It ain't where you're from, it's where you're at: hiring origins, firm heterogeneity, and wages.

Sabrina Di Addario, Patrick Kline, Raffaele Saggio, Mikkel Sølvsten[*]

June 30, 2020

Abstract

We develop a theoretically grounded extension of the two-way fixed effects model of Abowd et al. (1999) that allows firms to differ both in the wages they offer new hires and the wages required to poach their employees. Expected hiring wages are modeled as the sum of a worker fixed effect, a fixed effect for the "destination" firm hiring the worker, and a fixed effect for the "origin" firm, or labor market state, from which the worker was hired. This specification is shown to nest the reduced form for hiring wages delivered by semi-parametric formulations of the sequential auction model of Postel-Vinay and Robin (2002b) and its generalization in Bagger et al. (2014). Using Italian social security records that distinguish job quits from firings and layoffs, we demonstrate that our fixed effects model captures well differences in wage growth between workers involved in voluntary and involuntary job transitions. Bias correcting a variance decomposition of hiring wages, we find that origin effects explain only 0.7% of the variance of hiring wages among job movers, while destination effects explain more than 23% of the variance. Across firms, destination effects are more than 13 times as variable as origin effects. Interpreted through the lens of the sequential auction model, this finding requires workers to have implausibly strong bargaining strength. Studying a cohort of workers entering the Italian labor market in 2005, we find that differences in origin effects yield essentially no contribution to the evolution of the gender gap in hiring wages, while differences in destination effects explain the majority of the gap at the time of labor market entry. Our results suggest that where a worker is hired from is relatively inconsequential for his or her wages in comparison to where he or she is currently employed.

Keywords: Hiring Wages, Sequential Auction Models, Firm effects, Bargaining, Gender wage gap

In their pioneering study of the French wage structure Abowd et al. (1999, henceforth AKM) used high dimensional fixed effects methods to decompose wage inequality into components attributable to unobserved worker and firm heterogeneity. The AKM decomposition is motivated by the notion that there exists a stable wage hierarchy across firms. Hierarchies of this nature arise, for example, in the wage posting model of Burdett and Mortensen (1998), where each employer commits to a unique firm-wide wage premium. In practice, however, employers often possess information about workers' outside options, which they may use to craft personalized wage offers. Seminal work by Postel-Vinay and Robin (2002a,b) develops rigorously a notion of labor market competition where firms offer the lowest wage necessary to poach workers from an existing employer or unemployment. In these "sequential auction" models, hiring wages depend not only on the identity of the hiring firm but also the firm (or labor market state) from which a worker was hired. Price discrimination of this nature gives rise to a *dual* wage hierarchy: firms can be ranked both in terms of the wages required to poach their employees and the wage premia they offer new hires.

This paper studies empirically the relative importance of one's current employer ("where you're at") and the employer or labor market state from which a worker was hired ("where you're from") for the determination of wages. In contrast with previous estimates of the sequential auction framework (Postel-Vinay and Robin, 2002b; Dey and Flinn, 2005; Cahuc et al., 2006; Bagger et al., 2014) that jointly model worker mobility, hiring wages, and wage growth within a match, we confine ourselves to studying the evolution of *hiring* wages across jobs, leaving the adequacy of models for within-match wage growth and separation decisions to later research. We further depart from past work by studying this question using a generalization of the AKM fixed effects model that allows for a worker fixed effect, a fixed effect for the "destination" firm hiring the worker, and a separate fixed effect for the "origin" of the hire, which may include various forms of non-employment. Because the joint distribution of all three set of fixed effects is unrestricted, this "dual wage ladder" (DWL) specification accommodates very rich patterns of worker-firm sorting and allows firms that are high wage destinations to be high or low wage origins.

To clarify the link between the DWL model and the sequential auction framework, we show that our fixed effects specification nests the reduced form of hiring wages in the model of Postel-Vinay and Robin (2002b, henceforth PVR) when flow utility is logarithmic. Origin effects are increasing in productivity, as more productive firms can afford to counter more aggressive outside offers, while destination effects are decreasing in productivity because workers are willing to take wage cuts to join firms that offer greater prospects for future wage growth. Remarkably, the sum of a firm's origin and destination effects yields its productivity. Because workers in this model always accept offers from more productive firms, mobility is exogenous conditional on origin and destination fixed effects. The PVR model places sharp restrictions on the covariance structure of firms' origin and destination effects. The two sets of fixed effects must be negatively correlated because of their

opposite signed dependence on productivity, which is the only dimension along which firms are differentiated. Moreover, destination effects must be less variable across firms than origin effects, as the latter set of effects are more sensitive to firm productivity.

Extensions of the PVR model that allow workers to extract a positive share of the match surplus (Cahuc et al., 2006; Bagger et al., 2014) also turn out to admit a DWL representation where the sum of each firm's origin and destination effects corresponds to its productivity. When workers are able to extract all of the match surplus from hiring firms, the origin fixed effects disappear and an AKM style specification for hiring wages ensues. We show that the difference between the variances of firm destination and origin effects can be used to obtain a lower bound on worker bargaining power. When the variance of firm destination effects exceeds the variance of firm origin effects, the model additionally restricts the correlation between origin and destination firm effects to obey a positive lower bound that takes a simple analytic form. Finally, we derive some non-parametric shape restrictions on the relationship between a firm's origin and destination effects and its latent productivity level that can be scrutinized empirically with productivity proxies such as firm value added per worker.

Our empirical analysis relies on the INPS-INVIND panel of Italian social security earnings records. In addition to recording the annual earnings and months worked associated with each employer-employee match, these data contain information on the reason for each job separation. We use this information to distinguish voluntary job transitions involving worker quits from involuntary separations involving a firing, layoff, or contract non-renewal that are likely to substantially weaken a worker's outside options at the time of hiring. We find that workers involuntarily displaced from their first job experience less growth in hiring wages between their first two jobs than workers who quit their first job. Consistent with the log-additive DWL formulation of hiring wages, this displacement penalty appears to be invariant to the mean co-worker wage levels of those first two employers. We also find support for a key exclusion restriction suggested by the PVR/DWL framework: the identity of the firm from which a worker is involuntarily separated appears to have no effect on hiring wages. Evidently, it doesn't matter for hiring wage determination who fires you, just that you were fired.

Fitting the DWL model to a panel of workers with two or more jobs, we find an average wage penalty for being hired at one's first job of roughly 5% and a penalty for being involuntarily separated from one's previous job (i.e., hired from unemployment) of roughly 4%. To assess the overall contribution of origin and destination effects to hiring wage inequality, we conduct bias corrected variance decompositions using the methods developed in Kline et al. (2020). Adding origin fixed effects to a standard AKM specification explains only half of a percentage point of additional wage variance. Extending the traditional AKM variance decomposition, we find that person and destination effects respectively explain roughly 29% and 24% of the variance of hiring wages, while origin effects explain only 0.7% of the variance of hiring wages. We conclude that

where a worker was hired from exerts a quantitatively insignificant influence on his or her hiring wages in comparison to the identity of the hiring firm.

To tie our estimates more closely to the sequential auction framework, we investigate the co-variance structure of firms' origin and destination effects. The size weighted variance across firms of their destination effects is more than 13 times as large as that of their origin effects. Rationalizing this finding in the model of Bagger et al. (2014) requires that workers capture at least 88% of the rents in the employment relationship, far above the empirical estimates typically found in the literature (Card et al., 2018). Moreover, this level of bargaining strength would require a correlation between firm origin and destination effects of at least 0.84 to be rationalizable by the model, far above the empirical size weighted correlation we estimate of 0.25. Finally, we show that both origin and destination firm effects tend to increase with firm value added, and do so in a manner that violates the model's non-parametric shape restrictions.

Our key finding that firm destination effects are an order of magnitude more variable than firm origin effects echoes Postel-Vinay and Robin (2002a)'s early acknowledgment that "reality lies somewhere in between our complete information story and Burdett's and Mortensen's incomplete information assumption." One means of formalizing this middle ground comes from recent work that allows wage posting firms to coexist with firms that renegotiate wages as in the sequential auction framework (Flinn et al., 2017; Caldwell and Harmon, 2019). Consistent with the notion that firms differ in their wage setting strategies, we find substantial variability across sectors in the relative importance of firm origin and destination effects. For example, origin effects appear to play an especially inconsequential role in the restaurant sector but a fairly important role among law firms. Yet even among law firms, the empirical correlation between origin and destination firm effects is far too low to be rationalized by the model of Bagger et al. (2014), where firms are differentiated only by productivity. Our findings suggest it may be necessary to treat firms as differentiated along two or more dimensions, even within narrowly defined sectors, to match basic facts about the structure of hiring wages.

We conclude our analysis by investigating the extent to which Italian women face a dynamic disadvantage at the time of hiring attributable to the labor market state from which they were hired. Extending earlier results by Card et al. (2015), we find that both origin and destination firm effects differ by gender, with female hiring wages being less sensitive to measured firm productivity than male wages. We then study the evolution of the gender gap in hiring wages for Italians entering the labor market in 2005. The gender gap in hiring wages at labor market entry is almost entirely explained by gaps in destination effects. However, as workers age into the labor market, the hiring wage gap grows dramatically, while the gender gap in destination effects remains roughly constant. By contrast, the contribution of gender gaps in origin effects to gender hiring wage gaps is trivially small throughout the life cycle. For gender gaps, and for hiring wage inequality as a whole, the aphorism holds true "it ain't where you're from, it's where you're at."

4

1 The DWL model

Our analysis centers on the behavior of hiring wages, which we define as a worker's average daily wage in her first calendar year of employment with a given firm. For each worker $i \in \{1, \ldots, n\}$ in the sample, let $m \in \{1, \ldots, M_i\}$ index her job matches in chronological order. The dependent variable of interest is the log hiring wage of worker i in her m'th match, which we denote by y_{im}.

There are J firms in the labor market. We use $j(i, m) \in \{1, \ldots, J\}$ to denote identity of the firm employing worker i in her m'th match. The function $h(i, m) \in \{N, U, 1, \ldots, J\}$ gives the employer or labor market state from which worker i was hired into her m'th match. The state N corresponds to new labor market entrants, who have never been employed, while U corresponds to workers who were hired from unemployment. Empirically, we measure the state from which each worker was hired based upon whether she quit her previous job ($Q_{i,m-1} = 1$), was separated involuntarily ($Q_{i,m-1} = 0$), or has no labor market history ($m = 1$). Hence, we can write $h(i, m)$ as a function of m and $Q_{i,m-1}$ as follows:

$$
h(i, m) = \begin{cases}
j(i, m-1), & \text{if } Q_{i,m-1} = 1 \text{ and } m > 1, \\
U, & \text{if } Q_{i,m-1} = 0 \text{ and } m > 1, \\
N, & \text{if } m = 1.
\end{cases}
$$

Our dual wage ladder extension of the AKM model takes the form:

$$
y_{im} = \alpha_i + \underbrace{\psi_{j(i,m)}}_{\text{``destination effect''}} + \underbrace{\lambda_{h(i,m)}}_{\text{``origin effect''}} + X'_{im}\delta + \varepsilon_{im}, \tag{1}
$$

where X_{im} denotes time-varying covariates such as age and year, measured at the start of each job match. As in the traditional AKM model, the worker effect α_i captures a component of earnings ability that is transferable across firms, while the destination firm effect $\psi_{j(i,m)}$ gives the impact of the firm who is hiring worker i on her hiring wage – an effect she forfeits upon moving to a new job. What is new is the origin firm effect $\lambda_{h(i,m)}$, which gives the influence of the firm or state from which worker i was hired on her hiring wage. An important restriction of the DWL model is that the identity $j(i, m-1)$ of a worker's past employer does not affect her wage if she is hired from unemployment. We scrutinize this exclusion restriction later in our analysis. The coefficient vector δ governs the effects of age and calendar year at the time of hire, while the error term ε_{im} captures unobserved match specific factors determining hiring wages.

The closest analogue to (1) of which we are aware is the dynamic wage specification considered by Bonhomme et al. (2019), in which firms are assumed to fall into one of a finite number of classes that govern the wages of both new hires and incumbent workers. In their model, a worker's

wage may depend upon the firm class of both her current and past employers and her own latent type. However, they do not model the separate wage implications of voluntary and involuntary job transitions. By contrast, a key feature of our DWL specification is that past employers only influence the hiring wages that result from a voluntary transition, with involuntary transitions and labor market entry yielding distinct origin wage effects. Our fixed effects formulation additionally allows each firm to be its own two-dimensional hiring wage type.

1.1   Exogenous mobility

For the i'th worker, $y_i = \{y_{im}\}_{m=1}^{M_i}$ gives her wage history and $\mathcal{W}_i = \{j(i,m), h(i,m), X_{im}, \alpha_i\}_{m=1}^{M_i}$ collects her employment history, covariates, and the worker fixed effect α_i. We assume that $\{y_i, \mathcal{W}_i\}_{i=1}^n$ is an *i.i.d.* sample from a common unknown distribution. The wage types of the firms $\psi = (\psi_1, \ldots, \psi_J)'$, $\lambda = (\lambda_1, \ldots, \lambda_J)'$ and the values λ_N, λ_U, δ are treated as fixed parameters ("fixed effects") throughout.

Letting $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iM_i})'$ denote the history of hiring wage errors, our key identifying assumption is that

$$\mathbb{E}[\varepsilon_i \mid \mathcal{W}_i] = 0. \tag{2}$$

This is a strict exogeneity assumption, or as it is often referred to in this context, an "exogenous mobility" requirement. Equation (2) allows workers to base their mobility decision on any function of their own fixed effect α_i and the wage types (ψ, λ) of the firms in their economy. For instance, high skilled workers may be differentially likely to move from firms with lower λ values and towards those with higher ψ values. Equation (2) would be violated, however, if workers were to sort towards firms on the basis of an idiosyncratic match component of wages. We show below that a variety of sequential auction models imply the exogenous mobility requirement is satisfied for hiring wage specifications, despite the presence of a match effect in incumbent wages. We also find surprisingly little evidence that these wage errors predict separations over the first few years of a job match.

1.2   Implied dynamics

The wage dynamics implied by the DWL model are well illustrated by studying the hiring wage trajectories of a few career paths, distinguished by the sorts of transitions workers experience between their first three jobs. Workers following career path #1 are involuntarily separated from both of their first two jobs ($Q_{i1} = Q_{i2} = 0$). Workers following career path #2 quit both their first and second job ($Q_{i1} = Q_{i2} = 1$). Finally, workers following career path #3 are involuntarily separated from their first job but quit their second job ($Q_{i1} = 0$, $Q_{i2} = 1$).

The DWL model rationalizes the trajectory of hiring wages for these three career paths in

terms of a common set of origin and destination firm effects. First differencing equation (1) and suppressing for the moment the time varying covariates X_{im}, we can write the expected change in hiring wages between the second and third job for each career path as follows:

- Career Path #1 (two involuntary transitions)

$$\mathbb{E}[y_{i3} - y_{i2} \,|\, \mathcal{W}_i] = \psi_{j(i,3)} - \psi_{j(i,2)}$$

- Career Path #2 (two consecutive voluntary transitions)

$$\mathbb{E}[y_{i3} - y_{i2} \,|\, \mathcal{W}_i] = \psi_{j(i,3)} - \psi_{j(i,2)} + \lambda_{j(i,2)} - \lambda_{j(i,1)}$$

- Career Path #3 (an involuntary followed by a voluntary transition)

$$\mathbb{E}[y_{i3} - y_{i2} \,|\, \mathcal{W}_i] = \psi_{j(i,3)} - \psi_{j(i,2)} + \lambda_{j(i,2)} - \lambda_U$$

Inspecting these equations reveals that unemployment serves the role of a large firm from which workers can be poached. Because career path #1 involves being poached from the same firm twice, the origin effects cancel. Hence, it is as if the standard AKM model applies: expected wage growth depends entirely on the change in destination effects associated with the worker's second job transition.

The expected wage growth of a worker with career path #2 is substantially more complex, depending on the identities of each of her first three employers. Wage growth between such a worker's last two jobs will tend to be higher when her second job transition yields an improvement in destination effects or when her first job transition yielded an increase in origin effects.

The wage growth expected of a worker with career path #3 depends on the origin effect of her second job. However, it does not depend at all on the identity of her first employer $j(i,1)$, from which she was involuntarily separated. This exclusion restriction reflects a key non-parametric prediction of many search models: it doesn't matter which employer fires you. As described in the next section, this exclusion restriction can be motivated by the sequential auction framework, which envisions wages as being set via a bilateral negotiation between the worker and firm. We scrutinize this restriction empirically in a later section and find that it provides a good approximation to the wage dynamics found in our data.

2 Sequential auction models

In this section we develop a connection between the DWL specification and some popular variants of the sequential auction model. We start with the textbook PVR model of Postel-Vinay and Robin

([2002b](#)) and then progress to the extension of [Bagger et al. (2014)](#) that allows workers to extract a share of the match surplus from the poaching employer. Each model is shown to map into a variant of the DWL framework and to imply certain restrictions on the covariance structure of the origin and destination effects.

2.1 The PVR model

Workers are indexed by their productivity level ϵ and have flow utility over wages $\mathcal{U}(w)$. When unemployed, workers receive flow utility with wage equivalent value ϵb. Firms are indexed by their productivity p. Workers engage in search on and off the job, which leads them to encounter firm types drawn from a common distribution F with bounded support and survival function denoted $\bar{F}$. The marginal product of a worker of type ϵ when matched with a firm of type p is ϵp.

Though workers engage in random search, firms have full information regarding worker reservation wages. Upon meeting a worker, a firm will make a take it or leave it offer of a piece rate wage contract. Mobility is efficient: workers only accept offers from more productive firms. If a less productive firm contacts an employed worker, the incumbent firm offers the worker the smallest raise necessary to retain her. If a more productive firm contacts a worker, it offers her the lowest wage needed to compel her to leave the incumbent firm. PVR show that the "poaching wage" $\phi(\epsilon, p, q)$ required to compel a worker of type ϵ to quit a firm of type q for a firm of type $p > q$ solves:

$$\mathcal{U}(\phi(\epsilon, p, q)) = \mathcal{U}(\epsilon q) - \kappa \int_q^p \bar{F}(x) \mathcal{U}'(\epsilon x) \epsilon \, dx,$$

where the constant $\kappa \geq 0$ is an increasing function of the offer arrival rate and a decreasing function of the discount rate and an exogenous separation rate. In words, the flow utility of the poaching wage must equal the flow utility that would result if the incumbent firm were to pay the worker her full marginal product ϵq, minus a compensating differential for the future wage *growth* expected to result from moving to the more productive poaching firm (as it counters outside offers). When workers cannot search on the job then $\kappa = 0$ and this compensating differential disappears. The same equation turns out to govern the wage $\phi(\epsilon, p, b)$ required to hire a worker from unemployment, which is effectively a firm with productivity b – an idea that we generalize to other labor market states in the DWL specification.

We follow PVR in considering the case where $\mathcal{U}(x) = \ln x$, which yields a log-linear specification

for poaching wages:

$$\ln \phi\left(\epsilon, p, q\right) = \underbrace{\ln \epsilon}_{\text{person type}} + \underbrace{\ln q}_{\text{poached firm type}} \underbrace{- \kappa \int_{q}^{p} \frac{\bar{F}\left(x\right)}{x} \, dx}_{\text{type upgrade}}.$$

The log poaching wage is the sum of a person effect, a term summarizing the productivity of the poached firm, and a compensating differential for the upgrade in firm productivity. By the fundamental theorem of calculus $\kappa \int_{q}^{p} \bar{F}(x)/x \; dx = I(q) - I(p)$, where $I(z) = \kappa \int_{z}^{\infty} \bar{F}(x)/x \; dx$ gives the wage cut a worker is willing to take to move from a firm with productivity z to the most productive employer in the economy. This representation allows us to rewrite the poaching wage in the form of our earlier DWL specification:

$$\ln \phi\left(\epsilon, p, q\right) = \underbrace{\ln \epsilon}_{=\alpha(\epsilon)} + \underbrace{I(p)}_{=\psi(p)} + \underbrace{\ln q - I(q)}_{=\lambda(q)} \tag{3}$$

Here, poaching wages are the sum of a person effect $\alpha(\epsilon)$, a destination firm effect $\psi(p)$, and an origin firm effect $\lambda(q)$. For any given firm, the sum $\psi(p) + \lambda(p)$ of its origin and destination effects gives its log productivity $\ln p$. The assumption that both firm effects are driven by a common latent factor p is a strong restriction that the DWL framework relaxes by treating ψ and λ as potentially unrelated parameters.

The PVR model implies that ψ and λ are negatively correlated across *firms*: it takes high wages to poach from productive firms, while workers can be enticed to join productive firms at low wages. Formally, $\frac{d\psi(p)}{dp} < 0$ while $\frac{d\lambda(p)}{dp} > 0$ implying the two effects are (globally) negatively dependent. This dependence tends to be quite strong. For example, when firm productivity is uniform (i.e., $\bar{F}(x) = 1 - x$) the across-firm correlation between $\psi(p)$ and $\lambda(p)$ is bounded from above by -0.98. Moreover, the variance of destination effects across firms must be strictly smaller than the variance of origin effects. Intuitively, this ordering arises because destination effects capture only compensating differentials while origin effects capture both these differentials and employer productivity.

Because the PVR model requires a worker to always accept an offer from a more productive firm, the mobility decision depends entirely on p and q – or equivalently on $\psi(p)$ and $\lambda(q)$ – which is consistent with the exogenous mobility assumption in (2). Note that equation (3) does not include an error term specific to the worker-firm match. Such errors arise after the match has been consummated as workers begin to attract outside offers. Because we only apply the DWL specification to hiring wages, these within match errors do not generate a violation of the exogenous mobility requirement in (2).

9

2.2 Bargaining extensions

Cahuc et al. (2006, C-PVR) generalize the PVR model by allowing workers to negotiate a share $\beta \in [0,1]$ of the surplus in the employment relationship. Because the C-PVR model assumes linear utility, a DWL representation holds for wage levels rather than log wages.[1] Subsequent work by Bagger et al. (2014, BF-PVR) extends the C-PVR model to accommodate human capital accumulation while assuming flow utility is logarithmic.

Applying the fundamental theorem of calculus to equation 7 of Bagger et al. (2014) reveals that the deterministic solution to the BF-PVR model yields a DWL representation for log hiring wages of the form:

$$\ln \phi\left(\epsilon, p, q, \mathcal{X}, \mathcal{E} \mid \beta\right) = \alpha(\epsilon) + g(\mathcal{X}) + \mathcal{E} + \underbrace{\beta \ln p + I(p \mid \beta)}_{=\psi(p)} + \underbrace{(1 - \beta) \ln q - I(q \mid \beta)}_{=\lambda(q)}, \qquad (4)$$

where $\mathcal{X}$ represents labor market experience, which can be included in the DWL model's covariate vector X_{im}, and $\mathcal{E}$ is a transitory worker-specific productivity shock that provides a structural interpretation to the DWL errors ε_{im}. Because the productivity shock has not been realized at the time of negotiation, workers always accept offers from more productive firms, which implies ε_{im} satisfies our exogenous mobility requirement in (2).

The tail integral $I(z \mid \beta) = (1 - \beta)^2 \kappa \int_z^\infty \left(\bar{F}(x)/x\right) / \left(1 + \kappa \beta \bar{F}(x)\right) dx$ is decreasing in both its arguments. Note that $I(z \mid 0) = I(z)$; therefore, when $\beta = 0$, equation (4) specializes to the PVR reduced form in (3), albeit with additional covariates and a time varying error. When β is positive, workers are able to capture a share of the destination firm's log productivity, which becomes a part of the destination effect $\psi(p)$. When $\beta = 1$, the origin effects disappear and (4) collapses to an AKM style specification for log hiring wages.

As in the PVR model, the sum of a firm's origin and destination effects equals its log productivity. Unlike in the PVR model, however, the BF-PVR destination effects are increasing in the hiring firm's productivity whenever $\beta > 1/2$ because the direct wage effects of productivity overwhelm their indirect effects of $I(p \mid \beta)$ that are attributable to compensating differentials. Large values of β can therefore lead $\psi(p)$ and $\lambda(p)$ to covary positively and for the destination effects to exhibit greater variance than the origin effects. As described in the next section, these comparative statics imply some over-identifying restrictions on the covariance structure of firm origin and destination effects. Finally, because $I(p \mid \beta)$ is convex in $\ln p$ for any value of β (see Appendix B), the origin effects must be concave in log productivity, while the destination effects must be convex in $\ln p$, restrictions we examine empirically using data on firm value added per worker.

---

[1]See, for instance, Lemma 1 of Papp (2013) which establishes additive separability of origin and destination effects in the case where $\epsilon = 1$ for all workers. Introducing heterogeneity in the flow value b of non-employment (e.g., as in Postel-Vinay and Robin, 2002a) generates additively separable workers effects.

3   Variance components

The log-linear DWL specification in (1) admits a parsimonious summary of the model parameters in terms of variance components. A first set of variance components summarizes heterogeneity across firms and can be used to derive bounds on worker bargaining strength. A second set of variance components is useful for decomposing hiring wage variability across workers.

3.1   Variability across firms

We summarize the offered wage distribution with the following *firm*-level variance components:

$$\mathbb{V}_J[\psi], \quad \mathbb{V}_J[\lambda], \quad \mathbb{C}_J[\psi, \lambda],$$

where $\mathbb{V}_J[\cdot]$ and $\mathbb{C}_J[\cdot]$ denote, respectively, sample variances and covariances across the firms in our sample, weighted by average firm size over time. The covariance is only identified among the firms where both ψ_j and λ_j are identified, which requires both that some workers be hired by and quit from firm j over the sampling period. We therefore report variance components only for such firms.

The textbook PVR model implies that $\mathbb{V}_J[\psi] < \mathbb{V}_J[\lambda]$. By contrast, the BF-PVR model can rationalize destination effects that are more variable than origin effects, but only when workers have substantial bargaining strength. From (4) we have that ψ equals $\beta \ln p$ plus the variable $I(p \,|\, \beta)$ which is negatively correlated with $\ln p$. By standard omitted variables bias logic, the coefficient from a population projection of ψ onto $\ln p$ must therefore be smaller than β. Evaluating the expression for this projection coefficient and rearranging yields the following bound on worker bargaining power:

$$\beta \geq \frac{1}{2} + \frac{\mathbb{V}_J[\psi] - \mathbb{V}_J[\lambda]}{2\mathbb{V}_J[\psi + \lambda]}. \tag{5}$$

This bound reflects the intuition that as β grows large, the BF-PVR reduced form approaches an AKM specification, and the variance of destination firm effects must become large relative to the variance of origin firm effects. Note from (5) that for destination effects to be more variable than origin effects, workers must possess $\beta > 1/2$, implying very significant bargaining strength.

As shown in Appendix B, the BF-PVR model additionally restricts the derivative of $I(p \,|\, \beta)$ to be greater than $-(1 - \beta)^2/\beta$. When $\beta \geq 1/2$, this restriction can be exploited to derive the following lower bound on the correlation between origin and destination effects:

$$\rho_J(\psi, \lambda) \geq \sqrt{\frac{\mathbb{V}_J[\psi]}{\mathbb{V}_J[\psi + \lambda]}} \left(1 - \frac{3}{10} \sqrt{\frac{\mathbb{V}_J[\lambda]}{\mathbb{V}_J[\psi + \lambda]}} \right) \tag{6}$$

The logic of this bound can be described as follows. When destination firm effects are more variable

than origin firm effects, β must be large. But strong worker bargaining power requires both the origin and destination firm effects to be globally increasing in firm productivity, which is the only dimension along which firms differ. Hence, the origin and destination effects must be strongly positively correlated. Because the DWL model treats origin and destination effects as potentially unrelated parameters, we are able to evaluate whether this restriction is satisfied empirically in the data. When it is satisfied, an additional set of bounds, described in Appendix B, can be used to bracket β. When it is not, the model is rejected.

3.2 Variability across workers

We also consider *worker*-level variance components, which provide a summary of the distribution of accepted wages. For any two variables w and z, $\mathbb{C}_n[w, z]$ denotes sample covariance between w and z weighted by worker-match observations, while $\mathbb{V}_n[w] = \mathbb{C}_n[w, w]$ gives the corresponding sample variance of w. Letting $\mathcal{W} = \{\mathcal{W}_i\}_{i=1}^n$, the expected sample variance across workers of (covariate adjusted) log hiring wages can be decomposed as follows:

$$\mathbb{E}\big[\mathbb{V}_n\big[y - X'\delta\big] \mid \mathcal{W}\big] = \mathbb{V}_n[\alpha] + \mathbb{V}_n[\psi] + \mathbb{V}_n[\lambda] + 2\mathbb{C}_n[\alpha, \psi] + 2\mathbb{C}_n[\alpha, \lambda] + 2\mathbb{C}_n[\psi, \lambda] \qquad (7)$$
$$+ \mathbb{E}[\mathbb{V}_n[\varepsilon] \mid \mathcal{W}].$$

Here, exogenous mobility implies that all covariances between ε and the remaining variables are zero. The first three terms in this decomposition give the expected contributions to log hiring wage variance of variability in worker effects α, destination effects ψ, and origin effects λ. The first two terms are familiar from the standard AKM decomposition. The variance of the origin effects provides a metric of the contribution of state dependence to wage inequality.

The three covariances quantify different aspects of sorting. The first term $\mathbb{C}_n[\alpha, \psi]$ captures the extent to which high wage workers tend to be employed at high destination effect firms. This term is conceptually similar to the worker-firm effect covariance proposed by Abowd et al. (1999) as a measure of sorting. However, because we fit the model to hiring wages, the interpretation is potentially quite different. With random (i.e., undirected) search, a high wage worker is no more likely to draw an offer from a high wage firm. Hence, in the PVR model this covariance should be zero. Bagger and Lentz (2019) add endogenous search effort to the C-PVR framework, which allows for positive assortative matching between worker and firm productivities. Note however that productivity based sorting need not yield a positive correlation between worker effects and destination effects when workers exhibit low bargaining strength.

The next term, $\mathbb{C}_n[\alpha, \lambda]$ captures the extent to which high wage workers tend to be poached from firms with high origin effects. We are not aware of previous estimates of this parameter. Again, with random search, this covariance should be small. Finally, $\mathbb{C}_n[\psi, \lambda]$ captures the extent

to which workers poached from high origin effect firms tend to be hired by high destination effect firms. In the PVR model, only highly productive (and therefore low ψ) destination firms can poach from high λ sources, which implies this covariance will be negative when search is undirected.

The last line of (7) gives the "unexplained" variance in log hiring wages. Because little is known about the hiring wage errors, we avoid imposing that they are homoscedastic, instead allowing each error ε_{im} its own variance parameter. We provide evidence later in the paper that heteroscedasticity is empirically important.

The variance decomposition in (7) is only identified among worker-firm matches where the origin and destination effects, $\psi_{j(i,m)}$ and $\lambda_{h(i,m)}$, are separately identified. Furthermore, as established in Lemma 1 of Kline et al. (2020), unbiased estimators of the variance components only exist if identification holds when any single worker-firm match is dropped from the sample. We therefore restrict our estimation sample to ensure that these requirements are satisfied using an algorithm described in Appendix D. A discussion of the mobility patterns that yield identification of the DWL model is given in Appendix C. We note there that *pairwise differences* among workers who share certain aspects of their career path play a crucial role in identifying the parameters of the model.

4 Leave-out estimation

We now briefly review the leave-out estimation procedure of Kline et al. (2020), which enables consistent estimation of variance components in the presence of unrestricted heteroscedasticity. For expositional clarity, it is useful to map the observations in our data to a single index $\ell \in \{1, \ldots, L\}$ where $L = \sum_{i=1}^{n} M_i$ gives the total sample size. The DWL specification in (1) can then be written compactly as:

$$y_\ell = Z'_\ell \gamma + \varepsilon_\ell, \quad \text{for } \ell = 1, \ldots, L,$$

where $y_\ell = y_{im}$, $\varepsilon_\ell = \varepsilon_{im}$, and Z_ℓ collects the vectors of worker indicators, hiring firm indicators, hiring origin indicators, and time varying covariates for the worker-firm match (i, m). The unknown regression coefficients are collected in the vector γ.

Any of the variance components we study can be written as a quadratic form:

$$\theta = \gamma' A \gamma$$

for some square matrix A.[2] Let $S_{zz} = \sum_{\ell=1}^{L} Z_\ell Z'_\ell$ give the invertible design matrix. The OLS

---

[2]See Kline et al. (2020) or the appendix to Card et al. (2013) for examples.

estimator of γ is:

$$\hat{\gamma} = S_{zz}^{-1} \sum_{\ell=1}^{L} Z_{\ell}' y_{\ell} = \gamma + S_{zz}^{-1} \sum_{\ell=1}^{L} Z_{\ell}' \varepsilon_{\ell}.$$

The *plug-in* estimator of the variance component θ is $\hat{\theta}_{\mathrm{PI}} = \hat{\gamma}' A \hat{\gamma}$. When the errors ε_{ℓ} are mutually independent, this estimator exhibits a bias of

$$\mathbb{E}[\hat{\theta}_{\mathrm{PI}} \,|\, \mathcal{W}] - \theta = \mathrm{trace}\,(A \mathbb{V}[\hat{\gamma} \,|\, \mathcal{W}]) = \sum_{\ell=1}^{L} B_{\ell\ell} \sigma_{\ell}^2,$$

where $B_{\ell\ell} = Z_{\ell}' S_{zz}^{-1} A S_{zz}^{-1} Z_{\ell}$ measures the influence of the ℓ'th squared error ε_{ℓ}^2 on $\hat{\theta}_{\mathrm{PI}}$ and $\sigma_{\ell}^2 = \mathbb{V}[\varepsilon_{\ell} \,|\, \mathcal{W}]$ is the variance of the ℓ'th error.

   To remove this bias, we follow Kline et al. (2020) in constructing estimators of each σ_{ℓ}^2. Denote the leave-ℓ-out estimator of γ by $\hat{\gamma}_{-\ell} = \left(S_{zz} - Z_{\ell} Z_{\ell}'\right)^{-1} \sum_{l \neq \ell} Z_l' y_l$. An unbiased estimator of σ_{ℓ}^2 is

$$\hat{\sigma}_{\ell}^2 = y_{\ell} \left(y_{\ell} - Z_{\ell}' \hat{\gamma}_{-\ell}\right) = \frac{y_{\ell} \left(y_{\ell} - Z_{\ell}' \hat{\gamma}\right)}{1 - P_{\ell\ell}}, \tag{8}$$

where $P_{\ell\ell} = Z_{\ell}' S_{zz}^{-1} Z_{\ell}$ gives the statistical "leverage" of the ℓ'th observation on $\hat{\gamma}$. Our corresponding bias corrected estimator of θ can be written

$$\hat{\theta}_{\mathrm{KSS}} = \hat{\gamma}' A \hat{\gamma} - \sum_{\ell=1}^{L} B_{\ell\ell} \hat{\sigma}_{\ell}^2.$$

   Computation of $\hat{\theta}_{\mathrm{KSS}}$ requires evaluating the $\{B_{\ell\ell}, P_{\ell\ell}\}_{\ell=1}^{L}$. Because our baseline model contains more than 4 million parameters, brute force computation is intractable. We therefore rely on a variation of the random projection method (Johnson and Lindenstrauss, 1984; Achlioptas, 2003) described in Kline et al. (2020) to approximate $\hat{\theta}_{\mathrm{KSS}}$.

5   Data

Our data are derived from social security records spanning the years 1990-2015 maintained by the Italian Social Security Institute (*Istituto Nazionale Previdenza Sociale*, INPS). These records cover all private-sector workers who were employed at some point by a firm sampled by the Bank of Italy's INVIND survey and have featured in a number of recent studies of Italian wage inequality (Macis and Schivardi, 2016; Daruich et al., 2020).

   The INPS-INVIND dataset records the annual earnings, days worked, months of employment, and establishment and tax unit identifiers for each job-spell observed in a given year. We take

as our concept of a firm the tax unit identifier (*Codice Fiscale*).[3] Starting in 2005, the INPS-INVIND data also record the stated reason for the dissolution of each job match, which allows us to distinguish between job separations resulting from voluntary worker resignations and instances where a firm fires a worker, lays them off, or declines to renew their contract.[4] To take advantage of this information, we limit our analysis to the period 2005-2015; however, we use the records back to 1990 to determine whether a worker is entering the labor force for the first time. Appendix D provides details on our processing of the data.

To code employment histories, we extract the job start and end dates of all workers with two or more jobs. A job transition is coded as a voluntary quit ($Q_{i,m} = 1$) whenever a worker formally resigns from their job. When the reason for separation variable is missing, we code the separation as involuntary if the job start date comes more than a month after the separation date.[5] Because we seek to characterize the *sequence* of jobs each worker holds, we depart from the usual practice of restricting the sample to a single dominant earnings record in a year (e.g., as in Card et al., 2013). Instead, we assign to each worker-*month* a dominant employer (or unemployment) based upon the earnings records in that year. When workers transition between multiple dominant jobs in a year, each hiring event is entered as a separate record. Transitions between such jobs are coded according to the stated reason for separation in the usual way.

Roughly 31% of the transitions in our data are coded as voluntary. This estimate aligns closely with data from the Job Openings and Labor Turnover Survey (JOLTS): 29% of JOLTS separations were voluntary in May 2009 while 38% were voluntary in May 2019. Given that the Italian unemployment rate averaged 9% over our sample period, it is somewhat reassuring that our estimate is closer to the JOLTS figure for May 2009 than for May 2019.

We measure the hiring wage with the logarithm of the average daily wage in a worker's first calendar year on the job. If the worker transitions between multiple jobs in a year he or she will have multiple hiring wages for that year. In Italy, employment contracts are typically reassessed annually as established by sectoral collective bargaining agreements (*Contratti Collettivi Nazionali Del Lavoro*) and firm-level wage agreements; see for instance the discussion in Guiso et al. (2005) and the documentation provided in Assolombarda (2012). Because these institutions make it very difficult to change wages within the year, we expect the first year's average daily wage to closely

---

[3]These identifiers should be thought of as somewhat broader than the EIN definition used by the US Internal Revenue Service. Song et al. (2019) report that the 4,233 firms listed on the New York Stock Exchange possess 13,377 distinct EINs. By contrast, each publicly listed Italian firm has a unique Codice Fiscale.

[4]In Italy, firms are permitted to terminate permanent employment contracts for objective reasons (i.e, financial distress) or subjective reasons such as improper conduct by the worker. Firms can also allow temporary employment contracts to expire, which is a source of many involuntary separations (Cahuc et al., 2016; Daruich et al., 2020).

[5]Around one fourth of all observed transitions fail to report a reason for separation and roughly 70% of these transitions are coded as involuntary separations.

approximate the official wage offer negotiated at the time of hiring.

In later sections, we also leverage data from two additional sources that we link with INPS-INVIND. A file called *Anagrafica* contains national tax identifiers, firm size, and sector (2-Digit Ateco 2007 codes) for the universe of Italian employers. Using the tax identifiers, we merge in firm value added records from CERVED, a dataset which provides financial statements for the universe of Italian limited liability companies. CERVED is used in conjunction with Anagrafica to compute a measure of value added per worker.

6 Descriptive statistics

Table 1, Panel (a) shows some descriptive statistics on the worker-match panel derived from INPS-INVIND. The data contain roughly 13 million hiring events involving around 4.9 million individuals, 2.9 million of whom are men and 2.0 million of whom are women. Over the course of our study period, these workers quit jobs at 876 thousand distinct employers and are hired by roughly 1.5 million distinct employers. While most hires are from unemployment (i.e., involuntary transitions), roughly a third of hires involve voluntary transitions from another firm, and approximately 10% of hires are of workers new to the labor force. Women are slightly less likely to be poached from another firm than men, with 33% of male hires but only 29% of female hires involving a voluntary transition from a previous employer.

As mentioned in Section 3, unbiased estimation of the variance components associated with the DWL model requires that the origin and destination effects be estimable when any single person-job observation is dropped from the sample. Panel (b) of Table 1 shows the results of pruning the sample to enforce this requirement. The estimation sample has roughly a quarter fewer observations and workers than the starting sample. The number of origins and destinations falls by roughly a half in the pruned sample, primarily because many firms are associated with only a single hire. In the resulting estimation sample there are roughly 14 hires per destination firm and 8.6 quits per origin firm. Reassuringly, both the mean and variance of hiring wages change little with pruning.

Figure A.1 shows the distribution of months of non-employment between jobs by transition type for both our starting and estimation sample. The distributions of non-employment durations in the two samples is quite similar, with slightly longer tails present in the starting sample. The vast majority of voluntary transitions in our estimation sample involve very short bouts of non-employment between jobs, with fewer than 20% of such transitions entailing non-employment spells longer than three months. Interestingly, a non-trivial fraction of involuntary transitions involve only a month of non-employment between jobs. A disproportionate fraction of these cases correspond to workers that were subject to domestic outsourcing events.

Finally, Table A.1 provides summary statistics on the firms in our base and estimation sample

and compares them to the broader population of Italian firms monitored by INPS. While the sectoral mix of firms in our estimation sample is broadly representative of the Italian economy, smaller firms are under-represented.[6] However, the standard deviation of log firm size in our estimation sample is very close to that in the population INPS records, suggesting our firms are no more (or less) heterogeneous than the broader population of Italian firms.

7 Diagnostics

Before estimating the parameters of our main specification, we consider some diagnostics meant to probe the qualitative predictions of the DWL model. Our first diagnostic examines whether being hired from unemployment actually affects the hiring wage. Figure 1 plots the mean change in log hiring wages between the first and second job of workers who were involuntarily displaced from their first job against the mean change of those who voluntarily quit their first job to take the second job. Following Card et al. (2013), these means are broken out by the quartile of coworker wages at the first and second job, yielding 16 pairs of coworker wage groups in total.

The traditional AKM specification predicts that the labor market state from which a worker was hired is irrelevant, which implies the wage growth between jobs is attributable only to the difference in destination effects. Consequently, the plotted means should lie on a 45 degree line through the origin. By contrast, the DWL model predicts an approximately constant penalty for being hired from unemployment at the second job versus being poached from a job in the same quality quartile. Visually, this should lead the means to lie on a line below, and parallel to, the 45 degree line.

In practice, the DWL model appears to provide a very accurate approximation to mean wage changes. A linear fit to the mean wage changes yields a slope of 1.01 and an intercept of -0.06, suggesting that involuntary separations generate a penalty of roughly 6% on subsequent hiring wages. We explore in the next section whether ignoring such effects substantially biases conventional AKM estimates of destination effects.

Our second diagnostic probes a key restriction of the DWL model: upon being involuntarily displaced, a worker's prior employment history should not affect their hiring wage. To test this prediction, we examine the growth in hiring wages between the second and third jobs for workers displaced from *both* their first and second jobs. Recall that, without time-varying covariates, the DWL model predicts the wage growth of such individuals will obey the equation:

$$y_{i3} - y_{i2} = \psi_{j(i,3)} - \psi_{j(i,2)} + \varepsilon_{i3} - \varepsilon_{i2}.$$

---

[6]The median firm in the formal population INPS records has only 2 workers, as also reported by Akcigit et al. (2018), while the median firm in the INPS-INVIND data has 4 workers, and our pruned estimation sample has a median firm size of 8 workers.

Note that this is an AKM-style model that exhibits no dependence on the identity of the first employer $j(i, 1)$. To assess the excludability of the first employer, Figure 2 plots the wage growth of workers whose first job fell in the first tercile of coworker wages (a low wage employer) against the wage growth of workers whose first job fell in the last tercile of coworker wages (a high wage employer). The means are again classified into 16 groups, this time based upon the coworker wage quartiles of the second and third jobs. In accord with the DWL model, these means are tightly clustered around the 45 degree line, indicating that the identity of the first job does not affect mean wage growth between the second and third job.

For comparison, we also plot the corresponding mean wage growth for workers who voluntarily quit both their first and second jobs. Though these means are also clustered near the 45 degree line, the fit is somewhat worse. While projecting the mean wage change of workers initially employed at a high wage job onto the mean wage change of workers initially employed at a low wage job yields a slope of 0.999 for workers involuntarily separated from their first two jobs, the corresponding slope for workers who quit their first two jobs is only 0.942, indicating small violations of excludability. The latter violations, of course, present problems only for the AKM model, not the DWL model, which conditions on the prior employer of workers who quit their jobs.

8   Results

The above section suggests the DWL model provides a qualitatively accurate characterization of changes in hiring wages across jobs. We turn now to a quantitative assessment of the explanatory power of the DWL model. Table 2 reports bias corrected estimates of R^2 for three models: AKM, DWL, and an AKM variant with worker and origin (rather than destination) fixed effects.[7] Each model includes controls for a third order polynomial in age at hiring (centered at age 40) and a set of indicators for the calendar year of the hiring event. The AKM model explains 72% of the variation in log hiring wages in our sample.[8] Replacing the destination effects in the AKM model with origin effects lowers the R^2 by roughly 14 percentage points to 58%. Evidently, origin effects are much less predictive, unconditionally, of hiring wages than are destination effects.

Adding origin effects to the AKM model yields the DWL model, which achieves an R^2 of 72.5%. That adding origin effects to the AKM model explains only an additional 0.5% of the variance of wages suggests that where a worker is hired from is far less important for her wages than where she is currently employed. The subdued influence of origin effects is particularly evident for women,

---

[7]See Kline et al. (2019) for discussion of this fit measure, which can be thought of as a heteroscedasticity robust version of the conventional adjusted R^2.

[8]This R^2 estimate is lower than what has been found in past work using Italian wage records (Devicienti et al., 2019; Kline et al., 2020) because our sample does not include firm stayers, who mechanically enjoy a perfect fit to their match means.

for whom the added explanatory power of the origin effects is only 0.3 percentage points. Allowing the origin and destination fixed effects to vary by gender raises the pooled explanatory power of the DWL model by just under 2 percentage points. Interestingly, the DWL model's composite explanatory power is greater for the wages of men than for women, revealing that gender is a potentially important source of heteroscedasticity in the wage error variances. Figure A.2 shows that our leave out estimates of error variance $\hat{\sigma}_\ell^2$ vary systematically by worker gender, age at hiring, and employer value added.

A useful point of reference for the findings in Table 2 comes from Bonhomme et al. (2019) who find that moving from a static model of wage determination to a fully dynamic model with origin effects and within match dynamics raised the share of wage variance explained in Swedish administrative records from 74.9% to 77.9%. Though they included incumbent wages in their sample and used different methods to estimate wage decompositions, their static model explained roughly the same amount of wage variance as our AKM specification does for hiring wages in Italy. We conjecture that the greater increase in explanatory power Bonhomme et al. (2019) obtain with a dynamic model is primarily attributable to their inclusion of lagged wages as a predictor of wage growth rather than the inclusion of origin effects.

8.1 Worker-level variance decompositions

As a benchmark for our DWL estimates, Table 3 reports a standard AKM decomposition of the variance of log hiring wages into components attributable to worker and firm effects. After bias correction, we find that destination firm effects explain 24% of the variance of wages in our pooled sample, while worker effects explain 30%.[9] This relatively large contribution of firm effects to wage variability appears to be jointly attributable to our focus on hiring wages and job movers.[10] The bias corrected correlation between worker and firm effects is 0.31, indicating substantial worker-firm sorting. This correlation is estimated to be somewhat stronger among women than men.

Table 4 reports estimates of the DWL specification, which decomposes the variance of log hiring wages into components attributable to worker effects, destination effects, origin effects, and their covariances. After correction for over-fitting, the destination firm effects explain roughly 24% of

---

[9]Appendix Table A.2 reports the results of fitting a corresponding AKM specification to the set of firms that remain connected when leaving out all records associated with any single *worker*. Consistent with the findings of (Kline et al., 2020, Table A.1), bias correcting the variance of the firm effects by leaving out all records associated with a worker yields results nearly identical to those obtained by leaving out a single worker-firm match. This finding corroborates our maintained assumption that the DWL errors ε_{im} are approximately independent across matches.

[10]Appendix Table A.3 shows that including the within match wages of job movers lowers the firm effect variance share to roughly 19%. Additionally including job stayers in the sample reduces the variance share of firm effects to roughly 16%.

the variance of hiring wages, rivaling the worker fixed effects which explain 29% of the variance.[11] When disaggregated by gender, the destination and worker effects explain nearly the same shares of variance, with destination effects actually exhibiting slightly more variability than worker effects among women.

Comparing Tables 3 and 4 suggests that omitting origin effects yields little change in either set of fixed effect estimates, an impression corroborated by Figure 3. This finding allays to some extent the concerns of PVR who note regarding AKM decompositions that "Estimating a static error component model when the data generating process is dynamic will therefore attribute all historical differences (in the states of individual wage trajectories at the first observation date) to person effects." In practice, person effects are not especially sensitive to the omission of origin effects, both because origin effects are not particularly variable and because they exhibit weak correlation with the worker effects.

When included, the origin effects explain only 0.7% of the variance of hiring wages. While most of this variance is attributable to variation in origin effects among workers involved in a voluntary transition (i.e., "poached"), the wage penalties associated with job displacement or entering the labor force are non-trivial. New labor market entrants face an average hiring wage penalty of 5.6% relative to the average worker involved in a voluntary job transition. The wage penalty for being displaced in an involuntary job transition is estimated to average 4.0% for workers actually involved in such transitions and 3.9% for the workers who quit. The small difference in mean origin firm effects between workers who quit and those who are displaced suggests origin effects play little role in worker separation decisions.

As in the earlier AKM specification, we find that high wage workers sort to high wage destinations: the correlation between the worker effects and destination firm effects is 0.32. By contrast, the correlation between worker effects and origin effects is only 0.12, perhaps because skilled workers are often involuntarily displaced in our sample. Origin and destination effects are only weakly related with a correlation of 0.03. While women exhibit a stronger correlation between worker and destination effects than men, the correlation between worker effects and origin effects is stronger among male than female workers. Evidently, women are more assortatively matched to destinations, while men are more assortatively matched to origins. We examine in a later section what role these sorting differences may play in the evolution of the gender wage gap.

8.2 A firm level variance decomposition

Table 5 provides a variance decomposition across firms of the two dimensional fixed effect vector (ψ_j, λ_j). The correlation across firms between their origin and destination effects is 0.25, indicating that quitting a high wage firm tends to yield elevated wages at one's next job. Evidently, firms

---

[11]Uncorrected estimates of the DWL variance components are provided in Appendix Table A.4.

that are good to be at are also good to be from. As noted in section 2, rationalizing this pattern in the sequential auction framework requires that workers possess substantial bargaining strength.

Recall that in the BF-PVR model summing a firm's origin and destination effects yields an estimate of its log productivity. The size-weighted standard deviation across firms of the sum of origin and destination fixed effects is roughly 0.29. For comparison, the size-weighted standard deviation of log value added per worker is roughly 0.8. Since value added is likely a noisy measure of productivity, and should hypothetically be adjusted for input variation, this discrepancy need not pose a serious challenge to the model.

More troubling is that the size-weighted variance of destination effects is more than 13 times the size-weighted variance of origin effects. Ratios this large are difficult to rationalize in a sequential auction model without extremely strong worker bargaining power. From Table 5 we obtain an estimate for $\frac{\mathbb{V}_J[\psi] - \mathbb{V}_J[\lambda]}{\mathbb{V}_J[\psi + \lambda]}$ of 0.76. Plugging this number into (5) yields an estimated lower bound for β of 0.88! Conducting this computation separately by gender, the corresponding lower bound for men is 0.87 while the lower bound for women is 0.92. These lower bounds on the bargaining strength parameter substantially exceed rent sharing estimates in the literature reviewed by (Card et al., 2018), which typically finds estimates of β below 1/2. They also exceed BF-PVR's own indirect inference based estimates which average roughly 0.3.

Equation (6) provides another check on the plausibility of this bargaining power estimate. Rationalizing a worker bargaining share of 0.88 requires a correlation between origin and destination firm effects of at least 0.84, well above our empirical correlation estimate of 0.25. Correspondingly large violations of this model based correlation bound are present in both gender specific samples. Hence, the covariance matrix of origin and destination firm effects is incapable of being rationalized by the BF-PVR model.

One reason for these violations may be that our sample pools workers from the entire Italian economy. Figure 4 plots estimates of the variability of firm origin and destination effects among subsets of firms corresponding to selected sectors of the Italian economy.[12] A first finding is that substantial variability in firm origin and destination effects appears to be present even within narrow sectors of the Italian economy. Unsurprisingly, temp agencies have very small origin and destination effect variances, as workers are not meaningfully attached to these firms. However, the restaurant and hotel sector exhibits large variability in destination effects but relatively muted variability in origin effects. By contrast, law firms exhibit substantial variability in both origin and destination effects. Indeed, the two sets of effects are roughly equally variable.

Table 6 shows the corresponding lower bounds on bargaining power and the correlation between origin and destination firm fixed effects in these sectors. The general excess variation in destination effects across most of these sectors yields lower bounds on bargaining power that remain implausibly

---

[12]The fraction of hiring wage observations falling into each sector is reported in Table A.1.

high. The sole exception is law firms, which exhibit a lower bound on β of 0.54. However, law firms exhibit little correlation between origin and destination effects while the BF-PVR model requires a correlation of at least 0.58. This violation of the correlation bound extends to all of the sectors we study, implying the BF-PVR model cannot rationalize the structure of wages in any of these industries.

The inability of the BF-PVR model to rationalize destination effects that are so much more variable than origin effects is attributable to the assumption that both sets of effects are a common manifestation of a single latent factor: firm productivity. We now turn to investigating this assumption more directly by making use of data on firm value added.

8.3 Firm wage effects and productivity

Figure 5 plots means of the estimated destination and origin effects by centiles of log value added per worker. The destination effects are normalized to have mean zero in the bottom vingtile of value added, while the origin effects use the normalization $\lambda_N = 0$.

The sequential auction framework predicts that origin effects will be increasing functions of productivity, as more productive firms can offer higher wages to retain their workers. Destination wage effects, by contrast, may be decreasing in productivity if workers are willing to take pay cuts to join more productive firms. Panel (a) of Figure 5 shows that the estimated destination effects are in fact strongly increasing in value added per worker, exhibiting a "hockey stick" pattern of the sort first documented by Card et al. (2015). The origin effects are also increasing in value added, though only in the top half of the value added distribution.

The bottom panel of Figure 5 plots the mean origin and destination effect in each value added centile against the mean value of the sum of origin and destination effects, which should correspond to a firm's log productivity in the BF-PVR model. To quantify the relative sensitivity of origin and destination effects to productivity we fit a line to each series. Because both relationships appear somewhat nonlinear, these lines are fit separately to the top and bottom 50 value added bins. Note that the resulting slopes are equivalent to running two stage least squares regressions of each type of fixed effect on the sum of fixed effects and instrumenting with value added centiles in the relevant range.

Among the bottom 50 value added bins, the projection slope of average destination effects with respect to average log productivity is 0.92. Recall from equation (4) that in the BF-PVR model the derivative of the destination effects with respect to log productivity should provide a lower bound on β. Hence, if we take the projection slope as a weighted average derivative estimate, we arrive at an implausibly large lower bound for β of about 0.92. Among the top 50 value added bins, the projection slope falls to 0.78. Recall however that the destination effects should be convex in $\ln p$. The finding of a lower slope at higher productivity levels suggests the destination effects are instead

concave in log productivity, a pattern the sequential auction framework cannot rationalize.[13]

The origin effects are much less sensitive to productivity than the destination effects. The projection slope of the average origin effects with respect to average productivity rises from only 0.03 among the bottom 50 bins to 0.22 in the top 50 bins. This pattern suggests the origin effects are convex in log productivity, which contradicts the sequential auction model's prediction that this relationship should be concave.

A natural explanation both for the subdued sensitivity of origin effects to value added and the finding that firm destination effects are an order of magnitude more variable than firm origin effects is that many firms do not negotiate wages based on hiring origins, committing instead to uniform wage premia as in the classic wage posting framework of Burdett and Mortensen (1998). Fitting a version of the model of Flinn et al. (2017) to Danish data, Caldwell and Harmon (2019) find that only 31% of manual jobs and 51% of professional jobs engage in wage negotiation.[14] This finding is in line with an array of survey evidence indicating that most firms engage in ex-ante wage posting behavior (Hall and Krueger, 2012; Brenzel et al., 2014), especially for lower skilled jobs (Brenčič, 2012). Though a proper analysis of the ability of mixture formulations to match the covariance structure of origin and destination effects is beyond the scope of this paper, we suspect that rationalizing the estimates reported in Table 5 with plausible bargaining parameters would require even greater shares of firms engaged in wage posting than has been found in surveys.

A complementary explanation for the muted variance of origin effects is that even firms that do engage in negotiation have difficulty assessing the value of worker's outside options or fully exploiting that information when it is available. Consistent with that view, Jäger et al. (2018) find no evidence that an Austrian reform to the generosity of unemployment insurance affected hiring wages. Likewise, a growing experimental literature on pay transparency suggests firms face important horizontal equity constraints that may curtail their ability to price discriminate at the time of hiring (Card et al., 2012; Breza et al., 2018; Mas, 2017; Cullen and Pakzad-Hurson, 2019). Exploring how variation in pay transparency affects the tendency of firms to counter outside wage offers would seem to be a fertile area for future research.

8.4   Incumbent wage growth

We conclude this section by briefly investigating how our DWL estimates relate to incumbent wages and worker separation rates. Sequential auction models rationalize the wage growth of incumbent workers as resulting from the countering of outside offers. Figure 6 illustrates how wages evolve

---

[13]Appendix B formalizes the connection between concavity/convexity of the underlying firm effects in productivity and the patterns displayed in the bottom panel of Figure 5.

[14]Partial integration of equation 34 of Flinn et al. (2017) reveals that the model admits a DWL representation for wage levels among firms that engage in negotiation. Interestingly, the resulting origin and destination effects may be concave or convex in productivity depending on the region of evaluation.

over the first three years of a job match for job stayers by centiles of value added per worker of the hiring firm. The x-axis plots the mean prediction $Z'_\ell \hat{\gamma}$ of the DWL model in each value added centile. The triangles give the fraction of workers hired into each value added centile who separate from the job within three years. Separation rates hover near 80% at the least productive employers, which are often intensive in temporary jobs, and decline to around 35% at the most productive employers.

The circles plotted in Figure 6 report the mean hiring wage of the subpopulation of workers who remain employed at their destination firm for three or more years. Because the DWL model was fit only to hiring wages, these means would lie perfectly along the 45 degree line had the sample not been restricted to workers with three years of tenure. The fact that we see only very small discrepancies from the 45 degree line indicates that separations from the destination firm are not strongly related to the hiring wage error ε_{im}. This is good news for the BF-PVR model, which models these errors as transitory productivity disturbances. It is also good news for the DWL specification: if equation (1) was severely misspecified, the reduced form error ε_{im} would contain neglected worker and firm heterogeneity, which would likely predict separation behavior.

The squares plotted in Figure 6 depict the mean wage of incumbent workers in their third year on the job. Remarkably, these means are tightly clustered along a line segment parallel to the 45 degree line. A linear fit to these means reveals that, on average, workers in their third year of job tenure have wages roughly 18% higher than in the year they were hired, while the slope of the relationship with the mean DWL prediction is nearly one. Evidently, wage growth over the first three years of the match among job stayers bears little relationship to firm value added.

When mobility is efficient, a worker employed by the least productive firm leaves whenever she receives an outside offer; hence job stayers consist of offerless workers who should experience no wage growth. This prediction appears to be violated, as workers employed by firms in the bottom centiles of value added exhibit roughly the same wage growth as found at other levels of measured productivity. On the other hand, a worker employed by the most productive firm never leaves in response to an outside offer, and experiences a wage hike whenever she is approached by a sufficiently productive suitor. However, Figure 6 reveals that the wage growth of workers employed by firms in the top two value added centiles falls below that of workers employed at firms with much lower levels of value added per worker. A potential explanation for this depressed growth in the top two centiles comes from models with endogenous search effort (e.g., Bagger and Lentz, 2019), which predict that workers at the most productive firms have less incentive to engage in on the job search.

Between the poles of the productivity distribution, the predictions of the sequential auction framework are somewhat more difficult to assess in the absence of functional form restrictions, as a more productive firm can offer larger raises to underpaid employees but also retains more employees who would have otherwise separated with small raises. This discussion highlights the

24

importance of understanding the forces driving worker separations. Is the average pace of wage growth that workers should expect from a firm unrelated to its productivity? Or do workers with unusually low wage growth opportunities separate from less productive firms as in the PVR model? Convincing answers to this question would seem to require instrumental variables that plausibly shift separations but not potential wage growth. We leave such investigations to future research.

9   Gender differences

Table 4 indicated that gender specific DWL models fit somewhat better than pooled models. We turn now to investigating gender differences in DWL model parameters and the influence of origin and destination effects on the evolution of the gender wage gap.

9.1   Do origin effects differ by gender?

Figure 7 examines the relationship between origin and destination effects and measured productivity in models fit separately by gender. The destination effects are normalized to zero separately by gender in the bottom vingtile of log value added per worker, while the origin effects are normalized so that $\lambda_N = 0$ for each gender. Panel a) of Figure 7 plots mean female destination effects against mean male destination effects by centiles of firm value added. A linear fit to these means yields a slope of 0.90, remarkably close to the slope of 0.89 reported by Card et al. (2015) in Portuguese data and the slope of 0.85 reported by Casarico and Lattanzio (2019) using the universe of Italian social security records. The finding of a slope less than one reflects the tendency for female destination effects to rise less rapidly with productivity than male destination effects.[15]

Panel b) of Figure 7 plots mean female origin effects against mean male origin effects by centiles of firm value added. A linear fit to these means yields a slope of only 0.75, suggesting gender differences in origin effects are somewhat more pronounced than differences in destination effects. When interpreted through the lens of the BF-PVR model, the estimated intercept of 0.02 indicates that firms must offer women somewhat higher wages to convince them to leave the least productive employers. The linear fit suggests this gender difference fades at the most productive employers: firms at the 95th percentile of value added are predicted to have nearly the same origin effects for women and men.

Also displayed in this panel are the male and female values of λ_U, which captures the premium for being hired from unemployment relative to being hired into one's first job. λ_U is estimated to be larger for women than men, which could either indicate that it is harder to poach women than men from unemployment or that female labor market entrants face a hiring disadvantage

---

[15]Appendix Figure A.3 reports the direct relationships of these gender specific effects with value added, which turn out to be somewhat nonlinear.

relative to their male peers. Because the level of λ_N is not identified, the DWL estimates cannot be used to adjudicate between these explanations. However, the fact that the estimated λ_U lies below our fitted regression line unambiguously reveals that the wage costs of job displacement to unemployment are unambiguously larger for women than men. The size of this gap is largest at the least productive firms where women face a penalty roughly 2 log points greater than men.

9.2 Evolution of the gender wage gap

The finding of systematic gender differences in both origin and destination effects raises the question of how these effects contribute to gender wage inequality. Figure 8 illustrates the evolution of the gender gap in hiring wages for workers that enter the labor market in 2005, the first year of our data. Because not all of these workers experience job transitions in each year, we adjust the gender gap in hiring wages in each year for the change in each group's worker effects relative to the base year of 2005. For reference, unadjusted mean hiring wages by gender are provided in Figure A.4 along with the mean wages of all employed workers, including those who are not new hires.

At labor market entry, the composition adjusted gender gap in hiring wages hovers around 20% and is almost entirely explained by the gap in destination effects. By construction, the gender gap in origin effects is zero in 2005 because all workers are new labor market entrants and λ_N has been normalized to zero for both genders. As the cohort ages into the labor market, the gender gap in hiring wages grows, with little commensurate change in the destination effects gap. Perhaps surprisingly, the origin effects gap grows slightly negative with experience, but the magnitude of this gap is negligible. By 2015, the composition adjusted gender gap in hiring wages has increased to a staggering 35% with essentially none of this increase explained by origin or destination effects.

Past work suggests the dynamics of the gender wage gap are especially pronounced among highly skilled workers (Bertrand et al., 2010). Panel b) of Figure 8 plots results for the subsample of workers that were ages 25-27 when entering the labor market in 2005. Although our data do not allow us to directly measure education, the late entry of these workers to the labor force is likely due to educational delays. Late entry also puts these workers at prime fertility ages over the first ten years of their careers, a factor which recent research suggests is an important mediator of gender wage gaps (Kleven et al., 2019).[16] To illustrate these lifecycle effects, we plot the composition adjusted gaps for these workers by age at hiring. Upon entry, these workers exhibit a relatively small composition adjusted gender gap in hiring wages of roughly 12%, which is again almost entirely explained by destination effects. But as this cohort ages into the labor market, the gender gap in hiring wages explodes, reaching 40% by age 35. During this period, the gap in destination effects rises to 17%, while the gap in origin effects remains negligible. Hence, destination effects explain about 18% of the growth in hiring wage inequality for this cohort, while origin effects

---

[16]In 2018, the average age of an Italian woman giving birth to her first child was 31 (Istat, 2018).

explain none of the growth.

We conclude that women tend not to face a quantitatively important disadvantage in terms of where they are hired from. Rather, the gender gap in hiring wages is attributable in part to differences in where they are currently employed, differences that emerge early on. In later years the gender gap expands for reasons that likely have to do with childbearing and career interruption, rather than job ladders.

10   Conclusion

Sequential auction models provide a coherent and influential framework for interpreting wage dynamics in matched employer-employee data. The results of this paper demonstrate the potential for unrestricted fixed effects estimators of the sort pioneered by AKM to assist in evaluating semiparametric formulations of these models. A key finding of our analysis has been that the immense variation in destination firm effects relative to origin firm effects is difficult to rationalize with standard sequential auction models where firms are differentiated only by productivity. Whether sequential auction models featuring multi-dimensional firm heterogeneity can match the covariance structure of origin and destination firm effects is an interesting question for future research.

Our focus on hiring wages was motivated primarily as a means of circumventing endogeneity problems that arise in the study of within match dynamics. Surprisingly little is known about how the parameters governing hiring wages relate to those governing the wage growth of incumbent workers. In their original contribution, AKM (briefly) considered wage growth models allowing for firm specific tenure profiles (see also Margolis, 1996). How such tenure profile parameters relate to origin and destination effects in hiring wages awaits further study. Investigation of this relationship could be particularly helpful for better understanding the role of firm heterogeneity in mediating the earnings effects of job displacement (Lachowska et al., 2018; Schmieder et al., 2018).

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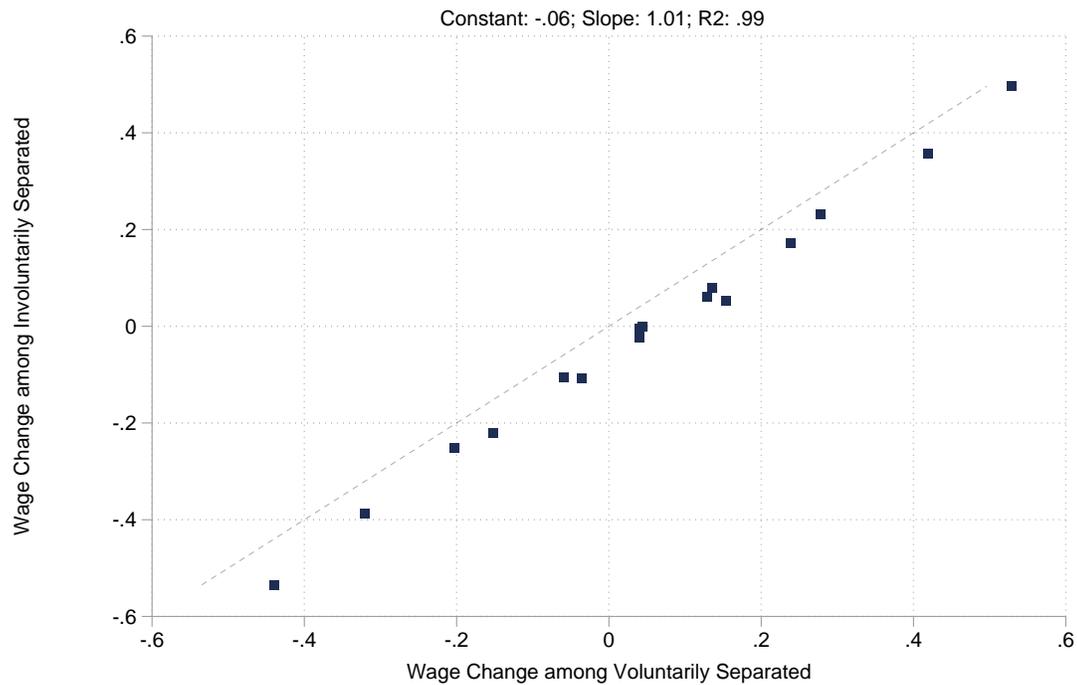
Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Achlioptas, D. (2003). Database-friendly random projections: Johnson-lindenstrauss with binary coins. *Journal of computer and System Sciences 66*(4), 671–687.

Ahn, H. and J. L. Powell (1993). Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *Journal of Econometrics 58*(1-2), 3–29.

Akcigit, U., S. Baslandze, and F. Lotti (2018). Connecting to power: political connections, innovation, and firm dynamics. Technical report, National Bureau of Economic Research.

Assolombarda (2012). Il lavoro nelle imprese dell'area milanese (milano, monza e brianza e lodi) - anno 2011. *Studi e Analisi*.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin (2014). Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review 104*(6), 1551–96.

Bagger, J. and R. Lentz (2019). An empirical model of wage dispersion with sorting. *The Review of Economic Studies 86*(1), 153–190.

Bank of Italy, I. (2018). Survey on industrial and service firms. Technical report.

Bertrand, M., C. Goldin, and L. F. Katz (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American economic journal: applied economics 2*(3), 228–55.

Bonhomme, S., T. Lamadon, and E. Manresa (2019). A distributional framework for matched employer employee data. *Econometrica 87*(3), 699–739.

Brenčič, V. (2012). Wage posting: evidence from job ads. *Canadian Journal of Economics/Revue canadienne d'économique 45*(4), 1529–1559.

Brenzel, H., H. Gartner, and C. Schnabel (2014). Wage bargaining or wage posting? evidence from the employers' side. *Labour Economics 29*, 41–48.

Breza, E., S. Kaur, and Y. Shamdasani (2018). The morale effects of pay inequality. *The Quarterly Journal of Economics 133*(2), 611–663.

Burdett, K. and D. T. Mortensen (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, 257–273.

Cahuc, P., O. Charlot, and F. Malherbet (2016). Explaining the spread of temporary jobs and its impact on labor turnover. *International Economic Review 57*(2), 533–572.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica 74*(2), 323–364.

Caldwell, S. and N. Harmon (2019). Outside options, bargaining, and wages: Evidence from coworker networks. *Unpublished manuscript, Univ. Copenhagen*.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics 36*(S1), S13–S70.

Card, D., A. R. Cardoso, and P. Kline (2015). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics 131*(2), 633–686.

Card, D., J. Heining, and P. Kline (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics 128*(3), 967–1015.

Card, D., A. Mas, E. Moretti, and E. Saez (2012). Inequality at work: The effect of peer salaries on job satisfaction. *American Economic Review 102*(6), 2981–3003.

Casarico, A. and S. Lattanzio (2019). What firms do: Gender inequality in linked employer-employee data.

Cullen, Z. B. and B. Pakzad-Hurson (2019). Equilibrium effects of pay transparency in a simple labor market. In *EC*, pp. 193.

Daruich, D., S. Di Addario, and R. Saggio (2020). The effects of partial employment protection reforms: Evidence from italy. Technical report, Mimeo.

Devicienti, F., B. Fanfani, and A. Maida (2019). Collective bargaining and the evolution of wage inequality in italy. *British Journal of Industrial Relations 57*(2), 377–407.

Dey, M. S. and C. J. Flinn (2005). An equilibrium model of health insurance provision and wage determination. *Econometrica 73*(2), 571–627.

Flinn, C., J. Mabli, J. Mullins, and O. London (2017). Firms' choices of wage-setting protocols in the presence of minimum wages.

Guiso, L., L. Pistaferri, and F. Schivardi (2005). Insurance within the firm. *Journal of Political Economy 113*(5), 1054–1087.

Hall, R. E. and A. B. Krueger (2012). Evidence on the incidence of wage posting, wage bargaining, and on-the-job search. *American Economic Journal: Macroeconomics 4*(4), 56–67.

Honoré, B. E. and J. L. Powell (1994). Pairwise difference estimators of censored and truncated regression models. *Journal of Econometrics 64*(1-2), 241–278.

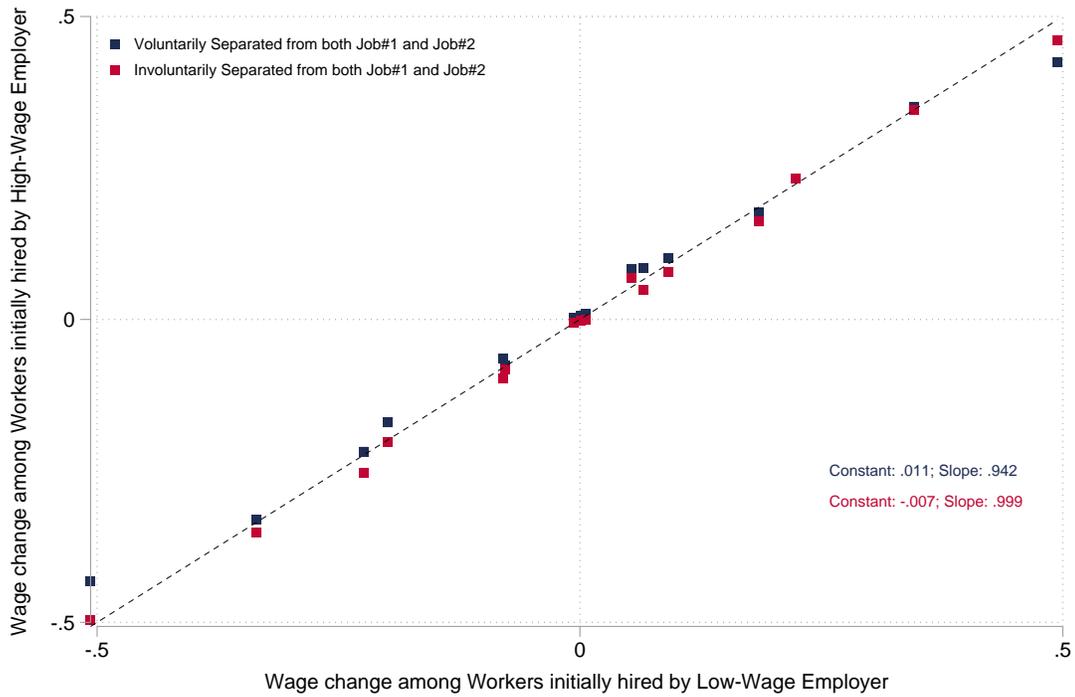Istat (2018). Natalità e fecondità della popolazione residente.

Jäger, S., B. Schoefer, S. G. Young, and J. Zweimüller (2018). Wages and the value of nonemployment. Technical report, National Bureau of Economic Research.

Johnson, W. B. and J. Lindenstrauss (1984). Extensions of lipschitz mappings into a hilbert space. *Contemporary mathematics 26*(189-206), 1.

Kleven, H., C. Landais, and J. E. Sogaard (2019). Children and gender inequality: Evidence from denmark. *American Economic Journal: Applied Economics 11*(4), 181–209.

Kline, P., R. Saggio, and M. Sølvsten (2019). Leave-out estimation of variance components. Technical report, National Bureau of Economic Research.

Kline, P., R. Saggio, and M. Sølvsten (2020). Leave-out estimation of variance components. *Econometrica, Forthcoming*.

Lachowska, M., A. Mas, and S. A. Woodbury (2018). Sources of displaced workers' long-term earnings losses. Technical report, National Bureau of Economic Research.

Macis, M. and F. Schivardi (2016). Exports and wages: Rent sharing, workforce composition, or returns to skills? *Journal of Labor Economics 34*(4), 945–978.

Margolis, D. N. (1996). Cohort effects and returns to seniority in france. *Annales d'Economie et de Statistique*, 443–464.

Mas, A. (2017). Does transparency lead to pay compression? *Journal of Political Economy 125*(5), 1683–1721.

Papp, T. K. (2013). Frictional wage dispersion with bertrand competition: An assessment. *Review of Economic Dynamics 16*(3), 540–552.

Postel-Vinay, F. and J.-M. Robin (2002a). The distribution of earnings in an equilibrium search model with state-dependent offers and counteroffers. *International Economic Review 43*(4), 989–1016.

Postel-Vinay, F. and J.-M. Robin (2002b). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica 70*(6), 2295–2350.

Schmieder, J., T. von Wachter, and J. Heining (2018). The costs of job displacement over the business cycle and its sources: evidence from germany. Technical report, UCLA, Mimeo.

Song, J., D. J. Price, F. Guvenen, N. Bloom, and T. Von Wachter (2019). Firming up inequality. *The Quarterly journal of economics 134*(1), 1–50.

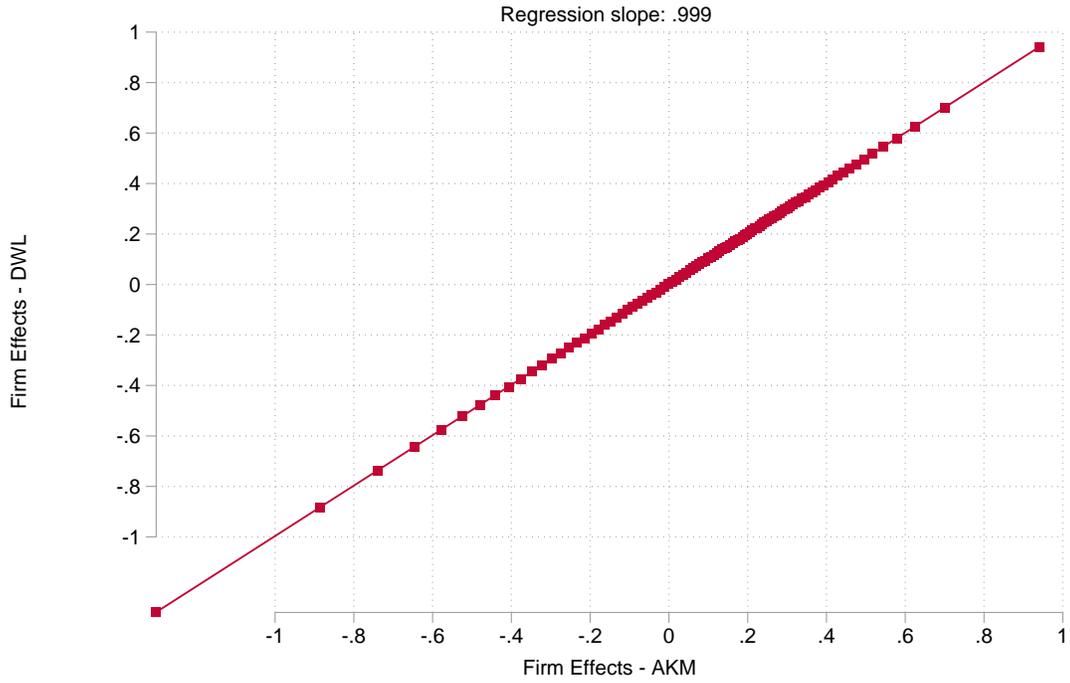## Figure 1: The Penalty of Being Hired from Unemployment

Note: Each dot represents the adjusted log hiring wage change from job#1 to job#2 for different combinations of origin/destination quartiles of mean-coworkers wages. These dots are computed for two groups of workers. The first group (x-axis) corresponds to workers that voluntarily quit their first job. The second group (y-axis) corresponds to workers that were involuntarily separated from their first job.
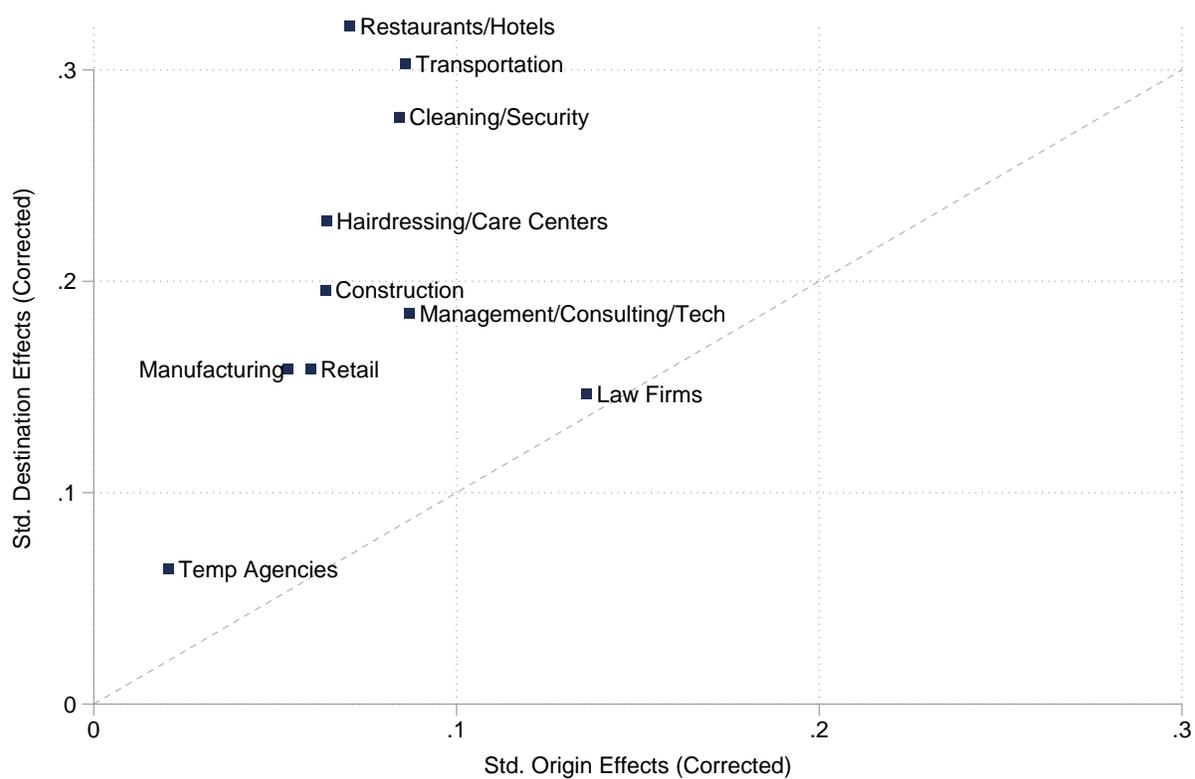
Figure 2: Does it matter who fires you?

Note: Each dot represents the adjusted log hiring wage change from job#2 to job#3 for different combinations of origin/destination quartiles of mean-coworkers wages. These dots are computed for three groups of workers. The first group (x-axis) corresponds to workers who involuntarily separated from both job#1 and job#2 and that had a low-wage employer in their first job. The second group (y-axis) corresponds to workers that were also involuntarily separated from both job#1 and job#2 but that had a high-wage employer in their first job. The last group corresponds to workers that were voluntarily separated from both job#1 and job#2 and that had a high-quality employer in their first job. The quality of the employer is based on terciles of the co-workers' wage distribution (low wage = first tercile, high wage = last tercile). We report the associated constant and slope associated for both scatter plots.

Figure 3: AKM firm effects vs. DWL firm effects

Note: For each firm we have an estimated firm effect according to either the AKM model or the DWL model. We then take centiles of the firm effects estimated from the AKM model. Within each centile of the AKM effects, we average the AKM effects and the corresponding DWL destination effects. The figure then shows these two means and we report the corresponding regression slope obtained from the micro-level regression. Both set of effects have been normalized to have mean zero in the lowest vingtile of the firm-size weighted distribution of mean value added per worker.
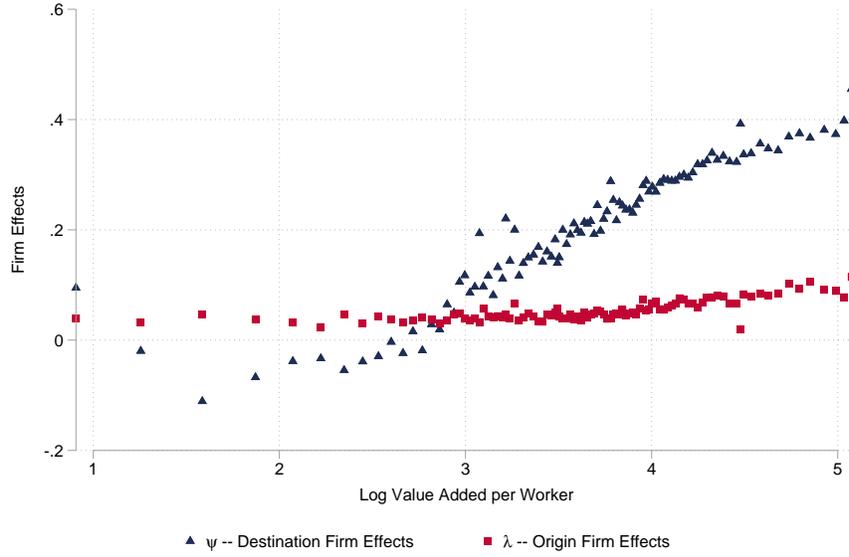
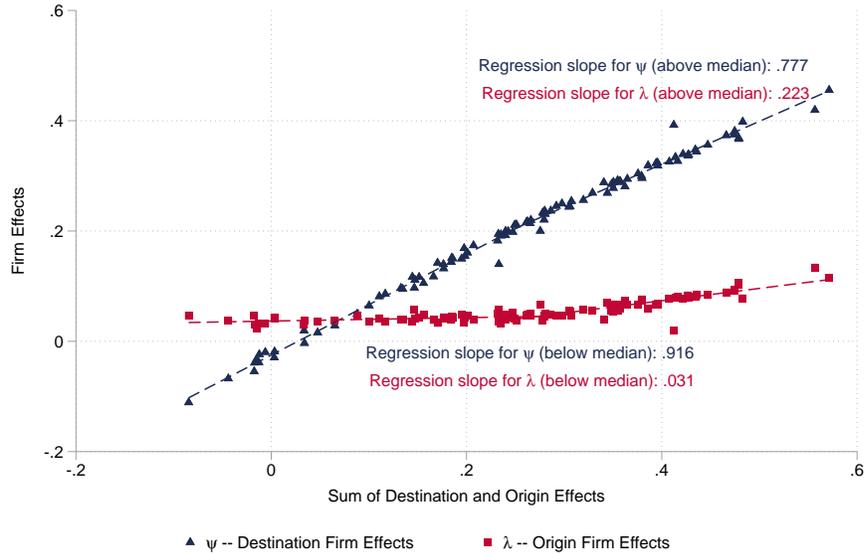Figure 4: Variability of Origin and Destination Effects by Sector

Note: This figure reports leave-out corrected standard deviations of destination and origin firm effects for selected sectors of the Italian economy (2-Digit 2007 Ateco codes). All variance components are firm-size weighted. The dashed line is the 45 degree line.

Figure 5: Origin and Destination Effects by Value Added
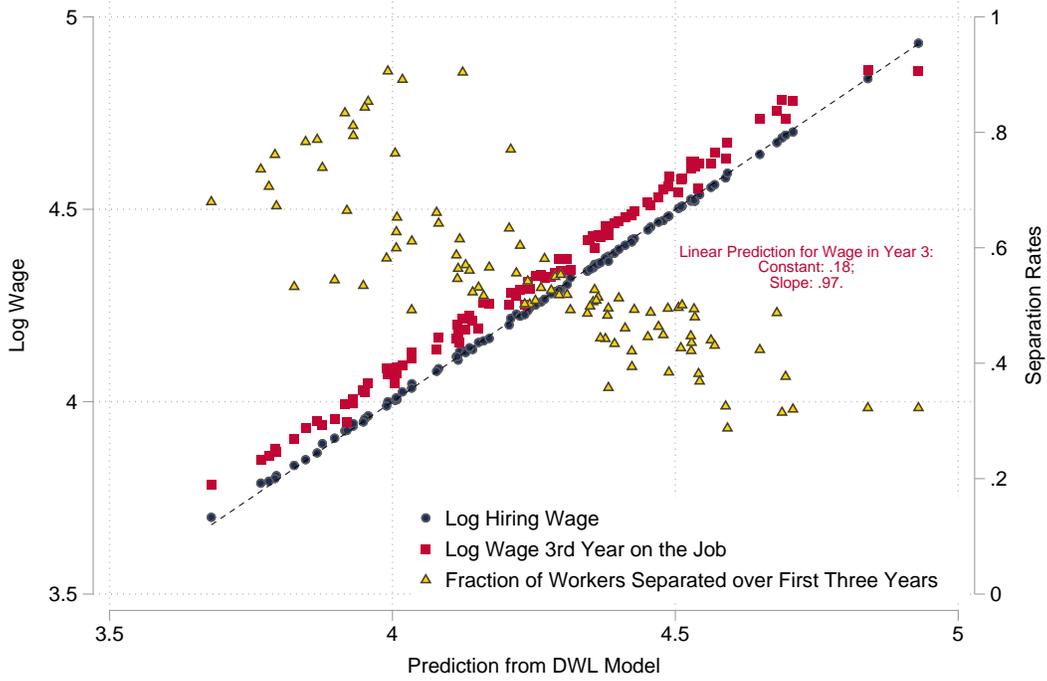
(a) Value Added per Worker

▲ ψ -- Destination Firm Effects   ■ λ -- Origin Firm Effects

(b) Sum of the Effects

Regression slope for ψ (above median): .777
Regression slope for λ (above median): .223

Regression slope for ψ (below median): .916
Regression slope for λ (below median): .031

▲ ψ -- Destination Firm Effects   ■ λ -- Origin Firm Effects

Note: In Panel (a), we start by computing log value added per worker averaged over the sample period and take firm-size weighted centiles of this measure. Within each centile, we average log value added per worker (x-axis), destination effects (the ψ_j's) and origin effects (the λ_j's). Origin effects have been normalized relative to λ_N. Destination effects have been normalized to have mean zero in the lowest vingtile of the firm-size weighted distribution of mean value added per worker. Panel (b) proceeds in a similar way but reports in the x-axis average of $(\psi_j + \lambda_j)$ within each centile of value added per worker. The reported slopes refer to the regression coefficient obtained in the binned data when fitting this regression to above/below median centiles of the value added per worker distribution.
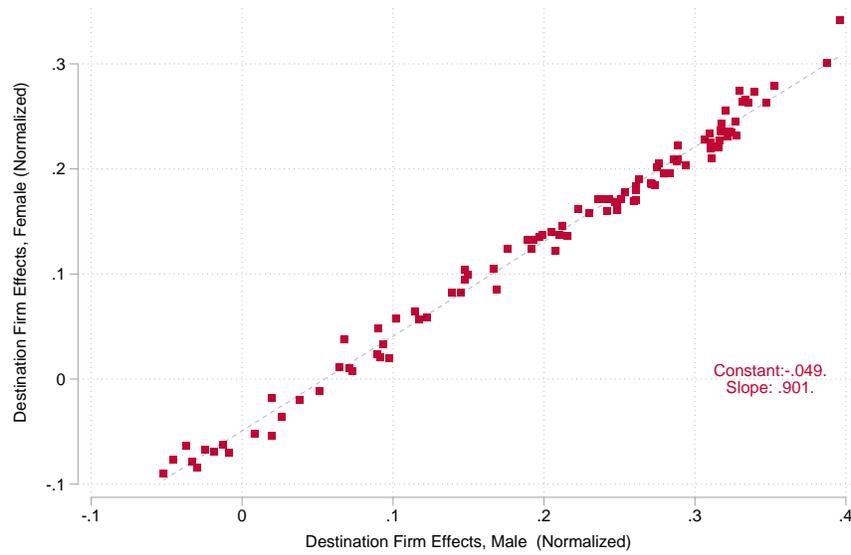
Figure 6: Fit to Hiring and Incumbent Wages

Linear Prediction for Wage in Year 3:
Constant: .18;
Slope: .97.

• Log Hiring Wage
■ Log Wage 3rd Year on the Job
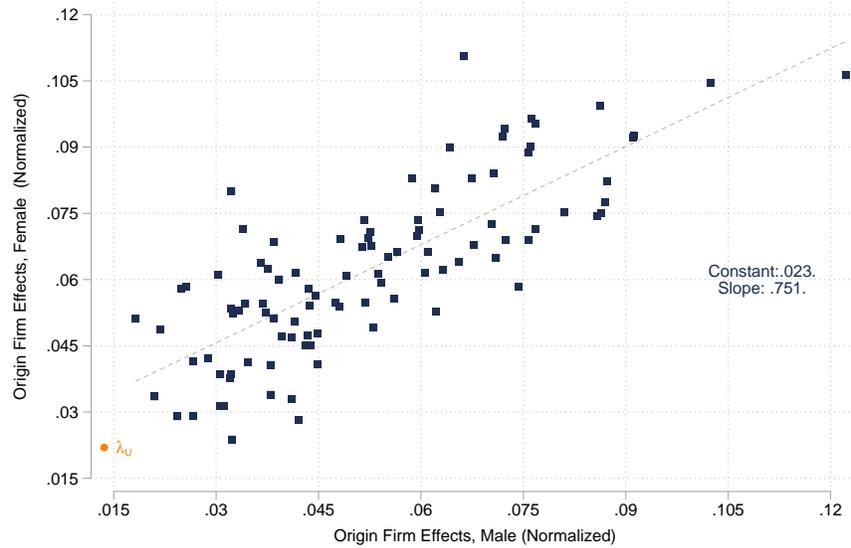▲ Fraction of Workers Separated over First Three Years

Note: The dashed line is the 45 degree line. For each job-spell in our starting sample that lasted at least three years, we report in the x-axis the mean prediction $Z'_\ell \hat{\gamma}$ of the DWL model in each value added per worker centile. The triangles give the fraction of workers hired into each value added centile who separate from the job within three years. The circles plotted report the mean hiring wage of the subpopulation of workers who remain employed at their destination firm for three or more years. The squares depict the mean wage of incumbent workers in their third year on the job. We report also the results of a linear fit estimated on the binned data of wages in the third year on the job onto $Z'_\ell \hat{\gamma}$.

Figure 7: Origin and Destination Effects by Gender and Value Added

(a) Destination Effects

Constant:-.049.
Slope: .901.

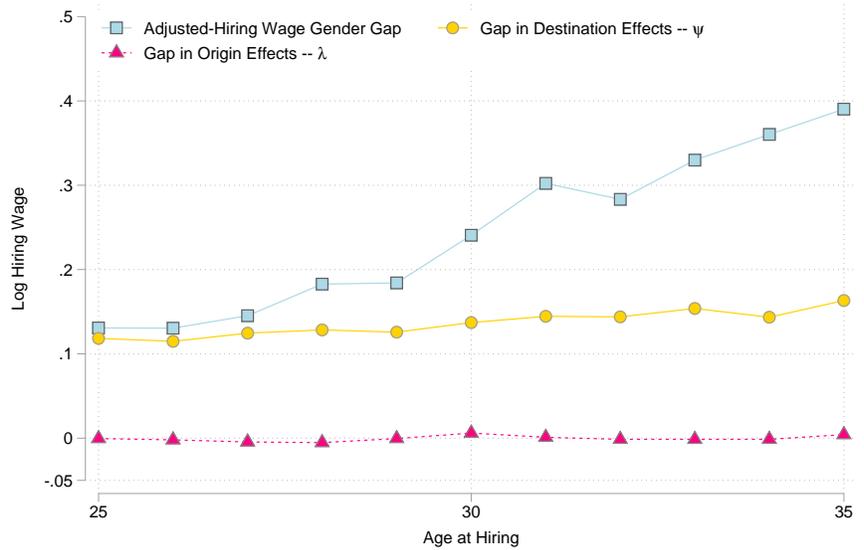### (b) Origin Effects

Constant:.023.
Slope: .751.

Note: This figure presents bin scatter plots of estimated origin and destination effects for female workers against estimated origin or destination effects for male workers. The estimated effects are grouped into 100 percentile bins based on mean log value added per worker at the firm. Estimated slope is estimated across percentile bins. Origin effects for each gender have been normalized relative λ_N^G. Destination effects have been normalized to have mean zero in the lowest vingtile of the firm-size weighted distribution of mean value added per worker specific to each gender firm effect.

Figure 8: Gender Wage Gap and the DWL Model

(a) Entered Labor Market in 2005

(b) Entered Labor Market in 2005 at Age 25-27

Note: Panel (a) focuses on individuals that entered the labor market in the year 2005, the first year of our data. Panel (b) focuses on individuals entering the labor market in 2005 for the first time and that were aged between 25-27 years at the moment of entry. In each panel, we plot the adjusted log hiring wages between men and women and the corresponding difference in average $(\psi_j^M - \psi_j^W)$ and $(\lambda j^M - \lambda_j^W)$. We adjust the gender gap in hiring wages in each year for the change in each group's worker effects relative to the base year of 2005.

Table 1: Summary Statistics

	Pooled	Men	Women
Panel (a): Starting Sample			
Number of Person-Job Observations	13,029,554	7,840,247	5,189,307
Number of Individuals	4,895,253	2,936,275	1,958,978
Share hired from unemployment	0.59	0.58	0.60
Share poached from another firm	0.31	0.33	0.29
Share new entrants	0.10	0.09	0.11
Number of origin fixed effects	876,395	623,478	432,317
Number of destination firm effects	1,493,788	1,070,614	836,018
Mean Log Hiring Wages	4.0826	4.2044	3.8986
Variance Log Hiring Wages	0.2939	0.2427	0.3151
Panel (b): Estimation Sample			
Number of Person-Job Observations	10,100,836	5,860,789	3,730,985
Number of Individuals	3,194,370	1,849,723	1,224,858
Share hired from unemployment	0.61	0.60	0.63
Share poached from another firm	0.28	0.29	0.24
Share new entrants	0.12	0.11	0.13
Number of origin fixed effects	328,377	223,156	111,606
Number of destination firm effects	701,459	477,923	295,890
Mean Log Hiring Wages	4.0753	4.1978	3.9001
Variance Log Hiring Wages	0.2794	0.2215	0.3162

Note: This table reports summary statistics for our analysis on three different samples: one sample formed only by female workers, one sample formed only by male workers and one sample that pools together workers of both genders. Each starting sample comprises all person-job observations contained in INPS-INVIND from 2005-2015 for individuals that held two or more jobs over this interval. Our estimation sample is represented by person-job observations where the associated statistical leverage is below one and for which we are able to identify both contemporaneous and lagged firm effects. See text for details. All statistics are person-job weighted.

Table 2: Goodness of Fit

	Pooled	Men	Women
AKM	0.7199	0.7311	0.6822
AKM (Gender-Interacted)	0.7349		
Origin Effects	0.5809	0.5660	0.5452
Origin Effects (Gender-Interacted)	0.5871		
DWL	0.7245	0.7370	0.6854
DWL (Gender-Interacted)	0.7427		

Note: This table presents the goodness of fit (R2) from various models for the three estimation samples described in Table 1. The model labeled as "Origin effects" corresponds to a DWL model with only origin effects and no destination effects. "DWL (Gender-interacted)" corresponds to a model where both contemporaneous and origin firm effects are interacted with a gender indicator. "AKM (Gender-Interacted)" interacts gender with destination firm effects while "Origin Effects (Gender-Interacted)" interacts gender with origin effects. All reported measures of the goodness fit computed using the leave-out bias correction of Kline, Saggio and Sølvsten (2020). See text for further details.

Table 3: Variance Decomposition of Poaching Wages - AKM Model

	Pooled	Men	Women
Std Dev of Log Hiring Wages	0.5286	0.4706	0.5623
Bias-Corrected Variance Components			
Std Dev of worker effects	0.2887	0.2558	0.2854
Std Dev of firm effects	0.2578	0.2431	0.2824
Correlation of worker, firm effects	0.3135	0.2311	0.3461
Percent of Total Variance Explained by			
Worker effects	29.83%	29.54%	25.77%
Firm effects	23.78%	26.68%	25.22%
Covariance of worker, firm effects	16.70%	12.98%	17.64%
$X'\delta$ and associated covariances	1.69%	3.91%	-0.41%
Residual	28.01%	26.89%	31.78%

Note: This table reports the variance decomposition after fitting an AKM model to hiring wages only using the estimation sample defined in Table 1, Panel (b). Corrected variance components are calculated using the leave out methodology of KSS (leaving a person-job out). AKM model controls for a cubic in age at hiring and year of hiring fixed effects.

Table 4: Variance Decomposition across Person-Job Observations --- DWL Model

	Pooled	Men	Women
Std Dev of log hiring wages	0.5286	0.4706	0.5623
Mean origin effect among involuntarily separated	0.0556	0.0536	0.0687
Mean origin effect among voluntarily separated	0.0561	0.0543	0.0690
Origin effect when hired from unemployment (λ_U)	0.0163	0.0136	0.0220
Bias-Corrected Variance Components			
Std Dev of worker effects	0.2823	0.2479	0.2798
Std Dev of destination firm effects	0.2580	0.2434	0.2828
Std Dev of origin effects	0.0439	0.0454	0.0431
Std Dev of origin effects (among poached workers)	0.0761	0.0782	0.0798
Correlation of worker, destination firm effects	0.3157	0.2351	0.3441
Correlation of worker, origin effects	0.1200	0.1629	0.0757
Correlation of destination firm, origin effects	0.0316	0.0308	0.0000
Percent of Total Variance Explained by			
Worker effects	28.52%	27.75%	24.77%
Destination firm effects	23.81%	26.74%	25.29%
Origin effects	0.69%	0.93%	0.59%
Covariance of worker, destination	16.46%	12.81%	17.23%
Covariance of worker, origin	1.06%	1.66%	0.58%
Covariance of destination, origin	0.26%	0.31%	0.00%
X'δ and associated covariances	1.66%	3.51%	0.09%
Residual	27.55%	26.30%	31.46%

Note: This table reports the variance decomposition based upon the DWL model across person-job observations. We also report the (firm-size weighted) corresponding average of the origin effects for individuals that were involuntary separated as well as the estimated origin effect when hired from unemployment. All origin effects are normalized relative to the origin effect in the first job, i.e. λ_N within each sample. Variance components are corrected using the leave-out bias correction of Kline, Saggio and Sølvsten (2020).

Table 5: Variance Decomposition across Firms

	Pooled	Men	Women
# of firms with identified destination and origin effect	297,865	201,080	99,508
Bias-Corrected Variance Components			
Std of Destination Effects	0.2590	0.2449	0.2724
Std of Origin Effects	0.0707	0.0721	0.0510
Correlation of destination, origin	0.2511	0.2491	0.3168
Std of Destination + Origin Effects	0.2851	0.2720	0.2926
Lower Bound on Bargaining Power	0.8819	0.8703	0.9182
Lower Bound on Correlation of Destination, Origin Effects	0.8409	0.8288	0.8824

Note: Here we report the variance decomposition across firms where each firm has an identified origin and destination firm effect. Variance components are weighted by average firm-size over 2005-2015 as recorded by official INPS records collected in the dataset _Anagrafica_, see text for details. Variance components corrected using the leave-out bias correction of Kline, Saggio and Sølvsten (2020). The lower bounds on the barganing power and correlation of destination and origin firm effects are based upon equation (5)-(6), see text for details.

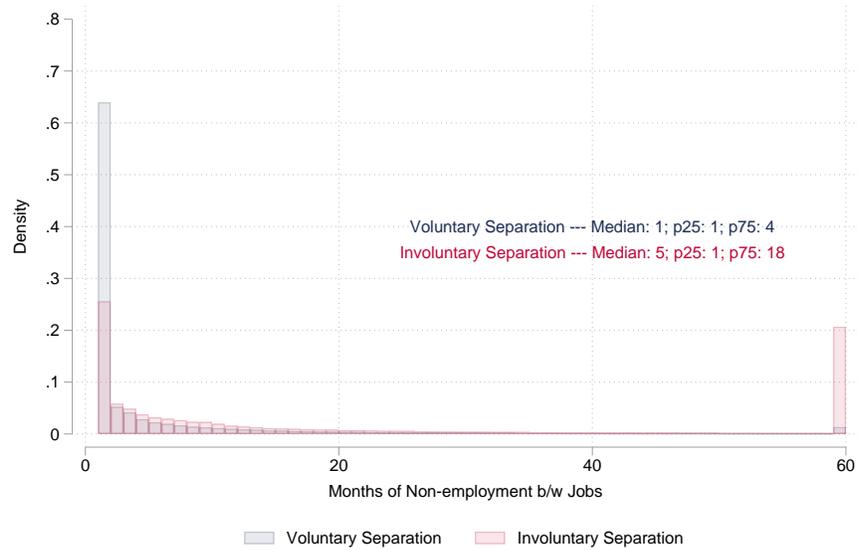**Table 6:** Variability of Origin and Destination Effects by Sector

	SD of Destination Effects	SD of Origin Effects	Correlation of Origin, Destination Effects	Lower Bound on Bargaining Power	Lower Bound on Correlation
Retail	0.1585	0.0597	0.2260	0.8269	0.7868
Construction	0.1959	0.0639	-0.0693	0.9211	0.8786
Restaurants / Hotels	0.3206	0.0706	0.0675	0.9413	0.9018
Hairdressing / Care Centers	0.2284	0.0641	0.1399	0.8979	0.8567
Law Firms	0.1468	0.1359	0.0399	0.5369	0.5758
Manufacturing	0.1585	0.0536	0.2737	0.8409	0.7992
Transportation	0.3028	0.0859	-0.0632	0.9401	0.8969
Cleaning / Security	0.2777	0.0842	0.0874	0.8966	0.8551
Temp Agencies	0.0639	0.0206	0.1651	0.8702	0.8291
Management / Consulting / Tech	0.1847	0.0870	0.4190	0.7406	0.6991

Note: This table reports leave-out corrected standard deviations of destination and origin firm effects within selected sectors of the Italian economy (2-Digit 2007 Ateco codes). All variance components are firm-size weighted. The lower bounds on the barganing power and correlation of destination and origin firm effects are based upon equation (5)-(6), see text for details.
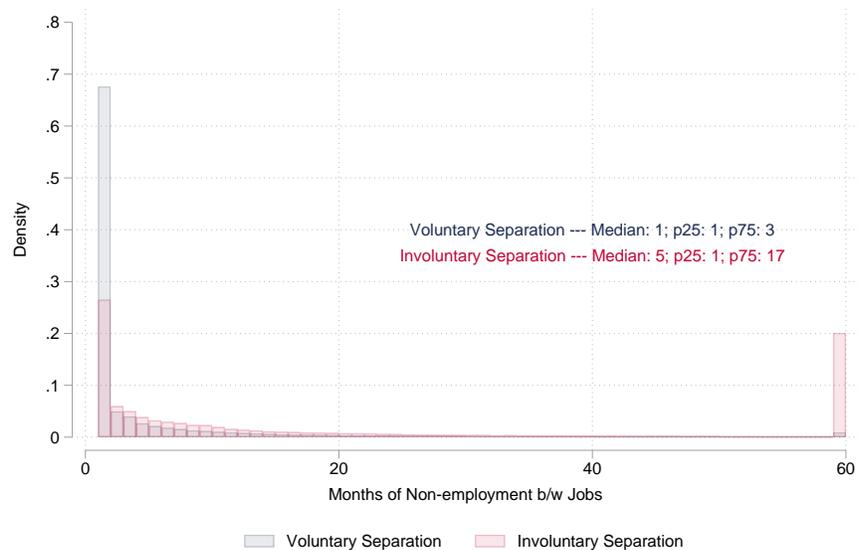
Appendix A   Additional results

Figure A.1: Months of Non-Employment Between Jobs
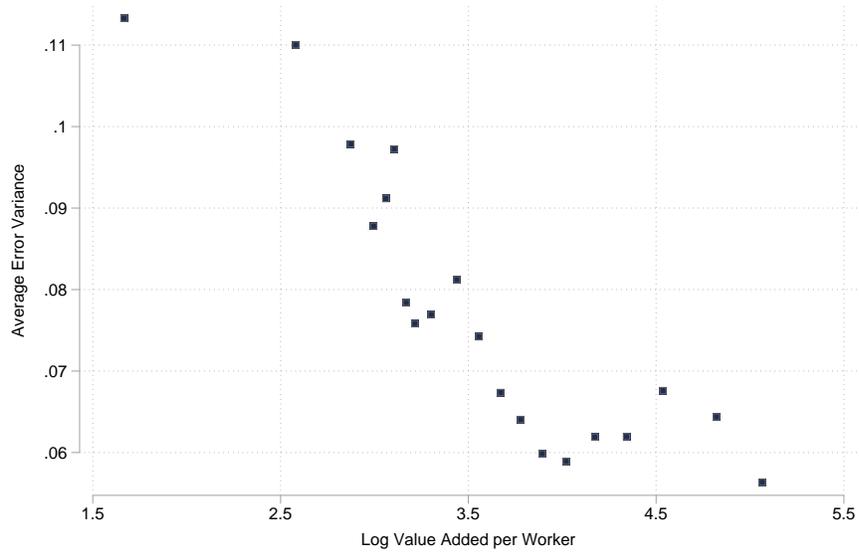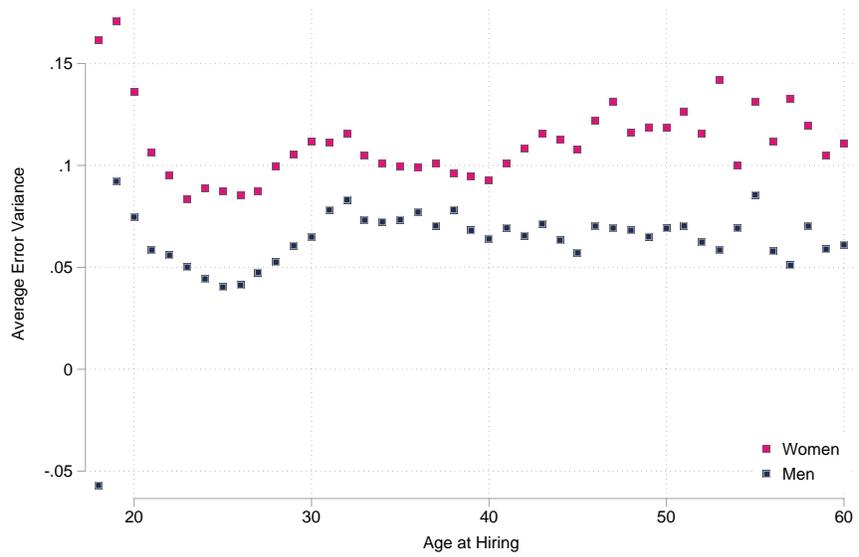
(a) Starting Sample

(b) Estimation Sample

Note: This figure provides the histogram of the months of non-employment spent between jobs in our starting sample (Panel a) and estimation sample (Panel b), see Table 1. The histogram is computed for two groups: those individuals that according to social security records voluntarily separated from the job vs. those that instead were involuntarily separated. Months of non-employment have been winsorized at 60 months.

Figure A.2: Heteroskedasticity in the DWL Model
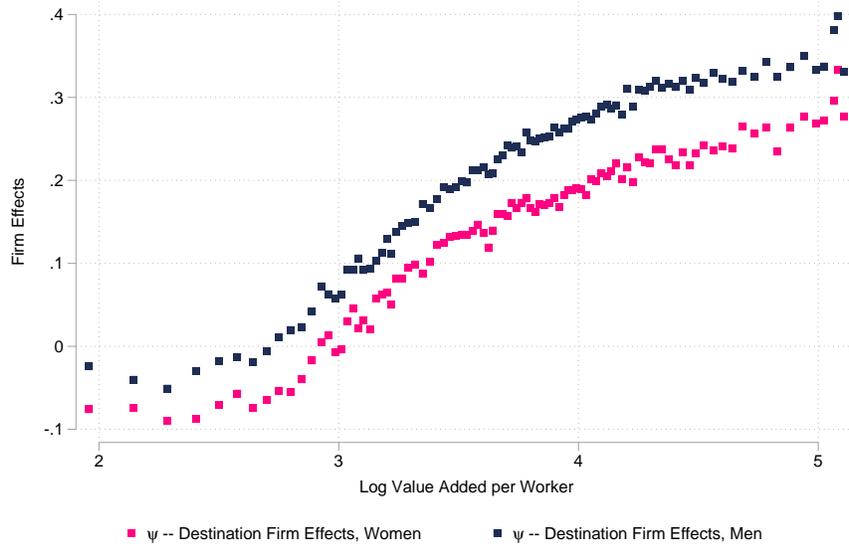
(a) Log Value Added per Worker
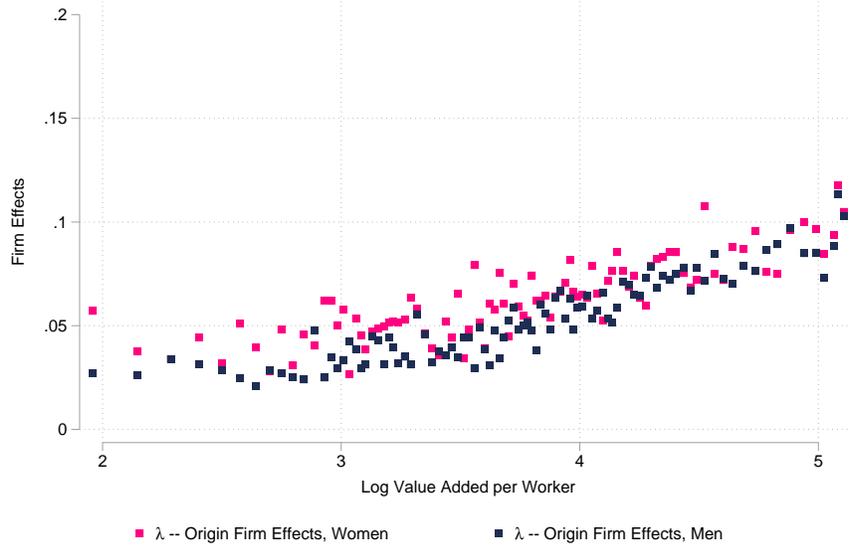

(b) Age and Gender

Note: Panel (a) displays average of $\hat{\sigma}_\ell^2$ defined in (8) by 20 bins of log value added per worker. Panel (b) reports average $\hat{\sigma}_\ell^2$ for different age at hiring and gender.

46

Figure A.3: Origin, Destination, Firms' Characteristics and Gender

(a) Destination Effects

ψ -- Destination Firm Effects, Women   ψ -- Destination Firm Effects, Men

(b) Origin Effects
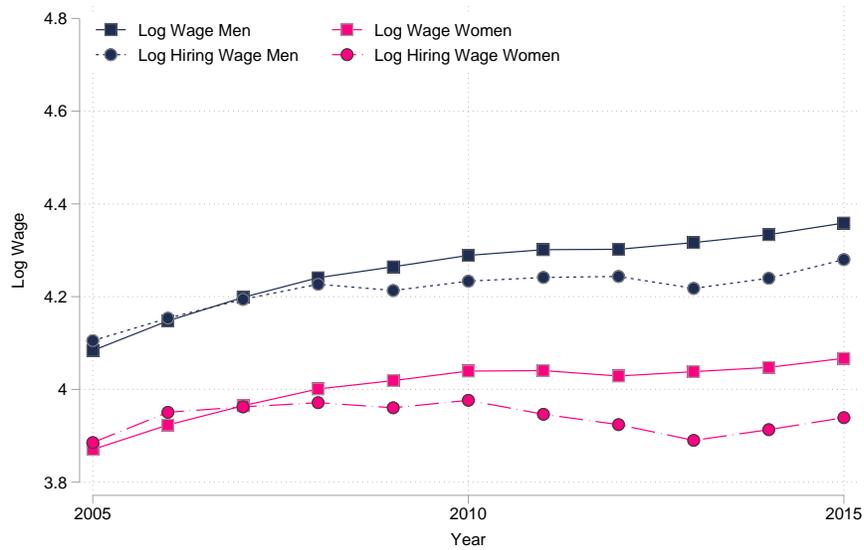

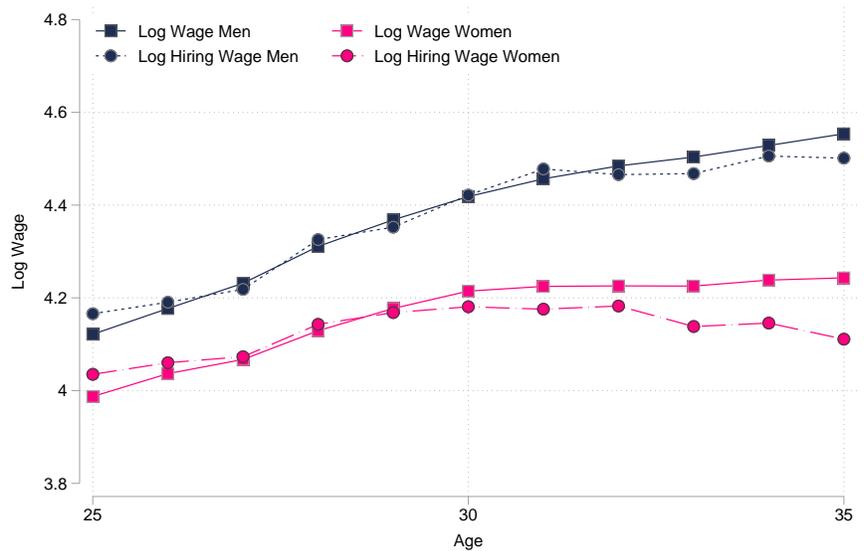λ -- Origin Firm Effects, Women   λ -- Origin Firm Effects, Men

Note: We start by considering firms that have origin and destination effects that are identified in both the female only and men only sample, i.e. $(\lambda_j^W, \psi_j^W, \lambda_j^M, \psi_j^M)$ are all identified, where (λ_j^G, ψ_j^G) correspond to gender specific $G \in \{W; M\}$ origin and destination effect according to the DWL model. Panel (a) and Panel (b) are then constructed by taking averages of log value added per worker and (ψ_j^W, ψ_j^M) (Panel (a)) or averages of log value added per worker and $(\lambda_j^W, \lambda_j^M)$ (Panel (b)). Origin effects for each gender have been normalized relative λ_N^G. Destination effects have been normalized to have mean zero in the lowest vingtile of the firm-size weighted distribution of mean value added per worker specific to each gender effect. First three centiles have been trimmed from both panels.

Figure A.4: Gender Gap in Wages and Hiring Wages

(a) Entered Labor Market in 2005

(b) Entered Labor Market in 2005, Birth Cohorts: 1978-1980

Note: "Log Wage" displays log real daily wages for men and women in their primary job across year (Panel a) or across the age profile (Panel b). "Log Hiring Wage" displays the *hiring* wage for jobs for individuals hired in a given year (Panel a) or hired at a particular age (Panel b). Panel (a) is computed only on the subpopulation of individuals that entered the labor market in 2005. Panel (b) is computed only on the subpopulation of individuals that entered the labor market in 2005 and were born between 1978-1980.

Table A1: Firm Characteristics across Samples

	Estimation Sample for DWL	Universe INPS-INVIND	Universe INPS
Number of firms	701,459	1,870,558	3,390,563
Summary Statistics on Firm Size			
Mean Log Firm Size	2.19	1.46	1.00
	(1.09)	(1.08)	(1.02)
Median Firm Size	8.00	3.75	2.00
Sector Breakdown (%)			
Retail	11.28	12.52	13.66
Construction	7.33	9.28	10.25
Restaurants / Hotels	9.19	9.93	9.98
Hairdressing / Care Centers	2.38	2.44	2.77
Law Firms	0.34	0.69	1.01
Manufacturing	5.26	4.84	4.43
Transportation	4.23	3.87	3.58
Cleaning / Security	5.74	4.89	4.46
Temp Agencies	2.09	1.62	1.42
Management / Consulting / Tech	2.06	2.00	1.94

Note: This table compares summary statistics for firms in three different samples over the interval 2005-2015. Firms in Column 1 correspond to the destination firms that are present in the pooled estimation sample described in Table 1, Panel B. Firms in Column 2 correspond the firms that we observe in the INPS-INVIND matched employer-employee data. Firm in Column3 correspond to the universe of firms observed in the Italian social security data as contained in dataset *Anagrafica* described in the text. Sectors correspond to 2-Digit ATECO (2007) codes and corresponding shares are firm-size weighted. Firm-size is calculated at the logarithm of mean firm size, where the mean is taken over the years in which the firm is active.

Table A2: Variance Decomposition in Poaching Wages - AKM Model

	Pooled	Men	Women
Estimation Sample			
Number of Job-Year Observations	10,067,164	5,839,976	3,714,261
Number of Individuals	3,173,400	1,838,010	1,215,720
Number of firms	696,815	474,529	292,883
Mean Log Hiring Wage	4.1361	4.2811	3.9285
Std Dev of Log Hiring Wages	0.5240	0.4611	0.5635
Variance Decomposition AKM Model - Uncorrected			
Std Dev of worker effects	0.3336	0.2952	0.3463
Std Dev of firm effects	0.2736	0.2591	0.3034
Correlation of Worker, Firm Effects	0.2213	0.1505	0.2156
Variance Decomposition AKM Model - Corrected			
Std Dev of worker effects	0.2886	0.2557	0.2853
Std Dev of firm effects	0.2578	0.2431	0.2824
Correlation of Worker, Firm Effects	0.3136	0.2316	0.3469
Std Dev of Firm Effects (Leaving Worker Out)	0.2561	0.2417	0.2798
Percent of Total Variance Explained by - Corrected			
Worker effects	30.33%	30.75%	25.65%
Firm effects	24.20%	27.79%	25.12%
Covariance of worker, firm effects	16.99%	13.54%	17.61%
X'δ and associated covariances	0.46%	1.03%	-0.16%
Residual	28.01%	26.89%	31.78%

Note: This table reports the variance decomposition based upon an AKM model applied to hiring wages only. The first panel reports summary statistics of the sample used in the analysis, i.e. the leave-worker-out connected set defined in Kline-Saggio-Sølvsten (2020 - KSS). The "uncorrected" panel reports variance components that are unadjusted for limited mobiliy biases. The "corrected" panel reports variance components corrected using the leave out methodology of KSS (leaving a job out). We also report the KSS-adjusted variance of firm effects when leaving the entire history of a worker out. See text for details.

Table A3: Comparing the Contribution of the Variance of Firm Effects

	DWL Estimation Sample	DWL Estimation Sample restricted to Dominant Jobs	Sample in Column (2) with Hiring and Within-Match Wages	Sample in Column (3) adding Firm-Stayers
	[1]	[2]	[3]	[4]
Summary Statistics on Leave-out-Sample				
Mean Log Wage	4.0753	4.0852	4.1765	4.3115
Std Dev of Log Wage	0.5286	0.5269	0.5443	0.5525
Number of Individuals	3,194,370	3,004,100	3,004,100	6,022,869
Number of firms	701,459	645,011	645,011	645,011
Number of observations	10,100,836	8,754,197	21,609,391	41,666,584
Contribution of Variance of Firm Effects according to AKM Model				
Std Dev of firm effects (Bias-Corrected)	0.2578	0.2555	0.2399	0.2217
Fraction of variance explained by firm effects	23.78%	23.52%	19.42%	16.10%

Note: This table summarizes how the contribution of firm effects varies across different estimation samples according to an AKM model. Sample in Column 1 corresponds to our pooled estimation sample described in Table 1, Panel (b). Our dependent variable there is therefore represented by hiring wages. In Column 2, we take our estimation sample of Column 1 but we restrict only to dominant jobs in the year. That is, we only retain person-job observations that correspond to the highest paying job of an individual in a particular year. Our dependent variable in Column 2 is still represented by hiring wages. In Column 3, we retain the worker-firm matches used in Column 2 but instead of looking at hiring wages we look at both hiring and within-match wages. Column 4 adds to the sample of Column 3 firm-stayers, i.e. individuals that remained always during the period 2005-2015 with one of the 645,011 employers characterizing the sample of Column 3. All summary statistics refer to the leave-out connected sample. All reported variance components are weighted by the number of observations present in each sample.

Table A4: Variance Decomposition (Uncorrected Estimates)

	Pooled	Men	Women
Std Dev of log poaching wages	0.5286	0.4706	0.5623
Variance Components			
Std Dev of worker effects	0.3316	0.2925	0.3464
Std Dev of destination firm effects	0.2759	0.2614	0.3066
Std Dev of origin effects	0.0782	0.0773	0.0866
Correlation of worker, destination firm effects	0.2087	0.1375	0.1981
Correlation of worker, origin effects	-0.0091	0.0046	-0.0455
Correlation of destination firm, origin effects	-0.0027	-0.0030	-0.0252
R2	0.8393	0.8533	0.8210
Adjusted R2	0.7238	0.7402	0.6818
Percent of Total Variance Explained by			
Worker effects	39.35%	38.63%	37.95%
Destination firm effects	27.24%	30.86%	29.74%
Origin effects	2.19%	2.70%	2.37%
Covariance of worker, destination	13.67%	9.50%	13.31%
Covariance of worker, origin	-0.17%	0.09%	-0.86%
Covariance of destination, origin	-0.04%	-0.06%	-0.42%
X'δ and associated covariances	1.70%	3.61%	0.01%
Residual	16.07%	14.68%	17.90%

Note: This table reports the variance decomposition based upon the DWL model as described in Table 4 of the main text. Variance components are not corrected and therefore are computed using a standard plug-in strategy.

Appendix B Shape constraints

Here, we establish the shape constraints on $\psi(\cdot)$ and $\lambda(\cdot)$ referenced in the main text and give a condition on the relationship between value added and p that ensures that these shape constraints are transferred to the conditional means reported in Figure 5, panel (b).

Lemma B.1. *Suppose that p is continuously distributed on an interval contained in the positive half-line that is bounded and bounded away from zero. Then,*

$$\frac{\partial I(z|\beta)}{\partial \ln z} = -\frac{(1-\beta)^2 \kappa \bar{F}(z)}{1 + \kappa \beta \bar{F}(z)} \in \left(-(1-\beta)^2/\beta, 0\right], \qquad \frac{\partial^2 I(z|\beta)}{\partial (\ln z)^2} = \frac{(1-\beta)^2 \kappa z f(z)}{\left(1 + \kappa \beta \bar{F}(z)\right)^2} \geq 0$$

where f is the density of p. Therefore, $I(p|\beta)$ is non-increasing and convex in $\ln p$, $\lambda(p)$ is increasing and concave in $\ln p$, and $\psi(p)$ is convex in $\ln p$ and increasing in $\ln p$ whenever $\beta \geq 1/2$.

Proof. For $z > 0$, a change of variables yields $I(z|\beta) = (1-\beta)^2 \kappa \int_{\ln z}^{\infty} \bar{G}(x)/\left(1 + \kappa \beta \bar{G}(x)\right) dx$ where $\bar{G}(x) = \bar{F}(\exp(x)) = \mathbb{P}(\ln p \geq x)$ is the survival function of $\ln p$. The Lemma follows by differentiation. $\qquad \square$

Lemma B.2. *Suppose that p is continuously distributed on an interval contained in the positive half-line that is bounded and bounded away from zero. Then,*

$$\beta \geq \frac{1}{2} + \frac{\mathbb{V}[\psi(p)] - \mathbb{V}[\lambda(p)]}{2\mathbb{V}[\psi(p) + \lambda(p)]}. \tag{B.1}$$

Furthermore, if $\beta \geq 1/2$, then

$$\left(1 + \sqrt{\frac{\mathbb{V}[\lambda(p)]}{\mathbb{V}[\psi(p) + \lambda(p)]}}\right)^{-1} \geq \beta \geq \frac{1}{4} + \sqrt{\frac{1}{4^2} + \frac{\mathbb{V}[\psi(p)] - \mathbb{C}[\psi(p), \lambda(p)]}{2\mathbb{V}[\psi(p) + \lambda(p)]}} \tag{B.2}$$

which implies that

$$\rho(\psi(p), \lambda(p)) \geq \sqrt{\frac{\mathbb{V}[\psi(p)]}{\mathbb{V}[\psi(p) + \lambda(p)]}} \left(1 - \frac{3}{10} \sqrt{\frac{\mathbb{V}[\lambda(p)]}{\mathbb{V}[\psi(p) + \lambda(p)]}}\right) \tag{B.3}$$

Proof. If two functions $f(p)$ and $g(p)$ are both increasing in $\ln p$ with $\partial f(p)/\partial \ln p \leq C_f$ and $\partial g(p)/\partial \ln p \leq C_g$, then $\mathbb{C}[f(p), g(p)] \in \left(0, C_f C_g \mathbb{V}[\ln p]\right)$. Since $I(p|\beta)$ is decreasing in $\ln p$, we therefore have that

$$\mathbb{V}[\psi(p)] - \mathbb{V}[\lambda(p)] = (2\beta - 1)\mathbb{V}[\psi(p) + \lambda(p)] + 2\mathbb{C}[I(p|\beta), \ln p] \leq (2\beta - 1)\mathbb{V}[\psi(p) + \lambda(p)]$$

and rearranging yields the lower bound in (B.1). As the derivative of $I(p|\beta)$ with respect to $\ln p$

is also bounded from below by $-(1-\beta)^2/\beta$, it additionally follows that

$$\mathbb{V}[\lambda(p)] = (1-\beta)^2\mathbb{V}[\psi(p) + \lambda(p)] + \mathbb{C}[I(p\,|\,\beta), I(p\,|\,\beta) - 2(1-\beta)\ln p]$$

$$\leq \left((1-\beta)^2 + \frac{(1-\beta)^4}{\beta^2} + \frac{2(1-\beta)^3}{\beta}\right)\mathbb{V}[\psi(p) + \lambda(p)] = \frac{(1-\beta)^2}{\beta^2}\mathbb{V}[\psi(p) + \lambda(p)]$$

and rearranging yields the upper bound in (B.2). When $\beta \geq 1/2$, then $4I(p\,|\,\beta) + 2(4\beta - 1)\ln p$ is increasing in $\ln p$ so that

$$\mathbb{V}[\psi(p)] - \mathbb{C}[\psi(p), \lambda(p)] = (2\beta^2 - \beta)\mathbb{V}[\psi(p) + \lambda(p)] + \mathbb{C}[I(p\,|\,\beta), 4I(p\,|\,\beta) + 2(4\beta - 1)\ln p]$$

$$\leq (2\beta^2 - \beta)\mathbb{V}[\psi(p) + \lambda(p)]$$

and rearranging yields the lower bound in (B.2).

Inserting the upper bound in (B.2) into the increasing function $2\beta^2 - \beta$ we obtain that

$$\mathbb{V}[\psi(p)] - \mathbb{C}[\psi(p), \lambda(p)] \leq \frac{\mathbb{V}[\psi(p)] + 2\mathbb{C}[\psi(p), \lambda(p)]}{\left(1 + \sqrt{\mathbb{V}[\lambda(p)]/\mathbb{V}[\psi(p) + \lambda(p)]}\right)^3}$$

and rearranging leads to the lower bound

$$\rho(\psi(p), \lambda(p)) \geq \sqrt{\frac{\mathbb{V}[\psi(p)]}{\mathbb{V}[\lambda(p)]}} \frac{\left(1 + \sqrt{\mathbb{V}[\lambda(p)]/\mathbb{V}[\psi(p) + \lambda(p)]}\right)^3 - 1}{2 + \left(1 + \sqrt{\mathbb{V}[\lambda(p)]/\mathbb{V}[\psi(p) + \lambda(p)]}\right)^3}.$$

The reported lower bound in (B.3) is smaller than the preceding one, as $[(1 + x)^3 - 1]/[2 + (1 + x)^3] - x + 3x^2/10 \geq 0$ for $x \in [0, 1]$. $\qquad\square$

Lemma B.3. *Suppose that p is continuously distributed on an interval contained in the positive half-line that is bounded and bounded away from zero. If $\ln p = m(V) + U$ where U is independent of V and $m(v) = \mathbb{E}[\ln p\,|\,V = v]$, then $\mathbb{E}[\lambda(p)\,|\,V = v]$ is increasing and concave in $m(v)$ while $\mathbb{E}[\psi(p)\,|\,V = v]$ is concave in $m(v)$ and increasing in $m(v)$ if $\beta \geq 1/2$.*

For V being log value added, Figure 5, panel (b), plots non-parametric estimates of $\mathbb{E}[\lambda(p)\,|\,V = v]$ and $\mathbb{E}[\psi(p)\,|\,V = v]$ against $m(v)$.

Proof. Independence between U and V implies that $\ln p$ conditional on $V = v$ is continuously

distributed with density $f_U(\cdot - m(v))$. Therefore,

$$\mathbb{E}[\lambda(p) \,|\, V = v] = \int \lambda(\exp(x)) f_{\ln p \,|\, V}(x \,|\, v) \; dx = \int \lambda(\exp(x)) f_U(x - m(v)) \; dx$$

$$= \int \lambda(\exp(m(v) + u)) f_U(u) \; du.$$

Differentiation under the integral sign reveals that the derivatives of $\mathbb{E}[\lambda(p) \,|\, V = v]$ with respect to $m(v)$ is a weighted average of the corresponding derivatives of $\lambda(p)$ with respect to $\ln p$. Therefore, the monotonicity and convexity of $\lambda(p)$ with respect to $\ln p$ implies monotonicity and convexity of $\mathbb{E}[\lambda(p) \,|\, V = v]$ with respect to $m(v)$. The argument is analogous for $\psi(p)$. $\qquad\square$

Appendix C   Identification of DWL model parameters

The use of *pairwise differences* has long been considered an intuitive and transparent way to establish identification and construct estimators in econometrics (Ahn and Powell, 1993; Honoré and Powell, 1994). Although we ultimately estimate the DWL model via OLS, the following discussions illustrates how the basis for identification of the DWL model involves pairwise differences and a generalization thereof to directed walks on a directed network.

To illustrate the type of worker mobility that allows us to identify the DWL model, we will suppress the time-varying regressors X_{im} and focus on a setting where each worker has two observed hiring wages and a known origin state for their first hiring wage. In this setting, the unique way to partial out the individual effects α_i is to consider a model of first differences

$$\Delta y_i = \underbrace{\psi_{j(i,2)} - \psi_{j(i,1)}}_{=\Delta F_i' \psi} + \underbrace{\lambda_{h(i,2)} - \lambda_{h(i,1)}}_{=\Delta H_i' \lambda} + \Delta \varepsilon_i \tag{C.1}$$

where, for any variable w, $\Delta w_i = w_{i2} - w_{i1}$. In this model, it is immediately clear that levels of the origin and destination firm effects are not identified. However, for identification of variances and covariances it suffices that first differences of the form $\psi_s - \psi_t$ and $\lambda_s - \lambda_t$ are identified, so we will focus on such differences. Moreover, as our argument is symmetric for ψ and λ, we will only discuss identification of $\psi_s - \psi_t$ for two arbitrary firms s and t that both hired a worker during the sampling frame.

The difference in firm effects $\psi_s - \psi_t$ is identified if and only if there exist a known vector of weights $v = (v_1, \ldots, v_n)' \in \mathbb{R}^n$ such that the weighted sum $\sum_{i=1}^n v_i \Delta y_i$ has a (conditional) mean of $\psi_s - \psi_t$ for any value of (ψ, λ). To understand when such a vector exists, it is useful to represent worker mobility as two directed networks where the firms are vertices and the workers' moves correspond to edges. There are two networks in play because the model in (C.1) includes two

moves for each worker: the mobility described by ΔH_i takes place on an "origin" network, while the mobility described by ΔF_i takes place on a "destination" network.

An example of such networks is visualized in Figure C.1. Here, there are five workers, three firms, and not yet in the labor force (N) as an origin state. The edges describing the first two workers' mobility are highlighted in red. In panel (a), which depicts the origin network, we see that these two workers have the same labor market experience in their first observed jobs as they both enter the labor market and are initially hired by firm #1. However, the destination network in panel (b) show that their subsequent employers differ, as the first (second) worker is hired by the second (third) firm. The shared experience of these workers on the origin network allows us to difference out the origin effects and establish identification of the destination effects difference among their second employers, i.e.,

$$\mathbb{E}[\Delta y_1 - \Delta y_2 \,|\, \mathcal{W}] = \underbrace{\psi_2 - \psi_1 - (\psi_3 - \psi_1)}_{\text{Destination difference}} + \underbrace{\lambda_1 - \lambda_N - (\lambda_1 - \lambda_N)}_{\text{Origin difference}}$$
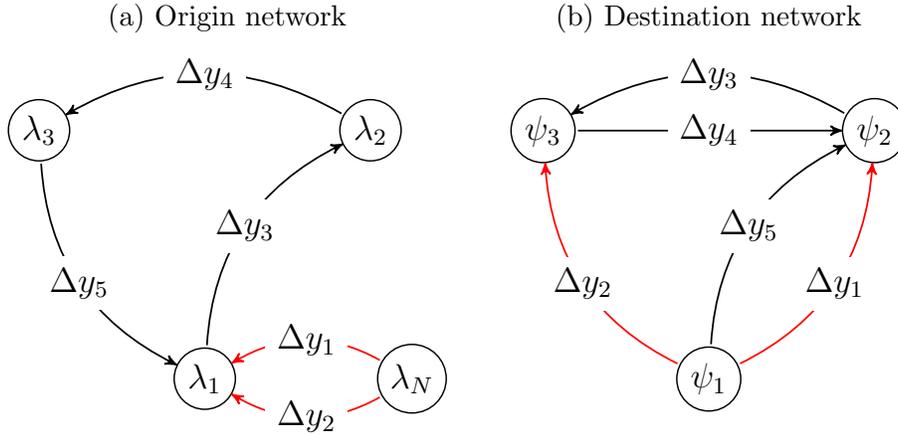
$$= \psi_2 - \psi_3.$$

This example illustrates how pairwise differences among workers who are hired into the labor force (or out of unemployment) by the same firm play a crucial role in identification of the DWL model. However, it is not only pairwise differences that contribute to identification of the model. The triwise sum $\Delta y_3 + \Delta y_4 + \Delta y_5$ can similarly be shown to yield identification of $\psi_2 - \psi_1$ by noting that

$$\mathbb{E}[\Delta y_3 + \Delta y_4 + \Delta y_5 \,|\, \mathcal{W}] = \underbrace{\psi_3 - \psi_2 + \psi_2 - \psi_3 + \psi_2 - \psi_1}_{\text{Destination difference}} + \underbrace{\lambda_2 - \lambda_1 + \lambda_3 - \lambda_2 + \lambda_1 - \lambda_3}_{\text{Origin difference}}$$

$$= \psi_2 - \psi_1.$$

The common features of the two weighted sums $\Delta y_1 - \Delta y_2$ and $\Delta y_3 + \Delta y_4 + \Delta y_5$ used to establish identification in the previous example are that they correspond to mobility that forms a *closed* walk on the origin network and an *open* walk on the destination network. Walks are common objects in the study of networks, but for completeness we give a brief description and a definition. A walk is a sequence of connected edges. When a walk starts and ends at the same place it is said to be closed and otherwise it is open. An open walk is said to be a walk between its endpoints. A collection of walks refers to multiple disjoint walks. A directed walk records the direction along which it traverses an edge.

Definition 1. Let $v = (v_1, \ldots, v_n)'$. We say that, (i) v is a collection of directed closed walks on the origin network if $\sum_{i=1}^{n} v_i \Delta L_i$ is equal to zero, and (ii) v is a collection of directed closed walks and a single directed open walk between firm k and κ on the destination network if $\sum_{i=1}^{n} v_i \Delta F_i$ is

Figure C.1: Identification in DWL Model

(a) Origin network   (b) Destination network

Note: Visualization of a network induced by data on five workers and three firms. The red edges correspond to the transitions of the first two workers. Those transitions form a closed walk on the origin network while they form an open walk from the third firm to the second firm on the destination network. The black edges similarly form a closed walk on the origin network and a open walk from the first to the second firm on the destination network. These observations imply the identification of destination effects.

equal to $e_s - e_t$ or $e_t - e_s$ where e_ℓ is the ℓ'th basis vector in $\mathbb{R}^J$.

The previous example did not need to consider collections of disjoint walks to establish identification. However, we end this section by noting that this is the right concept for establishing identification in general.

Theorem C.1. *The difference $\psi_s - \psi_t$ is identified if and only if there exist a vector v which is (i) a collection of directed closed walks on the origin network and (ii) a collection of directed closed walks and a single directed open walk between firm s and t on the destination network*

The preceding theorem discussed necessary and sufficient conditions for identification of firm effects differences. Kline et al. (2020) prove that estimating variance components without bias requires identification of the firm effect differences also hold when any single observation is dropped. However, in many datasets, including the one used in this paper, these identification conditions require that one "prunes" the data to find a subset of the data where identification holds. Appendix E.1 describes how we find this subset in practice.

Appendix D   Data

Our data come from the INPS-INVIND file which provides social security based earnings records on job spells for all private-sector workers who were employed at some point by a firm sampled by the Bank of Italy's INVIND survey. Since 2002, the INVIND survey has been representative of

firms with 20 or more employees in the manufacturing and service sector, see Bank of Italy (2018) for more details. Our job-level spell data is balanced, meaning that we have complete information on a worker's career even when this individual is not employed in a firm covered by the INVIND survey.

Each job-year spell in the INPS-INVIND lists a unique identifier of the employer and the employee, the start date, the end date, the number of days worked that year, and the total wage compensation received by the employee in that year. There is also information on which months during the year the employee was employed. The earnings records are top coded at 500,000 euros. We deflate earnings using the 2010 CPI. From 2005 and onwards, we have information on the reason why a particular job ended. Specifically, we have information on whether a worker has resigned from her job ("*Dimissioni*").

We consider data from the years 2005–2015. For our analysis, we include only spells where the worker is between 18 and 64 years of age. We omit spells with erroneous numbers of days worked or earnings. We also drop spells where the worker earned less than 2 euros per day. Finally, we dropped individuals that held more than 10 jobs per year or that entered the labor market before age 14 or after age 55.

After imposing these restrictions, we then use the monthly level employment information in INPS-INVIND to derive a person-job panel that contains information on a given job at the moment of hiring such as the hiring wage, age of the employee at hiring, reasons for separation from previous job, etc. Summary statistics for the resulting sample are given in Table 1, Panel (a).

Finally, our measure on value added comes from firms income statements collected by CERVED as described in Section 5. We winsorized information on value added at 5% and 95% in each year and then calculate for each firm its average log value added per worker over the years for which the firm's information is available in CERVED.

Appendix E   Implementation

E.1   Estimation sample

Estimation of the DWL model is conducted on a sample satisfying two conditions: (i) both the origin and the destination effect associated with a particular person-job observation is identified, and (ii) the statistical leverage, $P_{\ell\ell}$, of each person-job observation is less than unity. The latter requirement is equivalent to imposing (i) when any one person-job observation is dropped and is necessary for existence of unbiased estimators of variance components (Kline et al., 2020, Lemma 1).

In small samples or settings with few regressors, the unidentified origin and destination effects can be characterized through Gaussian elimination and the statistical leverages can be calculated

exactly. Thus, in small samples, one can easily prune away observations where the unidentified effects enter the conditional mean function $Z_\ell'\gamma$ and obtain a sample where S_{zz} is invertible. Afterwards, one can then drop observations with $P_{\ell\ell} = 1$ to also obtain a sample where $S_{zz} - Z_\ell'Z_\ell$ is invertible for any ℓ. However, in large samples with many regressors such as ours, Gaussian elimination and exact computation of the statistical leverages becomes computationally prohibitive. Therefore, we use the following iterative procedure to prune the sample.

In order to obtain a sample where $P_{\ell\ell} < 1$ for each $\ell \in \{1, \ldots, L\}$, we first note that $P_{\ell\ell} = 1$ implies that $\hat{y}_\ell = y_\ell$ where $\hat{y}_\ell = Z_\ell'\hat{\gamma}$ is the OLS prediction. We therefore remove all observations with $\hat{y}_\ell = y_\ell$ in a first step. In practice, we estimate the model using MATLAB's preconditioned conjugate gradient routine *pcg*, obtain the fitted values $\hat{y}_\ell$, and drop any observation with perfect fit, which we define as $|\hat{y}_\ell - y_\ell| < 1/1000$. Due to the slight numerical imprecision of *pcg*, we repeat this step until no observations are dropped.

We next prune the sample so that all person-job observations are associated with a pair of separately identified origin and destination effects. Rather than searching for collections of open and closed walks as described in Appendix C, we utilize another description of identification related to invertibility of S_{zz}. In our regression model of interest,

$$y_\ell = Z_\ell'\gamma + \varepsilon_\ell, \quad \text{for } \ell = 1, \ldots, L,$$

we have that γ is identified if and only if the OLS estimator $\hat{\gamma} = S_{zz}^{-1} \sum_{\ell=1}^{L} Z_\ell'y_\ell$ is equal to γ for any value of γ when all the error terms, ε_ℓ, are zero. To utilize this observation, we randomly draw γ^{sim} with *i.i.d.* standard normal entries. Then we compute the OLS estimate $\hat{\gamma}$ using *pcg* applied to the artificial data

$$\mathrm{y}_\ell = Z_\ell'\gamma^{\text{sim}}.$$

For any origin and destination effect where the corresponding entry of $|\hat{\gamma} - \gamma^{\text{sim}}|$ is greater than $1/100$, we drop the person-job observations where these effects enter the conditional mean function $Z_\ell'\gamma$.

This second step can possibly introduce new observations with statistical leverages of one, so we repeat the first step of the algorithm one more time and arrive at the estimation sample summarized in Table 1, Panel (b).

E.2   Computing the variance components

As detailed in Section 4, correcting for biases in the variance components requires knowledge of $B_{\ell\ell}$ and the statistical leverage $P_{\ell\ell}$. Both of these quantities are functions of available data. Specifically,

we have that

$$P_{\ell\ell} = Z_\ell' S_{zz}^{-1} Z_\ell, \qquad B_{\ell\ell} = Z_\ell' S_{zz}^{-1} A S_{zz}^{-1} Z_\ell, \qquad \text{for } \ell = 1, \dots, L.$$

However, *exact* computation of $(P_{\ell\ell}, B_{\ell\ell})$ is prohibitive in our context which involves tens of millions of observations and around 4 millions parameters. We therefore rely on the routine described in Kline et al. (2020) for computation. This methodology simplifies computation considerably by only requiring the solution of p systems (as opposed to k required by an exact solution) of k linear equations, where k is the total numbers of parameters associated with the DWL model. Specifically, using a variant of the random projection method of Achlioptas (2003) based on the Johnson-Lindenstrauss Approximation (JLA), one can approximate $(P_{\ell\ell}, B_{\ell\ell})$ using the columns of W_{JLA} in the system

$$\underset{k \times k}{S_{zz}} \underset{k \times p}{W_{JLA}} = \underset{k \times p}{(R_P Z)'},$$

where $Z = (Z_1, \dots, Z_L)'$ and R_P is a $p \times L$ matrix composed of mutually independent Rademacher random variables that are independent of the data, i.e., their entries take the values 1 and -1 with probability 1/2. As shown by Kline et al. (2020), the JLA algorithm reduces computation time dramatically when p is small relative to L while delivering very accurate estimates of the leave-out corrected variance component. The following algorithm describes in detail how to compute the JLA approximation of $P_{\ell\ell}$ and $B_{\ell\ell}$ for a given quadratic form matrix A. In what follows, we assume that the matrix A is positive semi-definite and can be written as $A = A_1' A_1$.[17]

Algorithm 1 Johnson-Lindenstrauss Approximation in the DWL Model

---

1: **function** JLA(Z, A_1)
2:     Generate $R_B, R_P \in \mathbb{R}^{p \times L}$, where (R_B, R_P) are composed of mutually independent Rademacher entries.
3:     Compute $(R_P Z)', (R_B A_1)' \in \mathbb{R}^{k \times p}$
4:     **for** $\kappa = 1, \dots, p$ **do**
5:         Let $r_{\kappa,0}, r_{\kappa,1}, \in \mathbb{R}^k$ be the κ-th columns of $(R_P Z)', (R_B A_1)'$.
6:         Let $w_{\kappa,l} \in \mathbb{R}^k$ be the solution to $S_{zz} w = r_{\kappa,l}$ for $l = 0, 1$.
7:     **end for**
8:     Construct $W_l = (w_{1,l}, \dots, w_{p,l}) \in \mathbb{R}^{k \times p}$ for $l = 0, 1$.
9:     Construct $\hat{P}_{\ell\ell} = \frac{1}{p} \lVert W_0' Z_\ell \rVert^2$, $\hat{B}_{\ell\ell} = \frac{1}{p} \lVert W_1' Z_\ell \rVert^2$ for $\ell = 1, \dots, L$.
10:    Return $\{\hat{P}_{\ell\ell}, \hat{B}_{\ell\ell}\}$
11: **end function**

---

[17]It is straightforward to extend the algorithm below to account for covariance components, such as the covariance in the origin and destination effects, see the computational appendix of Kline et al. (2020).

The solution to the linear system outlined in Line 6 of Algorithm 1 is performed via MATLAB's preconditioned conjugate gradient routine *pcg* and we used an incomplete Cholesky factorization of S_{zz} as the preconditioner with threshold dropping tolerance of 0.01.