Online Supplementary Material

F Omitted Model Details

We present here omitted details and proofs about the characterization of the model equilibrium.

F.1 Efficiency of Equilibrium in the One-Job Case

Although our model is one of imperfect competition in the labor market as firms are differentiated in their technologies, the equilibrium does not need to be inefficient. Intuitively, wages are determined by a second-price auction mechanism for a worker's services, which, as any such mechanism, has desirable efficiency properties. However, inefficiencies can arise because through employment, firms allow workers to acquire human capital and generate information about their ability, but neither human capital nor information about ability can be priced separately from a worker's labor services. Put differently, the pricing mechanism is not rich enough to align a worker and *all* firms' incentives to produce output, human capital, and information.

Yet, we are able to establish that equilibrium is efficient provided that the process of accumulation of human capital and information and the risk of exogenous separation are similar enough across firms that workers are able to internalize through their employment decisions the willingness of the third-best firm, fourth-best firm, and so on to employ them.[54] This result extends to the case in which firms consist each of many jobs, as shown in Appendix F.2 of the Online Supplementary Material.

Proposition F.1 (Efficiency)**.** *For any $\overline{D}$ such that $|\mathcal{D}| \leq \overline{D}$, whenever the process of human capital acquisition, the probabilities $\alpha(H_{n,1}, d, e_n)$ and $\beta(H_{n,1}, d, e_n)$ of high output for a high-ability worker and a low-ability worker, respectively, and the probability of exogenous separation $\varsigma(I_n^{t-1}, d)$ are sufficiently similar between any two firms d and d' for each realization of $H_{n,1}$, e_n, and I_n^{t-1}, where the required degree of similarity depends on $\overline{D}$, the equilibrium is efficient.*

Proof: Consider first the case in which $|\mathcal{D}| = 2$. The proof of the result in this case is an extension of Theorem 2 in Bergemann and Välimaki (1996). Although their result applies to a setting in which a worker's ability, unlike our case, is independent across firms—namely, a consumer's taste in their framework of two firms competing for a consumer is unknown to all and independent across firms—no steps of the argument is reliant on this assumption. We note that for this case, the assumption that the human capital process is sufficiently similar across firms can be dispensed with. When $|\mathcal{D}| > 2$, the assumptions of the proposition ensure that firms ranked third-best, fourth-best, and so on by the worker can always be chosen to have technologies that are sufficiently similar to that of the first- and second-best firms. Hence, the same argument as that under the case of $|\mathcal{D}| = 2$ applies. □

F.2 Multi-Job-Firm Case: Equilibrium, Efficiency, and Identification

We now argue that the probability of a worker's employment at any firm in the market and allocation to any job at the employing firm admits a pseudo-programming characterization, which only requires

[54] It is easy to construct counterexamples to efficiency when the market consists of three or more firms or each firm consists of multiple jobs, if the human capital or informational process are unrestricted across firms.

knowledge of a worker's wage in equilibrium. For this result, we consider the general case of our model in which firms consist of finitely many jobs indexed by j—without loss, we assume the set of jobs to be the same across firms. In this section, we denote by f a generic firm, by d the equilibrium employing (first-best) firm, and by d' the equilibrium second-best firm.

In this more general case, the definition of equilibrium extends straightforwardly; we note here only the important differences. In particular, we denote by $(w_{n,t,f}, j_{n,t,f})$, with $w_{n,t,f} := w_f(s_{n,t}, \epsilon_{n,t})$ and $j_{n,t,f} := j_f(s_{n,t}, \epsilon_{n,t})$, the wage and job offer strategy of a generic firm f and by $\{w_{n,t,f}, j_{n,t,f}\}_{f\in\mathcal{D}}$ the collection of all offer strategies. We denote by $l_{n,t,f} := l_f(s_{n,t}, \epsilon_{n,t}, \{w_{n,t,f}, j_{n,t,f}\}_{f\in\mathcal{D}})$ the acceptance strategy of worker n for firm f's offer—an indicator function, taking value one if f is the employing firm and zero otherwise at a given state—and by $\{l_{n,t,f}\}_{f\in\mathcal{D}}$ the collection of all acceptance strategies. Given firms' strategies, worker n's strategy satisfies

$$\begin{aligned}\tilde{W}(s_{n,t}, \epsilon_{n,t}, \{w_{n,t,f}, j_{n,t,f}\}_{f\in\mathcal{D}}) = &\max_{\{l_f\}_{f\in\mathcal{D}}} \sum_{f\in\mathcal{D}} l_f \times \Big\{ w_{n,t,f} + \delta[1 - \varsigma(H_{n,1}, I_n^{t-1}, f, j_{n,t,f})] \\ &\times \int_{\epsilon_{n,t+1}} \mathbb{E}\Big[\tilde{W}(s_{n,t+1}, \epsilon_{n,t+1}, \{w_{n,t+1,f}, j_{n,t+1,f}\}_{f\in\mathcal{D}}) | s_{n,t}, f, j_{n,t,f}\Big] dF_{\epsilon_{n,t+1}} \Big\}, \end{aligned} \tag{96}$$

where $j_{n,t,f}$ affects the exogenous separation rate $\varsigma(\cdot)$ and conditions the law of motion of the state, as is apparent from the conditional expectation in the last line of (96). Given worker n's strategy and its competitors' strategies, firm f's strategy satisfies

$$\begin{aligned}\Pi_f(s_{n,t}, \epsilon_{n,t}) = &\max_{w,j} \Big(l_{n,t,f} \times \Big\{ y(f, s_{n,t}, j, \epsilon_{n,t}(f,j)) - w \\ &+ \delta[1 - \varsigma(H_{n,1}, I_n^{t-1}, f, j)] \int_{\epsilon_{n,t+1}} \mathbb{E}\Big[\Pi_f(s_{n,t+1}, \epsilon_{n,t+1}) | s_{n,t}, f, j\Big] dF_{\epsilon_{n,t+1}} \Big\} \\ &+ \sum_{f'\in\mathcal{D}\setminus\{f\}} l_{n,t,f'} \Big\{ \delta[1 - \varsigma(H_{n,1}, I_n^{t-1}, f', j_{n,t,f'})] \int_{\epsilon_{n,t+1}} \mathbb{E}\Big[\Pi_d(s_{n,t+1}, \epsilon_{n,t+1}) | s_{n,t}, f', j_{n,t,f'}\Big] dF_{\epsilon_{n,t+1}} \Big\} \Big), \end{aligned} \tag{97}$$

which shows how each firm now chooses a wage and a job for the worker, taking into account the wage and job offers of all other firms—see the last line of (97). The cautious equilibrium refinement, namely, condition (*iv*) of our equilibrium definition, requires that if firm $d \in \mathcal{D}$ employs worker n at state $(s_{n,t}, \epsilon_{n,t})$, then for any other firm $f \in \mathcal{D}$,

$$\begin{aligned}\delta[1 - \varsigma(H_{n,1}, I_n^{t-1}, d, j_{n,t,d})] \int_{\epsilon_{n,t+1}} \mathbb{E}[\Pi_f(\cdot) | s_{n,t}, d, j_{n,t,d}] dF_{\epsilon_{n,t+1}} = \max_{w,j} \Big\{ y(f, s_{n,t}, j, \epsilon_{n,t}(f,j)) \\ - w + \delta[1 - \varsigma(H_{n,1}, I_n^{t-1}, f, j)] \int_{\epsilon_{n,t+1}} \mathbb{E}[\Pi_f(\cdot) | s_{n,t}, f, j] dF_{\epsilon_{n,t+1}} \Big\}. \end{aligned} \tag{98}$$

Two immediate implications of equilibrium are as follows. First, worker n must be indifferent between the offers of the first- and second-best firms, as discussed in the proof of Proposition 1, and weakly prefer either of these two offers to any other offers. Second, for each firm f, maximizing profits is equivalent to maximizing the sum of its own value of profits and worker n's value of wages from accepting its offer. We denote this value by $\tilde{W}(s_{n,t}, \epsilon_{n,t}|f)$, which is the value $\tilde{W}(s_{n,t}, \epsilon_{n,t}, \{w_{n,t,f'}, j_{n,t,f'}\}_{f'\in\mathcal{D}})$ conditional on the worker accepting firm f's offer, after suppressing the notation for firms' offers since they are function of the state $(s_{n,t}, \epsilon_{n,t})$.

That maximizing profits is equivalent to maximizing match surplus follows from: *(i)* any firm f different from the first-best firm d being indifferent between *not employing* the worker—in which case its payoff is $\Pi_f(s_{n,t}, \epsilon_{n,t}|d)$, the left side of (98)—and *employing* the worker—in which case its payoff is $\Pi_f(s_{n,t}, \epsilon_{n,t}|f)$, the right side of (98); and *(ii)* worker n weakly preferring employment at the first-best firm d to employment at any other firm f, ($\tilde{W}(s_{n,t}, \epsilon_{n,t}|d) \geq \tilde{W}(s_{n,t}, \epsilon_{n,t}|f)$). Summing the two conditions $\Pi_f(s_{n,t}, \epsilon_{n,t}|d) = \Pi_f(s_{n,t}, \epsilon_{n,t}|f)$ and $\tilde{W}(s_{n,t}, \epsilon_{n,t}|d) \geq \tilde{W}(s_{n,t}, \epsilon_{n,t}|f)$ term by term indeed gives that

$$\Pi_f(s_{n,t}, \epsilon_{n,t}|d) + \tilde{W}(s_{n,t}, \epsilon_{n,t}|d) \geq \Pi_f(s_{n,t}, \epsilon_{n,t}|f) + \tilde{W}(s_{n,t}, \epsilon_{n,t}|f),$$

so maximizing profits is equivalent to maximizing match surplus for any firm f different from the first-best firm d. As for the first-best firm d, since this firm must prefer employing to not employing the worker, that is, $\Pi_d(s_{n,t}, \epsilon_{n,t}|d) \geq \Pi_d(s_{n,t}, \epsilon_{n,t}|f)$, an analogous argument applies.

Thus, at any state at which a firm f is the employing firm d, we can rewrite the firm problem as

$$V_d(s_{n,t}, \epsilon_{n,t}) = \max_j \left\{ y(d, s_{n,t}, j, \epsilon_{n,t}(d,j)) + \delta[1 - \varsigma(H_{n,1}, I_n^{t-1}, d, j)] \int_{\epsilon_{n,t+1}} \mathbb{E}[V_d(\cdot)|s_{n,t}, d, j] dF_{\epsilon_{n,t+1}} \right\}, \tag{99}$$

whereas at any state at which firm f is *not* the employing firm d, we can rewrite the firm problem as

$$\begin{aligned} V_f(s_{n,t}, \epsilon_{n,t}) = \max \Bigg(& w_{n,t,d} + \delta[1 - \varsigma(H_{n,1}, I_n^{t-1}, d, j_{n,t,d})] \int_{\epsilon_{n,t+1}} \mathbb{E}[V_f(\cdot)|s_{n,t}, d, j_{n,t,d}] dF_{\epsilon_{n,t+1}}, \\ & \max_j \left\{ y(f, s_{n,t}, j, \epsilon_{n,t}(f,j)) + \delta[1 - \varsigma(H_{n,1}, I_n^{t-1}, f, j)] \int_{\epsilon_{n,t+1}} \mathbb{E}[V_f(\cdot)|s_{n,t}, f, j] dF_{\epsilon_{n,t+1}} \right\} \Bigg), \end{aligned} \tag{100}$$

where $w_{n,t,d}$ is worker n's wage at the first-best firm d. At each state, the solution to the subproblem in (99) determines worker n's probability of job assignment at the employing firm (firm d) conditional on employment at it, whereas the solution to the subproblem in (100) determines worker n's probability of employment at firms different from the first-best firm (any firm $f \neq d$) together with their job offer. By (99) and (100), the best-response problem of any firm f can be represented as

$$V_f(s_{n,t}, \epsilon_{n,t}) = \begin{cases} \max_j \Big\{ y(f, s_{n,t}, j, \epsilon_{n,t}(f,j)) \\ \quad + \delta[1 - \varsigma(H_{n,1}, I_n^{t-1}, f, j)] \int_{\epsilon_{n,t+1}} \mathbb{E}[V_f(\cdot)|s_{n,t}, f, j] dF_{\epsilon_{n,t+1}} \Big\} \text{ if } f = d \\ \max \Big(w_{n,t,d} + \delta[1 - \varsigma(H_{n,1}, I_n^{t-1}, d, j_{n,t,d})] \int_{\epsilon_{n,t+1}} \mathbb{E}[V_f(\cdot)|s_{n,t}, d, j_{n,t,d}] dF_{\epsilon_{n,t+1}}, \\ \quad \max_j \Big\{ y(f, s_{n,t}, j, \epsilon_{n,t}(f,j)) \\ \quad + \delta[1 - \varsigma(H_{n,1}, I_n^{t-1}, f, j)] \int_{\epsilon_{n,t+1}} \mathbb{E}[V_f(\cdot)|s_{n,t}, f, j] dF_{\epsilon_{n,t+1}} \Big\} \Big) \text{ if } f \neq d \end{cases}$$

which is a pseudo-planning problem up to the one-period match surplus value of firm f when it is not the employing firm, given by the wage $w_{n,t,d}$ paid to worker n by the first-best firm. The collection of all such match surplus values for all firms determines worker n employing firm and assigned job.

By the discrete choice logic in Kristensen et al. (2015), it is easy to see that up to δ and knowledge of $y(f, s_{n,t}, j, \epsilon_{n,t}(f,j))$ for one j, say $\tilde{j}$, the remaining $y(f, s_{n,t}, j), \epsilon_{n,t}(f,j)$, $j \neq \tilde{j}$, are identified once the remaining primitives are known—see Proposition A.8 for the formal statement of this result.

In terms of the efficiency of equilibrium, recall that in equilibrium, worker n must be indifferent between the first-best (d) and second-best (d') firm and weakly prefer their offers to any other; the employing firm must prefer employing the worker to not employing the worker; and each non-employing firm must prefer not employing the worker to employing the worker. Respectively, in equilibrium, for worker n to be employed at firm d, it must be that

$$\begin{aligned} w_{n,t,d} &+ \delta[1-\varsigma(H_{n,1}, I_n^{t-1}, d, j_{n,t,d})] \\ &\times \int_{\epsilon_{n,t+1}} \mathbb{E}\Big[W(s_{n,t+1}, \epsilon_{n,t+1}, \{w_{n,t+1,f'}, j_{n,t+1,f'}\}_{f'\in\mathcal{D}})|s_{n,t}, d, j_{n,t,d}\Big] dF_{\epsilon_{n,t+1}} \\ &\geq w_{n,t,f} + \delta[1-\varsigma(H_{n,1}, I_n^{t-1}, f, j_{n,t,f})] \\ &\times \int_{\epsilon_{n,t+1}} \mathbb{E}\Big[W(s_{n,t+1}, \epsilon_{n,t+1}, \{w_{n,t+1,f'}, j_{n,t+1,f'}\}_{f'\in\mathcal{D}})|s_{n,t}, f, j_{n,t,f}\Big] dF_{\epsilon_{n,t+1}}, \end{aligned}$$

with strict equality when f is the second-best firm d'; for the employing firm d, it must be that

$$\begin{aligned} &\max_{w,j} \Big\{ y(d, s_{n,t}, j, \epsilon_{n,t}(d,j)) - w + \delta[1-\varsigma(H_{n,1}, I_n^{t-1}, d, j)] \int_{\epsilon_{n,t+1}} \mathbb{E}[\Pi_d(\cdot)|s_{n,t}, d, j] dF_{\epsilon_{n,t+1}} \Big\} \\ &\geq \delta[1-\varsigma(H_{n,1}, I_n^{t-1}, f, j_{n,t,f})] \int_{\epsilon_{n,t+1}} \mathbb{E}[\Pi_d(\cdot)|s_{n,t}, f, j_{n,t,f}] dF_{\epsilon_{n,t+1}}; \end{aligned}$$

and for any other firm $f \neq d$, it must be that

$$\begin{aligned} &\delta[1-\varsigma(H_{n,1}, I_n^{t-1}, d, j_{n,t,d})] \int_{\epsilon_{n,t+1}} \mathbb{E}[\Pi_f(\cdot)|s_{n,t}, d, j_{n,t,d}] dF_{\epsilon_{n,t+1}}, \\ &\geq \max_{w,j} \Big\{ y(f, s_{n,t}, j, \epsilon_{n,t}(f,j)) - w + \delta[1-\varsigma(H_{n,1}, I_n^{t-1}, f, j)] \int_{\epsilon_{n,t+1}} \mathbb{E}[\Pi_f(\cdot)|s_{n,t}, f, j] dF_{\epsilon_{n,t+1}} \Big\}. \end{aligned}$$

By summing term by term these three sets of inequalities and recalling the definition of (social) welfare $S(s_{n,t}, \epsilon_{n,t})$ as the sum of values of worker n and all firms, namely, $S(\cdot) = W(\cdot) + \sum_f \Pi_f(\cdot)$, it is immediate that if the law of motion of the state and the risk of exogenous separation are similar enough across firms and jobs or firms' continuation profits are independent enough of the state (as under perfect competition when they are zero), then any MPE is efficient. Intuitively, in this case, the values of all players can be "passed through" the relevant maximization operators so that maximizing match surplus is equivalent to maximizing total surplus. We note that this argument applies to all MPEs, even those that do not satisfy the cautious requirement. In the general case, however, this argument does not apply, and it is easy to see that the equilibrium need not be efficient (see also Bergemann and Välimaki, 2006). □

G Additional Identification Results

G.1 Micro-Foundation of Assumption (ii) of Proposition D.1

Lemma G.1 shows that if the productivity shocks are "sufficiently independent," then Assumption (ii) of Proposition D.1 holds. This corresponds to Corollary 4.1 of D'Haultfoeuille and Maurel (2013).

Lemma G.1 (Moderate Dependence—Simplified Wage Equation (11))**.** *Let Assumption (i) of Proposition D.1 hold and, without loss, $\mathcal{D} \coloneqq \{0,1\}$. For some $q_t(1) \in (0,1]$, let*

$$\lim_{u\to\infty} \Pr\big(\epsilon_{n,t}(0) \le \epsilon_{n,t}(1) + a \,\big|\, \epsilon_{n,t}(1) \ge u\big) = q_t(1) \qquad \textit{for all } a \in \mathbb{R}. \tag{101}$$

Then, Assumption (ii) of Proposition D.1 holds in that $\lim_{w\to\infty} \Pr\big(D_{n,t} = 1 \mid s_{n,t} = s,\ w_{n,t}(1) \ge w\big) = q_t(1)$ for every $s \in \mathcal{S}_t$. Moreover, if $\epsilon_{n,t}(0)$ and $\epsilon_{n,t}(1)$ are independent, then $q_t(1) = 1$.

Proof: For simplicity, assume that the equilibrium is efficient. All the steps can be generalized to the non-efficient case by leveraging the pseudo-planner problem representation in Appendix F.2 of the Online Supplementary Material. Then. for any $s \in \mathcal{S}_t$ and $w \in \mathbb{R}$,

$$\begin{aligned}
&\Pr(D_{n,t} = 1 \,|\, s_{n,t} = s,\ w_{n,t}(1) \ge w) \\
&= \Pr\big(\Upsilon(1,s) + \epsilon_{n,t}(1) \ \ge\ \Upsilon(0,s) + \epsilon_{n,t}(0) \,\big|\, s_{n,t} = s,\ \varphi(1,s) + \epsilon_{n,t}(1) \ge w\big) \\
&= \Pr\big(\epsilon_{n,t}(0) \le \epsilon_{n,t}(1) + \Upsilon(1,s) - \Upsilon(0,s) \,\big|\, \epsilon_{n,t}(1) \ge w - \varphi(1,s)\big),
\end{aligned} \tag{102}$$

where $\Upsilon(d,s) + \epsilon_{n,t}(d)$ is the expected present discounted value of output for firm d in state s after productivity shocks have realised.

Set $u \coloneqq w - \varphi(1,s)$, so $w \to \infty$ iff $u \to \infty$. Applying (101) with $a = \Upsilon(1,s) - \Upsilon(0,s)$ gives

$$\lim_{w\to\infty} \Pr\big(\epsilon_{n,t}(0) \le \epsilon_{n,t}(1) + a(x) \,\big|\, \epsilon_{n,t}(1) \ge w - \varphi(1,s)\big) = q_t(1). \tag{103}$$

By (102), (103) is precisely

$$\lim_{w\to\infty} \Pr(D_{n,t} = 1 \mid s_{n,t} = s,\ w_{n,t}(1) \ge w) = q_t(1),$$

which is Assumption (ii) of Proposition D.1.

Now, suppose $\epsilon_{n,t}(1)$ and $\epsilon_{n,t}(0)$ are independent. Then, for any $a \in \mathbb{R}$ and $u \in \mathbb{R}$,

$$\begin{aligned}
\Pr(\epsilon_{n,t}(0) \le \epsilon_{n,t}(1) + a \mid \epsilon_{n,t}(1) \ge u) &= \mathbb{E}\Big[\Pr(\epsilon_{n,t}(0) \le \epsilon_{n,t}(1) + a \mid \epsilon_{n,t}(1)) \,\Big|\, \epsilon_{n,t}(1) \ge u\Big] \\
&= \mathbb{E}\Big[F_{\epsilon_{n,t}(0)}\big(\epsilon_{n,t}(1) + a\big) \,\Big|\, \epsilon_{n,t}(1) \ge u\Big],
\end{aligned} \tag{104}$$

where $F_{\epsilon_{n,t}(0)}$ is the CDF of $\epsilon_{n,t}(0)$ and the second line uses independence between $\epsilon_{n,t}(0)$ and $\epsilon_{n,t}(1)$. Since $F_{\epsilon_{n,t}(0)}$ is nondecreasing,

$$F_0(u+a) \ \le\ F_{\epsilon_{n,t}(0)}\big(\epsilon_{n,t}(1) + a\big) \ \le\ 1 \quad \text{on the event } \{\epsilon_{n,t}(1) \ge u\}.$$

Taking conditional expectations yields the bounds

$$F_{\epsilon_{n,t}(0)}(u+a) \;\le\; \mathbb{E}\Big[F_{\epsilon_{n,t}(0)}\big(\epsilon_{n,t}(1)+a\big)\,\Big|\,\epsilon_{n,t}(1)\ge u\Big] \;\le\; 1.$$

Letting $u \to \infty$ and using $\lim_{\tau\to\infty} F_{\epsilon_{n,t}(0)}(\tau) = 1$, we conclude that

$$\lim_{u\to\infty} \Pr(\epsilon_{n,t}(0) \le \epsilon_{n,t}(1) + a \mid \epsilon_{n,t}(1) \ge u) = 1,$$

so $q_t(1) = 1$. $\square$

Remark. Suppose that $\epsilon_{n,t}(1)$ and $\epsilon_{n,t}(0)$ are jointly normal—or lognormal. If $\mathrm{cov}(\epsilon_{n,t}(1), \epsilon_{n,t}(0)) < \mathrm{Var}(\epsilon_{n,t}(1))$—or if $\mathrm{cov}(\log(\epsilon_{n,t}(1)), \log(\epsilon_{n,t}(0))) < \mathrm{Var}(\log(\epsilon_{n,t}(1)))$—then (101) holds. Similar "sufficient independence" conditions can be given for many other parametric families, including both thin-tailed (for instance, Normal, Exponential, Gamma, Logistic, Gumbel) and fat-tailed (for instance, Pareto, Cauchy, Burr, Fréchet, log-logistic, and lognormal) distributions.

G.2 Proposition D.1 with Bounded Support

Proposition G.1 establishes identification of the deterministic wage components entering our simplified wage equation (11) in the case where the potential wages $w_{n,t}(d) \mid s_{n,t} = s$ and the observed, selected wages $w_{n,t} \mid (D_{n,t} = d, s_{n,t} = s)$ have different right endpoints. The identification result retains the spirit of Proposition D.1, but extra care is needed in taking limits because the two endpoints differ. We further show that, when finite, the lower and upper endpoints of the potential wages $w_{n,t}(d) \mid s_{n,t} = s$ and shock $\epsilon_{n,t}(d)$ are nonparametrically identified (Corollary G.1).

Proposition G.1 (Deterministic Wage—Simplified Wage Equation (11) with Bounded Supports)**.** *Let $t \in \{1,\ldots,T\}$ and $d \in \mathcal{D}$. Assume that the conditional wage distribution $\Pr\big(w_{n,t} \le w \,\big|\, D_{n,t} = d,\; s_{n,t} = s\big)$ is identified for each $w \in \mathbb{R}$ and $s \in \mathcal{S}_t$ (see Proposition A.5 for sufficient conditions), and that the conditional choice probability $\Pr(D_{n,t} = d \mid s_{n,t} = s)$ is identified for each $s \in \mathcal{S}_t$ (see Proposition A.4 for sufficient conditions). Moreover, assume:*

(i) (Supports.) For each $s \in \mathcal{S}_t$, $\omega_t(d,s) := \sup\{u : \Pr(w_{n,t}(d) \le u \mid s_{n,t} = s) < 1\} \le \infty$ and $\omega_t^{\mathrm{obs}}(d,s) := \sup\{u : \Pr(w_{n,t} \le u \mid D_{n,t} = d, s_{n,t} = s) < 1\} < \omega_t(d,s)$, with $\omega_t(d,s)$ and $\omega_t^{\mathrm{obs}}(d,s)$ potentially unknown.

(ii) (Relative Tail Decay.) For each $s \in \mathcal{S}_t$, define $r_{d,s,t}(u) := \Pr\big(D_{n,t} = 1 \mid s_{n,t} = s, w_{n,t}(d) > Q_{w_{n,t}(d)\mid s_{n,t}=s}(u)\big)$, $\quad u \in (0,1)$. There exists an (unknown) constant $q_t(d) \in (0,\infty)$ such that, for every $s \in \mathcal{S}_t$ and a fixed reference $\bar{s} \in \mathcal{S}_t$, $\lim_{u\to 1} \frac{r_{d,s,t}(u)}{r_{d,\bar{s},t}(u)} = q_t(d)$.

(iii) (Tail Regularity.) For each $s \in \mathcal{S}_t$, there exist (unknown) thresholds $a_t(d,s) < \infty$ and $a_t^{\mathrm{obs}}(d,s) < \omega_t^{\mathrm{obs}}(d,s)$ such that $F_{w_{n,t}(d)\mid s_{n,t}=s}$ and $F_{w_{n,t}\mid D_{n,t}=d, s_{n,t}=s}$ are continuous and strictly increasing *on $(a_t(d,s),\infty)$ and $(a_t^{\mathrm{obs}}(d,s), \omega_t^{\mathrm{obs}}(d,s))$, respectively. Moreover, $F_{w_{n,t}\mid D_{n,t}=d, s_{n,t}=s}$ is* continuous at the endpoint*:* $\lim_{w\to\omega_t^{\mathrm{obs}}(d,s)} F_{w_{n,t}\mid D_{n,t}=d, s_{n,t}=s}(w) = 1$.

(iv) (Normalization.) There exists a known $\bar{s} \in \mathcal{S}_t$ with $\varphi(d,\bar{s}) = 0$.

For each $s \in \mathcal{S}_t$, let $\{\tau^{(k)}_{d,\bar{s},t}\}_{k\geq 1} \subset (0,1)$ be any sequence with $\tau^{(k)}_{d,\bar{s},t} \to 1$ as $k \to \infty$ and define the sequence $\{\tau^{(k)}_{d,s,t}\}_{k\geq 1} \subset (0,1)$ where

$$\tau^{(k)}_{d,s,t} := 1 - \frac{\Pr(D_{n,t} = d \mid s_{n,t} = \bar{s})}{\Pr(D_{n,t} = d \mid s_{n,t} = s)} \left(1 - \tau^{(k)}_{d,\bar{s},t}\right).$$

Then,

$$\lim_{k\to\infty} \left[Q_{w_{n,t} \mid D_{n,t}=d, s_{n,t}=s}\left(\tau^{(k)}_{d,s,t}\right) \; - \; Q_{w_{n,t} \mid D_{n,t}=d, s_{n,t}=\bar{s}}\left(\tau^{(k)}_{d,\bar{s},t}\right) \right] \; = \varphi(d,s). \tag{105}$$

Hence, $\varphi(d,s)$ is identified for each $s \in \mathcal{S}_t$.

Proof: To facilitate reading, we divide the proof into steps. Without loss, consider firm $d = 1$.
Step 1: Bayes rule. Let $s \in \mathcal{S}_t$. For any real w, Bayes' rule gives

$$\begin{aligned} S_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(w) &= \frac{\Pr\big(w_{n,t}(1) > w \mid s_{n,t} = s\big) \; \Pr\big(D_{n,t} = 1 \mid s_{n,t} = s, \; w_{n,t}(1) > w\big)}{\Pr(D_{n,t} = 1 \mid s_{n,t} = s)} \\ &= S_{w_{n,t}(1)|s_{n,t}=s}(w) \; \frac{\Pr\big(D_{n,t} = 1 \mid s_{n,t} = s, \; w_{n,t}(1) > w\big)}{\Pr(D_{n,t} = 1 \mid s_{n,t} = s)}. \end{aligned} \tag{106}$$

Define

$$r_{1,s,t}(u) \; := \; \Pr\big(D_{n,t} = 1 \mid s_{n,t} = s, \; w_{n,t}(1) > Q_{w_{n,t}(1)|s_{n,t}=s}(u)\big), \qquad u \in (0,1).$$

Evaluating (106) at $w = Q_{w_{n,t}(1)|s_{n,t}=s}(u)$ yields

$$S_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}\big(Q_{w_{n,t}(1)|s_{n,t}=s}(u)\big) \; = \; (1-u)\,\frac{r_{1,s,t}(u)}{\Pr(D_{n,t} = 1 \mid s_{n,t} = s)}, \qquad u \in (0,1). \tag{107}$$

Step 2: Behavior of the composed survival near the observed endpoint. By Assumption (iii), there exist thresholds $a_t(1,s) < \omega_t(1,s)$ and $a^{\text{obs}}_t(1,s) < \omega^{\text{obs}}_t(1,s)$ such that $F_{w_{n,t}(d)|s_{n,t}=s}$ is continuous and strictly increasing on $(a_t(1,s), \omega_t(1,s))$, and $F_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}$ is continuous and strictly increasing on $(a^{\text{obs}}_t(1,s), \omega^{\text{obs}}_t(1,s))$. Define

$$u^*_{1,s,t} := F_{w_{n,t}(1)|s_{n,t}=s}(a_t(1,s)), \qquad \tau^*_{1,s,t} := F_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(a^{\text{obs}}_t(1,s)),$$

so $Q_{w_{n,t}(1)|s_{n,t}=s} : (u^*_{1,s,t}, 1) \to (a_t(1,s), \omega_t(1,s))$ and $Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=s} : (\tau^*_{1,s,t}, 1) \to (a^{\text{obs}}_t(1,s), \omega^{\text{obs}}_t(1,s))$ are strictly increasing. Because $\omega^{\text{obs}}_t(1,s) < \omega_t(1,s)$, set

$$\bar{u}_{1,s,t} := \sup\left\{u \in (u^*_{1,s,t}, 1) : \; Q_{w_{n,t}(1)|s_{n,t}=s}(u) < \omega^{\text{obs}}_t(1,s)\right\} \in (u^*_{1,s,t}, 1).$$

Then, $Q_{w_{n,t}(1)|s_{n,t}=s}(u) \to \omega^{\text{obs}}_t(1,s)$ as $u \to \bar{u}_{1,s,t}$. Since $Q_{w_{n,t}(1)|s_{n,t}=s}(u)$ is increasing and $\omega^{\text{obs}}_t(1,s)$ is finite, there exists $\tilde{u}_{1,s,t} \in (u^*_{1,s,t}, \bar{u}_{1,s,t})$ such that $Q_{w_{n,t}(1)|s_{n,t}=s}(u) \in (a^{\text{obs}}_t(1,s), \omega^{\text{obs}}_t(1,s))$ for all $u \in (\tilde{u}_{1,s,t}, \bar{u}_{1,s,t})$. On that interval, the map

$$u \; \mapsto \; S_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}\big(Q_{w_{n,t}(1)|s_{n,t}=s}(u)\big),$$

is a composition of a continuous, strictly increasing function (the potential quantile) with a continuous, strictly decreasing function (the observed survival on its tail), hence it is continuous and strictly decreasing on $(\tilde{u}_{1,s,t}, \bar{u}_{1,s,t})$. By the endpoint continuity in Assumption (iii),

$$\lim_{w \to \omega_t^{\text{obs}}(1,s)} F_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(w) = 1,$$

and therefore

$$\lim_{u \to \bar{u}_{1,s,t}} S_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}\big(Q_{w_{n,t}(1)|s_{n,t}=s}(u)\big) = 0. \tag{108}$$

Step 3: Exact tail matching. By the continuity and strict decrease of $u \mapsto S_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(Q_{w_{n,t}(1)|s_{n,t}=s}(u))$ on $(\tilde{u}_{1,s,t}, \bar{u}_{1,s,t})$ and its limit 0 as $u \to \bar{u}_{1,s,t}$, there exists $\tilde{\tau}_{1,s,t} \in (\tau^*_{1,s,t}, 1)$ such that for every $\tau \in (\tilde{\tau}_{1,s,t}, 1)$ there is a unique $u_{1,s,t}(\tau) \in (\tilde{u}_{1,s,t}, \bar{u}_{1,s,t})$ solving

$$S_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}\big(Q_{w_{n,t}(1)|s_{n,t}=s}(u_{1,s,t}(\tau))\big) = 1 - \tau.$$

Combining this with $S_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(\tau)) = 1 - \tau$ for all $\tau \in (\tau^*_{1,s,t}, 1)$ and the strict decrease of $w \mapsto S_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(w)$ on $(a_t^{\text{obs}}(1,s), \omega_t^{\text{obs}}(1,s))$ yields

$$Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(\tau) = Q_{w_{n,t}(1)|s_{n,t}=s}(u_{1,s,t}(\tau)) \qquad \text{for all } \tau \in (\tilde{\tau}_{1,s,t}, 1). \tag{109}$$

Moreover,

$$\lim_{\tau \to 1} u_{1,s,t}(\tau) = \bar{u}_{1,s,t}. \tag{110}$$

Step 4: Cross-x τ-alignment and the product identity. Fix $\bar{s} \in \mathcal{S}_t$ satisfying Assumption (iv). Let $\{\tau^{(k)}_{1,\bar{s},t}\}_{k\geq 1} \subset (0,1)$ be any sequence with $\tau^{(k)}_{1,\bar{s},t} \to 1$ as $k \to \infty$ and define the sequence $\{\tau^{(k)}_{1,s,t}\}_{k\geq 1} \subset (0,1)$ where

$$\tau^{(k)}_{1,s,t} := 1 - \frac{\Pr(D_{n,t} = 1 \mid s_{n,t} = \bar{s})}{\Pr(D_{n,t} = 1 \mid s_{n,t} = s)} \big(1 - \tau^{(k)}_{1,\bar{s},t}\big). \tag{111}$$

Let $u^{(k)}_{1,s,t} := u_{1,s,t}(\tau^{(k)}_{1,s,t})$ and $u^{(k)}_{1,\bar{s},t} := u_{1,\bar{s},t}(\tau^{(k)}_{1,\bar{s},t})$. Using (107) at $u = u^{(k)}_{1,s,t}$ and $u = u^{(k)}_{1,\bar{s},1}$,

$$1 - \tau^{(k)}_{1,s,t} = (1 - u^{(k)}_{1,s,t}) \frac{r_{1,s,t}(u^{(k)}_{1,s,t})}{\Pr(D_{n,t} = 1 \mid s_{n,t} = s)}, \qquad 1 - \tau^{(k)}_{1,\bar{s},t} = (1 - u^{(k)}_{1,\bar{s},t}) \frac{r_{1,\bar{s},t}(u^{(k)}_{1,\bar{s},t})}{\Pr(D_{n,t} = 1 \mid s_{n,t} = \bar{s})}.$$

Divide the two equalities and use (111) to obtain

$$\frac{(1 - u^{(k)}_{1,s,t})\, r_{1,s,t}(u^{(k)}_{1,s,t})}{(1 - u^{(k)}_{1,\bar{s},t})\, r_{1,\bar{s},t}(u^{(k)}_{1,\bar{s},t})} \;=\; 1. \tag{112}$$

Step 5: Aligning tail probabilities across states. Under Assumption (ii),

$$\lim_{k \to \infty} \frac{r_{1,s,t}(u^{(k)}_{1,s,t})}{r_{1,\bar{s},t}(u^{(k)}_{1,\bar{s},t})} = q_t(1) \in (0, \infty).$$

By (110), $u^{(k)}_{1,\bar{s},t} \to \bar{u}_{1,\bar{s},t}$ and $u^{(k)}_{1,s,t} \to \bar{u}_{1,s,t}$. Since $\bar{u}_{1,\bar{s},t}, \bar{u}_{1,s,t} < 1$ and $r_{1,\bar{s},t}(\cdot), r_{1,s,t}(\cdot)$ are continuous near those limits (by Assumption (iii)), (112) implies

$$\lim_{k\to\infty} \frac{1-u^{(k)}_{1,s,t}}{1-u^{(k)}_{1,\bar{s},t}} = 1,$$

and therefore

$$\lim_{k\to\infty} \left(u^{(k)}_{1,s,t} - u^{(k)}_{1,\bar{s},t}\right) = 0. \tag{113}$$

Step 6: Identification by differencing. By exogeneity of $\epsilon_{n,t}(1)$,

$$Q_{w_{n,t}(1)|s_{n,t}=s}(u) \;=\; \varphi(1,s) \;+\; Q_{\epsilon_{n,t}(1)}(u) \qquad \text{for all } u\in(0,1).$$

Apply (109) at $\tau=\tau^{(k)}_{1,s,t}$ and at $\tau=\tau^{(k)}_{1,\bar{s},t}$ to obtain

$$\begin{aligned} Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}\big(\tau^{(k)}_{1,s,t}\big) &= \varphi(1,s) + Q_{\epsilon_{n,t}(1)}(u^{(k)}_{1,s,t}),\\ Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=\bar{s}}\big(\tau^{(k)}_{1,\bar{s},t}\big) &= y(1,\bar{s}) + Q_{\epsilon_{n,t}(1)}(u^{(k)}_{1,\bar{s},t}). \end{aligned} \tag{114}$$

By (113) and continuity of $Q_{\epsilon_{n,t}(1)}$ near the upper tail,

$$\lim_{k\to\infty}\Big(Q_{\epsilon_{n,t}(1)}(u^{(k)}_{1,s,t}) - Q_{\epsilon_{n,t}(1)}(u^{(k)}_{1,\bar{s},t})\Big) = 0.$$

Subtracting the equations in (114) and using $\varphi(1,\bar{s})=0$ (Assumption (iv)) yields the identification result:

$$\lim_{k\to\infty}\Big[\,Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=s}\big(\tau^{(k)}_{1,s,t}\big) \;-\; Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=\bar{s}}\big(\tau^{(k)}_{1,\bar{s},t}\big)\,\Big] \;=\varphi(1,s). \quad \square$$

Remark. The only substantive difference between Proposition D.1 and Proposition G.1—apart from the support restrictions in Assumption (i)—is that Assumption (ii) in the unbounded case is replaced, in the bounded case, by a *relative tail decay* condition. For reference, Assumption (ii) of Proposition D.1 posits that there exists an (unknown) constant $q_t(d)\in(0,1]$ such that, for each $s\in\mathcal{S}_t$,

$$\lim_{w\to\omega_t(d,s)} \Pr\big(D_{n,t}=d \mid s_{n,t}=s, w_{n,t}(d)>w\big) = q_t(d). \tag{115}$$

The requirement (115) is too strong—and in fact necessarily violated—under a strict support gap $\omega^{\text{obs}}_t(d,s)<\omega_t(d,s)$ (Assumption (i) of Proposition G.1). To see this, Bayes' rule (Step 1 of the proof) implies, for any $s\in\mathcal{S}_t$ and any w,

$$S_{w_{n,t}|D_{n,t}=d,s_{n,t}=s}(w) = \; S_{w_{n,t}(d)|s_{n,t}=s}(w)\,\frac{\Pr(D_{n,t}=d\mid s_{n,t}=s,\; w_{n,t}(d)>w)}{\Pr(D_{n,t}=d\mid s_{n,t}=s)}.$$

For any $w\in(\omega^{\text{obs}}_t(d,s),\omega_t(d,s))$ we have $S_{w_{n,t}|D_{n,t}=d,s_{n,t}=s}(w)=0$ while $S_{w_{n,t}(d)|s_{n,t}=s}(w)>0$ and $\Pr(D_{n,t}=d\mid s_{n,t}=s)>0$, hence

$$\Pr(D_{n,t}=d\mid s_{n,t}=s,\; w_{n,t}(d)>w)=0 \quad \text{for all } w\in(\omega^{\text{obs}}_t(d,s),\omega_t(d,s)).$$

Therefore, the tail selection probability collapses to zero as $w \to \omega_t(d, s)$, forcing $q_t(d) = 0$ in (115). A positive limit could arise only in the case $\omega_t^{\text{obs}}(d, s) = \omega_t(d, s)$, which is excluded by Assumption (i). This is why we adopt a *relative* tail condition in place of (115), which governs the *rate* at which tail probabilities vanish across s (via ratios) rather than imposing a common nonzero limit that cannot hold under a support gap.

G.2.1 Identification of Support Endpoints

Corollary G.1 shows that, when finite, the lower and upper endpoints of the potential wages $w_{n,t}(d) \mid s_{n,t} = s$ and shock $\epsilon_{n,t}(d)$ are nonparametrically identified. Intuitively, for each s, the lower and upper endpoints of the observed, selected wage distribution, (denoted by $\underline{\omega}_t^{\text{obs}}(d, s)$ and $\omega_t^{\text{obs}}(d, s)$, respectively), are read directly from extremal quantiles: very low quantiles approach the lower endpoint and very high quantiles approach the upper endpoint. Because the deterministic part of wages $\varphi(d, s)$ is already known by Proposition G.1, we can "shift" the observed $\underline{\omega}_t^{\text{obs}}(d, s)$ and $\omega_t^{\text{obs}}(d, s)$ to learn about the latent lower and upper endpoints of both the shock $\epsilon_{n,t}(d)$ (denoted by $\underline{\omega}_\epsilon(d)$ and $\omega_\epsilon(d)$, respectively; time-invariant for simplicity) and the potential wage $w_{n,t}(d) = \varphi(d, s) + \epsilon_{n,t}(d)$ (denoted by $\underline{\omega}_t(d, s)$ and $\omega_t(d, s)$, respectively). Namely, selection trims the extremals, so the observed support sits inside the latent one: $\underline{\omega}_t^{\text{obs}}(d, s) \geq \underline{\omega}_t(d, s)$ and $\omega_t^{\text{obs}}(d, s) \leq \omega_t(d, s)$, with $\omega_t(d, s) = \varphi(d, s) + \omega_\epsilon(d)$ and $\underline{\omega}_t(d, s) = \varphi(d, s) + \underline{\omega}_\epsilon(d)$. Taking the best (tightest) such shifts across s gives bounds:

$$\sup_s\{\omega_t^{\text{obs}}(d, s) - \varphi(d, s)\} \leq \omega_\epsilon(d), \qquad \underline{\omega}_\epsilon(d) \leq \inf_s\{\underline{\omega}_t^{\text{obs}}(d, s) - \varphi(d, s)\},$$

and adding back $\varphi(d, s)$ yields corresponding tighest bounds for $\omega_t(d, s)$ and $\underline{\omega}_t(d, s)$. Moreover, if there exists a state value $s^\star$ where selection does *not* truncate the top ($\omega_t^{\text{obs}}(d, s^\star) = \omega_t(d, s^\star)$), the upper latent endpoint is revealed by the extremal quantile at $s^\star$:

$$\omega_\epsilon(d) = \lim_{\tau \to 1} \left\{ Q_{w_{n,t}|D_{n,t}=d, s_{n,t}=s^\star}(\tau) - \varphi(d, x^\star) \right\},$$

and then $\omega_t(d, s) = \varphi(d, s) + \omega_\epsilon(d)$ for every s. A symmetric argument applies to the lower endpoint if selection does not truncate the bottom at some $s^\dagger$.

Corollary G.1 (Identification of finite right and left endpoints of $\epsilon_{n,t}(d)$ and $w_{n,t}(d)$—Simplified Wage Equation (11))**.** *Let $t \in \{1, \ldots, T\}$ and $d \in \mathcal{D}$. Maintain the assumptions of Proposition G.1, implying that $\varphi(d, s)$ is identified for each $s \in \mathcal{S}_t$. In addition, assume finite and distinct endpoints, with two-sided tail regularity: for each $s \in \mathcal{S}_t$,*

$$\underline{\omega}_t(d, s) := \inf\{u : \Pr(w_{n,t}(d) \leq u \mid s_{n,t} = s) > 0\} > -\infty,$$
$$\omega_t(d, s) := \sup\{u : \Pr(w_{n,t}(d) \leq u \mid s_{n,t} = s) < 1\} < \infty,$$

$$\underline{\omega}_t^{\text{obs}}(d, s) := \inf\{u : \Pr(w_{n,t} \leq u \mid D_{n,t} = d, s_{n,t} = s) > 0\} > \underline{\omega}_t(d, s) > -\infty,$$
$$\omega_t^{\text{obs}}(d, s) := \sup\{u : \Pr(w_{n,t} \leq u \mid D_{n,t} = d, s_{n,t} = s) < 1\} < \omega_t(d, s) < \infty,$$

with $F_{w_{n,t}(d)|s_{n,t}=s}$ continuous and strictly increasing on $(\underline{\omega}_t(d, s),\, a_t(d, s)) \cup (\tilde{a}_t(d, s),\, \omega_t(d, s))$ for *some $a_t(d,s) < \tilde{a}_t(d,s)$, and $F_{w_{n,t}|D_{n,t}=d,s_{n,t}=s}$ continuous and strictly increasing on $(\underline{\omega}_t^{\text{obs}}(d,s),\ a_t^{\text{obs}}(d,s)) \cup ((\tilde{a}_t^{\text{obs}}(d,s)),\ \omega_t^{\text{obs}}(d,s))$ for some $a_t^{\text{obs}}(d,s) < (\tilde{a}_t^{\text{obs}}(d,s))$, as well as continuous at both endpoints:*

$$\lim_{w \to \underline{\omega}_t^{\text{obs}}(d,s)} F_{w_{n,t}|D_{n,t}=d,s_{n,t}=s}(w) = 0, \qquad \lim_{w \to \omega_t^{\text{obs}}(d,s)} F_{w_{n,t}|D_{n,t}=d,s_{n,t}=s}(w) = 1.$$

Define the shock $\epsilon_{n,t}(d)$ (finite) endpoints as:

$$\underline{\omega}_\epsilon(d) \coloneqq \inf\{u \in \mathbb{R} : F_{\epsilon_{n,t}(d)}(u) > 0\} > -\infty, \qquad \omega_\epsilon(d) \coloneqq \sup\{u \in \mathbb{R} : F_{\epsilon_{n,t}(d)}(u) < 1\} < \infty.$$

Then:

(a) (Observed wage endpoints are identified.) For every $s \in \mathcal{S}_t$, $\underline{\omega}_t^{\text{obs}}(d,s)$ and $\omega_t^{\text{obs}}(d,s)$ are identified:

$$\underline{\omega}_t^{\text{obs}}(d,s) = \lim_{\tau \to 0} Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=s}(\tau), \qquad \omega_t^{\text{obs}}(d,s) = \lim_{\tau \to 1} Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=s}(\tau).$$

(b) (Sharp bounds for latent endpoints.) For every $s \in \mathcal{S}_t$, a lower (resp. upper bound) bound for $\omega_\epsilon(d)$ (resp. $\underline{\omega}_\epsilon(d)$) and a lower (resp. upper bound) bound for $\omega_t(d,s)$ (resp. $\underline{\omega}_t(d,s)$) are identified:

$$L_\epsilon(d) \coloneqq \sup_s \left\{\omega_t^{\text{obs}}(d,s) - \varphi(d,s)\right\} \ \le\ \omega_\epsilon(d),\ U_\epsilon(1) \coloneqq \inf_s \left\{\underline{\omega}_t^{\text{obs}}(d,s) - \varphi(d,s)\right\} \ge \underline{\omega}_\epsilon(d),$$
$$\underline{\omega}_t(d,s) \le \min\Big\{\ \underline{\omega}_t^{\text{obs}}(d,s), \varphi(d,s) + U_\epsilon(d)\Big\}, \omega_t(d,s) \ge \max\Big\{\ \omega_t^{\text{obs}}(d,s), \varphi(d,s) + L_\epsilon(d)\ \Big\}.$$

Moreover, these bounds are sharp under the stated assumptions.

(c) (Upper endpoint point identification under no top truncation at some $s^\star$.) If there exists a known $s^\star \in \mathcal{S}_t$ with $\omega_t^{\text{obs}}(d,s^\star) = \omega_t(d,s^\star)$ (i.e., the finite upper endpoint of the selected observed wages equals the finite lower endpoint of the potential wages; in other words, selection does not affect the rightmost support of wages at s^*), then, for every $s \in \mathcal{S}_t$, $\omega_\epsilon(d)$ and $\omega_t(d,s)$ are identified:

$$\omega_\epsilon(d) = \lim_{\tau \to 1} \Big[Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=s^\star}(\tau) - \varphi(d,s^\star)\Big],$$
$$\omega_t(d,s) = \varphi(d,s) + \lim_{\tau \to 1} \Big[Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=s^\star}(\tau) - \varphi(d,s^\star)\Big].$$

(d) (Lower endpoint point identification under no bottom truncation at some $s^\dagger$.) If there exists a known $s^\dagger \in \mathcal{S}_t$ with $\underline{\omega}_t^{\text{obs}}(d,s^\dagger) = \underline{\omega}_t(d,s^\dagger)$ (i.e. the finite lower endpoint of the selected observed wages equals the finite left endpoint of the potential wages; in other words, selection does not affect the leftmost support of wages at $s^\dagger$), then, for every $s \in \mathcal{S}_t$, $\underline{\omega}_\epsilon(d)$ and $\underline{\omega}_t(d,s)$ are identified:

$$\underline{\omega}_\epsilon(d) = \lim_{\tau \to 0} \Big[Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=s^\dagger}(\tau) - \varphi(d,s^\dagger)\Big],$$
$$\underline{\omega}_t(d,s) = \varphi(d,s) + \lim_{\tau \to 0} \Big[Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=s^\dagger}(\tau) - \varphi(d,s^\dagger)\Big].$$

Proof: We present the proof for the upper endpoints; the argument for the lower endpoints is symmetric. Without loss, consider $d = 1$.

(a) Fix any $s \in \mathcal{S}_t$. By Assumption (iii) of Proposition G.1, $F_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}$ is continuous and strictly increasing on $(a_t^{\text{obs}}(1, s), \omega_t^{\text{obs}}(1, s))$ and continuous at the endpoint. Therefore, its upper quantiles converge to the endpoint, yielding *(a)*.

(b) Fix any $s \in \mathcal{S}_t$. It holds that

$$\omega_t(1, s) \; = \; \varphi(1, s) + \omega_\epsilon(1). \tag{116}$$

By Assumption (ii) of Proposition G.1, $\omega_t^{\text{obs}}(1, s) < \omega_t(1, s)$, so

$$\omega_t^{\text{obs}}(1, s) - \varphi(1, s) \; < \; \omega_t(1, s) - \varphi(1, s) \; = \; \omega_\epsilon(1).$$

Taking the supremum over s yields a lower bound for $\omega_\epsilon(1)$. Adding $\varphi(1, s)$ gives a lower bound for $\omega_t(1, s)$. These bounds are the best possible (sharp) without further restrictions.

(c) Fix any $s \in \mathcal{S}_t$. If there exists a known $s^\star \in \mathcal{S}_t$ with $\omega_t^{\text{obs}}(1, s^\star) = \omega_t(1, s^\star)$, then by *(a)*, $\omega_t(1, s^\star)$ is identified:

$$\omega_t(1, s^\star) = \lim_{\tau \to 1} Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=s^\star}(\tau).$$

Using (116) written for s^* and recalling that $\varphi(1, s^\star)$ is identified by Proposition G.1 gives $\omega_\epsilon(1) = \omega_t(1, s^\star) - \varphi(1, s^\star)$. We plug this into (116) and complete the proof. $\square$

G.3 Proposition D.1 with Location and Scale Parameters

Propositions D.1 and G.1 extend to wage specifications in which the shock $\epsilon_{n,t}(d)$ is multiplied by a scale parameter $\sigma(d, s_{n,t}) > 0$:

$$w_{n,t} = \sum_{d \in \{0,1\}} \mathbb{1}\{D_{n,t} = d\} \big[\varphi(d, s_{n,t}) + \sigma(d, s_{n,t}) \epsilon_{n,t}(d) \big]. \tag{117}$$

Proposition G.2 (Identification of $\varphi(d, \cdot)$ and $\sigma(d, \cdot)$—Simplified Wage Equation (117))**.** *Let $t \in \{1, \ldots, T\}$ and $d \in \mathcal{D}$. Assume that the conditional wage distribution $\Pr\big(w_{n,t} \leq w \,\big|\, D_{n,t} = d, \; s_{n,t} = s\big)$ is identified for each $w \in \mathbb{R}$ and $s \in \mathcal{S}_t$ (see Proposition A.5 for sufficient conditions), and that the conditional choice probability $\Pr(D_{n,t} = d \mid s_{n,t} = s)$ is identified for each $s \in \mathcal{S}_t$ (see Proposition A.4 for sufficient conditions). Moreover, assume:*

(i) (Supports.)[55] *For each $s \in \mathcal{S}_t$, $\omega_t(d, s) := \sup\{u : \Pr(w_{n,t}(d) \leq u \mid s_{n,t} = s) < 1\} = \infty$ and $\omega_t^{\text{obs}}(d, s) := \sup\{u : \Pr(w_{n,t} \leq u \mid D_{n,t} = d, s_{n,t} = s) < 1\} = \infty$.*

(ii) (Tail Limit.) There exists an (unknown) constant $q_t(d) \in (0, 1]$ such that for every $s \in \mathcal{S}_t$, $\lim_{w \to \infty} \Pr\big(D_{n,t} = d \mid s_{n,t} = s, w_{n,t}(1) > w\big) = q_t(d)$.

(iii) (Tail Regularity.) For each $s \in \mathcal{S}_t$, there exist (unknown) thresholds $a_t(d, s) < \infty$ and $a_t^{\text{obs}}(d, s) <$

[55] We focus on the case of unbounded supports. The bounded-support case follows analogously, with the technical modifications highlighted in Appendix G.2 of the Online Supplementary Material.

∞ such that the cumulative distribution functions $F_{w_{n,t}(d)|s_{n,t}=s}$ and $F_{w_{n,t}|D_{n,t}=d,s_{n,t}=s}$ are continuous and strictly increasing on $(a_t(d,s),\infty)$ and $(a_t^{\text{obs}}(d,s),\infty)$, respectively.

(iv) (Location and Scale Normalizations.) There exists a known $\bar{s} \in \mathcal{S}_t$ with $\varphi(d,\bar{s}) = 0$ and $\sigma(d,\bar{x}) = 1$.

For each $s \in \mathcal{S}_t$, fix the following sequences

$$\tau_{d,\bar{s},t}^{(k)} := 1 - 2^{-k}, \qquad \tilde{\tau}_{d,\bar{s},t}^{(k)} := 1 - 3^{-k}, \qquad k = 1, 2, \ldots$$

and let

$$\tau_{d,s,t}^{(k)} := 1 - \frac{\Pr(D_{n,t} = d \mid s_{n,t} = \bar{s})}{\Pr(D_{n,t} = d \mid s_{n,t} = s)}\big(1 - \tau_{d,\bar{s},t}^{(k)}\big), \qquad \tilde{\tau}_{d,s,t}^{(k)} := 1 - \frac{\Pr(D_{n,t} = d \mid s_{n,t} = \bar{s})}{\Pr(D_{n,t} = d \mid s_{n,t} = s)}\big(1 - \tilde{\tau}_{d,\bar{s},t}^{(k)}\big).$$

Then,

$$\sigma(d,s) = \lim_{k\to\infty} \frac{Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=s}\big(\tau_{d,s,t}^{(k)}\big) - Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=s}\big(\tilde{\tau}_{d,s,t}^{(k)}\big)}{Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=\bar{s}}\big(\tau_{d,\bar{s},t}^{(k)}\big) - Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=\bar{s}}\big(\tilde{\tau}_{d,\bar{s},t}^{(k)}\big)},$$

and

$$\varphi(d,s) = \lim_{k\to\infty} \Big[Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=s}\big(\tau_{d,s,t}^{(k)}\big) - \sigma(d,s)\, Q_{w_{n,t}|D_{n,t}=d,s_{n,t}=\bar{s}}\big(\tau_{d,\bar{s},t}^{(k)}\big)\Big].$$

Hence $\varphi(d,s)$ and $\sigma(d,s)$ are identified.

Proof: Let $s \in \mathcal{S}_t$, and without loss, consider firm $d = 1$. For any threshold w, Bayes' rule gives

$$\Pr(w_{n,t} > w \mid D_{n,t} = 1, s_{n,t} = s) = \frac{\Pr(w_{n,t}(1) > w \mid s_{n,t} = s)\ \Pr(D_{n,t} = 1 \mid s_{n,t} = s, w_{n,t}(1) > w)}{\Pr(D_{n,t} = 1 \mid s_{n,t} = s)}.$$

Letting $w \to \infty$ and using Assumption (ii),

$$\Pr(w_{n,t} > w \mid D_{n,t} = 1, s_{n,t} = s) \ \sim\ c_t(1,s)\ \Pr\big(w_{n,t}(1) > w \mid s_{n,t} = s\big) \qquad (w \to \infty), \quad (118)$$

where

$$c_t(1,s) \ :=\ \frac{q_t(1)}{\Pr(D_{n,t} = 1 \mid s_{n,t} = s)} \ \in (0,\infty),$$

and "$\sim$" denotes that the ratio of the two sides converges to 1.
Write $S_{1,s,t}(w) := S_{w_{n,t}(1)|s_{n,t}=s}(w)$ and $S_{1,s,t}^{\text{obs}}(w) := S_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(w)$. Then, (118) reads as

$$S_{1,s,t}^{\text{obs}}(w) \ \sim\ c_t(1,s)\, S_{1,s,t}(w) \qquad (w \to \infty). \quad (119)$$

By Assumption (i), both right wage endpoints are ∞; by Assumption (iii), the upper-tail CDFs $F_{w_{n,t}(1)|s_{n,t}=s}$ and $F_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}$ are continuous and strictly increasing beyond finite thresholds, so their tail quantile maps are the ordinary inverses on the corresponding index ranges near 1. Hence, by Lemma D.1 (stated and proved in Appendix D),

$$Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=s}(\tau) \ =\ Q_{w_{n,t}(1)\,|\,s_{n,t}=s}\Big(1 - \frac{1-\tau}{c_t(1,s)} \ +\ o_s(1-\tau)\Big) \qquad (\tau \to 1), \quad (120)$$

where $o_s(1-\tau)/(1-\tau) \to 0$. By $w_{n,t}(1) = \varphi(1,s) + \sigma(1,s)\epsilon_{n,t}(1)$ and the exogeneity of $\epsilon_{n,t}(1)$

$$Q_{w_{n,t}(1)\,|\,s_{n,t}=s}(u) \;=\; \varphi(1,s) + \sigma(1,s)Q_{\epsilon_{n,t}(1)}(u)$$

for all $u \in (0,1)$. Plugging into (120) gives

$$Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=s}(\tau) \;=\; \varphi(1,s) \,+\, \sigma(1,s)Q_{\epsilon_{n,t}(1)}\Big(1 - \frac{1-\tau}{c_t(1,s)} \,+\, o_s(1-\tau)\Big) \qquad (\tau \to 1). \quad (121)$$

Scale. Evaluate (121) at $\tau = \tau^{(k)}_{1,s,t}$ and $\tau = \tilde{\tau}^{(k)}_{1,s,t}$:

$$\begin{aligned} Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=s}\big(\tau^{(k)}_{1,s,t}\big) &= \varphi(1,s) + \sigma(1,s)Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tau^{(k)}_{1,s,t}}{c_t(1,s)} + o_s(1-\tau^{(k)}_{1,s,t})\Big), \\ Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=s}\big(\tilde{\tau}^{(k)}_{1,s,t}\big) &= \varphi(1,s) + \sigma(1,s)Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tilde{\tau}^{(k)}_{1,s,t}}{c_t(1,s)} + o_x(1-\tilde{\tau}^{(k)}_{1,s,t})\Big), \end{aligned} \qquad (k \to \infty). \quad (122)$$

Take the difference between the two equations in (122):

$$\begin{aligned} \Delta^{(k)}_{1,s,t} &\coloneqq Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(\tau^{(k)}_{1,s,t}) - Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(\tilde{\tau}^{(k)}_{1,s,t}) \\ &= \sigma(1,s)\Big[Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tau^{(k)}_{1,s,t}}{c_t(1,s)} + o_s(1-\tau^{(k)}_{1,s,t})\Big) - Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tilde{\tau}^{(k)}_{1,s,t}}{c_t(1,s)} + o_x(1-\tilde{\tau}^{(k)}_{1,s,t})\Big)\Big] \;(k \to \infty). \end{aligned} \quad (123)$$

Repeat analogous steps for $\tau = \tau^{(k)}_{1,\bar{s},t}$ and $\tau = \tilde{\tau}^{(k)}_{1,\bar{s},t}$ and use the normalizations in Assumption (iv):

$$\begin{aligned} \Delta^{(k)}_{1,\bar{s},t} &\coloneqq Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=\bar{s}}(\tau^{(k)}_{1,\bar{s},t}) - Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=\bar{s}}(\tilde{\tau}^{(k)}_{1,\bar{s},t}) \\ &= \Big[Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})} + o_{\bar{s}}(1-\tau^{(k)}_{1,\bar{s},t})\Big) - Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tilde{\tau}^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})} + o_{\bar{s}}(1-\tilde{\tau}^{(k)}_{1,\bar{s},t})\Big)\Big] \qquad (k \to \infty). \end{aligned} \quad (124)$$

By the definition of $\tau^{(k)}_{1,s,t}$ and $\tilde{\tau}^{(k)}_{1,s,t}$,

$$1 - \frac{1-\tau^{(k)}_{1,s,t}}{c_t(1,s)} = 1 - \frac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})}, \qquad 1 - \frac{1-\tilde{\tau}^{(k)}_{1,s,t}}{c_t(1,s)} = 1 - \frac{1-\tilde{\tau}^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})}.$$

Thus, as $k \to \infty$, (123) and (124) can be written as

$$\begin{aligned} \Delta^{(k)}_{1,s,t} &= \sigma(1,s)\Big[Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})} + o_s(1-\tau^{(k)}_{1,s,t})\Big) - Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tilde{\tau}^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})} + o_x(1-\tilde{\tau}^{(k)}_{1,s,t})\Big)\Big] \\ \Delta^{(k)}_{1,\bar{s},t} &= \Big[Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})} + o_{\bar{s}}(1-\tau^{(k)}_{1,\bar{s},t})\Big) - Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tilde{\tau}^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})} + o_{\bar{s}}(1-\tilde{\tau}^{(k)}_{1,\bar{s},t})\Big)\Big] \end{aligned}. \quad (125)$$

Take the ratio between the two equations in (125). Assumption (iii) implies that $Q_{\epsilon_{n,t}(1)}$ is continuous and strictly increasing near 1. Since $\tau^{(k)}_{1,\bar{s},t} = 1 - 2^{-k}$ and $\tilde{\tau}^{(k)}_{1,\bar{s},t} = 1 - 3^{-k}$ are distinct for all k, the denominator of the ratio is nonzero for all large k. By continuity and $o_s(1-\tau^{(k)}_{1,s,t}), o_{\bar{s}}(1-\tau^{(k)}_{1,\bar{s},t}) \to 0$,

$$\lim_{k\to\infty} \frac{Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})} + o_s(1-\tau^{(k)}_{1,s,t})\Big) - Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tilde{\tau}^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})} + o_x(1-\tilde{\tau}^{(k)}_{1,s,t})\Big)}{Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})} + o_{\bar{s}}(1-\tau^{(k)}_{1,\bar{s},t})\Big) - Q_{\epsilon_{n,t}(1)}\Big(1 - \tfrac{1-\tilde{\tau}^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})} + o_{\bar{s}}(1-\tilde{\tau}^{(k)}_{1,\bar{s},t})\Big)}.$$

Therefore,

$$\sigma(1,s)=\lim_{k\to\infty}\frac{Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(\tau^{(k)}_{1,s,t})-Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=s}(\tilde{\tau}^{(k)}_{1,s,t})}{Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=\bar{s}}(\tau^{(k)}_{1,\bar{s},t})-Q_{w_{n,t}|D_{n,t}=1,s_{n,t}=\bar{s}}(\tilde{\tau}^{(k)}_{1,\bar{s},t})}.$$

Location. Evaluate (121) at $\tau=\tau^{(k)}_{1,s,t}$ and, with $s=\bar{s}$, at $\tau=\tau^{(k)}_{1,\bar{s},t}$:

$$\left.\begin{aligned}Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=s}\big(\tau^{(k)}_{1,s,t}\big)&=\varphi(1,s)+\sigma(1,s)Q_{\epsilon_{n,t}(1)}\Big(1-\tfrac{1-\tau^{(k)}_{1,s,t}}{c_t(1,s)}+o_s(1-\tau^{(k)}_{1,s,t})\Big),\\ Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=\bar{s}}\big(\tau^{(k)}_{1,\bar{s},t}\big)&=Q_{\epsilon_{n,t}(1)}\Big(1-\tfrac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})}+o_{\bar{s}}(1-\tau^{(k)}_{1,\bar{s},t})\Big),\end{aligned}\right\}\quad(k\to\infty), \tag{126}$$

where we use the normalizations in Assumption (iv). By the definition of $\tau^{(k)}_{1,s,t}$,

$$1-\frac{1-\tau^{(k)}_{1,s,t}}{c_t(1,s)}=1-\frac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})}.$$

Therefore, (126) can be written as

$$\left.\begin{aligned}Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=s}\big(\tau^{(k)}_{1,s,t}\big)&=\varphi(1,s)+\sigma(1,s)Q_{\epsilon_{n,t}(1)}\Big(1-\tfrac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})}+o_s(1-\tau^{(k)}_{1,s,t})\Big),\\ Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=\bar{s}}\big(\tau^{(k)}_{1,\bar{s},t}\big)&=Q_{\epsilon_{n,t}(1)}\Big(1-\tfrac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})}+o_{\bar{s}}(1-\tau^{(k)}_{1,\bar{s},t})\Big),\end{aligned}\right\}\quad(k\to\infty). \tag{127}$$

As $k\to\infty$, subtracting the two equations in (127):

$$\begin{aligned}&Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=s}\big(\tau^{(k)}_{1,s,t}\big)-\sigma(1,s)Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=\bar{s}}\big(\tau^{(k)}_{1,\bar{s},t}\big)\\ &=\varphi(1,s)+\sigma(1,s)\Big[Q_{\epsilon_{n,t}(1)}\Big(1-\tfrac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})}+o_s(1-\tau^{(k)}_{1,s,t})\Big)-Q_{\epsilon_{n,t}(1)}\Big(1-\tfrac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})}+o_{\bar{s}}(1-\tau^{(k)}_{1,\bar{s},t})\Big)\Big].\end{aligned}$$

Also note that $o_s(1-\tau^{(k)}_{1,s,t})\to 0$ and $o_{\bar{s}}(1-\tau^{(k)}_{1,\bar{s},t})\to 0$ as $k\to\infty$. Therefore, by continuity of $Q_{\epsilon_{n,t}(1)}$ near 1 under Assumption (iii),

$$Q_{\epsilon_{n,t}(1)}\Big(u+o_s(1-\tau^{(k)}_{1,s,t})\Big)-Q_{\epsilon_{n,t}(1)}\Big(u+o_{\bar{s}}(1-\tau^{(k)}_{1,\bar{s},t})\Big)=o(1),\qquad u:=1-\frac{1-\tau^{(k)}_{1,\bar{s},t}}{c_t(1,\bar{s})}\qquad(k\to\infty).$$

Therefore, as desired,

$$\lim_{k\to\infty}\Big[\,Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=s}\big(\tau^{(k)}_{1,s,t}\big)\;-\;\sigma(1,s)Q_{w_{n,t}\,|\,D_{n,t}=1,s_{n,t}=\bar{s}}\big(\tau^{(k)}_{1,\bar{s},t}\big)\Big]=\varphi(1,s).\quad\square$$

H Extension to Search Models

The quantile approach in Proposition G.2 can be used to identify key parameters of equilibrium wage equations with inherent conditional heteroskedasticity as in standard search models. For example, in the spirit of Bagger et al. (2014), consider the potential wage of worker n at time t in firm $d\in\mathcal{D}$ is

$$w_{n,t}(d)=\omega\,\gamma_d^{\alpha_d}\,H_{n,t}^{1-\alpha_d}\,\epsilon_{n,t}(d)\;+\;(1-\omega)(1-\delta)\,U_d(H_{n,t}), \tag{128}$$

where $0 < \omega < 1$ is the worker's bargaining weight, $\gamma_d > 0$ is firm d's productivity, $\alpha_d \in [0,1)$ is the elasticity of wages with respect to γ_d, $\delta \in (0,1)$ is the discount factor, $H_{n,t}$ denotes a worker's human capital at time t with support $\mathcal{H}_t$, and $U_d(H_{n,t})$ is the value of unemployment,

$$U_d(H_{n,t}) \coloneqq z \ + \delta\, \mathbb{E}_{\epsilon_{n,t} \sim F_{\epsilon_{n,t}}} \left[f\big(S\big(H_{n,t}, \epsilon_{n,t};\, \omega, \alpha_d, \gamma_d, \delta\big)\big) \right],$$

where z is the flow value of unemployment; $S(\cdot; \omega, \alpha_d, \gamma_d, \delta)$ is the match surplus; $f(\cdot)$ is a functional of the surplus; and the expectation is taken with respect to the shocks $\epsilon_{n,t} \coloneqq (\epsilon_{n,t}(d) : d \in \mathcal{D})$ with distribution $F_{\epsilon_{n,t}}$. The index d on U_d reflects the dependence of S on (α_d, γ_d). As is standard, we treat δ and ω as known. The parameters to be identified are then γ_d, α_d, and z. The functions f and S are known up to $(\gamma_d, \alpha_d, F_{\epsilon_{n,t}})$. We consider two cases: (1) $H_{n,t}$ is observed (or unobserved with known distribution and support); (2) $H_{n,t}$ is unobserved with unknown distribution and support. Hereafter, we denote by $F_{\epsilon_{n,t}(d)}$ the marginal CDF of $\epsilon_{n,t}(d)$ and by $S_{\epsilon_{n,t}(d)}$ the survival function of $\epsilon_{n,t}(d)$.

Remark. Hereafter, for simplicity, we focus on the case in which observed and potential wages, as well as shocks, have unbounded support. The bounded-support case follows analogously, with the technical modifications highlighted in Proposition G.1 of Appendix G.2 of the Online Supplementary Material. In that case, it is also possible to nonparametrically identify the finite lower and upper endpoints of potential wages and shocks, as shown in Corollary G.1.

Case 1: $H_{n,t}$ is Observed or Unobserved with Known Distribution and Support. As a preview, Proposition H.1, Proposition H.2, and Corollary H.1 stated below follow immediately from Proposition G.2 and Proposition D.2. For Proposition H.1, replace $s_{n,t}$ with $H_{n,t}$ and define

$$y(d, H_{n,t}) \coloneqq (1-\omega)(1-\delta)\, U_d(H_{n,t}) \ \text{ and } \ \sigma(d, H_{n,t}) \coloneqq \omega\, \gamma_d^{\alpha_d}\, H_{n,t}^{1-\alpha_d}.$$

Then, Proposition G.2 identifies $y(d, H_{n,t})$ and $\sigma(d, H_{n,t})$. As for Proposition H.2, once $y(d, H_{n,t})$ and $\sigma(d, H_{n,t})$ are identified, the joint distribution of the shock vector, $F_{\epsilon_{n,t}}$, is identified under the conditions of Proposition D.2. As for Corollary H.1, once $y(d, H_{n,t})$, $\sigma(d, H_{n,t})$, and $F_{\epsilon_{n,t}}$ are identified, the parameters α_d, γ_d, and z are straightforward to recover.

Proposition H.1 (Identification of $y(d, H_{n,t})$ and $\sigma(d, H_{n,t})$)**.** *For each firm $d \in \mathcal{D}$ and period $t \in \{1, \ldots, T\}$, assume:*

(i) (Supports.) For each $h \in \mathcal{H}_t$, $\sup\{u : \Pr(w_{n,t}(d) \le u \mid H_{n,t} = h) < 1\} = \infty$, $\sup\{u : \Pr(w_{n,t} \le u \mid D_{n,t} = d, H_{n,t} = h) < 1\} = \infty$, and $0 < \Pr(D_{n,t} = d \mid H_{n,t} = h) \le 1$.

(ii) (Tail Limit.) There exists a constant $q_t(d) \in (0,1]$ such that for every $h \in \mathcal{H}_t$, $\lim_{w \to \infty} \Pr\big(D_{n,t} = d \mid H_{n,t} = h, w_{n,t}(d) > w\big) = q_t(d)$.

(iii) (Tail Regularity.) For each $h \in \mathcal{H}_t$, there exist thresholds $w_t(d,h) < \infty$ and $w_t^{\text{obs}}(d,h) < \infty$ such that the cumulative distribution functions $F_{w_{n,t}(d) \mid H_{n,t}=h}$ and $F_{w_{n,t} \mid D_{n,t}=d, H_{n,t}=h}$ are continuous and strictly increasing on $(w_t(d,h), \infty)$ and $(w_t^{\text{obs}}(d,h), \infty)$, respectively.

(iv) (Normalization.) There exists a known $\bar{h} \in \mathcal{H}_t$ with $y(d, \bar{h}) = 0$ and $\sigma(d, \bar{h}) = 1$.

Then, $y(d,h)$ and $\sigma(d,h)$ are identified for each $d \in \mathcal{D}$, $h \in \mathcal{H}_t$, and $t \in \{1,\dots,T\}$.

Proposition H.2 (Identification of $F_{\epsilon_{n,t}}$). *Suppose that Assumptions (i) to (iv) of Proposition H.1 hold so that $y(d,h)$ and $\sigma(d,h)$ are identified for each $d \in \mathcal{D}$, $h \in \mathcal{H}_t$, and $t \in \{1,\dots,T\}$. Moreover, for each firm $d \in \mathcal{D}$ and period $t \in \{1,\dots,T\}$, assume $\epsilon_{n,t}(d)$ belongs to a parametric family indexed by the $p_{t,d} \times 1$ vector of parameters $\mu_{t,d}$ that is tail-ratio identifiable. Namely, for any $p_{t,d}+1$ distinct large thresholds $0 < e_0 < e_1 < \cdots < e_{p_{t,d}+1}$, the map*

$$\mu_{t,d} \mapsto \left(\frac{S_{\epsilon_{n,t}(d)}(e_1;\mu_{t,d})}{S_{\epsilon_{n,t}(d)}(e_0;\mu_{t,d})}, \dots, \frac{S_{\epsilon_{n,t}(d)}(e_{p_{t,d}};\mu_{t,d})}{S_{\epsilon_{n,t}(d)}(e_0;\mu_{t,d})} \right),$$

is injective. Under these assumptions, for each period $t \in \{1,\dots,T\}$:[56]

(a) (Marginal Identification.) For each firm $d \in \mathcal{D}$, the parameter $\mu_{t,d}$ is identified.

(b) (Joint Identification.) If the shocks $\{\epsilon_{n,t}(d)\}_{d\in\mathcal{D}}$ are mutually independent across $d \in \mathcal{D}$, then the joint distribution of $\epsilon_{n,t}$ is identified as the product of the identified marginals. Alternatively, if a copula C_{μ_t} is specified so that

$$F_{\epsilon_{n,t}}(v_1,\dots,v_{|\mathcal{D}|}) \;=\; C_{\mu_t}\big(F_{\epsilon_{n,t}(1)}(v_1;\mu_{t,1}),\dots,F_{\epsilon_{n,t}(|\mathcal{D}|)}(v_{|\mathcal{D}|};\mu_{t,|\mathcal{D}|})\big) \quad \forall\,(v_1,\dots,v_{|\mathcal{D}|}) \in \mathbb{R}^{|\mathcal{D}|},$$

with μ_t known, then the joint distribution is identified via the identified marginals and C_{μ_t}. Absent further restrictions on the dependence among $\{\epsilon_{n,t}(d)\}_{d\in\mathcal{D}}$, the joint CDF is partially identified by the sharp Fréchet–Höffding bounds, namely, for all $(v_1,\dots,v_{|\mathcal{D}|}) \in \mathbb{R}^{|\mathcal{D}|}$,

$$\max\Big\{ \sum_{d\in\mathcal{D}} F_{\epsilon_{n,t}(d)}(v_d;\mu_{t,d}) - (|\mathcal{D}|-1),\; 0 \Big\} \;\le\; F_{\epsilon_{n,t}}(v_1,\dots,v_{|\mathcal{D}|}) \;\le\; \min_{d\in\mathcal{D}} F_{\epsilon_{n,t}(d)}(v_d;\mu_{t,d}).$$

Corollary H.1 (Identification of α_d, γ_d, and z). *Assume that $y(d,h)$ and $\sigma(d,h)$ are identified for each $d \in \mathcal{D}$, for every realization h of $H_{n,t}$, and for some period $t \geq 1$ (see Proposition H.1 for sufficient conditions). Assume also that the joint distribution of the shock vector $F_{\epsilon_{n,t}}$ is identified for the same period $t \geq 1$ (see Proposition H.2 for sufficient conditions). Then, the parameters α_d, γ_d, and z are identified for each $d \in \mathcal{D}$.*

[56] Note that our identification result in Proposition H.2 is established period by period. Accordingly, the shock vectors $\{\epsilon_{n,t}\}_{t=1}^{T}$ are allowed to be correlated across periods, and we impose no restrictions on their time dependence. If one is willing to assume that $\epsilon_{n,t}$ follows a (covariance-)stationary VAR(p) process,

$$\epsilon_{n,t} = A_0 + A_1\epsilon_{n,t-1} + \cdots + A_p\epsilon_{n,t-p} + u_{n,t},$$

and that the (population) mean μ and autocovariances Γ_j for $j = 0,\dots,p$ are known, then the autoregressive coefficient matrices $A_1,\dots,A_p$ (and hence $A_0 = \big(I - \sum_{j=1}^{p} A_j\big)\mu$) are identified under standard conditions such as orthogonality/whiteness of $u_{n,t}$ with respect to past $\epsilon_{n,t-h}$ and the usual rank conditions. Note, however, that without parametric assumptions on the distribution of $u_{n,t}$, the marginal distribution of each $\epsilon_{n,t}$ is not identified even given knowledge of μ, Γ_j, and A_j for $j = 0,\dots,p$. To recover these marginal distributions, one can instead rely on Proposition H.2.

Proof: The proof proceeds in three steps.
Step 1: Identification of α_d from $\sigma(d, H_{n,t})$. For any h, h' in $\mathcal{H}_t$,

$$\frac{\sigma(d,h)}{\sigma(d,h')} = \frac{\omega\,\gamma_d^{\alpha_d}\,h^{1-\alpha_d}}{\omega\,\gamma_d^{\alpha_d}\,(h')^{1-\alpha_d}} = \left(\frac{h}{h'}\right)^{1-\alpha_d}.$$

Taking logarithms and rearranging,

$$\alpha_d \;= 1 \;-\; \frac{\log\big(\sigma(d,h)/\sigma(d,h')\big)}{\log(h/h')}.$$

Note that α_d is identified *without* relying on the scale normalization $\sigma(d, \bar{h}) = 1$ in Assumption (iv) of Proposition H.1, because the ratio $\sigma(d, h)/\sigma(d, h')$ is identified without any such normalization, as shown in the proof of Proposition G.2. Also, to identify α_d, we do not need to know the distribution of $\epsilon_{n,t}$, namely, $F_{\epsilon_{n,t}}$.
Step 2: Identification of γ_d from $\sigma(d, H_{n,t})$. Recall that for any $h \in \mathcal{H}_t$, we have that $\sigma(d, h) = \omega\,\gamma_d^{\alpha_d}\,h^{1-\alpha_d}$. Solving for γ_d yields that

$$\gamma_d \;= \big(\sigma(d,h)\omega^{-1}\,h^{\alpha_d-1}\big)^{1/\alpha_d}.$$

Note that, unlike α_d, the identification of γ_d relies on knowing the *level* $\sigma(d, h)$ and therefore depends on the scale normalization $\sigma(d, \bar{h}) = 1$ in Assumption (iv) of Proposition H.1. As is the case with α_d, to identify γ_d, we do not need to know $F_{\epsilon_{n,t}}$.
Step 3: Identification of z from $y(d, H_{n,t})$ and $F_{\epsilon_{n,t}}$. For any $h \in \mathcal{H}_t$, we have

$$y(d,h) \;= (1-\delta)(1-\omega)\Big[z \;+\delta\,\mathbb{E}_{\epsilon_{n,t}\sim F_{\epsilon_{n,t}}}\big[\,f\big(S\big(h,\epsilon_{n,t};\,\omega,\alpha_d,\gamma_d,\delta\big)\big)\big]\Big].$$

Solving for z yields that z is identified as

$$z \;= \frac{y(d,h)}{(1-\delta)(1-\omega)} \;-\; \delta\,\mathbb{E}_{\epsilon_{n,t}\sim F_{\epsilon_{n,t}}}\big[\,f\big(S\big(h,\epsilon_{n,t};\,\omega,\alpha_d,\gamma_d,\delta\big)\big)\big]\,.$$

The identification of z relies on knowing the *level* $y(d, h)$ and therefore depends on the location normalization $y(d, \bar{h}) = 0$ in Assumption (v) of Proposition H.1. Differently from α_d and γ_d, to identify z we need to know $F_{\epsilon_{n,t}}$. $\square$

Case 1: Alternative Proof. Rather than relying on Proposition H.1, we can use a quantile-based approach that directly leverages the parametric structure of the wage equation in (128). We show how this approach works to identify α_d and γ_d in Proposition H.3. We discuss how the nonparametric approach illustrated so far and the parametric approach behind Proposition H.3 differ from each other after the proof of the proposition. We start by defining

$$y(d, H_{n,t}) := (1-\omega)(1-\delta)\,U_d(H_{n,t}) \quad\text{and}\quad M_{n,t}(d) := \omega\,\gamma_d^{\alpha_d} H_{n,t}^{1-\alpha_d}\,\epsilon_{n,t}(d).$$

Proposition H.3 (Identification of α_d and γ_d)**.** *For each firm $d \in \mathcal{D}$ and some period $t \in \{1, \ldots, T\}$, assume:*

(i) (Unbounded Upper Tail of Human Capital.) The upper tail of human capital $H_{n,t}$ is unbounded, that is, $\lim_{p\to 1} Q_{\log H_{n,t}|D_{n,t}=d}(p) = \infty$.

(ii) (Negligible Quantile Reminder Relative to Human Capital.) For each $p \in (0,1)$, define the conditional quantile reminder

$$R_{t,d}(p) := Q_{\log w_{n,t}|D_{n,t}=d}(p) - \Big\{ \log\omega + \alpha_d \log\gamma_d + (1-\alpha_d)\, Q_{\log H_{n,t}|D_{n,t}=d}(p) \Big\}. \quad (129)$$

Then, the contribution of this reminder to the selected wage $w_{n,t}|D_{n,t} = d$ grows strictly more slowly than the contribution of human capital $H_{n,t}$ as the wage increases, that is,

$$\lim_{p\to 1} \frac{R_{t,d}(p)}{Q_{\log H_{n,t}|D_{n,t}=d}(p)} = 0. \quad (130)$$

(iii) (Normalization.) The upper tail of the remainder $R_{t,d}(p)$ has a known finite limit. That is,

$$\lim_{p\to 1} R_{t,d}(p) \ = \ L_{t,d}, \quad (131)$$

with $L_{t,d}$ known.

In addition, assume that $H_{n,t} > 0$, $w_{n,t} > 0$, and $\epsilon_{n,t}(d) > 0$ almost surely, so that all logarithms above are well defined. Then, for each $d \in \mathcal{D}$, the parameters α_d and γ_d are identified.

Proof: The proof is articulated in two steps.
Step 1: Identification of α_d. Fix a firm $d \in \mathcal{D}$. Using the structure of the wage equation,

$$\log w_{n,t}(d) = \log M_{n,t}(d) + \log\Big(1 + \tfrac{y(d,H_{n,t})}{M_{n,t}(d)}\Big). \quad (132)$$

Using the definition of $M_{n,t}(d)$,

$$\log M_{n,t}(d) = \log\omega + \alpha_d \log\gamma_d + (1-\alpha_d)\log H_{n,t} + \log\epsilon_{n,t}(d). \quad (133)$$

Therefore, substituting (133) in (132),

$$\log w_{n,t}(d) = \log\omega + \alpha_d \log\gamma_d + (1-\alpha_d)\log H_{n,t} + \log\epsilon_{n,t}(d) + \log\Big(1 + \tfrac{y(d,H_{n,t})}{M_{n,t}(d)}\Big). \quad (134)$$

Now condition on $D_{n,t} = d$ and apply the conditional quantile operator $Q_{\cdot|D_{n,t}=d}(p)$ to both sides of (134). Using only that adding a constant shifts quantiles, we obtain that

$$Q_{\log w_{n,t}|D_{n,t}=d}(p) = \log\omega + \alpha_d\log\gamma_d + Q_{(1-\alpha_d)\log H_{n,t} + \log\epsilon_{n,t}(d) + \log\left(1+\frac{y(d,H_{n,t})}{M_{n,t}(d)}\right)\,|D_{n,t}=d}(p). \quad (135)$$

Define the conditional quantile remainder $R_{t,d}(p)$ as

$$R_{t,d}(p) := Q_{\log w_{n,t}|D_{n,t}=d}(p) - \Big[\log\omega + \alpha_d \log\gamma_d + (1-\alpha_d)\,Q_{\log H_{n,t}|D_{n,t}=d}(p)\Big]. \tag{136}$$

Plugging (135) into (136) yields that

$$\begin{aligned} R_{t,d}(p) = &\log\omega + \alpha_d \log\gamma_d + Q_{(1-\alpha_d)\log H_{n,t} + \log \epsilon_{n,t}(d) + \log\left(1+\frac{y(d,H_{n,t})}{M_{n,t}(d)}\right)\,|D_{n,t}=d}(p) \\ &- \Big[\log\omega + \alpha_d \log\gamma_d + (1-\alpha_d)\,Q_{\log H_{n,t}|D_{n,t}=d}(p)\Big], \end{aligned}$$

which implies that

$$Q_{(1-\alpha_d)\log H_{n,t} + \log \epsilon_{n,t}(d) + \log\left(1+\frac{y(d,H_{n,t})}{M_{n,t}(d)}\right)\,|D_{n,t}=d}(p) = (1-\alpha_d)\,Q_{\log H_{n,t}|D_{n,t}=d}(p) + R_{t,d}(p).$$

By substituting this expression into (135), we obtain that

$$Q_{\log w_{n,t}|D_{n,t}=d}(p) = \log\omega + \alpha_d \log\gamma_d + (1-\alpha_d)\,Q_{\log H_{n,t}|D_{n,t}=d}(p) + R_{t,d}(p). \tag{137}$$

Fix any $\bar{p} \in (0,1)$. Let $\Delta_W(p) := Q_{\log w_{n,t}|D_{n,t}=d}(p) - Q_{\log w_{n,t}|D_{n,t}=d}(\bar{p})$, $\Delta_H(p) := Q_{\log H_{n,t}|D_{n,t}=d}(p) - Q_{\log H_{n,t}|D_{n,t}=d}(\bar{p})$, and $\Delta_R(p) := R_{t,d}(p) - R_{t,d}(\bar{p})$. Subtracting (137) evaluated at p and at $\bar{p}$ yields that for all $p \in (0,1)$,

$$\Delta_W(p) = (1-\alpha_d)\,\Delta_H(p) + \Delta_R(p). \tag{138}$$

By Assumption (i), $\lim_{p\to 1} Q_{\log H_{n,t}|D_{n,t}=d}(p) = \infty$ so that

$$\lim_{p\to 1} \Delta_H(p) = \infty. \tag{139}$$

By combining (139) and Assumption (ii), we can show that

$$\lim_{p\to 1} \frac{\Delta_R(p)}{\Delta_H(p)} = 0. \tag{140}$$

Indeed, note that

$$\frac{\Delta_R(p)}{\Delta_H(p)} = \frac{R_{t,d}(p) - R_{t,d}(\bar{p})}{Q_{\log H_{n,t}|D_{n,t}=d}(p) - Q_{\log H_{n,t}|D_{n,t}=d}(\bar{p})} = \frac{\frac{R_{t,d}(p)}{Q_{\log H_{n,t}|D_{n,t}=d}(p)} - \frac{R_{t,d}(\bar{p})}{Q_{\log H_{n,t}|D_{n,t}=d}(p)}}{1 - \frac{Q_{\log H_{n,t}|D_{n,t}=d}(\bar{p})}{Q_{\log H_{n,t}|D_{n,t}=d}(p)}}. \tag{141}$$

As $p \to 1$, Assumptions (i) and (ii) imply that

$$\frac{R_{t,d}(p)}{Q_{\log H_{n,t}|D_{n,t}=d}(p)} \to 0, \qquad \frac{R_{t,d}(\bar{p})}{Q_{\log H_{n,t}|D_{n,t}=d}(p)} \to 0, \qquad \frac{Q_{\log H_{n,t}|D_{n,t}=d}(\bar{p})}{Q_{\log H_{n,t}|D_{n,t}=d}(p)} \to 0,$$

so that the numerator of (141) converges to 0 whereas the denominator of it to 1, which yields (140).

Now divide both sides of (138) by $\Delta_H(p)$ to obtain

$$\frac{\Delta_W(p)}{\Delta_H(p)} = 1 - \alpha_d + \frac{\Delta_R(p)}{\Delta_H(p)}. \tag{142}$$

Taking limits of both sides of (142) as $p \to 1$ and using (140), we obtain

$$\lim_{p\to 1} \frac{\Delta_W(p)}{\Delta_H(p)} = \lim_{p\to 1} \left\{ (1-\alpha_d) + \frac{\Delta_R(p)}{\Delta_H(p)} \right\} = 1 - \alpha_d,$$

which identifies $1-\alpha_d$ and hence α_d. Note that we have not used the normalization in Assumption (iv) to identify α_d. Assumption (iii) will be used below to identify γ_d.
Step 2: Identification of γ_d. Rearranging (137) gives

$$R_{t,d}(p) = Q_{\log w_{n,t}|D_{n,t}=d}(p) - (1-\alpha_d)\, Q_{\log H_{n,t}|D_{n,t}=d}(p) - (\log\omega + \alpha_d \log\gamma_d).$$

Taking limits as $p \to 1$ on both sides and using the tail normalization (131), we obtain

$$L_{t,d} = \lim_{p\to 1} R_{t,d}(p) = \lim_{p\to 1} \left\{ Q_{\log w_{n,t}|D_{n,t}=d}(p) - (1-\alpha_d)\, Q_{\log H_{n,t}|D_{n,t}=d}(p) \right\} - (\log\omega + \alpha_d \log\gamma_d)$$

and, by re-arranging terms,

$$\alpha_d \log\gamma_d = \lim_{p\to 1} \left\{ Q_{\log w_{n,t}|D_{n,t}=d}(p) - (1-\alpha_d)\, Q_{\log H_{n,t}|D_{n,t}=d}(p) \right\} - \log\omega - L_{t,d}.$$

Hence,

$$\gamma_d = \exp\left(\frac{1}{\alpha_d} \left[\lim_{p\to 1} \left\{ Q_{\log w_{n,t}|D_{n,t}=d}(p) - (1-\alpha_d)\, Q_{\log H_{n,t}|D_{n,t}=d}(p) \right\} - \log\omega - L_{t,d} \right] \right).$$

The right side of this expression is identified from the conditional joint distribution of $(w_{n,t}, H_{n,t})$ given $D_{n,t} = d$, which determines the limit of $Q_{\log w_{n,t}|D_{n,t}=d}(p) - (1-\alpha_d) Q_{\log H_{n,t}|D_{n,t}=d}(p)$ as $p \to 1$, the known bargaining parameter ω, the known constant $L_{t,d}$, and the already identified α_d. Thus, γ_d is identified. $\square$

Remarks. We now compare the approach of Proposition H.1, Proposition H.2, and Corollary H.1 (hereafter, the "*first approach*") with that of Proposition H.3 (hereafter, the "*second approach*") to recover α_d and γ_d. The *first approach* identifies the scale function $\sigma(d, H_{n,t})$ from the upper tail of the observed selected wage distribution $w_{n,t}$ conditional on $(D_{n,t}, H_{n,t})$. Given the structural relation $\sigma(d, H_{n,t}) := \omega\, \gamma_d^{\alpha_d} H_{n,t}^{1-\alpha_d}$, α_d is then identified from *ratios* of $\sigma(d,h)$ at different values of h, which do not depend on any normalization for $\sigma(d,\cdot)$. By contrast, γ_d is identified from the *level* of $\sigma(d,h)$ at some h and thus requires a level normalization, for example, that $\sigma(d,\bar h) = 1$ for some $\bar h$.

The *second approach* does not involve the intermediate identification of $\sigma(d, H_{n,t})$, but instead directly focuses on the log wage equation expressed as

$$\log w_{n,t}(d) = \log\omega + \alpha_d \log\gamma_d + (1-\alpha_d)\log H_{n,t} + \text{error},$$

in particular on the equilibrium relationship between the upper quantiles of $\log w_{n,t}$ and $\log H_{n,t}$ given $D_{n,t} = d$. Under Assumptions (i) and (ii), we obtain that as $p \to 1$,

$$Q_{\log w_{n,t}|D_{n,t}=d}(p) = \log \omega + \alpha_d \log \gamma_d + (1 - \alpha_d)\, Q_{\log H_{n,t}|D_{n,t}=d}(p) + o\big(Q_{\log H_{n,t}|D_{n,t}=d}(p)\big),$$

so that ratios of differences in this expression across different values of p identify the *slope* term $1 - \alpha_d$ and so the parameter α_d without the need for any normalization. The parameter γ_d is then recovered from an *intercept*- type tail normalization of the composite error term encoded in Assumption (iii). In this sense, the *second approach* resembles an asymptotic linear quantile regression of $Q_{\log w_{n,t}|D_{n,t}}(p)$ on $Q_{\log H_{n,t}|D_{n,t}}(p)$ at high levels of p: the slope $1 - \alpha_d$ is identified from the limiting ratio of quantile differences, whereas the intercept $\log \omega + \alpha_d \log \gamma_d$ is pinned down through a normalization of the extremal tail of the composite error.

Thus, both approaches are fundamentally based on *upper-tail identification*. The first approach relies on the upper tail of $w_{n,t}$ conditional on $(H_{n,t}, D_{n,t})$, whereas the second approach relies on the upper tail of $\log w_{n,t}$ and $\log H_{n,t}$ conditional on $D_{n,t}$. Under both approaches, α_d is identified through a slope argument. Instead, the recovery of γ_d requires a normalization condition.

Under the first approach, the key restriction is a *tail limit condition* (Assumption (ii) of Proposition H.1). It requires that, conditional on human capital $H_{n,t}$, the probability of working at firm d upon receiving a very high *potential* wage $w_{n,t}(d)$ converges to a firm-specific constant $q_t(d)$, as the lower bound for the potential wage to be considered large grows arbitrarily large ($w \to \infty$). This stabilisation of selection in the upper tail allows the tail of the *potential* wage distribution to be recovered from observed wages. Under the second approach, the key restriction is a *dominance condition* (Assumptions (ii) of Proposition H.3) on the quantile reminder of observed wages, which leads to a relation between $Q_{\log w_{n,t}|D_{n,t}=d}(p)$ and $Q_{\log H_{n,t}|D_{n,t}=d}(p)$ that is asymptotically ($p \to 1$) affine. From the identified slope and intercept of this relation, respectively, α_d and γ_d can be recovered.

Case 2: $H_{n,t}$ is Unobserved with Unknown Distribution and Support. In this case, we proceed in two steps. First, in Proposition H.4, we account for the fact that $\mathcal{H}_t$ is unknown and work in the human-capital *rank space* by mapping $H_{n,t}$ to its quantile (percentile) index via its CDF, namely,

$$y^\circ(d, U_{n,t}) \;\coloneqq\; y\big(d, F^{-1}_{H_{n,t}}(U_{n,t})\big) \;\text{ and }\; \sigma^\circ(d, U_{n,t}) \;\coloneqq\; \sigma\big(d, F^{-1}_{H_{n,t}}(U_{n,t})\big),$$

defined on the support of $U_{n,t}$ rather than on the support of $H_{n,t}$. Second, in Proposition H.2, we assume that $F_{\epsilon_{n,t}}$ is known and that two values of $H_{n,t}$, h_a and h_b, corresponding to the values u_a and u_b of $U_{n,t}$, are known to identify α_d, γ_d, and z.

Proposition H.4 (Identification of $y^\circ(d, U_{n,t})$ and $\sigma^\circ(d, U_{n,t})$)**.** *Given $d \in \mathcal{D}$ and $t \in \{1, \dots, T\}$, let $\mathcal{U}_{t,d} \subseteq (0, 1)$ be the set of realizations u of $U_{n,t}$ such that $\Pr(D_{n,t} = d \mid U_{n,t} = u) > 0$. For each firm $d \in \mathcal{D}$ and period $t \in \{1, \dots, T\}$, assume:*

(i) (Supports.) For each $u \in \mathcal{U}_{t,d}$, $\sup\{w : \Pr(w_{n,t}(d) \le w \mid U_{n,t} = u) < 1\} = \infty$ and $\sup\{w : \Pr(w_{n,t} \le w \mid D_{n,t} = d, U_{n,t} = u) < 1\} = \infty$.

(ii) (Tail Limit.) There exists an (unknown) constant $q_t(d) \in (0, 1]$ such that for every $u \in \mathcal{U}_{t,d}$,

$\lim_{w\to\infty} \Pr\big(D_{n,t} = d \mid U_{n,t} = u, w_{n,t}(d) > w\big) = q_t(d)$.

(iii) (Tail Regularity.) For each $u \in \mathcal{U}_{t,d}$, there exist (unknown) thresholds $w_{u,t,d} < \infty$ and $w^{\text{obs}}_{u,t,d} < \infty$ such that the cumulative distribution functions $F_{w_{n,t}(d)|U_{n,t}=u}$ and $F_{w_{n,t}|D_{n,t}=d,U_{n,t}=u}$ are continuous and strictly increasing on $(w_{u,t,d}, \infty)$ and $(w^{\text{obs}}_{u,t,d}, \infty)$, respectively.

(iv) (Normalization.) There exists a known $\bar{u} \in \mathcal{U}_{t,d}$ with $y^\circ(d, \bar{u}) = 0$ and $\sigma^\circ(d, \bar{u}) = 1$.

Then, the functions $y^\circ(d, u)$ and $\sigma^\circ(d, u)$ are identified for each $u \in \mathcal{U}_{t,d}$ and $d \in \mathcal{D}$.

Proof: The claim is an immediate consequence of Proposition G.2 after a change of conditioning variable from the latent value $H_{n,t}$ to its rank $U_{n,t} := F_{H_{n,t}}(H_{n,t})$. Note that this reparametrisation is without loss, because by the probability integral transform, $U_{n,t}$ is uniformly distributed on $(0, 1)$, and conditioning on $H_{n,t}$ is equivalent to conditioning on $U_{n,t}$. Since the support of $H_{n,t}$ is unknown, identification can only be stated for the *rank-indexed* objects $y^\circ(d, u)$ and $\sigma^\circ(d, u)$ rather than for the cardinal objects $y(d, h)$ and $\sigma(d, h)$ at the unknown levels h. $\square$

Proposition H.5 (Identification of $F_{\epsilon_{n,t}}$)**.** *Suppose that Assumptions (i) to (iv) of Proposition H.4 holds, which imply that $y^\circ(d, u)$ and $\sigma^\circ(d, u)$ are identified for each $d \in \mathcal{D}$, $u \in \mathcal{U}_{t,d}$, and $t \in \{1, \ldots, T\}$. Moreover, for each firm $d \in \mathcal{D}$ and period $t \in \{1, \ldots, T\}$, assume $\epsilon_{n,t}(d)$ belongs to a parametric family indexed by the $p_{t,d} \times 1$ vector of parameters $\mu_{t,d} \in M_{t,d} \subseteq \mathbb{R}^{p_{t,d}}$ that is tail-ratio identifiable. Namely, fix any $u \in \mathcal{U}_{t,d}$ and choose $p_{t,d}+1$ thresholds $0 < w_0 < w_1 < \cdots < w_{p_{t,d}}$. Define the function $\Phi_{t,d,u} : M_{t,d} \to \mathbb{R}^{p_{t,d}}$ as*

$$\Phi_{t,d,u}(\mu_{t,d}) := \left(\frac{S_{\epsilon_{n,t}(d)}\left(\frac{w_1 - y^\circ(d,u)}{\sigma^\circ(d,u)}; \mu_{t,d}\right)}{S_{\epsilon_{n,t}(d)}\left(\frac{w_0 - y^\circ(d,u)}{\sigma^\circ(d,u)}; \mu_{t,d}\right)}, \ldots, \frac{S_{\epsilon_{n,t}(d)}\left(\frac{w_{p_{t,d}} - y^\circ(d,u)}{\sigma^\circ(d,u)}; \mu_{t,d}\right)}{S_{\epsilon_{n,t}(d)}\left(\frac{w_0 - y^\circ(d,u)}{\sigma^\circ(d,u)}; \mu_{t,d}\right)} \right).$$

If $\Phi_{t,d,u}$ is injective, we say that $\epsilon_{n,t}(d)$ belongs to a parametric family that is tail-ratio identifiable. Under these assumptions, for each period $t \in \{1, \ldots, T\}$:

(a) (Marginal Identification.) The parameter $\mu_{t,d}$ is identified.

(b) (Joint Identification.) If the shocks $\{\epsilon_{n,t}(d)\}_{d\in\mathcal{D}}$ are mutually independent across $d \in \mathcal{D}$, then the joint distribution of $\epsilon_{n,t}(d)$ is identified as the product of the identified marginals. Alternatively, if a copula C_{μ_t} is specified so that

$$F_{\epsilon_{n,t}}(v_1, \ldots, v_{|\mathcal{D}|}) \;=\; C_{\mu_t}\big(F_{\epsilon_{n,t}(1)}(v_1; \mu_{t,1}), \ldots, F_{\epsilon_{n,t}(|\mathcal{D}|)}(v_{|\mathcal{D}|}; \mu_{t,|\mathcal{D}|})\big) \quad \forall\, (v_1, \ldots, v_{|\mathcal{D}|}) \in \mathbb{R}^{|\mathcal{D}|},$$

and the copula parameter μ_t is known, then the joint distribution is identified from the identified marginals and C_{μ_t}. The joint CDF is partially identified by the sharp Fréchet–Höffding bounds, namely, for all $(v_1, \ldots, v_{|\mathcal{D}|}) \in \mathbb{R}^{|\mathcal{D}|}$,

$$\max\Big\{ \sum_{d\in\mathcal{D}} F_{\epsilon_{n,t}(d)}(v_d; \mu_{t,d}) - (|\mathcal{D}| - 1),\; 0 \Big\} \;\leq\; F_{\epsilon_{n,t}}(v_1, \ldots, v_{|\mathcal{D}|}) \;\leq\; \min_{d\in\mathcal{D}} F_{\epsilon_{n,t}(d)}(v_d; \mu_{t,d}).$$

Corollary H.2 (Identification of α_d, γ_d, and z)**.** *For each firm $d \in \mathcal{D}$ and for some period $t \in \{1, \dots, T\}$, assume that:*

(i) $y^\circ(d,u)$ and $\sigma^\circ(d,u)$ are identified for each $u \in \mathcal{U}_{t,d}$ (see Proposition H.4 for sufficient conditions).

(ii) The distribution $F_{\epsilon_{n,t}}$ of $\epsilon_{n,t}$ is identified (see Proposition H.5 for sufficient conditions).

(iii) There exist two distinct ranks $u_a \neq u_b$ in $\mathcal{U}_{t,d}$ such that the corresponding levels of human capital $h_a := F^{-1}_{H_{n,t}}(u_a)$ and $h_b := F^{-1}_{H_{n,t}}(u_b)$ are known to the researcher.

Then, α_d, γ_d, and z are identified for each $d \in \mathcal{D}$.

Proof: The proof proceeds in three steps.
Step 1: Identification of α_d from $\sigma^\circ(d, U_{n,t})$. Recall that

$$\sigma^\circ(d,u) \;=\; \sigma\big(d, F^{-1}_{H_{n,t}}(u)\big) \;=\; \omega\, \gamma_d^{\alpha_d} \left(F^{-1}_{H_{n,t}}(u)\right)^{1-\alpha_d}.$$

Pick two ranks $u_a \neq u_b$ and the corresponding levels $h_a := F^{-1}_{H_{n,t}}(u_a)$ and $h_b := F^{-1}_{H_{n,t}}(u_b)$. Then

$$\frac{\sigma^\circ(d,u_a)}{\sigma^\circ(d,u_b)} \;=\; \frac{\omega\, \gamma_d^{\alpha_d}\, h_a^{1-\alpha_d}}{\omega\, \gamma_d^{\alpha_d}\, h_b^{1-\alpha_d}} \;=\; \Big(\frac{h_a}{h_b}\Big)^{1-\alpha_d}.$$

Taking logarithms and rearranging terms yields that

$$\alpha_d \;=\; 1 \;-\; \frac{\log\big(\sigma^\circ(d,u_a)/\sigma^\circ(d,u_b)\big)}{\log(h_a/h_b)}.$$

Hence, given knowledge of the two ranks u_a, u_b and their corresponding levels h_a, h_b, α_d is identified.
Step 2: Identification of γ_d from $\sigma^\circ(d, U_{n,t})$. Using any anchored pair (u_*, h_*) with $* \in \{a, b\}$,

$$\sigma^\circ(d,u_*) = \omega\, \gamma_d^{\alpha_d} h_*^{1-\alpha_d} \quad\Longrightarrow\quad \gamma_d = \left(\tfrac{\sigma^\circ(d,u_*)}{\omega\, h_*^{1-\alpha_d}}\right)^{1/\alpha_d},$$

which identifies γ_d.
Step 3: Identify z from $y^\circ(d, U_{n,t})$ and $F_{\epsilon_{n,t}}$. Pick any anchored pair (u_*, h_*) with $* \in \{a, b\}$. Since $y^\circ(d,u_*) = y(d,h_*)$ is identified, (α_d, γ_d) are now known, and $F_{\epsilon_{n,t}}$ is known by assumption, z is identified as

$$z = \frac{y^\circ(d,u_*)}{(1-\delta)(1-\omega)} - \delta\, \mathbb{E}_{\epsilon_{n,t}\sim F_{\epsilon_{n,t}}}\Big[f\big(S(h_*, \epsilon_{n,t}; \omega, \alpha_d, \gamma_d, \delta)\big)\Big],$$

which completes the proof. $\square$