

1 Introduction

Wage differentials across observationally equivalent workers are both sizable and persistent in reality, and empirical work has given stylized-fact status to several relationships between worker wages and employer characteristics. The employer size-wage effect is perhaps the strongest such stylized fact: larger firms or plants pay higher wages. The literature includes especially influential work by Krueger and Summers (1988) and Brown and Medoff (1989), and has been recently surveyed by Oi and Idson (1999). An apparent relationship between wages and employer size may be spurious in raw data, reflecting for example the fact that larger employers hire higher quality workers, or offer inferior working conditions (Masters, 1969), or face higher risk of unionization (Podgursky, 1986), or share high profits with workers (Weiss, 1966; Mellow, 1982), or need to offset a lower applicant-to-job vacancy ratio (Weiss and Landau, 1984), or reduce costly monitoring (Oi, 1983). The literature, however, has established that the employer's size remains significant in fixed-effect wage equations even when all variables relevant to such alternative explanations are controlled for.

Empirical evidence indicates that wage dispersion across production units is also related to employment levels. Davis and Haltiwanger's (1996) work on US manufacturing sector data, for example, finds that wage dispersion falls sharply with establishment size for non-production workers, and mildly for production workers. Like the mean effects, this second-moment wage-size effect is robust to the inclusion of control variables, such as establishment-level union density and indicators of worker quality.

This paper explores the wage-structure implications of matching models of the labor market, where job creation and destruction take place at firms of endogenously heterogeneous size, and shows that the joint distribution of wages and employment generated by such models can be consistent with the evidence on both the first and second moment of the wage-size distribution. We are certainly not the first to note that imperfect matching or "search" can rationalize wage differentials across identical workers employed by different firms. In the Burdett and Mortensen (1998) model, on-the-job search draws workers from low-pay to high-pay jobs and, in a "wage-posting" equilibrium where vacancies are costly and wages attached to a given job are never renegotiated, ex-ante identical employers can earn similar profits through different combinations of wage and employment levels. Consistently with evidence of wage-size effects, high-wage jobs are more numerous than lower-wage jobs in such an equilibrium. In that model and in recent extensions where optimal adjustment of match-level capital yields single-peaked wage-offer distributions (see Mortensen, 1998),

however, jobs could be spread in the economy at large rather than bundled in “firms.” In fact, since no renegotiation occurs in a wage-posting setting, jobs paying different wages could coexist within any legal entity in the absence of unmodeled additional constraints on within-firm wage dispersion.

In general, the association of high wages to a specific firm’s circumstances is lost in a constant-returns environment. Matching models with continuous renegotiation of wage rates also typically focus on job-level phenomena, and are similarly ill-suited to a study of firm-level phenomena (see Mortensen and Pissarides, 1999, for a survey). Observable characteristics of real-life firms, however, do appear empirically relevant. Dickens and Katz (1987), Katz and Summers (1989), and other references in Blanchflower et al. (1996) find that persistently high wages are associated not only to employment levels and rates of change, but also to larger profits and other indicators of employers’ “ability to pay.” Further, Davis et al. (1996) show that the intensity of job reallocation, or gross job creation and destruction, is lower at larger firms.

The class of models we study in this paper offers an interpretation of such evidence in terms of standard economic mechanisms. We assume that hiring workers entails costs for employers and, following Bertola and Caballero (1994) and Mortensen (1999), we let the recruiting technology have decreasing return to scale. Thus, optimal employment growth is faster at firms where the employer’s forward-looking surplus from new hiring is higher. We let wages be set by a process of continuously renegotiated surplus-splitting at the individual job level, to imply that relatively large values of the employer’s marginal surplus are also associated with high wage rates. Hence, fast employment growth is associated with high wage levels, and the resulting models yield a positive wage-size effect through stock-flow relationships similar to those featured by the Burdett-Mortensen model of worker search.

In the models we consider, however, productivity is not identical across firms. Rather, both productivity and wages are driven by exogenous firm-level stochastic processes, and depend on employment levels. Profits, employment, and wages are jointly endogenous and interrelated random variables in data generated by models with decreasing returns to recruiting effort, and the “matching” structure of the labor market readily rationalizes a variety of empirical facts. On the labor-supply side, arbitrage across different jobs’ wages is imperfect if moving from one job to another entails a spell of unemployment; hence, otherwise identical workers can and do earn different wages, even though they interact in the same market and have identical opportunity costs. On the labor-demand side, individual firms’ characteristics such as size and profits are relevant to wage bargains struck at the individual

job level, especially if—as in Bertola and Caballero (1994), Stole and Zwiebel (1996), Cahuc and Wasmer (2000)—the firm-level producer surplus depends nonlinearly on employment, as would be the case when the individual firm’s technology display decreasing returns not only in the recruiting, but also in the production technology.

2 Matching and firms

The object of our study is a stationary labor market, inhabited by a continuum of risk-neutral firms of fixed total mass and by an also fixed and continuously divisible number of risk-neutral workers. We normalize to unity the measure of both sets, and index firms by $f \in [0, 1]$. Since labor is assumed homogeneous in quality, it will not be necessary in what follows to index workers other than according to the identity of their employer.

The marginal revenue product of labor at each firm is a function $\pi(l, \eta)$ of its current employment level l and of η , a shifter of its labor demand. This function is decreasing in its first argument if each firm has exclusive use in production of specific non-traded factors (such as entrepreneurial skill), or enjoys monopoly power in the product market. In such cases, the identity of “firms” is quite clearly associated with the characteristics of their cost and product demand functions, and bears directly on their employment level: since returns to employment are decreasing, the simplest static model of wage-taking labor demand would pin down a unique optimal employment level. We let $\pi(l, \eta)$ be increasing in its second argument, so that larger values of η are associated with stronger labor demand at any given wage level.

We shall also consider cases where $\pi(l, \eta)$ is constant with respect to l , i.e., technological and market configurations where returns to labor are constant and the optimal scale of operation of the firm could not be characterized by simple static models of competitive labor demand. As in Cahuc and Wasmer (2000), constant returns to labor may be consistent with use of other factors (such as capital) as long as these are freely adjustable and the production function has constant returns to scale overall, and the dynamics of the firm’s employment may still be uniquely characterized if hiring takes time.

The pair (l, η) identifies “firms” in the context of the model. It will be convenient to let η jump discretely and infrequently between a finite number of values. In order to address cross-sectional empirical evidence, we view all such exogenous uncertainty as idiosyncratic, we assume that the exogenous process followed by η is stationary, and we suppose that the frequency distribution of the model’s endogenous variables is well approximated by the

ergodic distribution of firm-specific stochastic processes.

2.1 Labor demand

The endogenous dynamics of employment at each firm f are determined by the employer’s hiring activity and firing decisions. Both can be readily characterized in terms of the *shadow value* of firm f ’s employment, defined as the discounted expected value of the difference between marginal productivity and wage rates.

To model the hiring process, we let firm f post $v_f \geq 0$ vacancies. Positive vacancies should be posted whenever their cost is (weakly) lower than their expected contribution to the firm’s present discounted value, namely the shadow value of an additional worker times the probability that an open vacancy is matched to an unemployed worker. Following Bertola and Caballero (1994) and Mortensen (1999), we suppose that vacancies can be created costlessly, but entail more than proportionately increasing flow costs (equivalently, the matching productivity of vacancies could be modeled as a decreasing function of the firm’s search effort). At an interior optimum, the cost of the marginal vacancy posted must be equal to its expected contribution to the firm’s value, i.e., the value of a marginal addition to its labor force times the probability intensity of matching events for the marginal vacancy. We denote the latter with ϑ and view it as a given constant through most of our analysis (but Appendix B offers a brief discussion of its endogenous determination by way of a standard matching function specification). We let the marginal cost of posting v_f vacancies be given by cv_f , and we recognize that the shadow value of labor is a function $A(l, \eta)$ of the firm-specific state variables if the idiosyncratic process $\{\eta\}$ is Markov in levels and wages are continuously renegotiated. Thus, the optimality condition for vacancy-posting activities is

$$v_f = \frac{\vartheta}{c} A(l_f, \eta_f) \tag{1}$$

whenever $A(l_f, \eta_f) > 0$.

The value of additional employment need not be strictly positive for all levels of η and l . We let firms fire their workers instantaneously and costlessly whenever it is optimal for them to do so, i.e., whenever the marginal value of employment would otherwise be negative.¹

¹If firing were costly, voluntary quits would allow firms to obtain costless employment reductions after a negative shock as in Saint-Paul (1995); regulatory constraints may also delay layoffs as in Garibaldi (1998). Firing (or hiring) costs could make it optimal not to react fully to a state transition and opt instead for “optimal inaction” policies familiar from, e.g., Bentolila and Bertola (1990). These and other realistic extensions would greatly complicate the model but would not qualitatively affect the basic point of this paper.

Thus, firm f posts the positive vacancies determined by (1) whenever $A(l_f, \eta_f) > 0$. When $A(l_f, \eta_f) = 0$, it posts no vacancies and, if l_f is positive, is indifferent to firing the marginal worker.

In general, of course, $A(l_f, \eta_f)$ is itself endogenously determined by the firm's optimal labor demand policy, which is forward looking and, in turn, depends on the shape of the $\pi(\cdot)$ marginal productivity function, on the wage paid to the firm's employees, and on the dynamics of worker turnover. By its definition as the present discounted value of labor's marginal contribution to profits, the function $A(l_f, \eta_f)$ satisfies a standard asset-pricing relationship,

$$rA(\cdot) = \left[\pi(\cdot) - \frac{\partial(w(\cdot)l)}{\partial l} \right] - \rho A(\cdot) + \frac{\partial A(\cdot)}{\partial l} \dot{l} + E_f [\Delta A(\cdot)], \quad (2)$$

where we let workers quit into unemployment at an exogenous rate $\rho \geq 0$ and, to simplify notation, we omit the arguments (l_{ft}, η_{ft}) of the relevant functions and firm and time indexes. Equation (2) rules out arbitrage opportunities imposing that the product of the asset value $A(\cdot)$ and the rate of return on the firm's operation, denoted by $r \geq 0$, equal current dividends plus forward-looking capital-gain expectations. The term in square brackets on the right-hand side represents the marginal worker's contribution to the firm's current cash flow, i.e., the difference between labor's marginal productivity and the marginal cost of employing an additional worker. The latter is written as the derivative of the wage bill with respect to employment, which does not coincide with the wage if, as in the Stole and Zwiebel (1996) model of intra-firm bargaining and in the equilibrium model of Bertola and Caballero (1994), wages depend on the size of the labor pool attached to the firm.

The other terms on the right-hand side of (2) represent the capital gain components of the firm's return on the marginal worker employed. The employer experiences a flow $\rho A(\cdot)$ of marginal capital losses as workers, valued at $A(\cdot) \geq 0$, quit at rate $\rho \geq 0$. Further capital gains or losses result from exogenous shocks and from the firm's own hiring and firing decisions. In the absence of state transitions, the shadow value of labor evolves continuously if it depends on l and $\dot{l} \neq 0$. Since the firm's employment increases at rate ϑv_f as a result of the firm's hiring efforts, and declines at rate ρl as a consequence of natural turnover,

$$\dot{l}_f = \vartheta v_f - \rho l_f = \frac{\vartheta^2}{c} A(l_f, \eta_f) - \rho l_f \quad (3)$$

is not zero in general, and needs to be taken into account by our characterization of cases where $\partial A(\cdot)/\partial l \neq 0$.

The final forward-looking term on the right-hand side of (2), $E_f [\Delta A(\cdot)]$, represents the state-dependent expected rate of capital gain or loss per unit time resulting from shocks to η . Since employers may instantaneously react to shocks by firing some or all of their employees, models where $A(\cdot)$ depends on l_f as well as on η_f need to characterize such capital gains in terms of endogenous labor demand reactions as well as of exogenous productivity shocks.

2.2 Workers and wages

Turning next to wage determination, let the objective function of all workers be the risk-neutral discounted expectation of labor income flows. In the steady-state economy under consideration, the workers' thusly defined "human capital" depends on whether they are employed and, if so, on the current labor demand shifter η and employment l of their employer. Let unemployed workers exert constant search effort and enjoy an income-equivalent flow z from leisure and benefits when unemployed. In steady state, their human capital has constant value, denoted J^u in what follows and characterized below.

The wages and human capital of employed workers are functions $w(\cdot)$ and $J(\cdot)$ of the employing firm's labor demand index η and employment level l . They satisfy the asset-valuation relationship

$$rJ(\cdot) = w(\cdot) + \rho[J^u - J(\cdot)] + \frac{\partial J(\cdot)}{\partial l}l + E_f [\Delta J], \quad (4)$$

which uses the same notational conventions as (2) and also adds to the current cash-flow three conceptually distinct capital-gain component of flow returns. In (4), the worker expects to experience exogenous separations into unemployment, at rate ρ ; continuously evolving capital gains or losses, reflecting employment dynamics for unchanged business conditions; and infrequent gains or losses due to labor demand shocks, denoted $E_f [\Delta J]$.

Since forming matches is costly, existing ones enjoy economic rents. Following Mortensen and Pissarides (1994) and other recent contributions, we assume that the total expected discounted surplus from each match is split in fixed proportion at all times.

At firms where the marginal value of employment is zero, there is no surplus to be split. The workers (if any) which the employer finds it optimal to retain must be indifferent to the outside option:

$$J^u = J(l_f, \eta_f) \quad \forall f \text{ s.t. } A(l_f, \eta_f) = 0, l > 0. \quad (5)$$

Since each unemployed worker is matched to a random open vacancy with probability inten-

sity $\vartheta V/U$, we have the standard asset-valuation equation

$$rJ^u = z + \vartheta \frac{V}{U} (E_u[J] - J^u), \quad (6)$$

where $E_u[J]$ is the average value of jobs for which vacancies are posted.

At firms which are posting vacancies, conversely, the option of costlessly opening a new one is always available to the employer, while for their employees the outside option is a voluntary quit into unemployment. Denoting with β the employee's surplus share, we write:

$$\beta A(l_f, \eta_f) = (1 - \beta)(J(l_f, \eta_f) - J^u) \quad \forall f \text{ s.t. } A(l_f, \eta_f) > 0. \quad (7)$$

The relationships (5) and (7) between the levels of workers' human capital functions and employers' shadow-value functions hold identically in the absence of state transitions, and readily imply a linear relationship between the rates of change of the two forward-looking value functions. For any given η , and for all l consistent with the firms' dynamic labor demand policies, it must be the case that $(1 - \beta) \dot{J}(l, \eta) = \beta \dot{A}(l, \eta)$, or that

$$(1 - \beta) \frac{\partial J(l, \eta)}{\partial l} \dot{l} = \beta \frac{\partial A(l, \eta)}{\partial l} \dot{l}. \quad (8)$$

The assumption that the forward looking match surplus is split in fixed proportions between the parties further implies that capital gains or losses resulting from exogenous productivity shocks are also similarly split, and therefore that a linear relationship similar to those in (8) also holds between the workers' and employers' rates of expected capital gain or losses upon labor-demand transitions:²

$$(1 - \beta) E_f [\Delta J(\cdot)] = \beta E_f [\Delta A(\cdot)]. \quad (9)$$

²To show this, consider the implications of a transition from state (η, l) to state (η', l') , where $\eta' - \eta$ represents the exogenous productivity shock and $l' - l$ may be negative if the employer reacts to it by firing some employees. If $A(\eta', l) > 0$, the employer does not fire and the surplus-splitting rule implies that $J(\eta, l) - J(\eta', l) = \frac{\beta}{1-\beta} [A(\eta, l) - A(\eta', l)]$. If firing is optimal, then the employer's marginal value function jumps from $A(\eta, l) > 0$ to $A(\eta', l') = 0$, and the employees' human capital from $J(\eta, l)$ to J^u : since $J(\eta, l) = J^u + \frac{\beta}{1-\beta} A(\eta, l)$ by equation (7),

$$J(\eta, l) - J(\eta', l') = (J^u + \frac{\beta}{1-\beta} A(\eta, l) - J^u) = \frac{\beta}{1-\beta} (A(\eta, l) - A(\eta', l')).$$

Hence, all possible transitions imply that human capital losses entailed by exogenous disappearance of match surplus are a proportion $\beta/(1 - \beta)$ of the employer's loss of marginal value. Weighting all such transitions by their probability equation (9) follows.

Multiplying equation (2) by β and equation (4) by $1 - \beta$, recognizing that

$$\frac{\partial (w(\cdot)l)}{\partial l} = w(\cdot) - \frac{\partial w(\cdot)}{\partial l}l,$$

and subtracting term-by-term, the relationships in (7), (8), and (9) imply that all forward-looking terms cancel out. Hence,

$$w(l, \eta) = \beta \left(\pi(l, \eta) - \frac{\partial w(l, \eta)}{\partial l}l \right) + (1 - \beta) rJ^u \quad (10)$$

at any firm posting positive vacancies, independently of the probabilistic structure of the exogenous Markov process driving labor demand and of the process through which unemployed workers are matched to vacancy-posting firms. Wage rates are such as to split the marginal surplus in *current flow* terms between the employer and each of the employees whenever match-specific rents are positive in expected present value terms. As in Bertola and Caballero (1994) and Stole and Zwiebel (1996), continuous renegotiation of wages implies that the employer (like any monopsonist) should take into account the implications of additional employment for inframarginal wage payments when the wage paid by a vacancy-posting firm does depend on the current employment level. Hence, the term $[\partial w(\cdot)/\partial l]l$ is netted out of the employer's flow surplus share in (10), to yield a non-homogeneous ordinary differential equation for each value of η .

In canonical form, the equation reads

$$x'(l) + x(l)p(l) + q(l) = 0,$$

where $x(l) = w(l, \eta)$ denotes the unknown endogenous wage function; $p(l) \equiv \frac{1}{\beta l}$; and the function $q(l) \equiv -(\pi(l, \eta) + (1 - \beta) rJ^u / \beta) / l$ is parametrically given by the model's economic structure. By the method of variation of parameters, all solutions are in the form³

$$x(l) = \left(c - \int_0^l \frac{q(u)}{z(u)} du \right) z(l),$$

for c a constant of integration and $z(\cdot)$ a solution of the homogeneous differential equation

$$\frac{dz(l)}{dl} + p(l)z(l) = 0.$$

³See, e.g., Ince (1956), p.21. We are grateful to a referee for bringing this formula to our attention. We also benefited from reading Cahuc and Wasmer's (2000) discussion in the search-and-matching context of points made by Stole and Zwiebel (1996).

Since $p(l) \equiv \frac{1}{\beta l}$, homogeneous solutions are proportional to $z(l) = l^{-1/\beta}$, and diverge to infinity as l approaches zero. Hence, it must be the case that $c = 0$, for otherwise wage rates would become unbounded when small employment levels are considered.

Thus, employment levels and wages are related according to

$$\begin{aligned} w(l, \eta) &= x(l) = l^{-1/\beta} \int_0^l u^{\frac{1-\beta}{\beta}} \left[\pi(u, \eta) + \frac{(1-\beta)rJ^u}{\beta} \right] du \\ &= l^{-1/\beta} \int_0^l u^{\frac{1-\beta}{\beta}} \pi(u, \eta) du + (1-\beta)rJ^u \end{aligned} \quad (11)$$

at all firms posting positive vacancies. (If a firm posts no vacancies, the wage of its employees—if any—must be such as to let them be indifferent to separation into unemployment, as in equation (5). We shall encounter and discuss such labor-hoarding situations in Sections 3.2 and 3.3 below.)

The first term on the right-hand side (11) averages the marginal productivities of inframarginal workers with geometrically increasing (if $\beta < 1/2$) or decreasing weights. As Cahuc and Wasmer (2000) point out, this formula is a generalization to the case $\beta \neq 1/2$ of that derived by Stole and Zwiebel (1996) under the assumption that production surplus is evenly split by the employer with each employee, and that employees participating in such negotiations have access to a reservation wage and are irreplaceable. In the dynamic search-and-matching context we consider, the costly and time-consuming nature of the vacancy-posting technology indeed prevents the firm from locating replacement workers at the same (instantaneous) pace of continuous renegotiations, and the reservation wage is the flow equivalent of an unemployed worker's asset value, rJ^u , which depends endogenously on the structure of the labor market. The size of the insider pool participating in such negotiations, however, is not exogenously given: rather, it is determined endogenously by the firm's past history of forward-looking decisions to post vacancies.

The expression (11) for the wage negotiated by the firm with each of its employees is independent of l if $\pi(\cdot)$ is identically constant with respect to its first argument, and it is a decreasing function of l if marginal (hence, average) labor productivity declines as employment increases. This may at first sight appear to be at odds with plentiful evidence of a positive wage-size effect. In what follows, we study simple parameterizations of this general modeling structure, and discuss how explicit consideration of endogenous labor demand policies may support its realism.

3 Specialized models

Turning to the model's implications for the relationship between firm sizes and wages, we need to recognize that firm-specific employment levels are themselves endogenous. Even in cases where for each value of η the wage schedule in (11) is available in closed form, in fact, the observed joint distribution of employment and wage levels is a mixture of such wage schedules for different labor demand shocks. The character of such mixing plays a crucial role in generating a positive relationship between employer sizes and wage rates in the presence of constant or decreasing wage schedules at each firm.

Recall that during periods when no state transition occurs each firm's employment evolves according to (3) as a result of exogenous quits into unemployment at rate ρ and of the employer's hiring activity. To completely characterize the model, one would insert (3) and the wage expression from (11) in the asset-valuation equation (2), to obtain a system of differential equations in $A(l_f, \eta_f)$, with cross-equation linkages implied by the probabilistic structure of exogenous Markov transitions across labor demand states. A solution of this system of functional equations would in turn make it possible to characterize the implications of optimal labor demand policies for the endogenous distribution of employment levels. We do not attempt to characterize solutions of the class of models outlined above in such general terms. Rather, we postpone a brief discussion of numerical methods to Section 3.3 below, and focus attention on special cases which, while admittedly far from realistic, do offer simple theoretical results to be confronted with empirical evidence on the wage-size distribution.

To see which features of the general modeling perspective may be relevant to such evidence, it is useful to consider the long-run implications of the employment dynamics in (3). Setting $\dot{l} = 0$ in (3), for each value of η employment growth ceases when the firm's employment is $l_f = \bar{l}$ such that

$$\vartheta^2 A(\bar{l}, \eta) = c\rho\bar{l}. \tag{12}$$

On the basis of this simple relationship, it is easy to identify parametric configurations which are *not* conducive to a well-defined identity for the model's firms. If the marginal asset value $A(\bar{l}, \eta)$ is independent of l , then an equation in the form (12) can endogenously determine the firm's employment level only if $\rho > 0$. Conversely, if there is no exogenous turnover ($\rho = 0$), then the model can endogenize each firm's employment only if $A(\bar{l}, \eta)$ does depend on its first argument.

The two subsections that follow examine simple parameterizations of such special cases. We let the idiosyncratic shifter η take n possible values $\{\eta_i\}$, with $\eta_1 > \eta_2 > \dots > \eta_n$.

Denoting with δ_i the probability intensity of a transition out of state i , we let the row vector $\vec{p}_k = [p_{kj}]$ collect the probabilities of reaching each of the other states upon a transition out of state k (with $\sum_{j=1}^n p_{kj} \equiv 1$, and $p_{kk} \equiv 0$). It will become apparent in Section 3.3 that a fully general structure of transitions across productivity levels would greatly complicate the algebra, especially as regards the computation of the model's endogenous employer-size distribution. Thus, we initially restrict the \vec{p}_k probability vectors to attach unit weight to $j = n$ (the state associated with the lowest possible value of η) for $k > n$, and allow only \vec{p}_n to feature non-zero transition probabilities to all the higher labor-demand states: in words, we assume that all negative innovations to the $\{\eta\}$ process bring the firm's labor demand to its lowest level, while positive innovations are possible only from the latter.

3.1 Constant returns to scale in production

The first model we consider assumes constant returns to labor at the firm's level. Without loss of generality, let $\pi(l, \eta) = \eta$. Wage rates at firms with different productivity levels are readily computed from the relevant specialization of (11),

$$w(l, \eta) = \beta\eta + (1 - \beta)rJ^u, \quad (13)$$

and are quite intuitively constant with respect to employment.

The asset-valuation equations (2) for the marginal value of labor, where $E_f[\cdot]$ need only account for the possibility of a transition to the lowest possible labor productivity indicator, read

$$(r + \rho + \delta_i)A(l_i, \eta_i) = (1 - \beta)(\eta_i - rJ^u) + \delta_i A(l_n, \eta_n), \quad i = 1, \dots, n - 1. \quad (14)$$

Transitions out of state n can take the firm to any other state i , and the no-arbitrage condition for a firm in state n

$$(r + \rho + \delta_n)A_n = \max \left\{ (1 - \beta)(\eta_n - rJ^u) + \delta_n \vec{p}_n \vec{A}, 0 \right\},$$

recognizes that the marginal value of employing any workers may be negative at the lowest-productivity firms, to imply that employers should optimally fire all workers upon receiving a negative shock.

These relationships form a simple system of equations in the state-specific marginal asset values $\vec{A} = \{A_1, A_2, \dots, A_n\}$,

$$A_i = \max \left\{ \frac{(1 - \beta)(\eta - rJ^u) + \delta_i \vec{p}_i \vec{A}}{r + \rho + \delta_i}, 0 \right\}, \quad (15)$$

which depend only on each firm's current productivity and, in particular, are independent of employment levels. Thus, marginal producer surplus is unrelated to the employer's size, and jobs with the same productivity might as well be spread in the economy at large rather than organized in identifiable firms.

The firms we model, however, use a decreasing-returns vacancy-posting technology. As long as natural turnover occurs at a strictly positive rate, increasing vacancy-posting costs imply that a firm's stationary employment level, as defined by (12), is less than indefinitely large for each given η : at a firm whose productivity is η_i , employment increases or decreases according to whether it is smaller or larger than $\bar{l}_i = \vartheta^2 A_i / c\rho$.

If $A_n > 0$, the long-run employment level of low-productivity firms is positive, and no employer-initiated separations ever occur: continuation of all existing matches is efficient, since $A_n > 0$ implies that the employer posts vacancies and workers receive wages in excess of the flow equivalent rJ^u of entering unemployment (which they may only do as a result of the match-specific events modeled by exogenous separations at rate ρ). Employment is only reduced inasmuch as the job destruction induced by natural turnover exceeds job creation induced by vacancy posting.

If instead the low productivity level η_n and/or the probability intensity of exit from that state are so low as to imply that $A_n = 0$, no positive employment level can be stationary at firms in the lowest productivity state. In turn, this implies that all employees are fired when a firm receives a negative shock. In this case, the solution of the system (15) reads

$$\begin{aligned} A_n &= 0, \\ A_i &= (1 - \beta) \frac{\eta - rJ^u}{r + \rho + \delta_i}, \quad i = 1, \dots, n - 1, \end{aligned} \quad (16)$$

and it is not difficult to compute the steady-state size distribution of employment levels (see Appendix A).

We proceed to illustrate the character of the solution in the context of a numerical exercise, parameterized as in Table 1, and to discuss its general implications for the joint distribution of firm employment and wages. Figure 1 plots the size distribution of firms (derived in the Appendix) across positive-employment firms, partitioning the density according to labor-demand states consistent with each employment level. Each such partition reaches

zero at the point where the vacancies posted by firms in the corresponding state just offset natural turnover.

To illustrate the model's implications for the wage-size relationship, Figure 2 displays the theoretical wage functions of equation (13) together with the exact expectation of wages conditional on employment, with a scatter of Monte Carlo data drawn from the long-run distribution of employment and business conditions, and with a linear regression line estimated on such data. The model obviously implies a positive relationship between wages and employment. Conditional expectations are non-linear between the different states' long-run employment levels, but the linear regression approximated well the character of the wage-size relationship in this model. Intuitively, higher levels of labor productivity are associated with more intense vacancy-posting activity, hence faster employment growth towards firm-specific stationary employment levels,

$$\bar{l}_i = \left[\frac{(1 - \beta)(\eta_i - rJ_u)}{r + \rho + \delta_i} \right] \frac{\vartheta^2}{\rho c},$$

which are increasing in η_i and, therefore, positively related to wages. As increasingly large employment levels are considered, firms paying increasingly high wages are selected within the steady-state distribution. Further, wage heterogeneity decreases as increasingly large employment levels are considered and firms with low η_i values are selected out of the cross-sectional distribution. Hence, this simple model implies not only that mean of wages should increase, but also that their variance should decrease with employer size. Both implications are consistent with empirical evidence.

The model, however, is less realistic in other respects. In particular, within the labor market illustrated by Figure 2 employment increases at all continuing firms, and all dismissals are the result of plant closures. Transitions across positive-value states would of course be possible for more general probabilistic structures, but employment would not be reduced by firing decisions: in such a configuration of the model, employers receiving negative shocks would adjust their vacancy-posting activity, and experience employment reduction only insofar as hiring falls short of natural turnover. A positive wage-size association could still be expected, but a negative relationship between employer size and wage dispersion would not be a natural implication of the model. The model we consider next has quite different implications in such respects, also affords a closed-form solution, and provides a qualitatively realistic characterization of endogenous relationships between employer size and wages.

3.2 Decreasing returns to labor

In the second special parameterization of the model, we rule out natural turnover by setting $\rho = 0$ and, to ensure the existence of well-defined, finitely sized firms, we follow Bertola and Caballero (1994) in letting labor's marginal revenue product be a linearly decreasing function of employment,

$$\pi(l, \eta) = \eta - \sigma l. \quad (17)$$

The integral in equation (11) can again be solved explicitly to obtain a characterization of wages paid by hiring firms in terms of their employment level l and current labor demand strength η :

$$\begin{aligned} w(l, \eta) &= l^{-1/\beta} \int_0^l u^{\frac{1-\beta}{\beta}} (\eta - \sigma u + (1 - \beta) r J^u / \beta) du \\ &= \beta \eta + (1 - \beta) r J^u - \frac{\beta}{1 + \beta} \sigma l. \end{aligned} \quad (18)$$

This wage schedule, also derived by Bertola and Caballero (1994), implies that the wage paid by a hiring firm is *declining* in its size. As was the case in the context of the simpler model studied above, however, this wage schedule can be consistent with empirical evidence of a positive wage-size effect when the jointly endogenous nature of wage and employment levels is properly taken into account. Wages not only decrease with employment, but also increase with the idiosyncratic shifter η_i : at any given employment level, firms with stronger labor demand pay higher wages in an environment where matching is costly and surplus is shared with current employees. Since higher productivity induces employers to post more vacancies, and speeds up employment growth, high wages can readily be associated with high employment levels.

To study this mechanism formally in the present context, we need to compute the employer's asset values. This is a quite complex task in general, but is relatively simple under the assumptions of this section. Crucially, the linear form of the marginal productivity function and of the wage schedule (18) implies that

$$\pi(\cdot) - \frac{\partial(w(\cdot)l)}{\partial l} = (1 - \beta) \left(\eta - r J^u - \frac{\sigma}{1 + \beta} l \right)$$

is also linear in l . Further, in the absence of exogenous labor turnover, we can replace \dot{l} with $\frac{\vartheta^2}{c} A(\cdot)$ in (2). Finally, since all transitions from states $1, \dots, n - 1$ bring firms to the

lowest-productivity state indexed by n , the expression for expected shock-induced capital gains is simple:

$$E_f [\Delta A(\cdot)] = \delta_i [\max\{A(l, \eta_n), 0\} - A(l, \eta_i)], \quad i = 1, \dots, n - 1.$$

In words, firms which are currently in state i might in principle find it optimal to retain their current labor force l upon receiving a negative labor demand shock, but should fire as needed to ensure that $A(\eta_n, l) \not\leq 0$. Since such transitions are the only possible source of unemployment inflows in the present context, firms receiving negative labor demand shocks must indeed fire in the stationary configuration of the labor market, to imply that $A(\eta_n, l) = 0$.

Thus, the asset-valuation equations of firms in states $i = 1, \dots, n - 1$ read

$$(r + \delta_i)A(l, \eta_i) = (1 - \beta) \left(\eta - rJ^u - \frac{\sigma}{1 + \beta} l \right) + \frac{\vartheta^2}{c} A(l, \eta_i) \frac{\partial A_i(l, \eta_i)}{\partial l} \quad (19)$$

and, as in Bertola and Caballero (1994), have solutions in the form

$$\begin{aligned} A(\cdot) &= a_i + b_i l, \\ b_i &\equiv b(\delta_i) = \frac{c(r + \delta_i)}{2\vartheta^2} - \frac{1}{2} \sqrt{\frac{c^2(r + \delta_i)^2}{\vartheta^4} + \frac{4c(1 - \beta)\sigma}{\vartheta^2(1 + \beta)}} < 0, \\ a_i &\equiv a(\eta_i, \delta_i) = \frac{(1 - \beta)(\eta_i - rJ_u)}{r + \delta_i - \frac{\vartheta^2}{c} b_i}, \end{aligned} \quad (20)$$

if divergent homogenous terms are ruled out.

The simple linear form of marginal value functions at vacancy-posting firms makes it easy to characterize the relationship between the wage and employment level of a firm which is *not* hiring, and is indifferent to firing the marginal employee. Since $A(\eta_n, l_n) = 0$, current marginal cash flows and expected capital gains on the marginal employee must exactly offset each other at such a firm. Formally,

$$\eta_n - \sigma l_n - w(l_n, \eta_n) - \frac{\partial w(l_n, \eta_n)}{\partial l} l_n + \delta_n \vec{p}_n A(l_n, \vec{\eta}) = 0, \quad (21)$$

where l_n is the employment level of a firm that is posting no vacancies; the column vector $A(l_n, \vec{\eta})$ collects the marginal shadow values of the firm's current employment stock upon transitions to other (hiring) states, which occur with probability intensity δ_n ; and the row vector \vec{p}_n contains the transition probabilities from state n to the other states.

If firms where $A(l_n, \eta_n) = 0$ have strictly positive employment ($l_n > 0$), then the participation constraint (5) applies and the wage $w(l_n, \eta_n)$ should be such that employees are

as well off when retained as they would be in unemployment. In the event of positive labor demand shocks, which generate economic rents for existing matches when the matching process is costly and time-consuming, current employees of a firm that is posting no vacancies can negotiate a wage increase and share in their employer's good luck. Thus, the human capital of hoarded workers includes expected capital gains, amounting to the usual proportion $\beta/(1-\beta)$ of their employer's marginal gains. In flow terms, the hoarded worker's equilibrium requirement reads

$$rJ^u = w(l_n, \eta_n) + \frac{\beta}{1-\beta} \delta_n \vec{p}_n A(l_n, \vec{\eta}), \quad (22)$$

where $A(l_n, \vec{\eta}) = \vec{a} + \vec{b}l_n$ for $\vec{a} = [a_i]$ and $\vec{b} = [b_i]$ are the column vectors collecting the intercepts and slopes of the various hiring states' marginal shadow value functions from (20). The left-hand side of (22) is constant in steady state, and total differentiation yields

$$\frac{\partial w(l, \eta_n)}{\partial l} = -\frac{\beta}{1-\beta} \delta_n \vec{p}_n \frac{\partial A(l, \vec{\eta})}{\partial l} = -\frac{\beta}{1-\beta} \delta_n \vec{p}_n \vec{b} \equiv k_n > 0. \quad (23)$$

Thus, we can write

$$w(l_n, \eta_n) = h_n + k_n l_n \text{ for } h_n = rJ^u - \frac{\beta}{1-\beta} \delta_n \vec{p}_n a(\vec{\eta}, \vec{\delta}). \quad (24)$$

Like (18), this simple relationship offers insights of some generality. Since wages paid by hiring firms are negatively related to their employment level, workers hoarded by a firing firm can look forward to larger capital gains when their employer's labor force is smaller. Hence, a smaller wage flow suffices to keep fewer workers employed by a firing firm, and when characterizing the equilibrium it is necessary to take into account the fact that wages paid by such firms are a *positively* sloped function of their employment level.⁴

Using (24) in (21) and recalling the definition of k_n in (23) yields

$$l_n = \frac{\eta_n - h_n + \delta_n \vec{p}_n a(\vec{\eta}, \vec{\delta})}{\sigma + \frac{1+\beta}{\beta} k_n}. \quad (25)$$

This completes the closed-form solution of the special case where all negative shocks bring a firm's labor demand to its lowest possible level, and positive shocks can only occur from

⁴This is neglected by Bertola and Caballero (1994), who treat the wage rate paid by zero-vacancy firms as a constant in their derivations.

that same state. For this simple probabilistic structure, it is also possible to obtain closed-form solutions for the steady-state distribution of endogenous employment across firms (see Appendix A).

To illustrate the character of the solution and discuss its general implications, we again propose a parameterized example (see Table 1). Since the advantage of this model over that outlined in Section 3.1 is the ability to generate realistic hiring and firing by continuing firms, we let the probabilistic structure be such as to imply that most firms are small and subject to frequent but not very pronounced labor demand fluctuations, and that shocks are increasingly unlikely for firms with higher and higher labor demand levels. Hence, large firms very seldom experience state transitions, which would take them to the lowest possible labor demand level and cause their employment level to collapse. This structure of (unobservable) productivity shocks probabilities prevents intense turnover at large firms, ensuring that the model is qualitatively consistent with available evidence on the relationship between job flows and firm size.

Table 1 also reports the equilibrium values of V/U and J^u , computed on the basis of the ergodic distribution of firm sizes across labor demand levels (the relevant algebra is outlined in Appendix B). A discrete mass of labor is located at zero-vacancy firms (which employ about 30% of the labor force for the parameters listed in Table 1), and the remainder of employment is distributed continuously across hiring firms. Table 2 reports gross job flows for various firm sizes. These are statistics computed on Monte Carlo samples drawn from the analytic long-run distribution computed in the Appendix, and whose continuous portion is illustrated in Figure 3. The parameters in Table 1 imply a structure of job flows consistent with the empirical evidence on job flows and firm size reported in OECD (1994). Net employment changes (the difference between job creation and destruction) are negatively related to firm size, since the steady-state situation of interest features mean-reverting employment dynamics. The firm size distribution generated by our model and parameters is roughly consistent with the evidence reported in the OECD study: for example, the smallest size category accounts for a third of employment in our table, and the employment share of firms with less than 20 employees ranges from 27% in Canada to more than 40% in New Zealand in the OECD tables. More interestingly, we see in the Table that job reallocation (the sum of job creation and destruction) declines sharply as a function of size, which is also consistent with available evidence. Our model and parameters generate employment stability for large firms in the form of unrealistically large, if unlikely, job destruction events. On average, however, the model-generated job flows are qualitatively consistent with available

evidence.

Figure 4 illustrates the model’s implications for the wage-size relationship, plotting the downward-sloping wage/employment schedules for the various possible levels of labor demand. Since each hiring firm’s employment and wages evolve along a standard downward-sloping demand curve, a regression where the strength η of labor demand were controlled for would yield a negative relationship between wages and employment. If η_i is not observable, however, the model can generate a positive cross-sectional relationship between wages and employment. In Figure 4, the shape of the exact conditional expectation is rather complex, and non linear (we briefly discuss in Section 4 the possible empirical realism of its nonlinear shape). In the figure, Monte Carlo data drawn from the long-run distribution of employment and business conditions derived in the Appendix are scattered over the theoretical wage functions of equation (18). Clusters of points appear at the lower limit of the wage and employment distributions, and along the downward-sloping wage loci identified by equation (18) for the various values of η which induce hiring. The thick upward-sloping line in the figure plots predicted values from an OLS regression of wages on employer size. Since our model features homogeneous labor and uniform market power, this linear regression corresponds to what a researcher might find in data where all *worker* characteristics are controlled for, but wage differentials are generated by dynamic heterogeneity across *firms* under diminishing returns to labor. The dispersion of wage rates is also decreasing as larger employment levels are considered, and we return below to a discussion of the empirical relevance of this theoretical implication.

3.3 Numerical solution

It is certainly very restrictive to rule out direct transitions across productivity states where firms are hiring. In what follows, we outline a numerical solution procedure for a general structure of transition probabilities. The procedure, described in more detail by Garibaldi (1996), becomes unwieldy very quickly as the number of possible productivity states increases. However, the additional qualitative insights offered by the possibility of transitions across hiring states are satisfactorily illustrated by the case where η may take only three different values, $\eta_1 > \eta_2 > \eta_3$.

A firm with $\eta_f = \eta_3$ is not hiring, is indifferent to firing at the margin, and employs $l_n = l_3$ units of labor. If $\eta_f = \eta_1$, conversely, we can be sure that the firm is hiring, albeit at a slower and slower rate, as in the model of the previous section. What is new is that firms in the intermediate state, with $\eta_f = \eta_2$, may or may not be posting vacancies, depending

on the time and direction of the last transition they experienced or, equivalently, on their employment level.

When $\eta_f = \eta_2$, the firm should be hiring only if its employment is lower than the \bar{l}_2 level defined implicitly by

$$A(\bar{l}_2, \eta_2) = 0. \quad (26)$$

Any state transition from the highest productivity states induces the firm to fire as long as $l \geq \bar{l}_2$, and annihilates the shadow value of labor. In the $l \geq \bar{l}_2$ region, accordingly, the model features the same differential equation as that encountered above, with solution $A(l, \eta_1) = a_1 + b_1 l$ as in (20), and the wage schedule $w(l, \eta_1)$ is given by (18). The numerical solution for other regions can then proceed from higher to lower productivity states and from larger to smaller productivity levels. When employment is below the critical \bar{l}_2 level defined by (26), the firm is hiring (at different rates) both when $\eta = \eta_2$ and when $\eta = \eta_1$. If states 1 and 2 communicate directly, the relevant shadow value functions must be solved simultaneously, yielding a system of two functional equations:

$$(r + \delta_1)A(l, \eta_1) = (1 - \beta) \left(\eta_1 - rJ^u - \frac{\sigma}{1 + \beta} l \right) + \frac{\vartheta^2}{c} A(l, \eta_1) \frac{\partial A(l, \eta_1)}{\partial l} + \delta_1 p_{12} A(l, \eta_2), \quad (27)$$

$$(r + \delta_2)A(l, \eta_2) = (1 - \beta) \left(\eta_2 - rJ^u - \frac{\sigma}{1 + \beta} l \right) + \frac{\vartheta^2}{c} A(l, \eta_2) \frac{\partial A(l, \eta_2)}{\partial l} + \delta_2 p_{21} A(l, \eta_1).$$

For $l' \geq \bar{l}_2$, $A(l', \eta_2) \equiv 0$, and evaluating the second equation in (27) at \bar{l}_2 yields

$$(1 - \beta) \left(\eta_2 - rJ^u - \frac{\sigma}{1 + \beta} \bar{l}_2 \right) + \delta_2 p_{21} (a_1 + b_1 \bar{l}_2) = 0.$$

To extend the solution numerically below the \bar{l}_2 level identified by this equation, we evaluate the equations in (27) at two different values l and l' , subtract term by term, and write

$$A(l, \eta_i) \approx A(l', \eta_i) - \frac{\partial A(l, \eta_i)}{\partial l} \Delta l, \quad i = 1, 2, \quad (28)$$

where the approximation is satisfactory for $\Delta l = l' - l$ small. Given values of the levels and derivatives of the shadow value functions evaluated at $l' = l + \Delta l$, the equations in (27) may be rewritten using (28) and rearranged to obtain two equations in $\mathcal{X} = \partial A(l, \eta_i) / \partial l$ and

$\mathcal{Y} = \partial A(l, \eta_2) / \partial l :$

$$\begin{aligned} \frac{\vartheta^2}{c} \mathcal{X}^2 - \left((r + \delta_1) + \frac{\vartheta^2}{c} \frac{A(l', \eta_1)}{\Delta l} \right) \mathcal{X} - (\sigma + 2k_1) + p_{12} \delta_1 \mathcal{Y} + \frac{\vartheta^2}{c} \frac{A(l', \eta_1)}{\Delta l} \frac{\partial A(l', \eta_1)}{\partial l} &= 0, \\ \frac{\vartheta^2}{c} \mathcal{Y}^2 - \left((r + \delta_2) + \frac{\vartheta^2}{c} \frac{A(l', \eta_2)}{\Delta l} \right) \mathcal{Y} - (\sigma + 2k_2) + p_{21} \delta_2 \mathcal{X} + \frac{\vartheta^2}{c} \frac{A(l', \eta_2)}{\Delta l} \frac{\partial A(l', \eta_2)}{\partial l} &= 0. \end{aligned} \tag{29}$$

This quadratic system is easily solved numerically, and makes it possible to extend the solution to employment levels below \bar{l}_2 : one evaluates (29) for $l' = \bar{l}_2$; obtains a solution for the shadow value derivatives at $\bar{l} - \Delta l$ from (29); computes approximate values of the shadow value levels in the same point from (28); and proceeds recursively backwards to obtain shadow values for a sequence of employment levels separated by (arbitrarily small) discrete differences Δl . The lowest recurring employment level is approximated by the numerical procedure as the term in that sequence which approximately satisfies a zero-shadow-value condition when the firm is in the lowest possible labor-demand position, namely, the condition equivalent to (21) for the more complex model introduced here.

The steady-state distribution of wage and employment levels cannot be characterized analytically, but Figure 5 reports the results of Monte Carlo draws from the wage history of a single firm. The nonlinearity introduced by non-zero probabilities of transitions across hiring states affects both firms' marginal valuations of labor and employees' human capital functions, but not the wage functions, for which the linear form of (18) is valid in all cases. One feature is qualitatively different from the simpler case of Section 3.2: since large firms with $l > \bar{l}_2$ only fire $\Delta l = l - \bar{l}_2$ when their labor demand drops to the intermediate state 2, the Figure features mass points of employment at \bar{l}_2 as well as at \bar{l}_3 . The distribution is otherwise similar to that of Figure 4, and realistically implies a positive cross-sectional wage-size relationship.

The model also has implications for the relationship between wages and employment levels across firms where adjustment rents are completely dissipated, i.e., no vacancies are posted. These implications, which we review next, are akin to the “compensating differentials” notion of other empirical approaches to wage-size relationships. Since η is mean reverting, in fact, a particular dynamic sort of compensating differentials plays a role: the structure of the labor market offers different future outlooks to employees by firms at different points in the range of productivity states. To see this, consider that, by (5), all firms with employment \bar{l}_i such that $A(\bar{l}_i, \eta_i) = 0$ (i.e., firms which post no vacancies) should offer their employees the same *human capital* level J^u . The wage rate, however, is generally different across such firms:

while things can only get better for firms with the lowest possible labor demand level, the future outlook is grim for the largest firms—those approaching no-vacancy employment at the highest labor demand level. Hence, while at small zero-vacancy firms rJ^u results from a combination of a low wage $w(l^{min}, \eta_n)$ and an expectation of positive capital gains, the largest firms can offer no capital gains, and they pay a higher wage $w(\bar{l}_1, \eta_1) = rJ^u$. Formally,

$$w(\bar{l}_1, \eta_1) > w(l^{min}, \eta_n), \quad w(\bar{l}_1, \eta_1) \geq w(l_i, \eta_i) \quad \forall i > n. \quad (30)$$

In general, a positive association between wages and labor-demand strength (hence employment) can still be expected whenever a stationary process disturbs individual employers' demand schedules, and mobility is costly for workers.⁵

4 Discussion

The analytic and simulation results of the model specification of Section 3.2 illustrate the more general empirical implications of the modeling perspective we propose. While the model-generated data feature a sizable wage-size effect, a causal interpretation of the upward-sloping regression line is of course unwarranted. Wage-size effects reflect the interplay of exogenous shocks to labor demand with the labor-supply constraints introduced by slow matching. Table 3 reports correlation matrix in the Monte-Carlo data of Figure 4 for four observable variables (wages, employer size, job creation and profit per-capita) and the unobservable labor demand index η . Wages are correlated not only with employment, but also with average worker surplus (per-capita profits).⁶ Further, wages and job creation are strongly positively correlated in the model. This relationship highlights the dynamic nature of the model, which also implies that wages should be higher for faster growing firms.

⁵Garibaldi (1996) characterizes more generally models where transitions across positive-vacancy states are possible, and derives a set of conditions ensuring a monotonic relationship between wages and the “long-run” employment levels consistent with no vacancy posting.

⁶Per-capita profits are defined as the excess of revenue over labor-related costs, normalized by employment. Integrating the linear marginal surplus (17) and recalling that hiring costs are also a quadratic function of vacancies posted, each firm's thusly defined “per-worker operating surplus” reads

$$\Phi(l, \eta_i) = \eta_i - \frac{\sigma}{2}l - w(l, \eta_i) - \frac{cv^i(t)^2}{2l}.$$

4.1 Empirical realism

Table 4 reports the mean and the coefficient of variation within subsets of the same simulated data set. The data are partitioned in five size categories, scaled by the model’s minimum firm size l_n . Consistently with the exact conditional expectation plotted in Figure 4, average wages increase with small employment levels and decline slightly across the largest size categories. The model is also consistent with the evidence collected by Davis and Haltiwanger (1991,1996) on the relationship between wage dispersion and employer size. Table 4 reports the standard deviation of wages across model-generated plant observations within each size group. In the Davis and Haltiwanger (1996) data, the coefficient of wage variation is lower by 20 percentage points for nonproduction workers (and by 8 percentage points for production workers) across establishments with 1000 or more employees than across establishments with less than 25 employees. The model readily rationalizes such patterns of wage dispersion. Quite intuitively, heterogeneous labor demand levels and matching frictions in the adjustment process imply that firms can be in very different states when their employment is low: some small firms are hoarding labor and pay very low wages, while others have recently been hit by favorable business condition shocks, are paying very high wages, and are expected to grow larger in the future. Hence, wage dispersion is higher among small firms than among large firms.

Katz and Summers (1989), in their study on industry rents for the US manufacturing industry, report evidence of a growth effect in the average wage differentials. In their data set, after controlling for the average worker experience and for the level of human capital, wage premia are higher in industry with larger establishments *and* faster growth rates. Interestingly, and consistently with our theoretical perspective, industry growth rates are quantitatively more important than industry average sizes in determining the wage premium.

Our analysis suggests that empirical wage-size effects could be rationalized by focusing on dynamic firm-level factors, and that the same dynamic phenomena are relevant to the heteroskedastic patterns of wages across employer sizes. Empirical studies that estimate individual wage regressions, conversely, are unlikely to reduce the significance of wage-size effects by controlling for observable worker characteristics or invoking unobservable ones. In individual wage regressions, in fact, employer size should remain significant if—as in our model—at any given time it is positively correlated with the level of firm-specific labor demand; such wage-size effects are theoretically unrelated to individual tenure, human capital, and other worker characteristics. Similarly, wage dispersion among small firms is related to firm-specific labor demand factors.

Further analysis of such issues might benefit from recently developed data sets with matched information on employer and employee characteristics. Abowd *et al.* (1999) find in such a data set that worker characteristics account for a large proportion of the raw wage-size effect, but do leave unexplained about 25% of it. Their empirical model allows for cross-sectional heterogeneity of wage-size effects. Consistently with our model’s implications, the evidence indicates that the relationship between pure firm effects and employer size is positive on average, but not necessarily monotonic. Using individual and firm level data, Belzil (1997) investigates how individual wages vary with firm-level measures of job creation and worker turnover. Controlling for firm and worker characteristics, Belzil finds that male wages are higher at firms with contemporaneously higher job creation rates. While results vary across samples and estimation methods, the estimated effects are very large: 2% to 4% higher wages for each additional percentage point of annual net job growth.

A structural interpretation of significant size and growth effects in firm-level wage equations is difficult, however. In order to fit a structural wage equation in the form (18), in fact, one should take into account that firm-level employment—the regressor of interest—is endogenous and correlated to labor demand shifters, denoted η above, which may be unobservable and therefore omitted in typical empirical specifications. Observable firm-level profits, while also endogenous, are predicted to be significant in wage equations by our theoretical perspective. The point is relevant to evidence of persistent correlation between wages and various measures of an employer’s ability to pay. Blanchflower *et al.* (1996) merge data on individual wages with industrial data on profits, and show that in various wage equations industrial profits are persistently significant. Previous empirical work similarly found wages to be positively correlated with various measure of employer ability to pay (Dickens and Katz 1987, Katz and Summers 1989). Blanchflower *et al.* emphasize that a rise in sector’s profitability leads to a *long-run* increase in wages, and conclude that pay determination appears to exhibit elements of rent-sharing. Rent-sharing is implicit in our theoretical approach’s wage mechanism, and from its point of view it is far from surprising that profits should enter significantly in cross-sectional regressions.

Finally, Katz and Summers (1989) find that when the capital-output per worker is included in industry wage regressions, the size-wage effect is still significant, but quantitatively much smaller. While our simple model does not account explicitly for the role of capital investment in the adjustment process, the implications of such an extension would be consistent with the evidence if capital also faces adjustment costs, and is complementary to labor in the production function. In the appropriate extension of our model, capital inten-

sity would be another endogenous variable correlated to the unobservable demand shifter η_i . Thus, its empirical significance is not surprising, but it would be difficult to give a structural interpretation to its estimated coefficient.

4.2 Further research

The technical methods and results of this paper may be useful in further theoretical and empirical work. We show that theoretical matching models with decreasing returns to employers' recruiting effort can rationalize established evidence on the joint distribution of firm sizes and wage levels, and argue that decreasing returns to labor in production further increase the realism of such models—as they make it possible for job destruction to take the form of dismissals by continuing, labor-hoarding firms rather than of plant-closure events or partial replacement of exogenous quits. The models of interest tend to be quite complex. Complementing and extending research reported in Bertola and Caballero (1994), Mortensen (1999), and other recent papers, our work strives to highlight general theoretical insights and illustrate them by analytic solution of parametric examples. We also outline, however, how more complex and realistic configurations of the class of models proposed may be solved numerically in further research.

Our theoretical efforts are motivated by simple qualitative evidence on the relationship between the mean and variance of wages on the one hand, and the size of firms or establishments on the other. In turn, our theoretical perspective and results suggest interesting venues for further empirical work, particularly as regards the analysis of matched employer-employee data sets. Our preferred model predicts that wage dispersion among observationally equivalent workers should decline monotonically at increasingly large firms, and interprets the phenomenon in terms of dynamic selection of increasingly strong labor demand schedules within a distribution of (unobservable) firm types. The prediction deserves to be verified on other data than those studied by Davis and Haltiwanger (1996), such as e.g. those analyzed by Abowd *et al.* (1999), and the interpretation may be tested by dynamic econometric analysis of data panels with sufficiently long time dimension. Further, our structural perspective on the joint distribution of wage and employment levels suggests that wage-size effects, while positive on average, may be negative at large, decreasing-returns firms. The methods of Abowd *et al.* (1999), who allow for heterogeneity and nonlinearities in estimated wage-size effects, are quite relevant in this respect. They may be extended and refined in light of our theoretical insights, and the evidence may call for explicit theoretical treatment of heterogeneous capital adjustment across firms of different sizes.

Additional promising directions of further research may exploit the equilibrium nature of the models we propose. While this paper does not emphasize the endogenous character of market tightness, it would not be difficult to explore (numerically) the implications of different matching environments and technological structures. In general, less important matching frictions, smaller dispersion and/or higher stability of labor demand schedules across firms, and smaller surplus shares for workers should all tend to reduce the relevance of the dynamic phenomena we focus on. Any evidence on such features may in principle be brought to bear on the character of wage and firm-size distributions across microeconomic units within real-life labor markets.

A Steady-state distributions

What follows characterize steady-state firm size distributions in the context of the models proposed and solved in Sections 3.1 and 3.2.

In steady state, the probability mass of firms which flows into each productivity state must balance with the probability mass of firms that flows into the other states. If we indicate with $\vec{\pi}$ the ergodic distribution over the productive states, by virtue of our simple structure of productivity changes, this equilibrium requirements reads

$$\delta_n p_{ni} \pi_n = \delta_i \pi_i \quad \forall : i > n, \quad (31)$$

and

$$\sum_{i=1}^{n-1} \delta_i \pi_i = \delta_n \pi_n, \quad (32)$$

for the lowest labor demand state. The expressions in (31) and (32) form a homogeneous, rank-deficient system of n linear equations which, together with the summing-up condition $\sum_i \pi_i = 1$, easily yields a solution for $\vec{\pi}$.

In what follows we shall let $f(l, \eta)$ be the joint density of employment and business conditions within the subset of the model's state space where employment is positive. Transitions into and out of state (l, i) occur because of hiring and natural turnover when η is constant, and because of firing when a negative labor demand shock is realized.

In the model of Section 3.1, the joint distribution of employment and labor demand indexes features a mass π_n at the $(0, \eta_n)$ point. The shadow value of labor has constant value A_i within each state $i = 1, \dots, n - 1$ where vacancy posting (at an also constant rate) is optimal. Hence, the density across employment levels solves the differential equation

$$\frac{\partial f(l, \eta)}{\partial l} (k_i - \rho l) = -\delta_i f(l, \eta),$$

where $k_i = \vartheta^2 A_i / c$. Solutions of this equation have the form

$$f(l, \eta_i) = \lambda_i (k_i - \rho l)^{\frac{\delta_i}{\rho}},$$

where the constant of integration λ_i is determined by the summing up condition

$$\lambda_i \int_0^{l_i^{\max}} (k_i - \rho l)^{\frac{\delta_i}{\rho}} dl = \pi_i.$$

Integrating, $\lambda_i = \pi_i(\delta_i + \rho)k_i^{(-1-\frac{\delta_i}{\rho})}$, and the density and cumulative distribution functions are both available in closed form.

In the model of Section 3.2, the joint distribution of employment and business conditions features a mass of measure π_n at the (l_n, η_n) point. The Kolmogorov forward balance conditions

$$\frac{\partial f(l, \eta_i)}{\partial l} \dot{l} + \delta_i f(l, \eta_i) = 0, \quad (33)$$

where

$$\dot{l}_i = \frac{\vartheta^2}{c} A(l, \eta_i) = \frac{\vartheta^2}{c} (a_i + b_i l),$$

have solutions in the form

$$f(l, \eta_i) = \lambda_i (a_i + b_i l)^{-\mu_i}, \text{ for } \mu_i = \frac{\delta_i c}{\vartheta^2 b_i}. \quad (34)$$

The constant of integration λ_i is determined in closed form by the summing-up condition

$$\lambda_i \int_{l^{min}}^{\bar{l}_i} (a_i + b_i l)^{-\mu_i} dl = \pi_i, \quad (35)$$

where π_i is one element of the solution to (31) and (32) and \bar{l}_i , as defined in (12), is the maximum size attainable by a firm in state i .

B Equilibrium

We sketch in this Appendix how the models we study may be closed, as in Bertola and Caballero (1994), by a standard specification of the process by which vacancies and unemployed workers are matched. Under the convenient and fairly realistic assumption that the matching technology has constant returns to scale and a Cobb-Douglas functional form, the probability intensity of matching events for each vacancy is $\vartheta = \xi (V/U)^\nu$ per unit time, where $-1 < \nu < 0$ and $\xi > 0$. In the steady-state equilibrium of interest, the rate at which each vacancy is matched with an unemployed worker is constant but, in general, depends on the ratio of aggregate vacancies, $V \equiv \int_0^1 v_f df$, to aggregate unemployment, $U \equiv 1 - \int_0^1 l_f df$.

To compute the endogenous quantities U and V , note that, for a constant c and ϑ , each firm's vacancies are linearly related to its labor's marginal value $A(l, \eta)$, which in turn is

linearly related to employment in the models we can solve explicitly. In the model of Section 3.2,

$$v(\cdot) = (a_i + b_i l) \frac{\vartheta}{c}, \quad (36)$$

and a similar expression with $b_i = 0$ holds true in the model of Section 3.1. Aggregating,

$$V = \frac{\vartheta}{c} \sum_{i=1}^{n-1} \int_{l_n}^{\bar{l}_i} \lambda_i (a_i + b_i l)^{-\mu_i+1} dl, \quad (37)$$

Aggregate employment can be similarly obtained by first integrating the dynamics of each firm's employment, and then computing mean employment across all hiring and firing firms. In the model of Section 3.2, the relevant expression is

$$L = \sum_{i=1}^{n-1} \int_{l_n}^{\bar{l}_i} ((a_i + b_i l)^{-\mu_i} \lambda_i l) dl + \pi_n l_n. \quad (38)$$

An aggregate consistent equilibrium is obtained when the ϑ used to construct individual policies is consistent with the aggregate outcomes in (37) and (38). In equilibrium it must also be the case that the employees (if any) of firms posting no vacancies are indifferent to the outside (unemployment) option,

$$rJ^u = z + \vartheta \frac{V}{1-L} (E_u[J] - J^u). \quad (39)$$

Workers are in fact employed by firing firms in the model of Section 3.2 where, by (7) and (36), the capital gain term in (39) reads

$$E_u[J] - J^u = \frac{\vartheta V}{c U} \frac{\beta}{1-\beta} \frac{1}{1-\pi_n} \sum_{i=1}^{n-1} \int_{l_n}^{\bar{l}_i} \lambda_i A(l, \eta_i) (a_i + b_i l)^{-\mu_i+1} dl, \quad (40)$$

In (40) $1 - \pi_n$ is the probability that an unemployed worker meets a firm of quality i , conditional on finding a vacancy posting firm. Through this equilibrium relationship, the endogenous dynamics of individual-firm employment levels interact with each other in the aggregate labor market. As in Bertola and Caballero (1994), a search routine may be used to compute the model's fixed-point equilibrium.

References

- Abowd, John M., Francis Kramarz, and David N. Margolis, High Wage Workers and High Wage Firms, *Econometrica*, 67 (1999), 251-334.
- Belzil, Christian, Job Creation and Destruction, Worker Reallocation and Wages, working paper (1997), Concordia University Department of Economics.
- Bentolila, Samuel, and Giuseppe Bertola, Firing Costs and Labor Demand: How Bad is Euroclerosis?, *Review of Economic Studies* 57 (1990), 381-402.
- Bertola, Giuseppe, and Ricardo J. Caballero, Cross-Sectional Efficiency and Labour Hoarding in a Matching Model of Unemployment, *Review of Economic Studies* 61 (1994), 435-457.
- Blanchflower, David G., Andrew J. Oswald, and Peter Sanfey, Wages, Profits and Rent Sharing, *Quarterly Journal of Economics* 111 (1996), 226-251.
- Brown, Charles, and James Medoff, The Employer Size-Wage Effect, *Journal of Political Economy* 97 (1989), 1027-1057.
- Burdett, Kenneth, and Dale T. Mortensen, Wage Differentials, Employer Size, and Unemployment, *International Economic Review* 39 (1998), 257-73.
- Cahuc, Pierre, and Etienne Wasmer, Does Intrafirm Bargaining Matter in the Large Firm's Matching Model?, *Macroeconomic Dynamics*, forthcoming, (2000).
- Davis, Steve J., and John Haltiwanger Wage Dispersion between and within U.S. Manufacturing Plants, 1963-86, *Brooking Papers: Microeconomics* (1991), 115-180.
- Davis, Steven J., and John Haltiwanger, Employer Size and the Wage Structure in U.S. Manufacturing, *Annales d'Economie et de Statistique* 41-42 (1996) 323-67.
- Davis, Steven J., John Haltiwanger and Scott Schuh, "Job Creation and Destruction," MIT Press, Cambridge Mass, 1996.
- Dickens, William T., and Lawrence F. Katz, Inter-Industry Wage Differences and Industry Characteristics, in "Unemployment and the Structure of the Labor Markets" (K.Lang and J.Leonard, eds.), Basil Blackwell, Oxford, 1987.

- Garibaldi, Pietro, Job Flow Dynamics and Firing Restrictions, *European Economic Review* 42 (1998), 245-275.
- “Job Creation, Job Destruction and Employment Reallocation. Theory and Evidence,” Chapter 4. Unpublished Ph.D. dissertation, London School of Economics, 1996.
- Ince, E.L., “Ordinary Differential Equations,” Dover, New York, 1956.
- Katz, Lawrence F., and Lawrence H. Summers, Industry Rents: Evidence and Implications, *Brookings Papers on Economic Activity (Microeconomics)* 1 (1989), 209-275.
- Krueger, Alan B., and Lawrence H. Summers, Efficiency Wages and the Inter-Industry Wage Structure, *Econometrica*, 56 (1988), 259-293.
- Masters, Stanley H., An Interindustry Analysis of Wages and Plant Size, *Review of Economics and Statistics* 51 (1969), 495-501.
- Mellow, Wesley, Employer Size and Wages, *Review of Economics and Statistics* 64 (1982), 341-345.
- Mortensen, Dale T., Equilibrium Unemployment with Wage Posting: Burdett-Mortensen meet Pissarides, working paper (1998), Northwestern University.
- Mortensen, Dale T., Modeling Matched Job-Worker Flows, working paper (1999), Northwestern University.
- Mortensen, Dale T., and Christopher A. Pissarides (1994) “Job Creation and Job Destruction in the Theory of Unemployment, *Review of Economic Studies* 61, 397-415.
- Mortensen, Dale T., and Christopher A. Pissarides, Recent Development in Search Models of the Labor Market, in “Handbook of Labor Economics Vol. 3” (O. Ashenfelter and D. Card eds), 2567-2628, Elsevier Science, Amsterdam, 1999.
- OECD, Job Gains and Job Losses in Firms, in *Employment Outlook*, 103-166, OECD, Paris, 1994.
- Oi, Walter Y., The Fixed Employment Costs of Specialized Labor, in “The Measurement of Labor Cost” (Jack E. Triplett, ed.) University of Chicago Press, Chicago, 1983.

- Oi, Walter Y., and Todd L. Idson, Firm Size and Wages, in “Handbook of Labor Economics Vol. 3” (O. Ashenfelter and D. Card eds, 2165-2214, Elsevier Science, Amsterdam, 1999).
- Podgursky, Michael , Unions, Establishment Size, and Intraindustry Threat Effects, *Industrial and Labor Relations Review* 39 (1986), 277-284.
- Saint Paul, Gilles, The High Unemployment Trap, *Quarterly Journal of Economics* 110 (1995), 527-49.
- Stole, Lars A., and Jeffrey Zwiebel, Organizational Design and Technology Change under Intrafirm Bargaining, *American Economic Review* 86 (1996), 195-222.
- Weiss, Leonard W., Concentration and Labor Earnings, *American Economic Review* 56 (1966), 96-117.
- Weiss, Andrew and Henry J. Landau, Wages, Hiring Standards, and Firm Size, *Journal of Labor Economics* 2 (1984), 479-499.

Table 1: Parameter values for sample solution and Figures

(Common parameters:)						
	Matching function elasticity	ν		-0.400		
	Matching function scale	ξ		1.000		
	Vacancy cost slope	c		2.560		
	Discount rate	r		0.060		
	Employees' bargaining share	β		0.400		
	Income when unemployed	z		0.800		

(Model with Constant Return to Scale:)						
	$\eta_1=0.56$	$\eta_2=0.53$	$\eta_3=0.51$	$\eta_4=0.48$	$\eta_5=0.46$	
	$\delta_1=0.06$	$\delta_2=0.07$	$\delta_3=0.08$	$\delta_4=0.09$	$\delta_5=0.31$	
$\{p_{ni}\}_1^5$	0.098	0.164	0.276	0.463		
	Labor demand slope	σ		0.330		
	Natural Turnover	ρ		0.200		
	Market tightness	V/U		2.321		

(Linear Model with Decreasing Return to Labor;)						
(No transition across hiring states:)						
	$\eta_1=1.67$	$\eta_2=1.49$	$\eta_3=1.33$	$\eta_4=1.19$	$\eta_5=1.06$	$\eta_6=0.95$
	$\delta_1=0.06$	$\delta_2=0.07$	$\delta_3=0.09$	$\delta_4=0.10$	$\delta_5=0.12$	$\delta_6=0.15$
$\{p_{ni}\}_1^6$	0.032	0.053	0.089	0.150	0.252	0.424
	Labor demand slope	σ		0.330		
	Natural Turnover	ρ		0.000		
	Market tightness (equilibrium)	(V/U)		2.321		
	Unemployed welfare flow (equilibrium)	rJ^U		0.874		

(Linear Model with Decreasing Return to Labor;)						
(Transitions across hiring states allowed:)						
	$\eta_1=1.90$	$\eta_2=1.69$	$\eta_3=1.46$			
	$\delta_1=0.15$	$\delta_2=0.18$	$\delta_3=0.27$			
	Transition matrix:					
	0.00	0.50	0.50			
	0.50	0.00	0.50			
	0.36	0.64	0.00			
	Labor demand slope	σ		0.325		
	Natural Turnover	ρ		0.000		
	Market tightness	V/U		2.321		

Table 2: Job Flows by Firm Size

Firm Size	Job Creation	Job Destruction	Job Reallocation	Net Employment Change	Employment Share
$0l_n - 3l_n$	0.165	0.030	0.195	0.135	0.280
$3l_n - 5l_n$	0.231	0.074	0.305	0.158	0.144
$5l_n - 7l_n$	0.140	0.078	0.218	0.063	0.161
$7l_n - 12l_n$	0.103	0.071	0.174	0.033	0.203
$12l_n - \infty$	0.037	0.062	0.099	-0.025	0.212

Table 3: Correlation Matrix

	Labor d.	Wage	Employm.	Job cr.	Profits
Labor d.	1.00	0.93	0.72	0.84	0.31
Wage	0.93	1.00	0.40	0.92	0.10
Employm.	0.72	0.40	1.00	0.34	0.58
Job cr.	0.84	0.92	0.34	1.00	-0.05
Profits	0.31	0.10	0.58	-0.05	1.00

Table 4: Wage Differentials and Wage Dispersion by Firm Size

Firm Size	Mean Wage Differential	Between Plant Coefficient of Variation
$0l_n - 3l_n$	-0.034	0.436
$3l_n - 5l_n$	0.069	0.151
$5l_n - 7l_n$	0.069	0.086
$7l_n - 12l_n$	0.083	0.054
$12l_n - \infty$	0.044	0.018

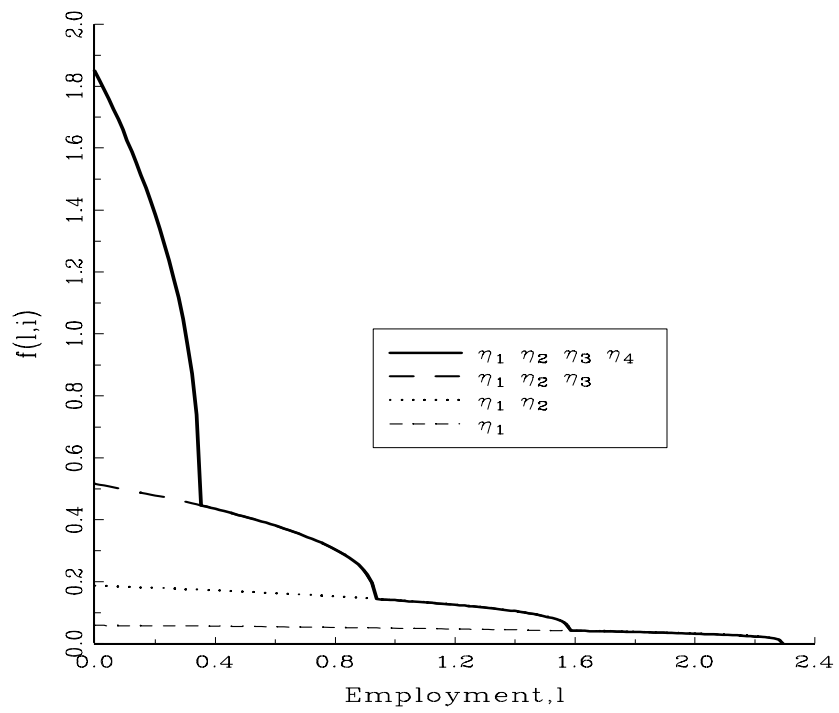


Figure 1: Probability density of employment at hiring firms: constant returns in production and positive turnover.

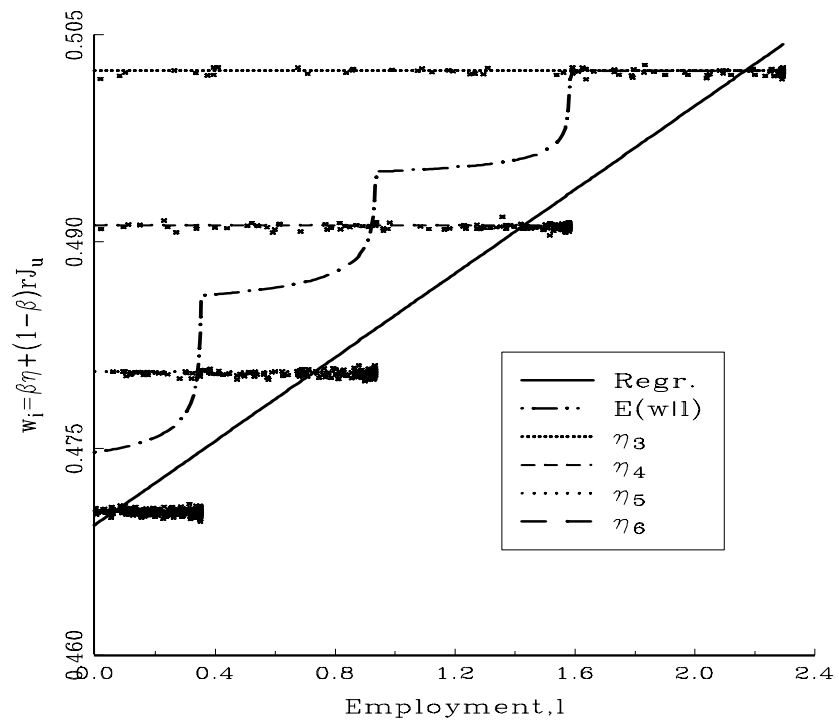


Figure 2: Wage functions: constant returns in production and positive turnover. Monte Carlo sample of 1000 observations, with linear regression.

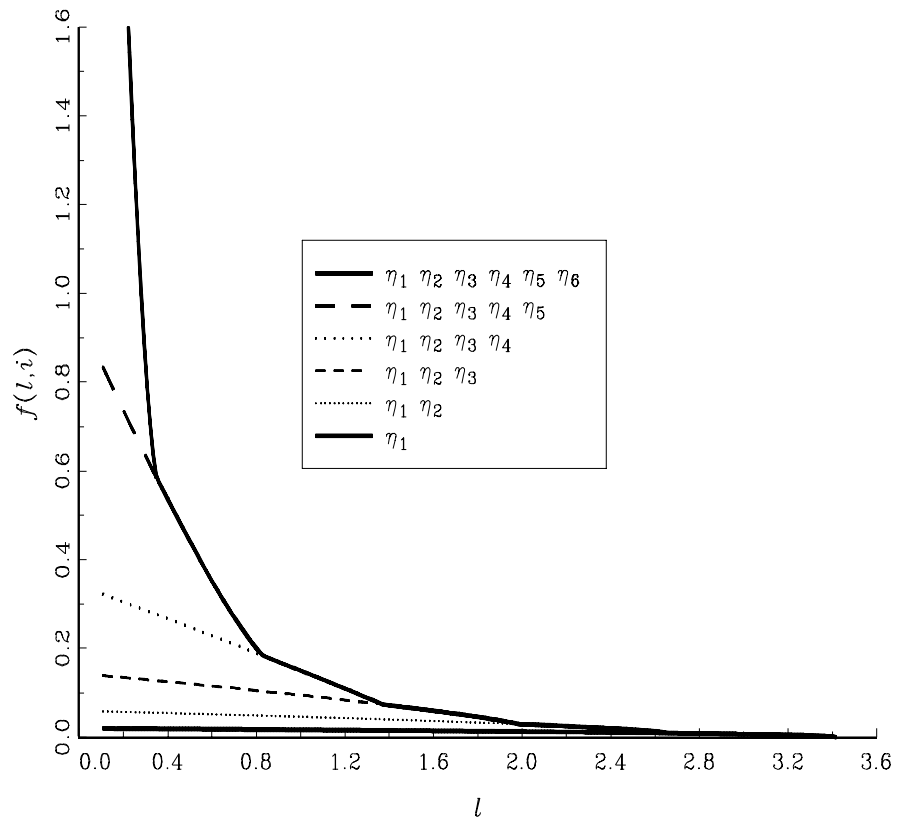


Figure 3: Probability density of employment at hiring firms: decreasing returns in production, no turnover.

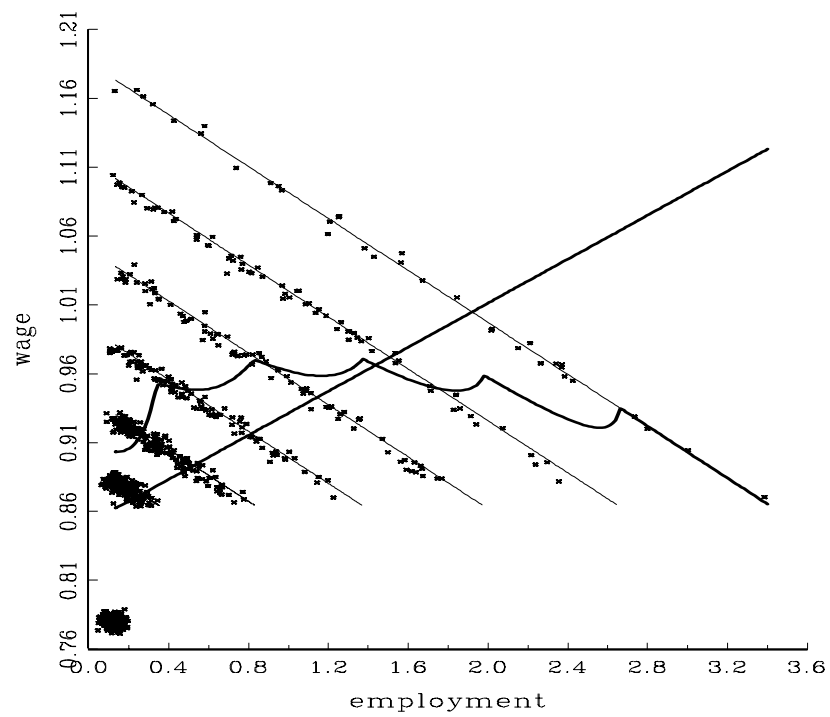


Figure 4: Wage functions: decreasing returns in production, no turnover. Exact conditional expectation and Monte Carlo sample of 1000 observations, with linear regression.

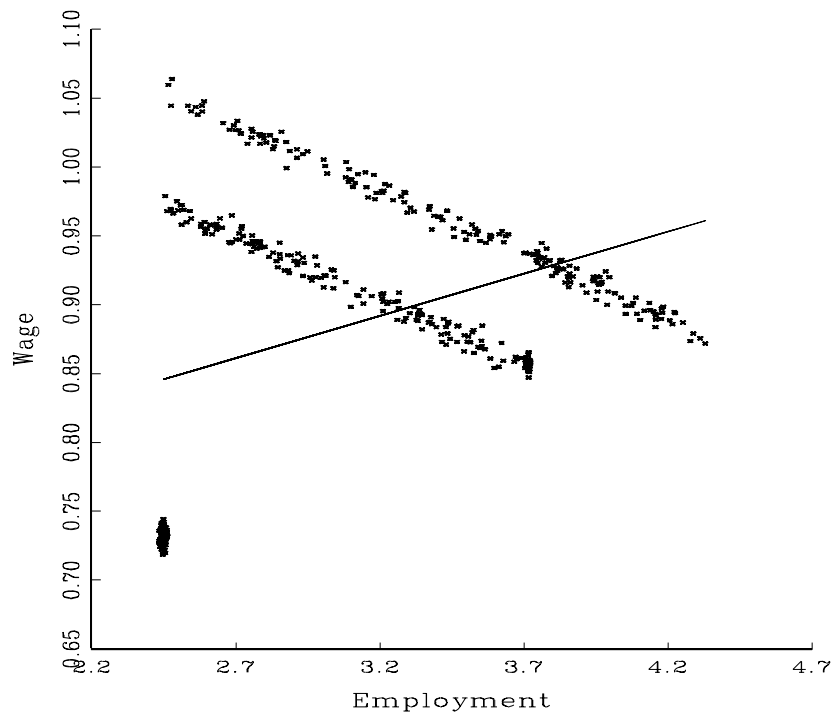


Figure 5: Wage functions: decreasing returns in production, no turnover; positive probability of transition across positive-vacancy states. Monte Carlo sample of 500 observations, with linear regression