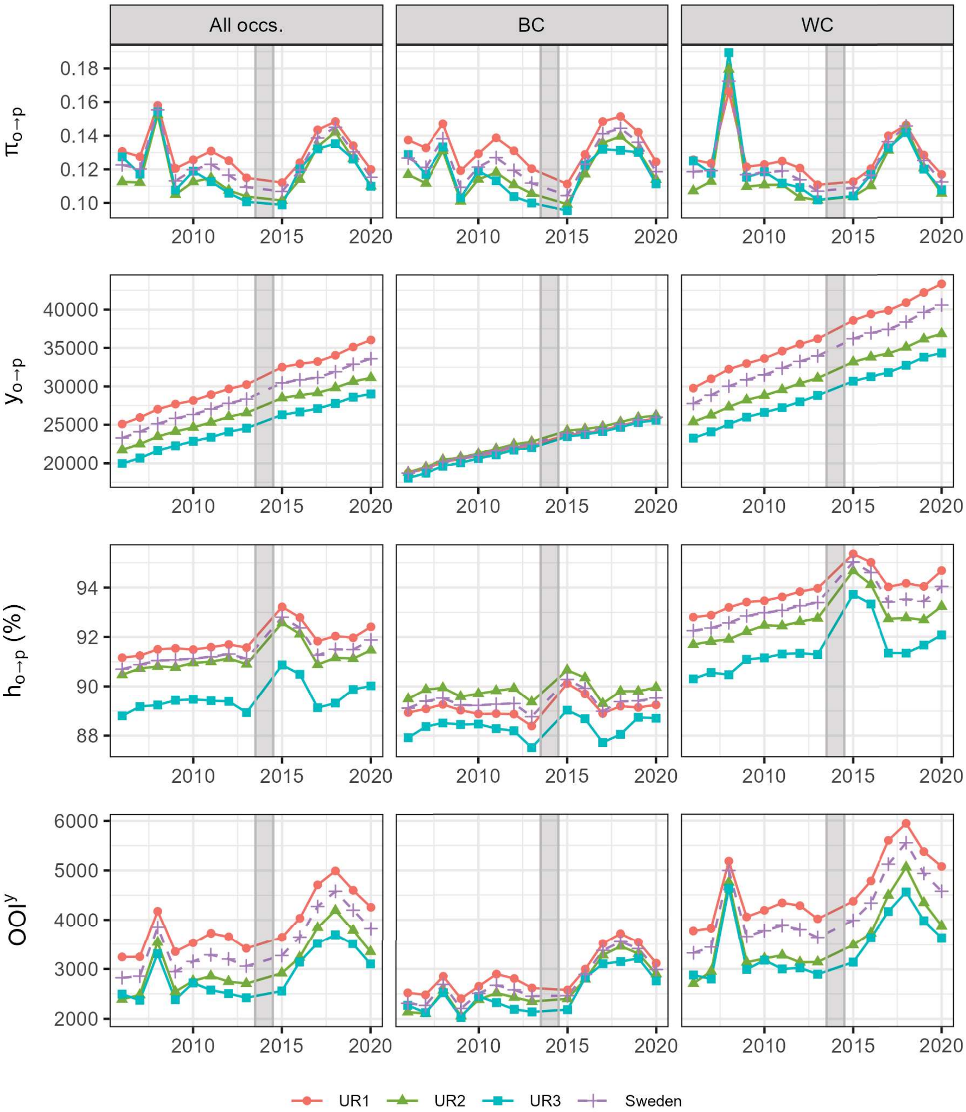

Figure 15: The figure illustrates the Outside Occupation Index for income's (OOI^y) components in a similar fashion as figure 8, replacing wages with the incomes of occupational movers ($y_{o\to p}$) and also includes hours of occupational movers ($h_{o\to p}$). 2014 is censored due to the abnormally high $\pi_{o\to p}$-value, resulting from occupational measurment errors in 2014 (see figure 16).

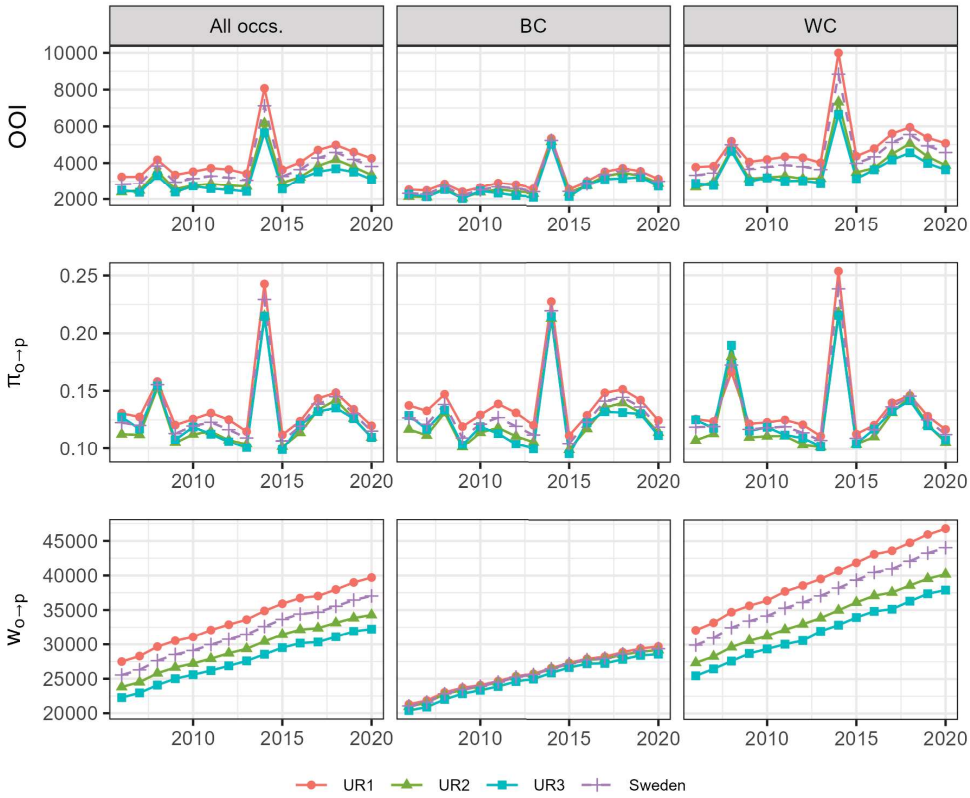

Figure 16: Uncensored version of figure 8, including the year 2014. We note that $\pi_{o \to p}$ approaches a probability of 0.25 in 2014, which is almost the double compared to other years.

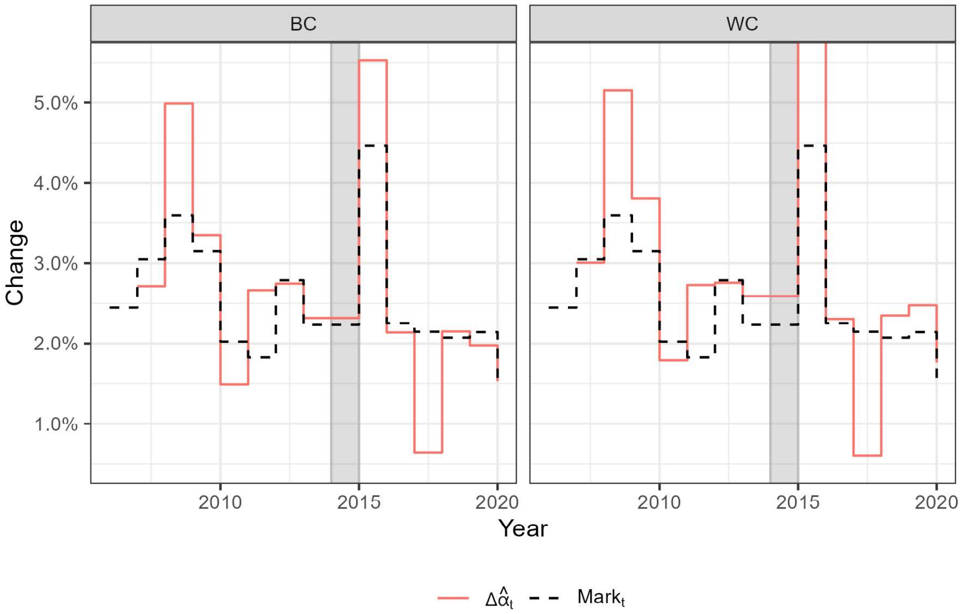

Figure 17: The diagram shows the yearly changes to fixed effects estimate summations for regression model III.4, which does not include instrumental variables. Thus, the model's ability to approximate collective bargaining effects is not an effect of including the Mark in the instrumental variables.