
Giuseppe Moscarini*
Yale University and NBER
June 2023

In this ambitious chapter, Berger, Herkenhoff, Kostol and Mongey make a significant contribution to a recent lively debate on the origins of rising income inequality. A growing literature, that they exhaustively discuss in the Introduction to the chapter and that includes important work by some of the same authors (Berger, Herkenhoff, and Mongey (2022)), argues that profit rates and price mark-ups, estimated to be rising in the US over the last four decades, are better understood as wage mark-downs, with market power manifesting itself more in labor than in product markets. The question is then: What gives firms labor market power? The answer is essential to understand what (if anything) policymakers can do about it.

A key step in this kind of analysis is how to delimit the confines of a labor market, to determine measures of employer concentration and labor market power. The statistical definition of a labor market proposed and implemented in this chapter is novel and quite interesting. It fully leverages the exceptional scope of the Norwegian data, especially the availability of occupation information, rare in matched employer-employee datasets that originate from administrative sources. Both the scope of a labor market, based on job-the-job transitions that remain within its confines, and the ordering of employers in a labor market, based on net poaching ranks, exploit compelling revealed-preference arguments.

Building on this definition of a labor market, the authors offer new empirical evidence on the cross-market correlation between employment concentration and: labor market flows, wage levels, and within-market wage dispersion. Noteworthy is the addition of labor market

*PO Box 208268, New Haven CT 06520-8268, giuseppe.moscarini@yale.edu
flows, central to the macroeconomic analysis of unemployment and wage dispersion in the last half century, and still largely absent from the monopsony literature. The authors then enrich a Sequential Auctions-cum bargaining search model, a la Cahuc, Postel-Vinay and Robin (2006), with vacancy posting and job creation, job amenities, and firm granularity. This flexible framework allows them to set up a horse race between neoclassical and frictional sources of wage compression and inequality. The estimated model addresses a classic question, what is the main source of residualized wage dispersion, this time across labor markets defined by space and occupations, rather than across individual workers.

The authors suggest that, of the three possible sources of market power, search frictions are the least tractable by policy, while amenities and firm granularity can be presumably addressed (resp.) by fiscal and anti-trust policy. But a natural mechanism to combat search and matching frictions are internal labor markets, certainly more developed within larger firms (e.g., Papageorgiou (2018)). Paradoxically, the emergence in a local labor market of a large, dominating company may cause wage compression, but also improve allocation, via information-sharing and the resulting career progression opportunities within the firm. If we are concerned about the level of wages, rather than just the labor share, the effects of employment concentration are a priori ambiguous.

More broadly, the general theme of my comment is that employer concentration is difficult to interpret in a world of search frictions. I will focus on a specific, possibly the most novel, aspect of the chapter, namely, the relationship between worker turnover rates and employment concentration. The empirical evidence on the association between employment concentration, job-to-job transitions and unemployment (Figures 2 and 3 and Table 3 in the paper) is new to the literature, and essential to identify frictional vs. neoclassical sources of labor market power in the structural estimation of the model. Worker flows between employment states and employers can originate either from search frictions or from preference shocks. The authors do not consider the latter: job-specific preferences/amenities do not change during an employment relationship. Therefore, only frictions remain to explain dynamics. Accordingly, I formally investigate the relationship between labor market flows and the HHI concentration index in a class of frictional models. Specifically, I study the role of several latent and unobserved fundamentals of a labor market that can cause the heterogeneity in worker flows and in employer concentration observed across labor markets in Norway. These fundamental, market-specific factors include the efficiency of job search, job stability, and TFP.

The basic insight I build upon is the role of outside offers as a counterweight to firms’ market power. While searching for microfoundations to tatonnement to competitive equilibrium, Diamond (1971) came to a paradoxical conclusion. In a wage-posting game, where workers have no exogenous bargaining power, but can only leverage their outside options, ar-
bitrarily small frictions in finding jobs from unemployment, barring any outside offer during employment, generate maximum monopsony power and wage compression. Put more simply, even negligible frictions before matching are sufficient to enforce a unique equilibrium where workers are at the complete mercy of the firms. Burdett and Judd (1983) and Burdett and Mortensen (1998) are the seminal frameworks that bring outside offers, (resp.) simultaneous or sequential, to break this Diamond Paradox. The Postel-Vinay and Robin (2002) sequential auctions framework, on which this chapter ultimately builds, further developed this line of thinking by introducing ex post competition for employed workers. Importantly, on-the-job search is more effective against firms’ bargaining power the more stable jobs are. If separations to unemployment are frequent, and the job ladder is slippery, workers can rarely climb it and are hired mostly from unemployment, at their reservation wages.

My main conclusions are as follows. A generic job ladder model generates an inverse relationship across markets between the pace of Employer-to-Employer (EE) transitions and employment concentration, as we observe in the data, under different parameter configurations that lead to opposite implications for market power. In this sense, the empirical evidence on this inverse relationship cannot reject the model, but does not fully discipline it either. Intuitively, a labor market can be highly concentrated, and job-to-job transitions suppressed, when frictions are small, jobs are stable, outside offers are very frequent, the market is nearly perfectly competitive, so workers concentrate in the most productive firm, which nearly drives the other firms out of the marker. There is no need for EE reallocation, as most workers are already on the top rung of the job ladder. Modest job instability thus reduces concentration and raises EE. Conversely, close to the Diamond (1971) paradoxical limit of full monopsony, EE transitions are low, because infeasible. If job ladder rungs capture firm-specific productivity, the top rungs earn the highest profits and post most vacancies, thus absorb most hires (from unemployment) and concentrate employment. In this case, modest on-the-job search opportunities raise EE and might reduce concentration.

In order to draw more definite conclusions, we need to consider additional local labor market indicators. The patterns of unemployment rates and of transition rates from employment to unemployment (EU), observed across labor markets of different concentration in Norway, support the view that high concentration is actually, and maybe counterintuitively, a symptom of small frictions and intense competition. The observed patterns of transition rates from unemployment to employment (UE) are uninformative, in the logic of job ladder models where only outside offers to employed workers matter. The observed patterns of wage levels and within-market wage dispersion support the view that concentration means monopsony, but can be also rationalized by additional forces that the model abstracts from, most notably implicit insurance contracts, also consistent with the other patterns.
1 A Job Ladder with Rank-Preserving Equilibrium

Much of the notation follows the chapter’s. There exist a measure of firms indexed by a fixed trait (including permanent amenities) \( z \in [\underline{z}, \overline{z}] \sim P(z) \). Each firm of index \( z \) posts \( v(z) \) vacancies; the total mass of vacancies is then \( v = \int_{\underline{z}}^{\overline{z}} v(z) dP(z) \).

Employed workers lose jobs and become unemployed with probability \( \delta \in (0,1) \). The \( u \) unemployed workers search full time; still employed workers \((1-\delta)(1-u)\) can search with probability \( \xi \in [0,1] \) each period. Aggregate job market tightness is \( \theta = \frac{v}{u+\xi(1-\delta)(1-u)} \). Searching worker contacts open vacancies with probability \( \lambda = \Lambda(\theta) \in (0,1) \), where \( \Lambda(\cdot) \) derives from a homothetic random meeting function.

The “sampling weight” of firms of index \( z \) is \( \gamma(z) = \frac{v(z)P'(z)}{v} = \frac{\Gamma'(z)}{\Gamma(z)} \), so \( \Gamma(z) \) is the probability that a worker, conditional on contacting a vacancy, draws a match with index less than \( z \).

I assume that the allocation reflected in the data is the outcome of a decentralized equilibrium (or of an optimal social plan) that is “Rank Preserving”: all workers have preferences for jobs monotonic in \( z \), thus agree on a vertical ranking, and climb the same “Job Ladder”. This is a strong restriction, that unlocks some useful properties. In this allocation, more desirable firms typically post more vacancies, \( v'(z) > 0 \).

Let \( L_t(z) \) denote the measure of workers employed at index less than \( z \) at time \( t = 0, 1, 2.. \). Rank-Preserving dynamics on the job ladder imply that workers, when given the opportunity, move from lower to higher \( z \) index jobs:

\[
L_t(z) = L_{t-1}(z)(1-\delta)[1-\xi\Lambda(\theta_t)(1-\Gamma(z))] + u_{t-1}\Lambda(\theta_t)\Gamma(z)
\]

where the unemployment rate is \( u_t = 1 - L_t(\overline{z}) \). The two terms multiplying past employment represent retention from (resp.) separations to unemployment and quits to preferred employers \( z' > z \). The last term is the inflow from unemployment. Note that there is no inflow from other firms, because this is a cumulated distribution. Quits between firms of index either both lower or both higher than \( z \) do not change the mass; only quits that cross \( z \) (in the second square brackets) impact this mass, and indeed reduce it.

In steady state, with \( \lambda = \Lambda(\theta) \) for notational convenience:

\[
L(z) = \frac{\lambda\Gamma(z)u}{\delta + (1-\delta)\xi\lambda[1-\Gamma(z)]}
\]

where

\[
u = 1 - L(\overline{z}) = \frac{\delta}{\delta + \lambda}\]
2 Worker Flows

The probability of observing a worker making an Employer-to-Employer (EE) transition is the product of the meeting probability and of the “Acceptance” probability (AC) that the new draw has a higher index:

$$EE = \left(1 - \delta \right) \xi \lambda \int_{\bar{z}}^{\bar{z}} \frac{[1 - \Gamma(z)]}{1 - u} \frac{dL(z)}{\text{Prob(offer)}} \cdot \int_{\bar{z}}^{\bar{z}} \frac{[1 - \Gamma(z)]}{1 - u} \frac{dL(z)}{\text{Prob(accept|offer)} \text{ “AC”}}$$

We now state and prove that in a Rank-Preserving equilibrium the specific sampling distribution $\Gamma$ is irrelevant for EE transitions: only quantiles matter, because workers always move up the ladder. This argument builds on Moscarini and Postel-Vinay (2023).

Let

$$\Delta = \frac{\delta}{(1 - \delta) \xi \lambda} \in [0, \infty)$$

This is a “frictional index” of search on-the-job, similar to the “$\kappa_1$” index in Burdett and Mortensen (1998)’s continuous time model. $\Delta$ is higher the lower the on-the-job search meeting probability $\xi \lambda$ and the higher the job losing probability $\delta$, both contributing to misallocation on low $z$ rungs. The case $\xi = 0$ yields the Diamond paradox of full monopsony. $\Delta$ is equal to zero if nobody ever loses a job and falls off the job ladder ($\delta = 0$). In this case, assuming workers do receive outside offers ($\xi \lambda > 0$), they all end up employed at the top rung $\bar{z}$, the competitive equilibrium outcome.

**Proposition 1** In steady state, the Acceptance probability only depends on the model parameters through the frictional index $\Delta$, and equals

$$AC = \Delta (1 + \Delta) \ln \left( \frac{1 + \Delta}{\Delta} \right) - \Delta$$

In particular, $AC$ does not depend on the distribution of job ranks $P$ nor (directly) on the sampling distribution $\Gamma$, thus on vacancy postings per rank $v(z)$. $AC$ is increasing and concave in frictions $\Delta$, with

$$\lim_{\Delta \to 0} AC = 0 \quad \text{and} \quad \lim_{\Delta \to \infty} AC = \frac{1}{2}$$
\textbf{Proof}

\[ AC = \int_{\hat{z}}^{z} [1 - \Gamma(z)] \frac{dL(z)}{1 - u} \]

\[ = \int_{\hat{z}}^{z} \frac{L(z)}{1 - u} d\Gamma(z) \]

\[ = \int_{\hat{z}}^{z} \frac{u\lambda\Gamma(z)}{\delta + (1 - \delta)\xi\lambda[1 - \Gamma(z)]} \frac{d\Gamma(z)}{1 - u} \]

\[ = \int_{0}^{1} \frac{u}{\delta + (1 - \delta)\xi\lambda(1 - \Upsilon)} \frac{\Upsilon}{1 - u} d\Upsilon \]

\[ = \int_{0}^{1} \frac{\delta}{\delta + (1 - \delta)\xi\lambda(1 - \Upsilon)} \Upsilon d\Upsilon \]

\[ = \Delta \int_{0}^{1} \frac{\Upsilon}{\Delta + 1 - \Upsilon} d\Upsilon \]

\[ = \Delta \int_{0}^{1} \left( \frac{\Delta + 1}{\Delta + 1 - \Upsilon} - 1 \right) d\Upsilon \]

\[ = \Delta(1 + \Delta) \ln \frac{1 + \Delta}{\Delta} - \Delta \]

where in the second line we integrate by parts, in the third we replace for \( L(z) \) from its expression, in the fourth we change variable to \( \Upsilon = \Gamma(z) \), in the fifth we replace for \( u \) from its expression, in the sixth for \( \Delta \), and the rest is standard. The other properties can be verified directly.

Now we return to the EE probability. The meeting probability of employed workers with competing firms decreases with the separation probability \( \delta \), because fewer workers can access outside offers if they are first separated into unemployment; it increases with the general contact probability per unit of search efficiency, \( \lambda \), and with the (relative) efficiency of on-the-job search \( \xi \). These parameters have exactly the opposite impact on the frictional index \( \Delta \) and thus, from the Proposition, on the acceptance probability AC. It is easy to verify that

\[ \lim_{\delta \downarrow 0} EE = \xi \lambda \cdot 0 = 0 \]

\[ \lim_{\delta \uparrow 1} EE = 0 \cdot \frac{1}{2} = 0 \]

\[ \lim_{\xi, \lambda \downarrow 0} EE = 0 \cdot \frac{1}{2} = 0 \]

\[ \lim_{\xi, \lambda \uparrow 1} EE = (1 - \delta) \cdot \frac{\delta}{1 - \delta} = \delta \]

In general, EE is non-monotonic (hump-shaped) in \( \delta \), a fact that poses an identification
challenge. When workers fall off the job ladder frequently, they have fewer chances to change jobs, so the fewer surviving employed are poorly distributed on the job ladder, thus accept more of the outside offers they receive, while firms have more labor market power. Therefore, a low observed EE probability is consistent with both very stable and very unstable jobs. The ‘competition’ parameter $\xi\lambda$ too has two opposing effects: more frequent outside offers allow more job switches but also less misallocation and acceptance. But, in this case, the direct effect dominates.

3 Employment Concentration

The same expression for the employment distribution $L(z)$ allows to understand the impact of frictions on employment concentration. First, normalizing $L(z)$ by total employment $1 - u$ yields the employment c.d.f.

$$L(z) = \frac{L(z)}{1 - u} = \frac{\delta \Gamma(z)}{\delta + (1 - \delta) \xi \lambda [1 - \Gamma(z)]} = \frac{\Delta \Gamma(z)}{\Delta + 1 - \Gamma(z)}$$

with $L(\bar{z}) = 1$ independently of frictions $\Delta$.

Let $\mathbb{I}$ denote the indicator function. In the frictionless limit

$$\lim_{\Delta \to 0} L(z) = \mathbb{I}\{z = \bar{z}\}$$

all employment concentrates on the top rung. If there is only one firm at the top, the HHI index, the sum of squared firms’ employment shares, equals one, its maximum feasible value. Conversely, as frictions grow unbounded:

$$\lim_{\Delta \to \infty} L(z) = \Gamma(z)$$

no reallocation takes place up the job ladder, workers remain where they are hired from unemployment, until separation, and the employment distribution converges to sampling distribution $\Gamma$. This is the Diamond (1971) paradoxical equilibrium, with full monopsony power. As long as $\Gamma$ is non degenerate, as it is usually the case, the HHI index is unambiguously lower than 1. If extreme frictions destroy any returns from vacancy postings, all firms post a vanishing measure of job openings, and the sampling distribution reduces to the population distribution $P(z)$. The same conclusion applies, the HHI index remains strictly below 1.

Therefore, concentration is eventually decreasing in frictions and market power. Concentration is low when firms have maximum market power and extract all rents from helpless job applicants, while it is highest when competition pushes all workers to the top firms, but
also wages up to productivity.

To understand what happens away from the limits, we can rewrite

\[ L(z) = \int_{\underline{z}}^{\bar{z}} s(z) dP(z) \]

where \( s(z) \) is the share of employment at each firm of type \( z \). Therefore,

\[ s(z) = \frac{L'(z)}{P'(z)} = \frac{\Delta(1 + \Delta)}{|\Delta + 1 - \Gamma(z)|^2} \frac{\Gamma'(z)}{P'(z)} = \frac{\Delta(1 + \Delta)}{|\Delta + 1 - \Gamma(z)|^2} \frac{v(z)}{v} \]

where we used (1). In a Rank-Preserving equilibrium, this share is typically increasing in rank \( z \). A sufficient condition is that firms higher on the ladder post more vacancies.

The HHI index is then

\[ \text{HHI} = \int_{\underline{z}}^{\bar{z}} s^2(z) dP(z) \]

How does it depend on our measure of labor market frictions \( \Delta \)? Taking a derivative, since the firm population distribution \( P(z) \) does not depend on the extent of labor market frictions:

\[ \frac{\partial \text{HHI}}{\partial \Delta} = \int_{\underline{z}}^{\bar{z}} 2s(z) \frac{\partial s(z)}{\partial \Delta} P'(z) dz = \int_{\underline{z}}^{\bar{z}} 2s(z) \frac{\partial (s(z)P'(z))}{\partial \Delta} dz = \int_{\underline{z}}^{\bar{z}} 2s(z) \frac{\partial L'(z)}{\partial \Delta} dz \]

\( L \) is a proper c.d.f., so \( L' \) is a proper density. When \( s(z) \) is increasing, the last integral is negative if an increase in frictions causes a downward shift in \( L \) in a First-Order Stochastic Dominance sense, that is, if for every \( z < \bar{z} \):

\[ 0 < \frac{\partial L(z)}{\partial \Delta} = \frac{\Gamma(z)[1 - \Gamma(z)] + \Delta(1 + \Delta)\frac{\partial \Gamma(z)}{\partial \Delta}}{|\Delta + 1 - \Gamma(z)|^2} \]

where we used again (1) and rearranged terms. The first term in the numerator is unambiguously positive: more severe frictions amplify misallocation and reduce concentration. The second term can be negative. It is zero when firms do not post vacancies, but stand ready to hire anybody who contact them, as in Burdett and Mortensen (1998), so that \( \Gamma = P \) is independent of frictions and employment concentration is globally, not just eventually, decreasing in frictions.
4 The Relationship between EE Flows and Employment Concentration

We are now in a position to understand how the EE transition probability and the HHI index of labor market concentration comove across labor markets characterized by different unobservable fundamentals, summarized in the frictional index $\Delta$. These fundamentals include not only job (in)stability $\delta$ and on-the-job search efficacy $\xi$, but also any other, such as local TFP, that can determine vacancy creation, thus job market tightness $\theta$ and average meeting probability $\lambda = \Lambda(\theta)$. We distinguish between two cases, and resulting possible interpretations of the empirical evidence:

1. **Job stability.** Labor markets differ by $\delta$. Figures 2C and 3C in the chapter shows an inverse relationship between the separation probability from employment into unemployment (EU) and the HHI index of concentration. If $\delta$ is small, as is average EU in the data, then the EE probability declines and the HHI index rises as $\delta$ falls. That is, in Figures 2 and 3 in the chapter, as we move from left to right, less fluid and more concentrated labor markets are actually those that feature more stable jobs, a less slippery job ladder, and more opportunities for workers to climb it and to extract rents from firms. Thus, more fluid and less concentrated labor markets are actually, and maybe counterintuively, more monopsonistic.

2. **On-the-job search frictions:** Labor markets differ by $\xi \lambda$. For example, in some markets, on-the-job search is made very difficult by long work hours, or non-competes are pervasive, or low market-specific TFP suppresses job creation. As $\xi \lambda$ decreases, and frictions $\Delta$ grow, markets become more monopsonistic, converging to the Diamond (1971) limit. As we showed, the EE probability declines. If concentration also declines, as in the example provided earlier, the resulting positive comovement between EE and concentration contradicts the evidence in Figures 1A and 2A. To fit this evidence, concentration must be non-monotonic in frictions. Since concentration is maximal near the frictionless benchmark $\Delta = 0$, where most employment joins the very top firms, it cannot increase in $\Delta$ near there. Therefore, Norwegian labor markets must be at the opposite, highly frictional end of the spectrum, and feature very low, and heterogeneous, on-the-job contact rates $\xi \lambda$. In this scenario, more concentrated labor markets are indeed more monopsonistic.

These observations highlight the identification challenge faced by the authors. Job ladder models are a natural environment in which higher concentration does not imply nor signal market power.
5 Additional Labor Market Indicators

In order to further discriminate between the two scenarios, consider other local labor market outcomes, which may provide evidence of high or low frictions and monopoly power.

The job-finding rate from unemployment UE rates appears to be decreasing in concentration (Figures 2B and 3B). While suggestive of higher monopsony power in more concentrated markets, this evidence is inconclusive. The logic of job ladder models with wage posting and on-the-job search indicates that UE flows are irrelevant to worker’s bargaining power, which originates entirely from outside offers that they receive only once employed.

The unemployment rate (Figures 2D and 3D) is also inconclusive, as its relationship with concentration differs across markets and for a given market over time. The cross-market negative correlation supports the first interpretation, that markets differ by job stability, and concentrated and less fluid markets are in fact more competitive. This is because unemployment declines despite falling job-finding UE rates from unemployment, so job-losing EU rates are the dominant force.

If job separations are endogenous, their heterogeneity across labor markets (Figures 2C and 3C) is naturally explained by local TFP: more productive labor markets feature more lenient retention standards and, as argued by Bilal (2023) for regions in France, lower EU separation probability. More productive labor markets are also more likely to generate, through firm entry, more vacancy openings and thus a higher meeting probability for job searchers. Whether separations to unemployment are exogenous or endogenous, the EU evidence supports then first interpretation.

What about wages? Those appear to provide direct evidence of mark-downs and wage compression in more concentrated labor markets (Figures 2E and 3E). There are, however, alternative considerations that can reconcile lower wages with lower unemployment. One is an insurance component in employment contracts. Labor markets that offer more job stability, for example because they are more productive and resilient to idiosyncratic productivity shocks, also offer more stable and compressed wages, as also indicated by the evidence on within-market wage inequality (Figures 2F and 3F). Workers accept lower wages to enjoy more job stability and be spared unemployment. Any empirical test of insurance in wage contracts must rely on dynamics, specifically on pass-through of firm revenues shocks to wages. A second possibility, inspired by the chapter itself, is that the job ladder rung is dominated by amenities, and not by productivity; in this case, high poaching-index firms could be low-paying, but extremely pleasant to work for.

The empirical evidence on wages calls for one final comment. Wages are not allocative in the model. If one takes a stand on wages, as the authors do, then the model predicts not only wage inequality across and within local labor markets, but also individual wage career paths. Specifically, like any job ladder model, it generates a distribution of wage changes.
upon job switches, scarring effect of layoffs, etc. These dynamic properties are the litmus test of search models, which were built to explain the cross-section of earnings, and can help to identify also the correct model of labor market power, thus to advance this exciting research agenda.

References


