

Stochastic Search Equilibrium*

Giuseppe Moscarini[†]
Yale University
and
NBER

Fabien Postel-Vinay[‡]
University of Bristol
and
Sciences Po, Paris

November 2011

Abstract

We study equilibrium wage and employment dynamics in a class of search models where firms post wage contracts, workers search randomly for such contracts both on and off the job, while the economy is subject to aggregate shocks. Our exercise provides the first dynamic stochastic general equilibrium analysis of a popular class of search models, drawing in part from the literature on recursive contracts under moral hazard. Firms offer and commit to (Markov) contracts, which specify a wage contingent on all payoff-relevant states, but must pay equally all of their workers, who have limited commitment and are free to quit at any time. We find sufficient conditions for the existence and uniqueness of a stochastic search equilibrium in such contracts, which is Rank Preserving [RP]: larger and more productive firms offer a larger value to their workers in all states of the world. On the RP equilibrium path turnover is always efficient as workers always move from less to more productive firms, and the stochastic dynamics of firm size provide an intuitive explanation for the empirical finding that large employers have more cyclical job creation (Moscarini and Postel-Vinay, 2011). Finally, computation of RP equilibrium contracts is tractable.

Keywords: Equilibrium Job Search, Dynamic Contracts, Stochastic Dynamics.

JEL codes: J64, J31, D86.

*Earlier versions of this paper circulated under the title “Non-stationary search equilibrium”. We acknowledge useful comments to earlier drafts of this paper from seminar and conference audiences at numerous venues. We also wish to thank Ken Burdett, Dale Mortensen and Robert Shimer for constructive discussions of earlier versions of this paper. The usual disclaimer applies.

[†]Address: Department of Economics, Yale University, PO Box 208268, New Haven CT 06520-8268. Tel. +1-203-432-3596. E-mail giuseppe.moscarini@yale.edu. Web <http://www.econ.yale.edu/faculty1/moscarini.htm>

[‡]Address: Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, UK. Tel: +44 117 928 8431. E-mail Fabien.Postel-Vinay@bristol.ac.uk. Web www.efm.bris.ac.uk/economics/staff/vinay/. Postel-Vinay is also affiliated with CEPR (London) and IZA (Bonn).

1 Introduction

The continuous reallocation of employment across firms, sectors and occupations, mediated by various kinds of frictions, is a powerful source of aggregate productivity growth.¹ Workers try to move in response to various reallocative shocks, and search on and off the job to take advantage of the large wage dispersion that they face, even conditional on their observable characteristics. A popular class of search wage-posting models, originating with Burdett and Mortensen (1998, henceforth BM), aims to understand these phenomena. The BM model provides a coherent formalization of the hypothesis that cross-sectional wage dispersion is largely a consequence of labor market frictions, but also allows for permanent heterogeneity in firms' productivities. The BM model started a fruitful line of research in the analysis of wage inequality and worker turnover, as the vibrant and empirically very successful literature organized around that hypothesis continues to show (see Mortensen, 2003 for an overview).

This job search literature, however, is invariably cast in deterministic steady state. Ever since the first formulation of the BM model, job search scholars have regarded the characterization of its out-of-steady-state behavior as a daunting problem, essentially because one of the model's state variables, which is also the main object of interest, is the endogenous *distribution* of wage (or job value) offers. This is an infinite-dimensional object, endogenously determined in equilibrium as the distribution across firms of offer strategies that are mutual best responses, and evolving stochastically with the aggregate impulse.

The restriction to steady state analysis is not costless. The ongoing reallocation of employment across firms has a cyclical pattern. Moscarini and Postel-Vinay (2011, henceforth MPV11) document that the net job creation of larger, higher-paying firms is more positively (negatively) correlated with detrended GDP (resp., unemployment rate) than at smaller firms. Essentially, the firm size/growth relationship "tilts" up and down with the business cycle, although it is still flat on average, consistently with Gibrat's Law.² Any theory of turnover and wage dispersion based on frictional worker reallocation among firms, and allowing for aggregate dynamics, speaks directly to these facts, thus must confront them.

We provide the first analysis of aggregate stochastic dynamics in wage-posting models with random job search. We study a frictional labor market where firms offer and commit to employment contracts, workers search randomly on and off the job for those contracts, while aggregate productivity is subject to persistent shocks. In our economy, both in the

¹See Foster, Haltiwanger and Krizan (2000) and Lentz and Mortensen (2008) for recent evidence.

²Haltiwanger, Jarmin and Miranda (2010) present the most comprehensive study to date, based on the full longitudinal census of US employers (Longitudinal Business Database, 1976-2005), the same data underlying MPV11's evidence. They find that a firm's growth is unrelated to its size, after controlling for its age. They do not address business cycle patterns.

constrained efficient allocation and in equilibrium, smaller firms contribute more to net job creation when unemployment is high, consistently with MPV11’s empirical observations.

Our key contribution overcomes the technical hurdle that stunted progress of the job search literature beyond steady state analysis. We find sufficient conditions for a unique, constrained-efficient equilibrium, where the distribution of wage contracts is actually very easy to track: the workers’ ranking of firms is the same in all aggregate states — what we call a *Rank-Preserving Equilibrium* (RPE). The sufficient conditions are that firms either are equally productive, or differ in the permanent component of their productivity and the more productive they are, the (weakly) more workers they initially employ — for example, they all start empty. In the latter case, in RPE more productive firms offer a larger value and employ more workers at all points in time: when given a chance, a worker always moves from a less into a more productive firm, so that the equilibrium allocation of employment is constrained efficient. This parallels a similar property of BM’s static equilibrium.

In our economy, infinitely lived and risk neutral firms and workers come in contact infrequently. Firms produce homogenous output with labor in a linear technology, which may permanently differ across firms. Aggregate multiplicative TFP shocks affect labor productivity as well as the job contact rates, on and off the job, the exogenous job destruction rate, and the value of leisure. A social planner constrained by search frictions, when given the opportunity, moves an employed worker from a less productive to a more productive firm. This efficient turnover gives rise to a simple Markov process for the evolution of the firm size distribution, which can be solved for analytically, given any history of aggregate shocks. The solution replicates the MPV11 facts: larger firms grow relatively faster when aggregate TFP is high. If we shut down aggregate shocks, this process converges deterministically to BM’s stationary size distribution.

To study equilibrium, we assume that firms offer and commit to a (Markov) contracts where the wage is allowed to depend on all payoff-relevant states. We impose only one further restriction, in order to obtain a well-defined notion of a firm. Following BM, we define a firm as a wage policy, thus impose an *equal-treatment constraint*: it must pay the same wage in a given period to all of its employees, whether incumbent, newly hired from unemployment or from employment. Workers cannot commit not to quit to other jobs when the opportunity arises, or to unemployment whenever they please, so commitment is one-sided and firms face a standard moral hazard problem. Contract offers are privately observed only by the recipients, thus deviations cannot be detected by other players.

We look for an equilibrium of this contract-posting game in Markov contracts. Those contracts offer values that are a function of the only four current payoff-relevant states: two exogenous, firm-specific and aggregate productivity, one endogenous to the firm, its current

size, and one endogenous to the economy but exogenous to the firm, the distribution of employment across all firms. A firm must track the latter, infinitely-dimensional object in order to know the distributions of competing offers and of values earned by currently employed workers, thus how much recruitment and retention its own contract will generate. We establish that at most one Markov contract-posting equilibrium exists, characterize it, and show that it decentralizes the constrained efficient allocation, thus is consistent with MPV11's evidence.

Key to our analysis is the following comparative dynamics property of the best response contract offer: at any node in the game and for any distribution of offers made by other firms and of values earned by employed workers, a more productive and/or larger firm optimally offers a contract that pays its existing and new workers a larger value. Therefore, if firms are homogeneous, or if more productive firms are initially no smaller, then no firm wants to break ranks in the distribution of competing offers, which then coincides with the given distribution of firm productivity or initial size. This immediately implies our main result that equilibrium, if it exists, is unique, and is also RP, thus constrained efficient.

The intuition behind this comparative dynamics property parallels a single-crossing property of the static BM model. There, a more productive firm gains more from employing a worker, hence wants to (and can) pay a higher wage. In addition, under the equal treatment constraint, the effect of the wage on retention is proportional to own size, while that on hiring is independent of size. Finally, size increases in the wage due to its effect on recruitment and retention. Thus, *ceteris paribus*, a larger firm also wants to pay more. This intuition does not extend immediately to our dynamic stochastic setting, because firm size is an evolving state variable with a given initial condition. However if more productive firms are initially weakly larger, this initial size ranking self-perpetuates on the equilibrium path, as required by RPE, so that more productive firms always pay and employ more. This intuitive and natural outcome is unique despite the strategic complementarity of a wage-posting game.

As a by-product of this analysis we offer a methodological contribution. We formulate the first (to the best of our knowledge) theory of Monotone Comparative Dynamics in a dynamic stochastic decision problem. We show how the optimal choice on the entire decision tree changes with a parameter of the model (firm productivity) which affects initial conditions and current payoffs, but could also affect the law of motion of the state variables.³ In our setting, firms solve a fully dynamic problem in a changing environment. In the sequential formulation of this problem, the choice set is an infinite sequence (a stochastic process), a case that the theory of Monotone Comparative Statics (Topkis, 1998) does not cover. The

³The literature tackled the different question whether the optimal policy is monotonic in a state variable (Stokey and Lucas, 1989), sometimes using our same tools (Gonzalez and Shi, 2010).

objective function of the one-step Bellman maximization contains the value function of the problem, whose properties are ex ante unknown. We establish that the Bellman operator of the contract-posting problem is a contraction on the space of functions that satisfy some single-crossing and convexity properties, which are then inherited by the value function. We can then apply Monotone Comparative Statics to the Bellman equation and a forward induction argument to prove the Monotone Comparative Dynamics properties of the best response illustrated above. The same logic can be applied in many other settings.⁴

One final, and crucial, benefit of RPE is its tractability and computability. The (unique) path of the equilibrium and efficient employment allocation is solved analytically. We obtain an algorithm to compute equilibrium contracts (prices), which also allows us to prove their uniqueness and to find a sufficient condition for their existence.

The rest of the paper is organized as follows. In Section 2 we place our contribution in the context of the relevant literature. In Section 3 we lay out the basic environment. In Section 4 we characterize the constrained efficient allocation. In Section 5 we describe and formally define an equilibrium, introduce the notion of Rank Preserving Equilibrium, characterize RPE contracts, and present uniqueness and existence results. In Section 6 we revisit our assumptions and discuss the robustness and interpretation of our results. Section 7 concludes and describes future research.

2 Related Literature

Besides its intrinsic theoretical interest, our characterization of the dynamics of the BM model opens the analysis of aggregate labor market dynamics as a whole potential new field of application of search/wage-posting models. Explaining the evidence in MPV11 is one such application. More generally, we hope to contribute to a synthesis between the BM contract-posting approach and the “other”, equally successful side of the search literature, organized around the matching framework (Pissarides, 1990; Mortensen and Pissarides, 1994), initially designed for the understanding of labor market flows and equilibrium unemployment.

The analysis of equilibrium wage and employment dynamics in equilibrium search models with wage dispersion has recently become a subject of keen investigation. First, as a stepping stone to the present paper, Moscarini and Postel-Vinay (2008, MPV08) and its discussion by Shimer (2008) study the deterministic transitional dynamics of the BM model. Rudanko

⁴For example, one may ask in a stochastic growth model whether the socially optimal level of investment is decreasing at all states and dates in the initial stock of capital, in the labor share, or in risk aversion; in each case, one must track the effect of the parameter change on the endogenous state variable (capital) along the entire optimal path. In this analogy, the existing literature investigates instead when is investment monotonic in the current level of capital.

(2011) and Menzio and Shi (2011) analyze wage contract-posting models with aggregate productivity shocks, where job search is directed. The latter assumption greatly simplifies the analysis by severing the link between the individual firm’s contract-posting problem and the distribution of contract offers. This is the main hurdle that we face, and that we resolve by exploiting the emergence of Rank-Preserving Equilibria, while maintaining BM’s assumption of random search common to the majority of the search literature. While we see both programs as fruitful directions of theoretical exploration, from a quantitative viewpoint the directed search approach is focused on the response of the job-finding rate to aggregate shocks, and does not generate a well-defined notion of employer size. Hence, it does not speak to MPV11’s facts, that we envision as central to our understanding of the propagation of aggregate shocks in labor markets. One exception is Kaas and Kircher (2011)’s extension of Menzio and Shi’s model, to allow for firm size. They obtain interesting and empirically accurate predictions on firm growth, pay and recruitment strategies, but do not allow for on-the-job search and do not address MPV11’s business cycle facts.

Finally, Coles and Mortensen (2011) build on our notion of RPE to characterize the dynamic equilibrium of Coles’ (2001) version of the BM model, which closely resembles ours, except for the assumption that firms *cannot* commit to wage contracts. They endogenize firms’ hiring behavior using a standard matching-function approach, only specifying a firm’s recruitment cost *per current employee* as a function of new hires per current employee.⁵ This assumption makes a firm’s wage policy size-independent and guarantees a unique equilibrium within the RP class. In our setup, the RPE is unique among all (not just RP) equilibria. When firms are identical, our result requires no further assumptions. When firms differ by productivity, we can only prove it under an additional sufficient condition that, although far from implausible, may fail. Coles and Mortensen make progress in a richer setup by replacing our sufficient condition with their ingenious, albeit still knife-edge, assumption.⁶ Because they only study deterministic transitional dynamics, their model cannot be confronted yet with MPV11’s business cycle facts.

⁵Formally, if a firm employs n workers and looks at hiring H new workers, it will incur a total cost of $nc(H/n)$ where $c(\cdot)$ is an increasing and convex function. They also consider an additional linear cost of vacancy posting, but without the convexity in $c(\cdot)$ the most productive firm would post infinite vacancies.

⁶In Section 6, we argue that the RP property of equilibrium, thus constrained efficiency, is likely to hold also with endogenous hiring effort, but we also show that it can fail with either entry by very productive firms or idiosyncratic shocks to firm-level TFP. In contrast, Coles and Mortensen (2011) accommodate firm entry “at the top”, as well as specific idiosyncratic productivity processes.

3 The economy

We study a stochastic economy where firms commit to employment contracts and workers search randomly for those contracts. The special case of a stationary and deterministic economy where contracts are restricted to a constant wage is the BM wage posting model. We work in discrete time as it affords more clarity in the presentation of the contract posting problem under one-sided commitment as a recursive problem.

The labor market is populated by a unit-mass of workers, who can be either employed or unemployed, and by a unit measure of firms.⁷ Workers and firms are risk neutral, infinitely lived, and maximize payoffs discounted with factor $\beta \in (0, 1)$. Firms operate constant-return technologies with labor as the only input and with productivity scale $\omega\theta$, where ω is an aggregate component, evolving within some bounded set of values $\Omega \subset \mathbb{R}_+$ according to a discrete-time stationary first-order Markov process $H(d\omega' | \omega)$, and θ is a fixed, firm-specific component, distributed across firms θ according to a cdf Γ over $[\underline{\theta}, \bar{\theta}] \subset \mathbb{R}_+$.

The labor market is affected by search frictions in that unemployed workers can only sample job offers sequentially with some probability $\lambda_0^\omega \in (0, 1)$ each period. Employed workers earn a wage, are allowed to search on the job, and face a per-period sampling chance of job offers of $\lambda_1^\omega \in (0, 1)$. For notational simplicity we will assume uniform sampling of firms by workers, in that any worker receiving a job offer draws the type of the firm from which the offer emanates from the distribution $\Gamma(\cdot)$.⁸ All firms of equal productivity θ start out with the same labor force. We denote by $\Lambda_0(\theta)$ the measure of employment initially at firms of productivity at most θ . Each employed worker is separated from his employer and enters unemployment every period with probability $\delta^\omega \in (0, 1)$. Note that all these transition probabilities, although exogenous, are allowed to depend on the aggregate state ω .

In each period, the timing is as follows. Given a current state ω of aggregate labor productivity and size (measure of workers employed) L :

1. production and payments take place at all firms in current state ω ; the flow benefit b^ω accrues to unemployed workers;
2. the new state ω' of aggregate labor productivity is realized;
3. employed workers can quit to unemployment;

⁷We implicitly fix the measure of active firms, thus remaining mostly silent on the question of entry and exit. A simple extension of the model to make it capture entry and exit of firms over the business cycle is illustrated in MPV08. Finally, that the mass of firms and workers both have measure one is obviously innocuous and only there to simplify the notation.

⁸When calibrating and simulating the model in Moscarini and Postel-Vinay (2010a) we allow for non-uniform sampling, in that different types- θ firms have different chances of being sampled by jobseekers. This extension is theoretically straightforward and useful in quantitative applications.

4. jobs are destroyed exogenously with chance $\delta^{\omega'}$;
5. the remaining employed workers receive an outside offer with chance $\lambda_1^{\omega'}$ and decide whether to accept it or to stay with the current employer;
6. each previously unemployed worker receives an offer with probability $\lambda_0^{\omega'}$.

Finally, in order to avert unnecessary complication, we will assume throughout the paper that the distribution of firm types, Γ , has continuous and everywhere strictly positive density over $[\underline{\theta}, \bar{\theta}]$, and that the initial measure of employment across firm types, Λ_0 , is continuously differentiable in θ . Combining those two assumptions, we obtain that the initial average size of a type- θ firm, which is given by $L_0(\theta) = \frac{d\Lambda_0(\theta)/d\theta}{\gamma(\theta)}$, is a continuous function of θ .

4 The constrained efficient allocation

A social planner constrained by the same search frictions as private agents only has to decide which transition opportunities to take up and which ones to ignore. Recall that opportunities to move from unemployment to employment or from job-to-job only arise infrequently due to search frictions, while the option to move workers into unemployment is always available. Here we only consider the simple case where the planner never finds it optimal to exercise the latter option, because the value of leisure b^ω is sufficiently lower than the productivity of any existing firm in all states and/or because employed search is sufficiently effective relative to unemployed search. We return later to the important issue of employer entry and exit.

The constrained efficient allocation is then simple enough to characterize: the planner will take up any opportunity to move an unemployed worker into employment, and the unemployment rate $u := 1 - \Lambda(\bar{\theta})$ evolves according to $u' = \delta^{\omega'}(1 - u) + (1 - \lambda_0^{\omega'})u$, where primes are used to denote next-period values. Moreover, the planner always seeks to move employed workers from less productive toward more productive firms. This induces the following simple Markov process for the evolution of the measure $L^*(\theta)$ of workers efficiently allocated to a typical type- θ firm:

$$\begin{aligned} L^*(\theta)' &= L^*(\theta) \left(1 - \delta^{\omega'}\right) \left(1 - \lambda_1^{\omega'} \bar{\Gamma}(\theta)\right) + \lambda_0^{\omega'} u + \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \Lambda^*(\theta) \\ &= L^*(\theta) \left(1 - \delta^{\omega'}\right) \left(1 - \lambda_1^{\omega'} \bar{\Gamma}(\theta)\right) + \lambda_0^{\omega'} u + \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \int_{\underline{\theta}}^{\theta} L^*(x) d\Gamma(x). \end{aligned} \quad (1)$$

Given new aggregate state ω' , of the $L^*(\theta)$ workers initially employed by this firm, a fraction $(1 - \delta^{\omega'})$ are not separated exogenously into unemployment. Of these survivors, a fraction $\lambda_1^{\omega'}$ receive an opportunity to move to another firm. The planner exercises that option if and only if the new firm is more productive than θ , which is the case with probability

$\bar{\Gamma}(\theta) := 1 - \Gamma(\theta)$. The initially unemployed u find jobs with chance $\lambda_0^{\omega'}$. Workers employed at other firms who have not lost their jobs draw with chance $\lambda_1^{\omega'}$ an opportunity to move to the type- θ firm, that the planner exploits if and only if the firm they currently work at has productivity $x < \theta$. The measure of such workers in the optimal plan is $\Lambda^*(\theta)$.

Equation (1) combines an ordinary differential equation and a first-order difference equation in Λ^* , a function of time and θ . Multiplying through by $\gamma(\theta)$ in (1) and integrating with respect to θ yields:

$$\Lambda^*(\theta)' = \lambda_0^{\omega'} u \Gamma(\theta) + (1 - \delta^{\omega'}) (1 - \lambda_1^{\omega'} \bar{\Gamma}(\theta)) \Lambda^*(\theta).$$

To solve this equation forward in time we introduce a time index t . For any initial condition $\Lambda_0^*(\theta) = \Lambda_0(\theta)$ at some (renormalized) initial date 0 such that the aggregate state last switched to ω at time 0 and then remained at ω between 0 and t , the latter law of motion is a first-order difference equation which solves as:

$$\Lambda_t^*(\theta) = [(1 - \delta^\omega) (1 - \lambda_1^\omega \bar{\Gamma}(\theta))]^t \Lambda_0(\theta) + \lambda_0 \Gamma(\theta) \sum_{s=1}^t [(1 - \delta^\omega) (1 - \lambda_1^\omega \bar{\Gamma}(\theta))]^{s-1} u_{t-s}. \quad (2)$$

By inspection, $\Lambda_t^*(\theta)$ is differentiable in θ at all dates t , and one obtains a closed-form expression for the workforce of any type- θ firm:

$$L_t^*(\theta) = \frac{d\Lambda_t^*(\theta)/d\theta}{\gamma(\theta)} = (1 - \delta^\omega)^t (1 - \lambda_1^\omega \bar{\Gamma}(\theta))^{t-1} [(1 - \lambda_1^\omega \bar{\Gamma}(\theta)) L_0^*(\theta) + t \lambda_1^\omega \Lambda_0(\theta)] + \lambda_0 \left\{ u_{t-1} + \sum_{s=2}^t (1 - \delta^\omega)^{s-1} (1 - \lambda_1^\omega \bar{\Gamma}(\theta))^{s-2} [1 - \lambda_1^\omega + \lambda_1^\omega s \Gamma(\theta)] u_{t-s} \right\}, \quad (3)$$

where $L_0^*(\theta)$ was the value of this solution under state $\hat{\omega}$ at the time of the last state switch from $\hat{\omega}$ to ω . If the aggregate state forever stays at ω , the solutions to (2) and (3) converge to:

$$\Lambda_\infty^*(\theta) = \frac{\delta^\omega \lambda_0^\omega}{\delta^\omega + \lambda_0^\omega} \cdot \frac{\Gamma(\theta)}{1 - (1 - \delta^\omega) (1 - \lambda_1^\omega \bar{\Gamma}(\theta))}$$

and

$$L_\infty^*(\theta) = \frac{\delta^\omega \lambda_0^\omega}{\delta^\omega + \lambda_0^\omega} \cdot \frac{1 - (1 - \delta^\omega) (1 - \lambda_1^\omega)}{[1 - (1 - \delta^\omega) (1 - \lambda_1^\omega \bar{\Gamma}(\theta))]^2} \quad (4)$$

which are the familiar steady-state expressions found in the BM model.

As is well known and immediately verifiable from (4), the distribution of employment across firm types is increasing in λ_1^ω and decreasing in δ^ω in the sense of stochastic dominance. Intuitively, workers upgrade to higher- θ firms in larger numbers if they receive more opportunities to do so (higher λ_1^ω) or if they get thrown off the job ladder into unemployment less

often (lower δ^ω). This comparative statics property is reflected in the dynamic behavior of the firm size distribution if we assume, as is consistent with empirical evidence on job-to-job quits and job separations, that λ_1^ω is increasing, and δ^ω decreasing in the state of aggregate productivity ω , and also that more productive firms initially employ more workers, as is suggested by empirical evidence on the size-productivity relationship (and as is necessarily the case in the model's steady state). Then, hitting the economy with a randomly drawn sequence of aggregate shocks, in MPV10a we find that large employers are more cyclically sensitive, because they gain workers faster over an aggregate expansion as job upgrading accelerates, and vice versa in a slump. This property of the efficient allocation replicates the new empirical evidence that we document in MPV11.

5 Equilibrium

5.1 Definition

Each firm chooses and commits to an employment contract, namely a state-contingent wage depending on some state variable ζ , to maximize the present discounted value of profits, given other firms' contract offers. The firm is further subjected to an *equal treatment constraint*, whereby it must pay the same wage to all its workers. This is the sense in which we generalize the BM restrictions placed on the set of feasible wage contracts to a non-steady-state environment.⁹ Under commitment, such a wage function implies a value V for any worker to work for that firm, also a function of the state ζ . For reasons that will become clear shortly, we assume that a contract offered by a firm to its workers is observable only by the parties involved.

Let Z be the (Borel-)measurable set of all histories of play in the game, and \mathcal{V}_Z the set of measurable functions $[\underline{\theta}, \bar{\theta}] \times Z \rightarrow \mathbb{R}$. A behavioral strategy of the contract-posting game is a function $V \in \mathcal{V}_Z$ such that, when the state of the game is $\zeta \in Z$, each firm $\theta \in [\underline{\theta}, \bar{\theta}]$ offers value $V(\theta, \zeta)$ to all of its workers.

As V is measurable, the c.d.f.

$$F(W | \zeta, V) := \int_{\underline{\theta}}^{\bar{\theta}} \mathbb{I}\{V(\theta, \zeta) \leq W\} d\Gamma(\theta) \quad (5)$$

⁹We thus rule out, beyond contracts that condition wages on tenure (Burdett and Coles, 2003) and employment status (Carrillo-Tudela, 2009), also offer-matching and individual bargaining (Postel-Vinay and Robin, 2002; Dey and Flinn, 2005; Cahuc, Postel-Vinay and Robin, 2006). Note, however, that the model can be generalized to allow for time-varying individual heterogeneity under the assumption that firms offer the type of piece-rate contracts described in Barlevy (2008). In that sense experience and/or tenure effects can be introduced into the model.

is well-defined for every $\zeta \in Z$, $W \in \mathbb{R}$ and \mathbb{I} an indicator function. This is the probability that a randomly drawn firm offers value no greater than W , given history ζ and given that all firms follow strategy V . Let $\bar{F} = 1 - F$ denote the survival function.

Again, let $\Lambda(\theta)$ be the measure of workers currently employed at all firms of productivity up to θ , so $\Lambda(\bar{\theta})$ is total employment. For any increasing $\Lambda : [\underline{\theta}, \bar{\theta}] \rightarrow [0, 1]$, $\zeta \in Z$, $W \in \mathbb{R}$, the following c.d.f.

$$G(W | \zeta, \Lambda, V) := \frac{1}{\Lambda(\bar{\theta})} \cdot \int_{\underline{\theta}}^{\bar{\theta}} \mathbb{I}\{V(\theta, \zeta) \leq W\} d\Lambda(\theta) \quad (6)$$

is also well-defined. This is the probability that a randomly drawn worker is currently earning value no greater than W after history ζ .

Given a strategy $V \in \mathcal{V}_Z$ followed by all firms and the resulting F , an unemployed worker earns a value solving:

$$U(\zeta | V) = b^\omega + \beta \mathbf{E}_{\zeta' | \zeta} \left[\left(1 - \lambda_0^{\omega'} \right) U(\zeta' | V) + \lambda_0^{\omega'} \int \max\langle v, U(\zeta' | V) \rangle dF(v | \zeta', V) \right], \quad (7)$$

because she collects a flow value b^ω and, one period later, when the aggregate state becomes ω' , she draws with chance $\lambda_0^{\omega'}$ a job offer from the equilibrium distribution of offered values F , which she accepts if the associated value exceeds that of staying unemployed.

A firm of current size L which posts a value W in state ζ has size zero next period if $W < U$, otherwise, invoking a large numbers approximation, new firm size is:

$$L' = \mathcal{L}(\zeta, W | V) := L \left(1 - \delta^{\omega'} \right) \left(1 - \lambda_1^{\omega'} \bar{F}(W | \zeta, V) \right) + \lambda_0^{\omega'} [1 - \Lambda(\bar{\theta})] \mathbb{I}\{W \geq U(\zeta | V)\} + \lambda_1^{\omega'} \left(1 - \delta^{\omega'} \right) \Lambda(\bar{\theta}) G(W | \zeta, V). \quad (8)$$

After the new aggregate state ω' is drawn, of the measure L of workers currently employed by this firm, a fraction $(1 - \delta^{\omega'})$ are not separated exogenously into unemployment. Of these survivors, a fraction $\lambda_1^{\omega'} \bar{F}(W | \zeta, V)$ quit because they draw from F an outside offer which gives them a value larger than W . The currently unemployed $1 - \Lambda(\bar{\theta})$ find jobs with chance $\lambda_0^{\omega'}$, and accept an offer W from a firm if this is better than unemployment. By random matching, each firm offering more than U receives the same inflow from unemployment. The employed who have not lost their jobs $(1 - \delta^{\omega'}) \Lambda(\bar{\theta})$ receive an offer with chance $\lambda_1^{\omega'}$, and accept it if the value W they draw is larger than what they were earning before (probability $G(W | \zeta, V)$), in which case they quit to this firm offering W .

Adding up, the cumulated firm size evolves as the sum of individual firm sizes on the equilibrium path. For any $\theta \in [\underline{\theta}, \bar{\theta}]$:

$$\Lambda(\theta | \zeta, V)' = \int_{\underline{\theta}}^{\theta} \mathcal{L}(\zeta, V(x, \zeta) | V) d\Gamma(x) \Rightarrow \Lambda(\cdot | \zeta, V)' := \mathcal{F}(\zeta | V). \quad (9)$$

The map \mathcal{F} denotes next period's employment distribution given current state ζ and offer strategy V . The support of Λ is contained in that of Γ , because no worker can be at a firm of type θ if there exists no such firm. By induction, starting from the initial distribution of employment and for every history of the game, Λ has a (possibly nil) Radon-Nikodym derivative $d\Lambda(\theta | \zeta, V) / d\Gamma(\theta)$ everywhere in θ , and (9) requires this derivative to be $\mathcal{L}(\zeta, V(\theta, \zeta) | V)$. Therefore, F and G also exist at all nodes of the game when firms play the strategy V .

A value strategy $W \in \mathcal{V}_Z$ can also be implemented by a wage strategy $w \in \mathcal{V}_Z$ such that the worker's Bellman equation is solved by W given that all other firms play V : the worker receives the wage and, next period, the expected value of being either displaced, or retained at the same firm, or poached by a higher-paying firm.

$$W(\theta, \zeta) = w(\theta, \zeta) + \beta \mathbf{E}_{\zeta' | \zeta} \left[\delta^{\omega'} U(\zeta' | V) + (1 - \delta^{\omega'}) \left(W(\theta, \zeta') + \lambda_1^{\omega'} \int_{W(\theta, \zeta')}^{+\infty} [v - W(\theta, \zeta')] dF(v | \zeta', V) \right) \right] \quad (10)$$

We are now going to define an equilibrium of the contract-posting game. Each firm plays a game against other firms as well as vis-à-vis its current and prospective workers. Workers act sequentially, as they are always free to quit. Firms choose once and for all at time 0 a strategy V (a state-contingent value policy), and commit to it. So firms can only deviate at time 0, when they choose the strategy. Our assumption that a contract offer is only observed by the firm and by the workers who receive it implies that any deviation by a firm to a different pre-committed contract will be observed at most by a measure zero of workers (those that the firm will hire over the countable infinite horizon) and a measure zero of firms (those that will hire workers who worked in the past at, and observed, the deviating firm). So the firm anticipates that any deviation will trigger no relevant change in other firms' behavior, and, when choosing its strategy, takes the distributions of offers F and earned contracts G as given at any future point in time and state.

We assume that firms commit for the entire future. The reputational underpinnings of the firm's commitment power have been explored in the wage-posting literature (Coles, 2001; Coles and Mortensen, 2011). The constraint of delivering the promised value to the workers once hired is binding because, after hiring a worker with a promise of W , the firm would like to renege and to squeeze the worker against the participation constraint $W = U$.

Our first task is to find the state space Z on which equilibrium strategies can be conditioned. By assumption, past play by other firms is unobservable, hence cannot be part of Z . Each firm can only observe the set Q of public histories of $\{\omega, F, G, \Lambda\}$ as well, of course, as calendar time t .

We look for the smallest subset $Z \subseteq Q$ which is sufficient for Q . For every $\zeta \in Z$ and $V(\cdot, \zeta)$, the current offer distribution $F(\cdot | \zeta, V)$ is uniquely determined from (5), so it contains no independent information about ζ . The same is true, from (6), of $G(\cdot | \zeta, V)$, given ζ, V and Λ . Next, each individual firm takes the strategy V chosen by others as given, whether or not this firm is maximizing, given ζ . Therefore, for every V , a firm can calculate the history of Λ based only on the history of ω . That is, each firm takes the path of employment at other firms Λ as an exogenous stochastic process. Hence, for every value-offering strategy defined on the history of ω and Λ , there exists an equivalent value-offering strategy defined on the history of ω only, given Λ_0 , which produces the same payoff relevant variables for firm θ . For the purpose of calculating firm θ 's best response, the history of ω is sufficient for the history of Λ .

The only other independent piece of information that is relevant to a firm's profit maximization is own size L , that is directly controlled by the firm and has a direct impact on the firm's continuation payoffs. Because the history of own size $\{L_s\}_{s=0}^t$ is private information, it cannot affect values offered by other firms. Hence only current size L_t can affect the firm's best response, because of its direct impact on profits. We conclude that the only strategically relevant history for a firm can be $\zeta_t = \{\omega_0, \dots, \omega_t, t, L_t\}$. Clearly, past values of ω cannot be ruled out of the state space Z , as they are exogenous and public events that firms can use to coordinate actions.

Definition 1 A CONTRACT-POSTING EQUILIBRIUM is a measurable function $V \in \mathcal{V}_Z$ of the set Z of histories of the aggregate productivity state and current size, such that V maximizes the present discounted value of profits, given that all other firms also play V . Formally, at least one solution $w \in \mathcal{V}_Z$ to (10) with $W = V$ also solves:

$$w(\theta, \zeta) = \arg \max_{\tilde{w}} \mathbf{E} \left[\sum_{t=0}^{+\infty} \beta^t (\omega_t \theta - \tilde{w}(\theta, \zeta_t)) \mathcal{L}(\zeta_{t-1}, W | V) | \zeta_0 = \zeta \right],$$

where $\Lambda_t(\theta) = \int_{\underline{\theta}}^{\theta} \mathcal{L}(\zeta_{t-1}, V(x, \zeta_{t-1}) | V) d\Gamma(x)$, and F, G, U, W are defined by (5), (6), (7), (10) with $\zeta = \zeta_t$.

The equilibrium strategy V is a fixed point, a Nash equilibrium at time 0: if all firms follow V and workers act optimally, then given the implied evolution of the cross-section distributions of values offered F and earned G and of the value of unemployment U , each firm θ 's best response is to follow the same strategy $W = V$.

To further reduce the state space to a tractable dimensionality, we restrict attention to strategies that depend only on *current values of payoff-relevant variables*:

$$\hat{\zeta} = \{\theta, L, \omega', \Lambda\}. \tag{11}$$

This restriction is in the spirit of Maskin and Tirole (2001)'s Markov Perfect Equilibrium, although that is defined for games of observable actions without commitment. Importantly, this restriction excludes calendar time, because not payoff-relevant, from the state space. In this sense, our Markov strategies are “stationary”. Let \hat{Z} the space where $\hat{\zeta}$ lives and $\mathcal{V}_{\hat{Z}}$ be the space of measurable functions $\hat{Z} \rightarrow \mathbb{R}$:

Definition 2 *A MARKOV CONTRACT-POSTING EQUILIBRIUM is a contract-posting equilibrium V in the set $\mathcal{V}_{\hat{Z}} \subset \mathcal{V}_Z$, a measurable function of $\hat{\zeta}$ defined in (11).*

An issue arises with the initially promised value, before the firm chooses and commits to the contract at time 0. If the firm has some employees at time 0 and all the bargaining power, it optimally offers $U(\omega_0, \Lambda_0)$ and extracts all rents from its initial workforce, by making them just indifferent between staying or quitting into unemployment. It cannot be always optimal, however, to pre-commit to offer again the same value of unemployment in the future whenever the state returns to $\hat{\zeta}_0$. So the best response by any firm in state $\hat{\zeta}_0$ is different at time 0 and at a later date, thus is time-dependent, or non Markovian. This is true even if all other firms offer Markov contracts. Hence, strictly speaking, a Markov contract-posting equilibrium does not exist if the firm has initial employees. (If all firms start empty, and all workers are initially unemployed, then this issue is moot, because there is no one to promise anything to.) As we will see shortly, this turns out to be a technical issue of little economic substance, with multiple possible and equivalent resolutions.

Making strategies independent of past values of aggregate productivity comes at the cost of introducing in the state the current distribution of employment Λ . This is also an infinitely dimensional object, but it turns out to be much more tractable than the entire history of ω , as we will see next.

From now on, we let the new distributions $F(\cdot | \omega', \Lambda)$ and $G(\cdot | \omega', \Lambda)$, firm size $\mathcal{L}(L, \omega', \Lambda, W)$, employment distribution $\mathcal{T}(\omega', \Lambda)$, and value of unemployment $U(\omega', \Lambda)$ be defined as in (5) - (9), with the state $\hat{\zeta}$ in (11) replacing ζ . Notice that only new firm size \mathcal{L} depends on L ; the other objects only depend on the aggregate components of the state, ω' and Λ , that each firm takes as given stochastic processes on and off the equilibrium path. That is, $\hat{\zeta}$ contains only one endogenous (to the firm) state variable, its own size L .

5.2 The firm's contract-posting problem

Suppose all competing firms offer a value $V(\theta, L, \omega', \Lambda)$ which depends on own productivity θ , beginning-of-period own size L and distribution of employment Λ , and new state of aggregate productivity ω' . Because these four objects evolve according to a Markov process, it

is natural to seek a recursive formulation of the firm's problem. As is standard in the contracting literature (Spear and Srivastava, 1987), the firm's sequential contracting problem is equivalent to a recursive problem, in which the firm takes the value currently promised to its workers as a state variable, and faces a promise-keeping constraint.

We fix the Markov strategy of other firms V and omit it from the notation for simplicity. The firm can always guarantee itself zero flow profits by making the participation constraint $W(\omega') \geq U$ bind and dismissing all workers, so offering any value lower than U is equivalent to an offer $W(\omega') = U$. Using the law of motions of own employment and of the aggregate employment distribution, the firm solves:

$$\begin{aligned} \Pi(\theta, L, \omega, \Lambda, \bar{V}) = & \sup_{w, W(\omega') \geq U(\omega', \mathcal{T}(\omega', \Lambda))} \left\langle (\omega\theta - w)L \right. \\ & \left. + \beta \int_{\Omega} \Pi[\theta, \mathcal{L}(L, \omega', \Lambda, W(\omega')), \omega', \mathcal{T}(\omega', \Lambda), W(\omega')] H(d\omega' | \omega) \right\rangle \quad (12) \end{aligned}$$

subject to a Promise-Keeping (PK) constraint to deliver the promised \bar{V} :

$$\begin{aligned} \bar{V} = & w + \beta \cdot \int_{\Omega} \left\{ \delta^{\omega'} U(\omega', \mathcal{T}(\omega', \Lambda)) + (1 - \delta^{\omega'}) \right. \\ & \left. \cdot \left[\left(1 - \lambda_1^{\omega'} \bar{F}(W(\omega') | \omega', \Lambda)\right) W(\omega') + \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF(v | \omega', \Lambda) \right] \right\} H(d\omega' | \omega), \quad (13) \end{aligned}$$

where the continuation value on the RHS comes from (10) after a small algebraic manipulation. In (12), given the timing of events, the firm collects flow revenues, equal to per worker productivity $\omega\theta$ times firm size L , then pays the flow wage w to each worker, then observes the new state of aggregate productivity ω' , and finally chooses the continuation contract (promised value) $W(\omega')$, so that wage and continuation values deliver at least the current expected value \bar{V} to the workers.

Note that we wrote the future promised value $W(\omega')$ as a function of the future aggregate state only, although strictly speaking this Markov strategy should be a function of the entire state $\hat{\zeta} = \{\theta, L, \omega', \Lambda\}$. The reason is that, in this recursive formulation, θ , L , and Λ are known when choosing the future value, and only ω' is a random variable, so our notation highlights the state-dependence of the promise on the only state that is not known at the moment it is chosen.

To characterize the best response contract, we first describe an equivalent unconstrained recursive formulation of the contract-posting problem. We define the joint value of the firm-worker collective as:

$$S = \Pi + \bar{V}L.$$

Next solving for the wage from (13) and replacing it into the firm's Bellman equation (12) we see that the joint value function S solves:

$$\begin{aligned}
S(\theta, L, \omega, \Lambda) &= \omega\theta L + \beta \int_{\Omega} \left\{ \delta^{\omega'} U(\omega', \mathcal{F}(\omega', \Lambda)) L \right. \\
+ \sup_{W(\omega') \geq U(\omega', \mathcal{F}(\omega', \Lambda))} &\left\langle S(\theta, \mathcal{L}(L, \omega', \Lambda, W(\omega')), \omega', \mathcal{F}(\omega', \Lambda)) + L(1 - \delta^{\omega'}) \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF(v | \omega', \Lambda) \right. \\
&\left. \left. - W(\omega') \left(\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda(\bar{\theta}) G(W(\omega') | \omega', \Lambda) \right) \right\rangle \right\rangle H(d\omega' | \omega). \quad (14)
\end{aligned}$$

Crucially, the promised value \bar{V} does not appear as an argument of S in the above equation, so the DP problem in (14) is independent of \bar{V} . Along the optimal path, the level of current promised utility \bar{V} only affects the distribution of payoffs between the firm and its workers, not their overall level S , nor the choice of tomorrow's promised values $W(\omega')$. The intuition is clear. The workers' turnover decisions only depend on continuation values $W(\omega')$ promised by the firm, so the same applies to firm continuation profits $\Pi(\omega')$. The firm thus chooses $W(\omega')$ to maximize $\Pi(\omega')$ independently of the currently promised value \bar{V} . Then, to deliver \bar{V} as promised without distorting the optimally set future turnover, the firm adjusts the current wage w .

We now return to the time 0 issue. Assume all other firms offer a Markov contract. Then, as just shown, at each time $t = 1, 2, \dots$ the best-response value that the firm promises in each state next period does not depend on the current promised value, inherited from the previous period $t - 1$. By induction, the best-response value that the firm promises at each future date and state does not depend on the value earned by its initial employees. Therefore, we can calculate a Markov best-response from time $t = 1$ by solving (14) and ignoring the time 0 problem. Then we can fix arbitrarily the value paid to initial employees, subject to participation constraints of all parties. For example, we can slightly modify the definition of Markov equilibrium and relax its measurability requirement just at time 0. Or, we can assume that initial workers bargain with the firm over the initial value, and each of them obtains exactly the same value $V(\hat{\zeta}_0)$ that the firm promises to deliver if the state returns later to the initial one $\hat{\zeta}_0$, so that the whole best response to a Markov contract is itself Markov. The continuation contract promised in each state from time 1 on, thus the equilibrium predictions of the model, are unaffected by this initial choice, so we do not delve into it.

The optimal policy solving the unconstrained DP problem (14) also solves (12) subject to (13). We therefore focus on the analysis of the simpler problem (14). An equilibrium is a solution V that coincides with the one followed by the other firms. To find the equilibrium,

we proceed as follows. First, we show that the equilibrium distribution of values offered (F) and earned (G), whatever they are, must be atomless on a connected support. Next, under certain sufficient conditions a best response value to any strategy followed by all other firms must be strictly increasing in own productivity θ and size L . Finally, in that smaller set of monotonic functions we construct an equilibrium.

5.3 Properties of the equilibrium distributions of contracts

The distributions of offered and accepted worker values, F and G , must satisfy certain general properties in equilibrium, which parallel similar properties of the corresponding wage distributions in the original BM model.

Proposition 1 (F and G are atomless) *In equilibrium F and G must be atomless at all dates and in all states, with their common support being compact and convex.*

To see why there cannot be an atom in F or G , observe that, by the equal treatment constraint, if F had an atom at some value W , then so would G . But an atom in G would open the way to a profitable deviation, as in BM. A firm that is part of the atom that offers the same W in some state could deviate, offer an epsilon more, win the competition for employed workers against all other competitors offering W , and poach an additional positive measure of workers at a negligible marginal cost. This deviation is unprofitable only if the firm was already offering its workers so much as to break even in expected present discounted terms. But then a deviation toward offering, e.g., $W = U$ in all states is profitable as all unemployed workers accept this offer and stay for a while, generating strictly positive profits for all but the zero measure of firms with marginal type that break even with $W = U$.

To see why the support of offered and paid values is convex, observe that if there was a gap then the lower and upper bounds of this gap would generate the same hiring and retention, so the same firm size, but the upper bound would cost the firm more in terms of wages, so no firm would post such an upper bound. To see why the support is compact, observe that $\bar{W} = \max \omega \theta / (1 - \beta)$ is a natural upper bound to the offered value: the firm can always do weakly better by offering less than \bar{W} , as it can hope to make some profits. So the support is a convex and bounded subset of \mathbb{R}_+ , which we can therefore take to be compact WLOG.

5.4 Rank-Preserving Equilibrium (RPE)

While solving for equilibrium directly is an intractable problem because the size distribution of firms Λ is an infinitely-dimensional state variable, we can still define a tractable and

natural class of equilibria, which have the following property. Let $L(\theta)$ denote employment size of a type- θ firm along the equilibrium path, i.e. the size attained by that firm given the initial size distribution at date 0 and given that all firms have played the equilibrium strategy from date 0 up to the current date. Then:

Definition 3 *An equilibrium is RANK-PRESERVING (RP) if a more productive firm always pays its workers more: $\theta \mapsto V(\theta, L(\theta), \omega, \Lambda)$ is increasing in θ .*

A direct consequence of the above definition is that in a Rank Preserving Equilibrium (RPE) workers rank their preferences to work for different firms according to firm productivity at all dates. The following two properties thus hold true at all dates in any RPE: the proportion of firms that offer less than $V(\theta, L(\theta), \omega, \Lambda)$ is simply that proportion of firms that are less productive than θ

$$F(V(\theta, L(\theta), \omega', \Lambda) \mid \omega', \Lambda) \equiv \Gamma(\theta), \quad (15)$$

and the number of employed workers who earn a value that is lower than that offered by θ equals employment at firms less productive than θ :

$$\Lambda(\bar{\theta}) G(V(\theta, L(\theta), \omega', \Lambda) \mid \omega', \Lambda) = \Lambda(\theta). \quad (16)$$

As we will see those restrictions will decisively simplify the calculations involved in solving for equilibrium in the stochastic model. Moreover, the RP property is theoretically appealing for at least two more reasons. First, it parallels a well-known property of the static equilibrium characterized by BM, which is to have a unique equilibrium where workers rank firms according to productivity. Second, RPE feature constrained-efficient labor reallocation at all dates: if workers consistently rank more productive firms higher than less productive ones, then job-to-job moves will always be up the productivity ladder. That is, in any RPE $L(\theta) = L^*(\theta)$ and the allocation is unique.

It is therefore natural to ask how general Rank-Preserving Equilibria are. We now show that under some weak sufficient conditions on the initial size distribution of employment, all Markov equilibria must have this property. This is the central result of the paper. We assume that Ω is finite only for simplicity of exposition and proof, to avoid dealing with measurability issues, but nothing conceptually depends on this restriction.

Proposition 2 (Ranked initial firm size implies rank-preserving equilibrium) *Assume Ω is finite. If at the initial date 0 the initial state of the economy is such that L_0 is non-decreasing in θ (i.e. higher- θ firms start out no smaller), then any symmetric Markov contract-posting Equilibrium is necessarily Rank-Preserving, and the initial ranking of firms'*

relative sizes is maintained on the equilibrium path. If Γ is degenerate and firms are equally productive, then the same conclusion holds and initially larger firms offer more and remain larger on any equilibrium path.

Although the proof, in Appendix A, is technically quite involved, the proposition has a simple economic intuition. In BM's steady-state model, more productive firms offer higher wages due to a single-crossing property of their steady state profits, which in turn reflects two very basic economic forces. First, a higher wage implies a larger firm size, as a more generous offer makes it easier to poach workers and to fend off competition. Second, a larger firm size is more valuable to a more productive firm, because each worker produces more. Therefore, by a simple monotone comparative statics argument, it must be the case that more productive firms offer more, employ more workers, and earn higher profits. Simply put, a productive firm can afford paying more, and is willing to do so to attract workers, because its opportunity cost of not producing is higher. Key to this argument is the fact that firm size is an endogenous object, and BM look for an appropriate firm size distribution which guarantees a stationary allocation.

In our dynamic model, firm size is a state variable, and its *initial* value is a parameter of the model, arbitrarily fixed, not an endogenous object. Therefore, in order to get a start on monotone comparative statics, it is sufficient (but not necessary) that the initial size distribution shares the key property of BM's steady state distribution; namely, it is increasing in productivity. In the proof, we essentially invoke a single-crossing property of the maximand (the term in $\langle \cdot \rangle$) in the Bellman equation of the modified but equivalent value-posting problem (14).¹⁰ A more productive firm still wants and can afford to pay more, now in terms of values accruing to workers. If larger, this firm has a further motive to offer more, namely more workers to retain, independently of its productivity. In contrast, the effect of a higher offer on successful poaching from other firms is independent of current size, because of CRS in production. Therefore, the initial ranking of sizes by productivity is preserved throughout, and so is the ranking by productivity of values offered to workers. This condition is only sufficient, not necessary. It aligns two separate motives to pay workers more, firm productivity and size, so clearly there is some slack. If firms are equally productive and only differ in their initial size, then only the size motive operates and all equilibria are RP, with no additional conditions.

We stress that this is a characterization result, which neither establishes nor requires existence, let alone uniqueness, of a RPE. Our main result says that, if a Markov contract-posting Equilibrium V exists, then V can only be a best response to itself if it is increasing

¹⁰In a way similar to that in which Caputo (2003) appeals to single-crossing properties of the Hamiltonian in his analysis of comparative dynamics for *deterministic* optimal control problems.

in θ , including the effect of endogenous size on the posted value. So ours is a general monotonicity result, which does not require to either propose or calculate a particular value-offer strategy. In the next section, we show by construction existence and uniqueness of a RPE, which must then be the unique Markov equilibrium of the contract-posting game.

To characterize a RPE we need to describe how allocation and prices depend on exogenous states. The allocation is easy because constrained efficient. We already know from Section 4 how the size of each firm evolves in equilibrium. Indeed, the same logic applies to any job ladder model in which a similar concept of RPE can be defined. Nothing in the dynamics of L^* or Λ^* depends on the particulars of the wage setting mechanism, so long as this is such that employed jobseekers move from lower-ranking into higher-ranking jobs in the sense of a time-invariant ranking. Therefore, this model's predictions about everything relating to firm sizes are in fact much more general than the wage- (or value-) posting assumption retained in the BM model. We now turn to supporting prices.

5.5 Existence and uniqueness of (Rank-Preserving) equilibrium

Our aim in this section is to characterize equilibrium contracts in a way that will provide a constructive proof of uniqueness and — subject to a sufficient condition — of existence of RP equilibrium contracts. We begin by establishing some important properties of optimal contracts. Equation (3) combined with the assumption that initial firm size, $L_0(\theta)$, is a continuous function of θ (see Section 3) ensures that $L^*(\theta)$ is a continuous function of θ at all dates in a RPE. With that in mind, we can establish the following additional properties of the joint value function S and worker value function V in a RPE:

Proposition 3 (Differentiability of Value Functions in RPE) *The following properties hold in a RPE:*

1. $L \mapsto S(\theta, L, \omega, \Lambda^*)$ is convex in L and differentiable in L at $L^*(\theta)$, i.e. $S_L(\theta, L^*(\theta), \omega, \Lambda^*)$ exists for all θ . Moreover, $S_L(\theta, L, \omega, \Lambda^*)$ is continuous in L at $L^*(\theta)$;
2. $\theta \mapsto S_L(\theta, L^*(\theta), \omega, \Lambda^*)$ is continuously differentiable in θ ;
3. $\theta \mapsto V(\theta, L^*(\theta), \omega, \Lambda^*)$ is continuously differentiable in θ .

The proof is in Appendix B. While most of that proof is essentially technical, it begins by establishing continuity of $\theta \mapsto V(\theta, L^*(\theta), \omega, \Lambda^*)$, which is intuitive by a simple improvement argument. If V jumps up at some value of θ , the right and left limits of this value at θ generate the same transitions and firm size, but the right limit costs the firm more, and revenues are continuous in θ .

The third statement in Proposition 3 allows us to differentiate (15) and (16) w.r.t. θ :

$$f(V | \omega, \Lambda^*) \cdot \frac{dV}{d\theta} = \gamma(\theta) \quad \text{and} \quad g(V | \omega, \Lambda^*) \cdot \frac{dV}{d\theta} = L^*(\theta) \gamma(\theta). \quad (17)$$

at $V = V(\theta, L^*(\theta), \omega, \Lambda^*)$.

These differentiability properties allow the use in (14) of first-order conditions, which, for each state ω' , write down as (using the definition of $\mathcal{L}(\cdot)$ again and using subscripts to denote partial derivatives):

$$\begin{aligned} & \lambda_0^{\omega'} (1 - \Lambda^*(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda^*(\bar{\theta}) G(W(\omega') | \omega', \Lambda^*) \\ & = [S_L(\theta, \mathcal{L}(L^*(\theta), \omega', \Lambda^*, W(\omega')), \omega', \mathcal{T}(\omega', \Lambda^*)) - W(\omega')] \\ & \times (1 - \delta^{\omega'}) \lambda_1^{\omega'} [L^*(\theta) f(W(\omega') | \omega', \Lambda^*) + \Lambda^*(\bar{\theta}) g(W(\omega') | \omega', \Lambda^*)] - m^{\omega'} \end{aligned} \quad (18)$$

where $m^{\omega'}$ is the Lagrange multiplier for the workers' participation constraint $W(\omega') \geq U(\omega', \mathcal{T}(\omega', \Lambda^*))$, and where complementary slackness $m^{\omega'} [W(\omega') - U(\omega', \mathcal{T}(\omega', \Lambda^*))] = 0$ applies. In a RPE, (18) is solved by $W = V(\theta, L^*(\theta), \omega, \Lambda^*)$. Next, in the firm's problem (14), the Envelope condition w.r.t. firm size writes as:

$$\begin{aligned} & S_L(\theta, L^*(\theta), \omega, \Lambda^*) = \omega\theta + \beta \int_{\Omega} \left\{ \delta^{\omega'} U(\omega', \mathcal{T}(\omega', \Lambda^*)) + (1 - \delta^{\omega'}) \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF(v | \omega', \Lambda^*) \right. \\ & \left. + S_L(\theta, \mathcal{L}(L^*(\theta), \omega', \Lambda^*, W(\omega')), \omega', \mathcal{T}(\omega', \Lambda^*)) (1 - \delta^{\omega'}) (1 - \lambda_1^{\omega'} \bar{F}(W(\omega') | \omega', \Lambda^*)) \right\} H(d\omega' | \omega). \end{aligned} \quad (19)$$

We now introduce a time index t again. With a slight notational abuse, we denote:

$$V_{t+1}(\theta | \omega) := V(\theta, L_t^*(\theta), \omega, \Lambda_t^*) \quad \text{and} \quad U_t(\omega) := U(\omega, \Lambda_t^*).$$

We further define the costate variable:

$$\mu_{t+1}(\theta | \omega) := S_L(\theta, \mathcal{L}(L_t^*(\theta), \omega, \Lambda_t^*, V_{t+1}(\theta | \omega)), \omega, \mathcal{T}(\omega, \Lambda_t^*)),$$

which measures the shadow value to the worker-firm collective of the marginal worker, given the aggregate state, along the equilibrium path. Note that the dependence of V and μ on the state variables L^* and Λ^* is subsumed into the time index in the above notation, which is licit as those two variables evolve deterministically conditional on ω . Combining (18) and the various restrictions (15), (16), and (17) that hold in a RPE, we obtain the RPE version of the FOC (18):

$$\begin{aligned} & \lambda_0^{\omega} u_t + \lambda_1^{\omega} (1 - \delta^{\omega}) \Lambda_t^*(\theta) \\ & = \lambda_1^{\omega} (1 - \delta^{\omega}) [\mu_{t+1} - V_{t+1}] [L_t^*(\theta) f(V_{t+1} | \omega, \Lambda_t^*) + (1 - u_t) g(V_{t+1} | \omega, \Lambda_t^*)] - m_t^{\omega} \\ & = 2\lambda_1^{\omega} (1 - \delta^{\omega}) \frac{L_t^*(\theta) \gamma(\theta)}{dV_{t+1}/d\theta} (\mu_{t+1} - V_{t+1}) - m_t^{\omega}, \end{aligned} \quad (20)$$

and the RPE version of the Euler equation (19):

$$\begin{aligned} \mu_t(\theta | \omega) = \omega\theta + \beta \int_{\Omega} \left\{ \delta^{\omega'} U_{t+1}(\omega') + (1 - \delta^{\omega'}) \lambda_1^{\omega'} \int_{\theta}^{+\infty} V_{t+1}(x | \omega') d\Gamma(x) \right. \\ \left. + \mu_{t+1}(\theta | \omega') (1 - \delta^{\omega'}) (1 - \lambda_1^{\omega'} \bar{\Gamma}(\theta)) \right\} H(d\omega' | \omega), \end{aligned} \quad (21)$$

Note that now the shadow marginal value μ only depends on the distribution of employment Λ^* through total employment in all firms of productivity up to θ , $\Lambda^*(\theta)$ and the corresponding density $L^*(\theta)\gamma(\theta)$. Both are scalars, and the state reduces from $\hat{\zeta} = (\theta, L, \omega', \Lambda)$, which is infinite-dimensional due to the relevance of the entire firm size distribution Λ , to the four-dimensional vector $\mathbf{z} = (\theta, L, \omega', \Lambda(\theta))$: in order to make its decisions, the firm only needs to know the mass of employment at less productive firms $\Lambda(\theta)$ and not the entire size distribution Λ .

Finally, a Transversality Condition (TVC) requires that the discounted joint value of the marginal worker vanishes in expectation w.r. to the stochastic path of ω

$$\lim_{t \rightarrow \infty} \mathbf{E} [\beta^t \mu_t(\theta | \omega) L_t^*(\theta) | \mathbf{z}_0] = 0. \quad (22)$$

We now assume that $\omega\underline{\theta} \geq b^\omega \geq 0$ for all ω , so that $U \geq 0$, because a worker has always the option of staying unemployed to collect positive payoffs. We also assume that $\lambda_0^\omega - \lambda_1^\omega$ is small enough for every ω that the worker participation constraint never binds in equilibrium.¹¹ A RPE is then a value V increasing in θ , a shadow value of employment μ , and a value of unemployment U positive and smaller than V , obeying the boundary condition $V_t(\underline{\theta} | \omega) = U_t(\omega)$ and solving the FOC (20), the Euler equation (21) and the unemployment Bellman equation (7) given the RPE employment dynamics (1), subject to the TVC (22). Let

$$Q_t(\theta | \omega) := (\lambda_0^\omega u_t + \lambda_1^\omega (1 - \delta^\omega) \Lambda_t^*(\theta))^2.$$

We can verify by direct substitution that the following value function satisfies the FOC (20) and the boundary condition:

$$V_t(\theta | \omega) = \frac{Q_t(\underline{\theta} | \omega)}{Q_t(\theta | \omega)} U_t(\omega) + \int_{\underline{\theta}}^{\theta} \mu_t(x | \omega) \frac{\partial Q_t}{\partial \theta}(x | \omega) \frac{dx}{Q_t(\theta | \omega)} := \mathbf{T}_V[\mu_t, U_t](\theta | \omega).$$

Let \mathcal{E}_Θ be the space of continuous cdf's over $\Theta = [\underline{\theta}, \bar{\theta}]$, $\mathcal{F}_{\Theta \times \Omega \times \mathcal{E}}$ be the space of positive functions $\Theta \times \Omega \times \mathcal{E}_\Theta \rightarrow \mathbb{R}_+^2$ such that the first component is θ -integrable and the second

¹¹Notice that when $\lambda_0 = \lambda_1$ the worker has no reason to decline any offer, and with $\omega\underline{\theta} > b^\omega$ even the least productive firm can hire some unemployed workers and obtain positive profits. We abstract from entry and exit of firms to focus on the poaching competition and job ladder.

component is constant as we vary $\theta \in \Theta$, and $\mathbf{T} = \begin{pmatrix} \mathbf{T}_\mu \\ \mathbf{T}_U \end{pmatrix}$ be the linear function on $\mathcal{F}_{\Theta \times \Omega \times \mathcal{E}}$ defined by

$$\mathbf{T}_\mu [\mu_t, U_t] (\theta | \omega) = \int_{\Omega} \left\{ \delta^{\omega'} U_{t+1} (\omega') + \mu_{t+1} (\theta | \omega') (1 - \delta^{\omega'}) (1 - \lambda_1^{\omega'} \bar{\Gamma} (\theta)) \right. \\ \left. + (1 - \delta^{\omega'}) \lambda_1^{\omega'} \int_{\underline{\theta}}^{\bar{\theta}} \mathbf{T}_V [\mu_{t+1}, U_{t+1}] (x | \omega') d\Gamma (x) \right\} H (d\omega' | \omega)$$

$$\mathbf{T}_U [\mu_t, U_t] (\omega) = \int_{\Omega} \left\{ (1 - \lambda_0^{\omega'}) U_{t+1} (\omega') + \lambda_0^{\omega'} \int_{\underline{\theta}}^{\bar{\theta}} \mathbf{T}_V [\mu_{t+1}, U_{t+1}] (x | \omega') d\Gamma (x) \right\} H (d\omega' | \omega)$$

Note that by definition of these mappings, \mathbf{T} preserves positivity of its arguments and the second component is independent of θ , so that \mathbf{T} maps $\mathcal{F}_{\Theta \times \Omega \times \mathcal{E}}$ into itself, whenever the function is well defined (the integrals exist).

Then a RPE is a solution $\begin{pmatrix} \mu \\ U \end{pmatrix} \in \mathcal{F}_{\Theta \times \Omega \times \mathcal{E}}$ of

$$\begin{pmatrix} \mu \\ U \end{pmatrix} = \begin{pmatrix} \omega\theta \\ b\omega \end{pmatrix} + \beta \mathbf{T} \begin{pmatrix} \mu \\ U \end{pmatrix}, \quad (23)$$

which satisfies the TVC (22) and has $0 \leq U \leq \mathbf{T}_V [\mu, U] \leq \mu$ and $\mathbf{T}_V [\mu, U]$ increasing in θ .

We are now in a position to prove the following result:

Proposition 4 (Uniqueness and Existence) *There exists at most one equilibrium, which is Rank-Preserving. If it exists, the optimal contract in this unique RPE is the wage policy that pays the worker a value $\mathbf{T}_V [\mu^*, U^*]$ where:*

$$\begin{pmatrix} \mu^* \\ U^* \end{pmatrix} (\theta | \omega) := \lim_{n \rightarrow \infty} \sum_{j=0}^n \beta^j \mathbf{T}^j \begin{pmatrix} \omega\theta \\ b\omega \end{pmatrix} (\theta | \omega). \quad (24)$$

Existence is guaranteed under the sufficient conditions $\omega\underline{\theta} \geq b\omega$ and $\lambda_0^\omega \leq \lambda_1^\omega (1 - \delta^\omega) \forall \omega$.

The proof, in Appendix C, simply proceeds through forward substitution and induction, and establishes also that this limit exists. While we have not been able to derive conditions on parameters that are both necessary and sufficient for equilibrium existence (i.e. for (μ^*, U^*) to be a RPE), this is not an issue in applications. In fact, we proved that there is only one possible equilibrium set of contracts, that we can compute (see MPV10a) and then check ex post whether in fact it satisfies all equilibrium conditions.

6 Discussion

6.1 Contracts

There is no unique way to generalize the steady-state BM model to an environment that is subjected to aggregate shocks. Our proposed extension of BM's model features contracts that implement the efficient allocation and generate equilibrium dynamics that preserves all desirable properties from BM, but also explain MPV11's facts. We now revisit some of our assumptions, both to explore the robustness of our theoretical results and to prepare the ground for future research.

While our proposed contracts are “general”¹² within the bounds imposed by commitment and the equal treatment constraint, the RP property and constrained efficient allocation of employment are also implemented in more restricted contract-posting games. For example, suppose firms can post wages that are conditioned on the aggregate state ω and their own type θ , but not on L or Λ because, say, L is too difficult to verify for the worker, and Λ is too difficult to measure for either the worker or the firm. The proof of Proposition 2 can easily be adapted (indeed, simplified) to show that a firm's best response to any strategy adopted by other firms is increasing in θ as long as it can be conditioned on the aggregate state ω . Therefore, this game in restricted contracts that can only depend on θ and ω must also have a unique equilibrium which is RP and constrained efficient.

Next, the assumption of commitment to state-contingent wages is standard in the literature, where it is well understood that commitment may both be beneficial to the firm and sustainable in long term employment relationships. Nonetheless, this assumption could be relaxed in many different ways. For example, the firm may be allowed a choice of whether or not commit to a specific strategy (as in Postel-Vinay and Robin, 2004). Or, it may only be committed to end-of-period payments with (as in Coles and Mortensen, 2011) or without an equal-treatment constraint. Then, a firm may choose to deliver value to its workers upon hiring them, and then squeeze them against the participation constraint thereafter. This strategy of extremely front-loaded payments (pure sign-up bonus and then rip-off) has the cost of accelerating turnover. Even without equal treatment, it is then plausible that an individual firm may not want to play this strategy if no other firm does, but rather choose to backload wages sequentially, even without commitment, due to the need to retain workers once hired. This is an interesting avenue for future research.

¹²In the sense that they are conditioned on the largest possible state space consistent with fairly standard assumptions about observability and the Markov requirement.

6.2 Sufficient conditions for RPE

The conditions for uniqueness and RP property of equilibrium are only sufficient and can be relaxed to an extent. However it is also possible to modify the economy so that those conditions fail and our unique, efficient RPE breaks down. One of the key restrictions here is that we have treated firm productivity as a fixed, time-invariant parameter. Shocks to firm productivity create obvious issues for RPE, as a very large and productive firm may suddenly become unproductive, and then face contrasting incentives to offer its employees a high value, its sheer size and retention needs against low productivity. Similarly, highly profitable business opportunities may arise and cause entry of highly productive firms, which by definition start out with a size of zero.

To illustrate this using our formal results, and to provide an example situation in which equilibrium fails to be RP, we follow Hopenhayn (1992) and assume that a large measure of potential entering firms can pay an entry cost to draw θ from Γ , and then decide whether to be active or not. An entering firm of size $L = 0$ and productivity $\theta + \varepsilon$ offers strictly less than an incumbent firm of productivity θ and size $L(\theta) > 0$, for $\varepsilon > 0$ small enough. In fact, by Claim 3 of Lemma 1 in Appendix A, the best response contract $V^*(\theta, L)$ to *any* distributions of offers F and payments G that are consistent with equilibrium is increasing in θ and L , so $V^*(\theta, L(\theta)) > V^*(\theta, 0)$. By Claim 1 of Lemma 5 in Appendix B, $V^*(\theta, L)$ is continuous in θ , so there exists $\varepsilon > 0$ small enough that $V^*(\theta, L(\theta)) > V^*(\theta + \varepsilon, 0) > V^*(\theta, 0)$. In this case the productivity disadvantage of firm θ is swamped by the stronger motive to offer a more generous contract to retain a stock of employment $L(\theta) > 0$.

Clearly a similar phenomenon occurs when the productivity of *existing* firms is subject to idiosyncratic shocks. The relevant empirical question, then, is how variable is a typical firm's productivity at business cycle frequencies. The available empirical evidence using longitudinal business microdata is mostly limited to the manufacturing sector. Summarizing the results of the early literature, Bartelsman and Doms (2000) conclude that firm-level productivity is best characterized as a unit-root process. Haltiwanger et al. (2008, Table 3) find that establishment-level TFP is very persistent, about as much as aggregate TFP. MPV11 show with data from a few countries that several correlated features of a firm, such as its size, the average wages it pays, and its revenue-based productivity, when measured at one point in time strongly predict how job creation by the same firm responds to business cycle shocks that hit it *over two decades later*. This striking phenomenon suggests that our assumption of fixed firm productivity might be a reasonable approximation for our purposes, at least at business cycle frequencies. We note that the assumption of fixed and heterogeneous firm-level TFP has become commonplace in the International Trade literature (Eaton and Kortum 2002, Melitz 2003).

From a theoretical standpoint, however, the question remains whether one can analyze aggregate dynamics in wage-posting models in the presence of idiosyncratic shocks to firm productivity. As discussed in Section 2, Coles and Mortensen (2011) answer this question by assuming a specific hiring technology, which makes contracts size-independent. In contrast, models that allow ex-post competition between employers for employed workers always implement efficient turnover, as those models all share the auction-flavored property that the employer with the highest valuation of the worker’s services (i.e., the most productive employer) always succeeds in hiring/retaining the worker.¹³

6.3 Endogenous job creation

We have treated job-contact probabilities as exogenous, albeit state-dependent, objects. More natural and common is to endogenize them through a matching function. We envision the following natural extension of the model. A firm can post vacancies, or spend hiring effort, at a convex cost. Own vacancies determine the firm’s sampling weight in workers’ job search. The firm now has two tools at its disposal to recruit workers, promised contract value and vacancies (hiring effort).¹⁴ We expect the dynamic single-crossing property that we uncover in the value function of the optimal contract-posting problem to imply that not only the value offered to the worker, but also the intensity of hiring effort, increase with firm productivity and size. This result would only reinforce the mechanism that we highlight and which gives rise to the unique RPE. Quantitatively, it would greatly help to explain the empirical inequality in firm sizes based on labor turnover frictions alone. On the time domain, in order for large firms to exhibit more cyclical job creation rates in equilibrium, they would have to post in equilibrium a measure of vacancies that is procyclical relative to that posted by small firms.

Finally, we are aware that multiple factors, beyond employment frictions, contribute to determine the size of a firm, most notably capital adjustment costs, including financial frictions, and diminishing marginal revenues from hiring, due to either technology (decreasing returns to labor), span of control frictions, or price-making power. Diminishing returns in wage-posting models have been partially explored in a steady state context, and can invalidate some of the equilibrium properties, such as the absence of atoms in the offer distribution. We cannot identify, though, obvious reasons why they would overturn the main result that equilibrium must be RP. To violate this property, a more productive firm

¹³Postel-Vinay and Robin, 2002; Dey and Flinn, 2005; Cahuc, Postel-Vinay and Robin, 2006 — see also Postel-Vinay and Turon, 2010, for an example with idiosyncratic productivity shocks

¹⁴As mentioned in Section 2, Coles and Mortensen (2011) pursue this avenue under the assumptions of no commitment to wage offers and a specific hiring cost function, which makes wages size-independent, so recruiting is entirely controlled by hiring investment.

would have to optimally hire so many more workers as to drive its marginal revenue of labor below that of a less productive firm. This is another avenue for future investigation.

7 Conclusion

This paper is the first to characterize stochastic equilibrium of an economy where the Law of One Price fails due to random search frictions and monopsony power, a problem that was long held to be intractable. Specifically, we introduce aggregate productivity shocks in a wage-posting model a la Burdett and Mortensen (1998), and we allow for rich state-contingent employment contracts. By extending the theory of Monotone Comparative Statics to a Dynamic Programming environment, we identify sufficient conditions under which the equilibrium is unique, constrained efficient, and very tractable. The second best is decentralized by contracts that do not respond to outside offers. The equilibrium stochastic dynamics of this model economy exhibit qualitative properties that are in line with the new business cycle facts that we illustrate in MPV08, MPV11 and Moscarini and Postel-Vinay (2010b), most notably, that small firms as a group exhibit less cyclical net job creation and returns to capital than large firms.

Future research will pursue a full quantitative analysis of this model, to illustrate its practicality as a tool for business cycle analysis. In MPV10a we illustrate an algorithm to solve quickly and efficiently for equilibrium contracts and we present some preliminary quantitative results. The constrained efficient allocation in the stochastic economy already naturally explains why larger firms have more cyclical net job creation. We believe that the slow propagation of aggregate shocks to average labor productivity and wages, due to the slow upgrading of labor through job-to-job quits, is an important feature of actual business cycles which is missing altogether from existing quantitative business cycle models. The extensions mentioned above, as well as possibly others, are bound to help the quantitative performance of the model.

References

- [1] Barlevy, G. (2008) “Identification of Search Models using Record Statistics”, *Review of Economic Studies*, 75(1) 29-64.
- [2] Bartelsman, E. and M. Doms, (2000) “Understanding Productivity: Lessons From Longitudinal Microdata”, *Journal of Economic Literature*, 38, 569-94.

- [3] Burdett, K. and M. Coles (2003) “Equilibrium Wage-Tenure Contracts”, *Econometrica*, 71(5), 1377-404.
- [4] Burdett, K. and D. T. Mortensen (1998) “Wage Differentials, Employer Size and Unemployment”, *International Economic Review*, 39, 257-73. [BM]
- [5] Cahuc, P., F. Postel-Vinay and J.-M. Robin (2006) “Wage Bargaining with On-the-job Search: Theory and Evidence”, *Econometrica*, 74(2), 323-64.
- [6] Caputo, M. R. (2003) “The Comparative Dynamics of Closed-loop Controls for Discounted Infinite-Horizon Optimal Control Problems”, *Journal of Economic Dynamics and Control*, 27, 1335-65.
- [7] Carrillo-Tudela, C. (2009) “An Equilibrium Search Model when Firms Observe Workers’ Employment Status”, *International Economic Review*, 50(2), 485-506.
- [8] Coles, M. G. (2001) “Equilibrium Wage Dispersion, Firm Size and Growth”, *Review of Economic Dynamics*, 4, 159-87.
- [9] Coles, M. G., and D. T. Mortensen (2011) “Equilibrium Wage and Employment Dynamics in a Model of Wage Posting Without Commitment”, Northwestern U. *mimeo*.
- [10] Corb e, D., M. B. Stinchcombe, and J. Zeman (2009), *An Introduction to Mathematical Analysis for Economic Theory and Econometrics*, Princeton University Press.
- [11] Dey, M. S. and C. J. Flinn (2005) “An Equilibrium Model of Health Insurance Provision and Wage Determination,” *Econometrica*, 73(2), 571-627.
- [12] Eaton, J., and S. Kortum (2002) “Technology, Geography, and Trade,” *Econometrica*, 70, 1741-1779.
- [13] Foster, L., J. Haltiwanger and C. Krizan (2000) “Aggregate Productivity Growth: Lessons From Microeconomic Evidence”, in C. R. Hulten, E. R. Dean and M. J. Harper, eds., *New Developments in Productivity Analysis*, Chicago, IL: U. of Chicago Press.
- [14] Gonzalez, F. and S. Shi (2010) “An Equilibrium Theory of Learning, Search, and Wages”, *Econometrica*, 78(2), 509-538.
- [15] Haltiwanger, J., L. Foster and C. Syverson (2008), “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?”, *American Economic Review*, 98:1, 394-425.

- [16] Haltiwanger, J., R. Jarmin and J. Miranda (2010), “Who Creates Jobs? Small vs. Large vs. Young”. NBER WP 16300.
- [17] Hopenhayn, H. (1992) “Entry, Exit, and firm Dynamics in Long Run Equilibrium”, *Econometrica*, 60(5), 1127-50
- [18] Kaas, L. and P. Kircher (2010), “Efficient Firm Dynamics in a Frictional Labor Market”, *mimeo*.
- [19] Lentz, R. and D. T. Mortensen (2008) “An Empirical Model of Growth Through Product Innovation”, *Econometrica*, 76(6), 1317-73.
- [20] Maskin, E. and J. Tirole (2001) “Markov Perfect Equilibrium: I. Observable Actions”, *Journal of Economic Theory*, 100(2), 191-219.
- [21] Melitz, M. (2003) “The Impact of Trade on Intra-industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, 71(6), 1695-1725.
- [22] Menzio, G. and S. Shi (2011) “Efficient Search on the Job and the Business Cycle”, *Journal of Political Economy*, 119(3), 468-510.
- [23] Mortensen, D. (2003) *Wage Dispersion: Why Are Similar Workers Paid Differently?*, MIT Press.
- [24] Moscarini, G. and F. Postel-Vinay (2008) “The Timing of Labor Market Expansions: New Facts and a New Hypothesis”, in D. Acemoglu, K. Rogoff and M. Woodford, eds., *NBER Macroeconomics Annual*, 23, 1-51.
- [25] Moscarini, G. and F. Postel-Vinay (2010a) “Stochastic Search Equilibrium”, Cowles Foundation Discussion Paper 1754. [MPV10a]
- [26] Moscarini, G. and F. Postel-Vinay (2010b) “Unemployment and Small Cap Returns: the Nexus”, *American Economic Review*, Papers and Proceedings, May 2010, 100(2), 333-337. [MPV10b]
- [27] Moscarini, G. and F. Postel-Vinay (2011) “The Contribution of Large and Small Employers to Job Creation in Times of High and Low Unemployment”, forthcoming *American Economic Review*. [MPV11]
- [28] Postel-Vinay, F. and J.-M. Robin (2002) “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity”, *Econometrica*, 70(6), 2295-350.

- [29] Postel-Vinay, F. and J.-M. Robin (2004) “To Match Or Not To Match? Optimal Wage Policy with Endogenous Worker Search Intensity”, *Review of Economic Dynamics*, 7(2), 297-331.
- [30] Postel-Vinay, F. and H. Turon (2010) “On-the-job Search, Productivity Shocks, and the Individual Earnings Process”, *International Economic Review*, 51(3), 599-629.
- [31] Royden, H. L. (1988) *Real Analysis – Third Edition*, Prentice-Hall.
- [32] Rudanko, L. (2011) “Aggregate and Idiosyncratic Risk in a Frictional Labor Market”, *American Economic Review*, 101(6), 2823-2843.
- [33] Rudin, W. (1976) *Principles of Mathematical Analysis – Third Edition*, McGraw-Hill.
- [34] Shimer, R. (2008) “Discussion of Moscarini and Postel-Vinay (2008)”, in D. Acemoglu, K. Rogoff and M. Woodford, eds., *NBER Macroeconomics Annual*, 23, 59-71.
- [35] Spear, S. and S. Srivastava, (1987), “On Repeated Moral Hazard With Discounting”, *Review of Economic Studies*, 54, 599-617.
- [36] Stokey, N. and R. Lucas Jr. (1989), “Recursive Methods in Economic Dynamics”, Harvard University Press.
- [37] Topkis, D. M. (1998), *Supermodularity and Complementarity*, Princeton, N.J.: Princeton University Press.

Appendix

A Proof of Proposition 2

For convenience, we repeat the firm’s DP problem (14):

$$\begin{aligned}
S(\theta, L, \omega, \Lambda) &= \omega\theta L + \beta \int_{\Omega} \left\{ \delta^{\omega'} U(\omega', \mathcal{T}(\omega', \Lambda)) L \right. \\
+ \sup_{W(\omega') \geq U(\omega', \mathcal{T}(\omega', \Lambda))} &\left\langle S(\theta, \mathcal{L}(L, \omega', \Lambda, W(\omega')), \omega', \mathcal{T}(\omega', \Lambda)) + L(1 - \delta^{\omega'}) \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF(v | \omega', \Lambda) \right. \\
&\left. \left. - W(\omega') \left(\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda(\bar{\theta}) G(W(\omega') | \omega', \Lambda) \right) \right\rangle \right\} H(d\omega' | \omega),
\end{aligned}$$

and the claim: if this problem has a solution, then any measurable selection $V(\theta, L, \omega, \Lambda)$ from the optimal correspondence is such that $V(\theta, L^*(\theta), \omega, \Lambda)$ is increasing in θ . We introduce the following notation:

$$\begin{aligned}
A(\theta, L, \omega, \Lambda) &:= \omega\theta L + \beta \int_{\Omega} \delta^{\omega'} U(\omega', \mathcal{F}(\omega', \Lambda)) LH(\omega' | \omega), \\
B(L, \omega', \Lambda; W(\omega')) &:= L \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF(v | \omega', \Lambda) \\
&\quad - W(\omega') \left(\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda(\bar{\theta}) G(W(\omega') | \omega', \Lambda) \right).
\end{aligned}$$

Our proof strategy is as follows. First, we define certain supermodularity properties SM of a value function that imply that the maximizer V in (14) is increasing in θ . Then, we fix an arbitrary Λ and show that the Bellman operator in (14) for the restricted problem with fixed Λ is a contraction mapping from the space of SM functions into itself, and that this space is Banach and closed under the sup norm. Therefore, for any fixed Λ (14) has a unique solution. Finally, *if* there exists a solution S to (14) when Λ is not fixed, then S must also solve the restricted problem (14) for any fixed Λ . By uniqueness and SM of the solution to the restricted problem any solution to the unrestricted problem must also have the SM properties. We cannot extend the same logic to show existence of S with variable Λ because Blackwell's sufficient conditions for a contraction mapping apply only to functions over \mathbb{R}^n .

So fix Λ to be some given CDF over $[\underline{\theta}, \bar{\theta}]$. Then, for any function $\mathcal{S}(\theta, L, \omega)$, we define the following operator \mathbf{M}^Λ :

$$\begin{aligned}
\mathbf{M}^\Lambda \mathcal{S}(\theta, L, \omega) &:= A(\theta, L, \omega, \Lambda) \\
&\quad + \beta \int_{\Omega} \max_{W(\omega')} \left\langle \mathcal{S}[\theta, \mathcal{L}(L, \omega', \Lambda, W(\omega')), \omega'] + B(L, \omega', \Lambda; W(\omega')) \right\rangle H(d\omega' | \omega). \quad (25)
\end{aligned}$$

The following additional consideration simplifies the proof: the worker participation constraint $W \geq U$ can be ignored in the proof. To see why, observe the following. Once we establish that an interior solution is increasing in θ , we can conclude that any set of firms that offers a corner solution $W = U$ and shuts down must be the set of the least productive firms. But then, the global solution, including the corner, is weakly increasing in θ as claimed. Incidentally, if all firms offered U , from the previous reasoning (and barring the trivial case where all firms are too unproductive to operate) the most productive firms would deviate and profitably offer more, so there exist always some firms that have an interior solution where PC does not bind.

Lemma 1 *Let $\mathcal{S}(\theta, L, \omega)$ be bounded, continuous in θ and L , increasing and convex in L and with increasing differences in (θ, L) over $(\underline{\theta}, \bar{\theta}) \times (0, 1)$. Then:*

1. $\mathbf{M}^\Lambda \mathcal{S}$ is bounded and continuous in θ and L ;
2. There exists a measurable selection $V(\theta, L, \omega, \Lambda)$ from the maximizing correspondence associated with $\mathbf{M}^\Lambda \mathcal{S}$;
3. Any such measurable selection V is increasing in θ and L ;
4. $\mathbf{M}^\Lambda \mathcal{S}$ is increasing and convex in L and with increasing differences in (θ, L) over $(\underline{\theta}, \bar{\theta}) \times (0, 1)$.

Proof. In this proof, wherever possible without causing confusion, we will make the dependence of all functions on aggregate state variables ω and Λ implicit to streamline the notation.

Points 1 and 2 of this lemma are immediate: continuity of $\mathbf{M}^\Lambda \mathcal{S}$ is a direct consequence of Berge's Theorem. Boundedness of $\mathbf{M}^\Lambda \mathcal{S}$ is obvious by construction. Existence of a measurable selection from the maximizing correspondence associated with $\mathbf{M}^\Lambda \mathcal{S}$ is a direct consequence of the Measurable Selection Theorem.

To prove point 3, we first establish that the maximand in (25) has increasing differences in (θ, W) and (L, W) . Monotonicity of V in θ and L will then follow from standard monotone comparative statics arguments. Proving that the maximand in (25) has increasing differences in (θ, W) is immediate as B is independent of θ : letting $\tau > 0$, differences in θ of the maximand equal $\mathcal{S}(\theta + \tau, \mathcal{L}(L, W)) - \mathcal{S}(\theta, \mathcal{L}(L, W))$ which is increasing in W because \mathcal{L} is increasing in W by construction and \mathcal{S} has increasing differences in (θ, L) by assumption. We thus now fix θ and focus on establishing that the maximand in (25) has increasing differences in (L, W) . To this end, first note that, since \mathcal{S} is assumed to be continuous and convex in L , it has left and right derivatives everywhere (and those two can at most differ at countably many points). Now take L and $h > 0$ and define the difference in L of the maximand in (25):

$$\begin{aligned} \mathcal{D}(W) := & \mathcal{S}(\theta, \mathcal{L}(L + h, W)) - \mathcal{S}(\theta, \mathcal{L}(L, W)) \\ & + h \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_W^{+\infty} v dF(v | \omega'). \end{aligned}$$

(The dependence of \mathcal{D} on θ is kept implicit.) We want to establish that $\mathcal{D}(W)$ is increasing in W . We do not know whether F and G , thus \mathcal{D} , are differentiable, so we proceed by showing that the upper-right Dini derivative of $\mathcal{D}(W)$, which

we denote as $D^+ \mathcal{D}(W)$ and which exists everywhere (although possibly equalling $\pm\infty$), is everywhere positive. Take $x > 0$:

$$\begin{aligned}
& \mathcal{D}(W+x) - \mathcal{D}(W) \\
&= \mathcal{S}(\theta, \mathcal{L}(L+h, W+x)) - \mathcal{S}(\theta, \mathcal{L}(L, W+x)) \\
&\quad - [\mathcal{S}(\theta, \mathcal{L}(L+h, W)) - \mathcal{S}(\theta, \mathcal{L}(L, W))] \\
&\quad - h \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_W^{W+x} v dF(v | \omega') \\
&= \{\mathcal{S}_{L,r}(\theta, \mathcal{L}(L+h, W)) + \varepsilon_1 [\mathcal{L}(L+h, W+x) - \mathcal{L}(L+h, W)]\} \\
&\quad \times \{\mathcal{L}(L+h, W+x) - \mathcal{L}(L+h, W)\} \\
&- \{\mathcal{S}_{L,r}(\theta, \mathcal{L}(L, W)) + \varepsilon_2 [\mathcal{L}(L, W+x) - \mathcal{L}(L, W)]\} \\
&\quad \times \{\mathcal{L}(L, W+x) - \mathcal{L}(L, W)\} \\
&\quad - h \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_W^{W+x} v dF(v | \omega')
\end{aligned}$$

where ε_1 and ε_2 are functions that have limit 0 at 0, $f_{x,\ell}$ [$f_{x,r}$] is used to designate the left [right] partial derivative of any function f w.r.t. x , and $\mathbf{M}^\Lambda \mathcal{S}$ has one-sided derivatives because it is convex. Majorizing the last integral:

$$\begin{aligned}
& \mathcal{D}(W+x) - \mathcal{D}(W) \\
&\geq \{\mathcal{S}_{L,r}(\theta, \mathcal{L}(L+h, W)) + \varepsilon_1 [\mathcal{L}(L+h, W+x) - \mathcal{L}(L+h, W)]\} \\
&\quad \times \{\mathcal{L}(L+h, W+x) - \mathcal{L}(L+h, W)\} \\
&- \{\mathcal{S}_{L,r}(\theta, \mathcal{L}(L, W)) + \varepsilon_2 [\mathcal{L}(L, W+x) - \mathcal{L}(L, W)]\} \\
&\quad \times \{\mathcal{L}(L, W+x) - \mathcal{L}(L, W)\} \\
&\quad - h \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} (W+x) [F(W+x | \omega') - F(W | \omega')].
\end{aligned}$$

Dividing through by x and taking the limit superior as $x \rightarrow 0^+$ (using the definition of \mathcal{L} , the fact that $\mathcal{S}_{L,r} \geq 0$ by assumption, continuity of F and G , and some basic properties of Dini derivatives), we obtain:

$$\begin{aligned}
D^+ \mathcal{D}(W) &\geq \\
& \mathcal{S}_{L,r}(\theta, \mathcal{L}(L+h, W)) \cdot \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \{(L+h) D^+ F(W) + \Lambda(\bar{\theta}) D^+ G(W)\} \\
& - \mathcal{S}_{L,r}(\theta, \mathcal{L}(L, W)) \cdot \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \{L D_+ F(W) + \Lambda(\bar{\theta}) D_+ G(W)\} \\
& \quad - h \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} W D_+ F(W),
\end{aligned}$$

where, in standard fashion, $D_+ F$ denotes the lower-right Dini derivative of F (and likewise for G). Because F and G are increasing, their Dini derivatives are

such that $D^+F \geq D_+F \geq 0$ (and likewise for G). Because \mathcal{S} is convex in L by assumption, $\mathcal{S}_{L,r}$ is increasing in L . Combining all those properties, the latter inequality implies:

$$D^+ \mathcal{D}(W) \geq [\mathcal{S}_{L,r}(\theta, \mathcal{L}(L, W)) - W] \cdot (1 - \delta^{\omega'}) \lambda_1^{\omega'} h D_+ F(W). \quad (26)$$

The only way the RHS in this last inequality can be negative is if $\mathcal{S}_{L,r}(\theta, \mathcal{L}(L, W)) - W < 0$. We now show that this cannot be if W is an optimal selection. Let V be an optimal selection and let $x > 0$. Optimality requires that:

$$\begin{aligned} 0 \geq & \mathcal{S}(\theta, \mathcal{L}(L, V-x)) + L \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V-x}^{+\infty} v dF(v | \omega') \\ & - (V-x) \left(\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda(\bar{\theta}) G(V-x | \omega')\right) \\ & - \mathcal{S}(\theta, \mathcal{L}(L, V)) - L \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_V^{+\infty} v dF(v | \omega') \\ & + V \left(\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda(\bar{\theta}) G(V | \omega')\right). \end{aligned}$$

Collecting terms and again majorizing the integral term as we did for \mathcal{D} :

$$\begin{aligned} 0 \geq & \{\mathcal{S}_{L,\ell}(\theta, \mathcal{L}(L, V)) + \varepsilon [\mathcal{L}(L, V-x) - \mathcal{L}(L, V)]\} \cdot \{\mathcal{L}(L, V-x) - \mathcal{L}(L, V)\} \\ & + L \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} (V-x) [F(V-x | \omega') - F(V | \omega')] \\ & - V \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \Lambda(\bar{\theta}) [G(V-x | \omega') - G(V | \omega')] \\ & + x \left(\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda(\bar{\theta}) G(V | \omega')\right). \end{aligned}$$

Now again taking the limit superior as $x \rightarrow 0^+$ (in what follows D^-F and D_-F designate the upper and lower left Dini derivative of F , respectively, and likewise for G):¹⁵

$$\begin{aligned} 0 \geq & -\mathcal{S}_{L,\ell}(\theta, \mathcal{L}(L, V)) \cdot \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \{LD_-F(V) + \Lambda(\bar{\theta}) D_-G(V)\} \\ & + V \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \{LD^-F(V) + \Lambda(\bar{\theta}) D^-G(V)\} \\ & + \lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda(\bar{\theta}) G(V | \omega'). \end{aligned}$$

Finally recalling that $D^-F \geq D_-F \geq 0$ (and likewise for G), the latter inequality implies:

$$\mathcal{S}_{L,\ell}(\theta, \mathcal{L}(L, V)) - V \geq \frac{\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda(\bar{\theta}) G(V | \omega')}{\lambda_1^{\omega'} (1 - \delta^{\omega'}) \{LD_-F(V) + \Lambda(\bar{\theta}) D_-G(V)\}} \geq 0. \quad (27)$$

¹⁵This uses the facts that $\mathcal{S}_{L,\ell} \geq 0$, that F and G are continuous, and that $D^-[-f] = -D^-f$ for any function f .

This, together with (26), shows that $D^+ \mathcal{D}(V) \geq 0$ at all V which is an optimal selection, i.e. at all V in the support of F . To finally establish that \mathcal{D} is increasing over the support of F , recall that, as F and G are continuous by Proposition 1, so is $W \mapsto \mathcal{L}(L, W)$. Moreover, as \mathcal{S} is convex in L (by assumption), it is continuous w.r.t. L . Thus by inspection, \mathcal{D} is a continuous function of W . Continuity plus the fact that $D^+ \mathcal{D}(V) \geq 0$ are sufficient to ensure that \mathcal{D} is strictly increasing (see, e.g., Proposition 2 p99 in Royden, 1988). Point 3 of the lemma is thus proven.

Now on to point 4. Take $(\theta_0, L_0) \in (\underline{\theta}, \bar{\theta}) \times (0, 1)$ and $h > 0$ such that $(\theta_0 + h, L_0 + h)$ are still in $(\underline{\theta}, \bar{\theta}) \times (0, 1)$. We first consider right-differentiability of $\mathbf{M}^\Lambda \mathcal{S}$ w.r.t. L at L_0 . Again fixing an arbitrary selection V from the optimal policy correspondence, we note that, while V may have a discontinuity at (θ_0, L_0) , the fact that it is increasing in L ensures that $V(\theta_0, L_0^+, \omega') := \lim_{h \rightarrow 0^+} V(\theta_0, L_0 + h, \omega')$ exists everywhere (and likewise for $V(\theta_0^+, L_0, \omega')$). By point 3, $V(\theta_0, L_0^+, \omega')$ is increasing in L_0 . Then:

$$\begin{aligned}
& \mathbf{M}^\Lambda \mathcal{S}(\theta_0, L_0 + h) - \mathbf{M}^\Lambda \mathcal{S}(\theta_0, L_0^+) = A(\theta_0, L_0 + h) - A(\theta_0, L_0) \\
& + \beta \int_{\Omega} \left\langle \mathcal{S}[\theta_0, \mathcal{L}(L_0 + h, V(\theta_0, L_0 + h, \omega'))] - \mathcal{S}[\theta_0, \mathcal{L}(L_0, V(\theta_0, L_0^+, \omega'))] \right. \\
& \quad \left. + B(L_0 + h; V(\theta_0, L_0 + h, \omega')) - B(L_0; V(\theta_0, L_0^+, \omega')) \right\rangle H(d\omega' | \omega) \\
& \geq A(\theta_0, L_0 + h) - A(\theta_0, L_0) \\
& + \beta \int_{\Omega} \left\langle \mathcal{S}[\theta_0, \mathcal{L}(L_0 + h, V(\theta_0, L_0^+, \omega'))] - \mathcal{S}[\theta_0, \mathcal{L}(L_0, V(\theta_0, L_0^+, \omega'))] \right. \\
& \quad \left. + B(L_0 + h; V(\theta_0, L_0^+, \omega')) - B(L_0; V(\theta_0, L_0^+, \omega')) \right\rangle H(d\omega' | \omega) \\
& = \left(\omega \theta_0 + \beta \int_{\Omega} \delta^{\omega'} U(\omega') H(d\omega' | \omega) \right) \cdot h \\
& + \beta \int_{\Omega} \left\langle \mathcal{S}[\theta_0, \mathcal{L}(L_0 + h, V(\theta_0, L_0^+, \omega'))] - \mathcal{S}[\theta_0, \mathcal{L}(L_0, V(\theta_0, L_0^+, \omega'))] \right. \\
& \quad \left. + h \cdot (1 - \delta^{\omega'}) \lambda_1^{\omega'} \int_{V(\theta_0, L_0^+, \omega')}^{+\infty} v dF(v | \omega') \right\rangle H(d\omega' | \omega), \quad (28)
\end{aligned}$$

where the last equality follows from the definitions of A and B . Then again:

$$\begin{aligned}
& \mathbf{M}^\mathbf{A} \mathcal{S}(\theta_0, L_0 + h) - \mathbf{M}^\mathbf{A} \mathcal{S}(\theta_0, L_0^+) = A(\theta_0, L_0 + h) - A(\theta_0, L_0) \\
& + \beta \int_{\Omega} \left\langle \mathcal{S}[\theta_0, \mathcal{L}(L_0 + h, V(\theta_0, L_0 + h, \omega'))] - \mathcal{S}[\theta_0, \mathcal{L}(L_0, V(\theta_0, L_0^+, \omega'))] \right. \\
& \quad \left. + B(L_0 + h; V(\theta_0, L_0 + h, \omega')) - B(L_0; V(\theta_0, L_0^+, \omega')) \right\rangle H(d\omega' | \omega) \\
& \leq A(\theta_0, L_0 + h) - A(\theta_0, L_0) \\
& + \beta \int_{\Omega} \left\langle \mathcal{S}[\theta_0, \mathcal{L}(L_0 + h, V(\theta_0, L_0 + h, \omega'))] - \mathcal{S}[\theta_0, \mathcal{L}(L_0, V(\theta_0, L_0 + h, \omega'))] \right. \\
& \quad \left. + B(L_0 + h; V(\theta_0, L_0 + h, \omega')) - B(L_0; V(\theta_0, L_0 + h, \omega')) \right\rangle H(d\omega' | \omega) \\
& = \left(\omega \theta_0 + \beta \int_{\Omega} \delta^{\omega'} U(\omega') H(d\omega' | \omega) \right) \cdot h \\
& + \beta \int_{\Omega} \left\langle \mathcal{S}[\theta_0, \mathcal{L}(L_0 + h, V(\theta_0, L_0 + h, \omega'))] - \mathcal{S}[\theta_0, \mathcal{L}(L_0, V(\theta_0, L_0 + h, \omega'))] \right. \\
& \quad \left. + h \cdot \left(1 - \delta^{\omega'} \right) \lambda_1^{\omega'} \int_{V(\theta_0, L_0 + h, \omega')}^{+\infty} v dF(v | \omega') \right\rangle H(d\omega' | \omega). \quad (29)
\end{aligned}$$

Now dividing through by h in (28) and (29), and invoking continuity w.r.t. V of $\mathcal{L}_L(L, V) = (1 - \delta^{\omega'}) (1 - \lambda_1^{\omega'} \overline{F}(V))$ (by continuity of F), everywhere right-differentiability of \mathcal{S} w.r.t. L (by convexity of \mathcal{S}), and existence of a right limit of V at any L_0 (by monotonicity of V established in point 1 of this lemma), we see that the lower and upper bounds of $\frac{1}{h} [\mathbf{M}^\mathbf{A} \mathcal{S}(\theta_0, L_0 + h) - \mathbf{M}^\mathbf{A} \mathcal{S}(\theta_0, L_0^+)]$ exhibited in (28) and (29) both converge to the same limit as $h \rightarrow 0^+$, which, together with continuity of $\mathbf{M}^\mathbf{A} \mathcal{S}$ in L at L_0 which implies $\mathbf{M}^\mathbf{A} \mathcal{S}(\theta_0, L_0^+) = \mathbf{M}^\mathbf{A} \mathcal{S}(\theta_0, L_0)$, establishes right-differentiability of $\mathbf{M}^\mathbf{A} \mathcal{S}$ w.r.t L with the following expression for $[\mathbf{M}^\mathbf{A} \mathcal{S}]_{L,r}(\theta, L)$

$$\begin{aligned}
[\mathbf{M}^\mathbf{A} \mathcal{S}]_{L,r}(\theta, L) &= \omega \theta + \beta \int_{\Omega} \delta^{\omega'} U(\omega') H(d\omega' | \omega) \\
& + \beta \int_{\Omega} \left\langle \mathcal{S}_{L,r}[\theta, \mathcal{L}(L, V(\theta, L^+, \omega'))] \cdot \mathcal{L}_L(L, V(\theta, L^+, \omega')) \right. \\
& \quad \left. + \left(1 - \delta^{\omega'} \right) \lambda_1^{\omega'} \int_{V(\theta, L^+, \omega')}^{+\infty} v dF(v | \omega') \right\rangle H(d\omega' | \omega). \quad (30)
\end{aligned}$$

Straightforward inspection shows that $[\mathbf{M}^\mathbf{A} \mathcal{S}]_{L,r}(\theta, L) > 0$, so that $\mathbf{M}^\mathbf{A} \mathcal{S}$ is increasing in L . We now show that $[\mathbf{M}^\mathbf{A} \mathcal{S}]_{L,r}(\theta, L)$ is increasing in L and θ . It is sufficient to show that the term under the \int in (30) is increasing in L and θ

for all $\omega' \in \Omega$. We begin with L . Let $L_1 < L_2 \in [0, 1]^2$. To lighten the notation, let $V_k = V(\theta, L_k^+, \omega')$ for $k = 1, 2$. Because V is increasing in L , $V_2 \geq V_1$. Then:

$$\begin{aligned}
& \mathcal{S}_{L,r}[\theta, \mathcal{L}(L_2, V_2)] \cdot \mathcal{L}_L(L_2, V_2) - \mathcal{S}_{L,r}[\theta, \mathcal{L}(L_1, V_1)] \cdot \mathcal{L}_L(L_1, V_1) \\
& \quad - \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V_1}^{V_2} v dF(v \mid \omega') \\
& = [\mathcal{L}_L(L_2, V_2) - \mathcal{L}_L(L_1, V_1)] \cdot \mathcal{S}_{L,r}[\theta, \mathcal{L}(L_2, V_2)] \\
& \quad + \mathcal{L}_L(L_1, V_1) \cdot (\mathcal{S}_{L,r}[\theta, \mathcal{L}(L_2, V_2)] - \mathcal{S}_{L,r}[\theta, \mathcal{L}(L_1, V_1)]) \\
& \quad - \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V_1}^{V_2} v dF(v \mid \omega') \\
& = \mathcal{L}_L(L_1, V_1) \cdot (\mathcal{S}_{L,r}[\theta, \mathcal{L}(L_2, V_2)] - \mathcal{S}_{L,r}[\theta, \mathcal{L}(L_1, V_1)]) \\
& \quad + \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V_1}^{V_2} (\mathcal{S}_{L,r}[\theta, \mathcal{L}(L_2, V_2)] - v) dF(v \mid \omega'),
\end{aligned}$$

where the last equality stems from the definition of \mathcal{L}_L . Because $\mathcal{S}_{L,r}$ and \mathcal{L} are both increasing in L , and because \mathcal{L} is also increasing in V , the first term in the r.h.s. of the last equality above is positive. Finally, convexity of \mathcal{S} combined with the first-order condition (27) implies that $\mathcal{S}_{L,r}[\theta, \mathcal{L}(L_2, V_2)] \geq \mathcal{S}_{L,\ell}[\theta, \mathcal{L}(L_2, V_2)] \geq V_2$, so that $\mathcal{S}_{L,r}[\theta, \mathcal{L}(L_2, V_2)] \geq v$ for all $v \leq V_2$, implying that the integral term is nonnegative. This shows that $[\mathbf{M}^\Lambda \mathcal{S}]_{L,r}$ is (strictly) increasing in L . The proof that $[\mathbf{M}^\Lambda \mathcal{S}]_{L,r}$ is strictly increasing in θ proceeds along similar lines (details available upon request). Thus $\mathbf{M}^\Lambda \mathcal{S}$ is a continuous function whose right partial derivative w.r.t. L exists everywhere, is increasing in L — which proves convexity w.r.t. L —, and increasing in θ — which proves increasing differences in (θ, L) . \square

Now consider the set of functions defined over $[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega$ that are continuous in (θ, L) and call it $C_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$. That set is a Banach space when endowed with the sup norm. As Lemma 1 suggests we will be interested in the properties a subset $C'_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega} \subset C_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$ of functions that are increasing and convex in L and have increasing differences in (θ, L) . We next prove two ancillary lemmas, which will establish as a corollary (Corollary 1) that $C'_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$ is closed in $C_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$ under the sup norm.¹⁶

Lemma 2 *Let X be an interval in \mathbb{R} and $f_n : X \rightarrow \mathbb{R}$, $N \in \mathbb{N}$ such that $\{f_n\}$ converges uniformly to f . Then:*

¹⁶While for the purposes of this proof (which is concerned with closedness under the sup norm) both lemmas are stated for sequences that converge uniformly, it is straightforward to extend them to the case of pointwise convergent sequences.

1. if f_n is nondecreasing for all n , so is f ;
2. if f_n is convex for all n , so is f .

Proof. For point 1, take $(x_1, x_2) \in X^2$ such that $x_2 > x_1$. Fix $k \in \mathbb{N}$. By uniform convergence, $\exists n_k \in \mathbb{N} : \forall n \geq n_k, \forall x \in X, |f_n(x) - f(x)| < \frac{1}{2k}$. Then:

$$f(x_2) - f(x_1) = \underbrace{f(x_2) - f_{n_k}(x_2)}_{> -1/2k} + \underbrace{f_{n_k}(x_2) - f_{n_k}(x_1)}_{\geq 0 \text{ by monotonicity of } f_{n_k}} + \underbrace{f_{n_k}(x_1) - f(x_1)}_{> -1/2k} > -\frac{1}{k}.$$

As the above is valid for an arbitrary choice of $k \in \mathbb{N}$ and $(x_1, x_2) \in X^2$, it establishes that f is nondecreasing. For point 2, uniform convergence of $\{f_n\}$ to f implies pointwise convergence, so that Theorem 6.2.35 p282 in Corbae, Stinchcombe and Zeman (2009) can be applied. \square

Lemma 3 *Let $X \subset \mathbb{R}^2$ be a convex set and $f_n : X \rightarrow \mathbb{R}$, $N \in \mathbb{N}$ be functions with increasing differences such that $\{f_n\}$ converges uniformly to f . Then f has increasing differences.*

Proof. Let $\{(x_1, y_1), (x_2, y_2)\} \in X^2$ such that $x_2 > x_1$ and $y_2 > y_1$. Fix $k \in \mathbb{N}$. By uniform convergence, $\exists n_k \in \mathbb{N} : \forall n \geq n_k, \forall (x, y) \in X, |f_n(x, y) - f(x, y)| < \frac{1}{4k}$. Then:

$$\begin{aligned} & f(x_2, y_2) - f(x_1, y_2) \\ &= \underbrace{f(x_2, y_2) - f_{n_k}(x_2, y_2)}_{> -1/4k} + \underbrace{f_{n_k}(x_2, y_2) - f_{n_k}(x_1, y_2)}_{> f_{n_k}(x_2, y_1) - f_{n_k}(x_1, y_1) \text{ by ID of } f_{n_k}} + \underbrace{f_{n_k}(x_1, y_2) - f(x_1, y_2)}_{> -1/4k} \\ &> -\frac{1}{2k} + f_{n_k}(x_2, y_1) - f_{n_k}(x_1, y_1) \\ &= -\frac{1}{2k} + \underbrace{f_{n_k}(x_2, y_1) - f(x_2, y_1)}_{> -1/4k} + f(x_2, y_1) - f(x_1, y_1) + \underbrace{f(x_1, y_1) - f_{n_k}(x_1, y_1)}_{> -1/4k} \\ &> -\frac{1}{k} + f(x_2, y_1) - f(x_1, y_1). \end{aligned}$$

As the above is valid for an arbitrary choice of $k \in \mathbb{N}$ and $\{(x_1, y_1), (x_2, y_2)\} \in X^2$, it establishes that f has increasing differences. \square

Corollary 1 *The set $C'_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$ of functions defined over $[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega$ that are increasing and convex in L and have increasing differences in (θ, L) is a closed subset of $C_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$ under the sup norm.*

The latter corollary establishes that, given a fixed Λ , the set of functions that are relevant to Lemma 1 is a closed subset of a Banach space of functions under the sup norm. The following lemma shows that the operator considered in Lemma 1 is a contraction under that same norm.

Lemma 4 *The operator \mathbf{M}^Λ defined in (25) maps $C'_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$ into itself and is a contraction of modulus β under the sup norm.*

Proof. That \mathbf{M}^Λ maps $C'_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$ into itself flows directly from a subset of the proof of Lemma 1. To prove that \mathbf{M} is a contraction, it is straightforward to check using (25) that \mathbf{M}^Λ satisfies Blackwell's sufficient conditions with modulus β . \square

We are now in a position to prove the proposition. Given the initially fixed Λ , the operator \mathbf{M}^Λ , which by Lemma 4 is a contraction from $C_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$ into itself, and has a unique fixed point \mathcal{S}_Λ in that set (by the Contraction Mapping Theorem). Moreover, since $C'_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$ is a closed subset of $C_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$ (Lemma 2) and since \mathbf{M}_Λ also maps $C'_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$ into itself (Lemma 1), that fixed point \mathcal{S}_Λ belongs to $C'_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$.

Summing up, what we have established thus far is that for any fixed $\Lambda \in C_{[\underline{\theta}, \bar{\theta}]}$, the operator \mathbf{M}^Λ over functions of (θ, L, ω) has a unique, bounded and continuous fixed point $\mathcal{S}_\Lambda^* = \mathbf{M}_\Lambda \mathcal{S}_\Lambda^* \in C'_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega} \subset C_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$.

We finally turn to the Bellman operator \mathbf{M} which is relevant to the firm's problem. That operator \mathbf{M} applies to functions $\overline{\mathcal{S}}$ defined on $[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega \times C_{[\underline{\theta}, \bar{\theta}]}$ and is defined as the following "extension" of \mathbf{M}^Λ :

$$\begin{aligned} \mathbf{M}\overline{\mathcal{S}}(\theta, L, \omega, \Lambda) &:= A(\theta, L, \omega, \Lambda) \\ &+ \beta \int_{\Omega} \max_{W(\omega')} \left\langle \overline{\mathcal{S}}[\theta, \mathcal{L}(L, \omega', \Lambda, W(\omega')), \omega', \mathcal{T}(\omega, \Lambda)] + B(L, \omega', \Lambda; W(\omega')) \right\rangle H(d\omega' | \omega). \end{aligned}$$

If an equilibrium exists, then a firm has a best response and a value S which solves $S = \mathbf{M}S$. For every $\Lambda \in C_{[\underline{\theta}, \bar{\theta}]}$, by definition of \mathbf{M} and \mathbf{M}_Λ this implies $S = \mathbf{M}_\Lambda S$. Since the fixed point of \mathbf{M}_Λ is unique, if $S = \mathbf{M}S$ exists then for every fixed $\Lambda \in C_{[\underline{\theta}, \bar{\theta}]}$ we have for all $(\theta, L, \omega) \in [\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega$: $S(\theta, L, \omega, \Lambda) = \mathcal{S}_\Lambda^*(\theta, L, \omega)$. Therefore, if the value function S and an equilibrium of the contract-posting game exist, then $S \in C'_{[\underline{\theta}, \bar{\theta}] \times [0, 1] \times \Omega}$: the typical firm's value function is continuous in θ and L , increasing and convex in L and has increasing differences in (θ, L) . By the same standard monotone comparative statics arguments that we invoked in the proof of Lemma 1, the maximizing correspondence is increasing in θ and L in the strong set sense, hence all of its measurable selections are weakly increasing in θ and L .

The proposition is finally established by the following simple induction. Consider two firms with $\theta_2 > \theta_1$. By assumption, at date 0, $L_2 \geq L_1$. Because any selection $V(\theta, L, \omega, \Lambda)$ from the maximizing correspondence of the typical firm's problem is increasing in θ and L , the values posted by those two firms at date 0 are such that $V_2 \geq V_1$. Then because \mathcal{L} is

strictly increasing in both L and V , firm 2 is again larger than firm 1 at date 1. The same reasoning applies again at date 1, and at all subsequent dates, so that $V_2 \geq V_1$ holds true at all dates.

Finally, if firms are equally productive the RP property follows as a simple corollary of the convexity of S in L , by the assumption that the initial Λ is continuous. \square

We conclude with a remark on atoms in the initial size distribution, including the symmetric case of identical firms that are equally productive and start out with the same size. That case would require mixed strategies in the first period. After the mixing plays out, in the second period of play firms would differ in size, and the previous case would apply from then on. We leave the computation of the equilibrium mixed strategies to future research.

B Proof of Proposition 3

In an attempt to simplify the notation without causing confusion, we define:

$$V^*(\theta, \omega) := V(\theta, L^*(\theta), \omega, \Lambda^*)$$

for use throughout this proof. This notation keeps the dependence of $V(\cdot)$ on Λ implicit.

The main purpose of Proposition 3 is actually to establish claim 2, continuous differentiability of V^* . Our proof strategy is as follows. We know from Proposition 2 that the optimal policy V^* is increasing in θ , hence differentiable a.e. It remains to show that it is differentiable everywhere. To do so, first, we establish continuity properties of $V(\theta, L, \omega, \Lambda^*)$ in θ , both for fixed L and for $L = L^*(\theta)$, and in L at $L = L^*(\theta)$ for fixed θ . Using these properties, we show that any solution to the Bellman equation $S(\theta, L, \omega, \Lambda)$ when all other firms are playing a RPE is continuously differentiable in L at $L = L^*(\theta)$; that is, on the equilibrium path the shadow marginal value of one worker always exists and is continuous in firm size. Next, we exploit this property and the implications of RPE to show that the optimal policy V^* is Lipschitz continuous in θ . This implies that V^* is differentiable everywhere.

We begin with an ancillary lemma, which is interesting in its own right.

Lemma 5 *V has the following continuity properties along the (RP) Equilibrium path:*

1. $\theta \mapsto V(\theta, L^*(\theta), \omega, \Lambda^*) = V^*(\theta, \omega)$ is continuous;
2. $L \mapsto V(\theta, L, \omega, \Lambda^*)$ is continuous at $L = L^*(\theta)$;
3. $\tau \mapsto V(\tau, L^*(\theta), \omega, \Lambda^*)$ is continuous at $\tau = \theta$.

Proof. $\theta \mapsto V^*(\theta, \omega)$ is increasing by Proposition 2, so V^* can only have (countably many) jump discontinuities. But then a jump discontinuity in V^* would imply a gap in the support of F , which is inconsistent with equilibrium as argued in Appendix A. This proves claim 1 of the lemma.

For claim 2, fix θ and $\varepsilon > 0$. Then by continuity of V^* (point 1 of this lemma), $\exists \alpha > 0 : \forall \eta \in (0, \alpha], V^*(\theta, \omega) \leq V^*(\theta + \eta, \omega) \leq V^*(\theta, \omega) + \varepsilon$. But then monotonicity of V in L and in θ (see Appendix A) further implies: $V^*(\theta, \omega) \leq V(\theta, L^*(\theta + \eta), \omega, \Lambda^*) \leq V^*(\theta + \eta, \omega) \leq V^*(\theta, \omega) + \varepsilon$, so that $\forall L \in [L^*(\theta), L^*(\theta + \eta)]$, $V(\theta, L, \omega, \Lambda^*) - V(\theta, L^*(\theta), \omega, \Lambda^*) \leq \varepsilon$, which establishes right-continuity of V in L at $L^*(\theta)$. Left-continuity is established in the same way, and so is continuity of $\tau \mapsto V(\tau, L^*(\theta), \omega, \Lambda^*)$ at $\tau = \theta$. \square

We now go on to establish point 1 of the proposition. In so doing, to avoid notational overload, we will keep the dependence of all value functions and laws of motion on Λ^* implicit. Now first, convexity of S w.r.t. L was established as a by-product of Proposition 2 (see Appendix A), and implies that S is everywhere left- and right-differentiable w.r.t. L , and that the right and left derivatives $S_{L,r}$ and $S_{L,\ell}$ are both increasing functions of L . As such they have right and left limits everywhere. We can thus define $S_{L,r}(\theta, L^+, \omega) = \lim_{h \rightarrow 0^+} S_{L,r}(\theta, L + h, \omega)$, and symmetrically $S_{L,\ell}(\theta, L^-, \omega) = \lim_{h \rightarrow 0^+} S_{L,\ell}(\theta, L - h, \omega)$. Now following exactly the same steps as in (28) and (29) (see the proof of Lemma 1 in Appendix A), only applied to S , we establish:

$$\begin{aligned} S_{L,r}(\theta, L^+, \omega) &= \omega\theta + \beta \int_{\Omega} \delta^{\omega'} U(\omega') H(d\omega' | \omega) \\ &\quad + \beta \int_{\Omega} \left\langle S_{L,r}[\theta, \mathcal{L}(L, \omega', V(\theta, L^+, \omega')), \omega'] \cdot \mathcal{L}_L(L, \omega', V(\theta, L^+, \omega')) \right. \\ &\quad \left. + (1 - \delta^{\omega'}) \lambda_1^{\omega'} \int_{V(\theta, L^+, \omega')}^{+\infty} v dF(v | \omega') \right\rangle H(d\omega' | \omega). \end{aligned}$$

Next, the facts that V is increasing in L (see the proof of Proposition 2) and continuous in L at $L = L^*(\theta)$ (from Lemma 5), combined with continuity of \mathcal{L} and \mathcal{L}_L w.r.t. V (by continuity of F), imply that $\mathcal{L}_L(L, \omega', V(\theta, L^+, \omega')) = \mathcal{L}_L(L, \omega', V(\theta, L, \omega'))$ and $S_{L,r}[\theta, \mathcal{L}(L, \omega', V(\theta, L^+, \omega')), \omega'] = S_{L,r}[\theta, \mathcal{L}(L, \omega', V(\theta, L, \omega'))^+, \omega']$ at $L = L^*(\theta)$. As a

further consequence:

$$\begin{aligned}
S_{L,r}(\theta, L^*(\theta)^+, \omega) &= \omega\theta + \beta \int_{\Omega} \delta^{\omega'} U(\omega') H(d\omega' | \omega) \\
&\quad + \beta \int_{\Omega} \left\langle S_{L,r}[\theta, \mathcal{L}(L, \omega', V^*(\theta, \omega'))^+, \omega'] \cdot \mathcal{L}_L(L, \omega', V^*(\theta, \omega')) \right. \\
&\quad \quad \left. + (1 - \delta^{\omega'}) \lambda_1^{\omega'} \int_{V^*(\theta, \omega')}^{+\infty} v dF(v | \omega') \right\rangle H(d\omega' | \omega). \quad (31)
\end{aligned}$$

A symmetric expression can be arrived at in the same way for $S_{L,\ell}(\theta, L^*(\theta)^-, \omega)$, so that defining $\mathcal{D}_{S_L}(\theta, L, \omega) := S_{L,r}(\theta, L^+, \omega) - S_{L,\ell}(\theta, L^-, \omega)$, which is positive by convexity of S in L , we have:

$$\begin{aligned}
0 \leq \mathcal{D}_{S_L}(\theta, L^*(\theta), \omega) &= \beta \int_{\Omega} \mathcal{D}_{S_L}[\theta, \mathcal{L}(L, \omega', V^*(\theta, \omega')), \omega'] \cdot \mathcal{L}_L(L, \omega', V^*(\theta, \omega')) H(d\omega' | \omega) \\
&< \beta \int_{\Omega} \mathcal{D}_{S_L}[\theta, \mathcal{L}(L, \omega', V^*(\theta, \omega')), \omega'] H(d\omega' | \omega).
\end{aligned}$$

At this point, if we can prove that \mathcal{D}_{S_L} is uniformly bounded above by some $K > 0$, then iterating the last inequality will show that $0 \leq \mathcal{D}_{S_L}(\theta, L^*(\theta), \omega) < \beta^n K$ for all $n \in \mathbb{N}$, which implies that $\mathcal{D}_{S_L}(\theta, L^*(\theta), \omega) = 0$ for all (θ, ω) and that S_L exists everywhere. Since S is convex, S_L is increasing, hence it can only have jumps up. But we just concluded that its right and left limit are equal everywhere, so S_L is continuous for all $L \in [0, \bar{L}]$, thus proving point 1 of the proposition.

We still need to show that \mathcal{D}_{S_L} is uniformly bounded above. Because $S_{L,\ell} \geq 0$, it suffices to show that $S_{L,r}$ is bounded above. The following series of inequalities use the facts that $S_{L,r}$ is increasing in L (by convexity of S), that $L \leq 1$ (since the total mass of workers in the economy is 1), and that $S_{L,r}(\theta, L^*(\theta), \omega) \geq S_{L,\ell}(\theta, L^*(\theta), \omega) \geq V^*(\theta, \omega) \geq U(\omega)$ again invoking convexity in conjunction with the FOC (27):

$$\begin{aligned}
S_{L,r}(\theta, L^+, \omega) &\leq S_{L,r}(\theta, 1, \omega) \\
&\leq \omega\theta + \beta \int_{\Omega} \delta^{\omega'} U(\omega') H(d\omega' | \omega) + \beta \int_{\Omega} \left\langle S_{L,r}(\theta, 1, \omega') \cdot (1 - \delta^{\omega'}) (1 - \lambda_1^{\omega'} \bar{F}(V^*(\theta, \omega'))) \right. \\
&\quad \left. + (1 - \delta^{\omega'}) \lambda_1^{\omega'} \int_{V^*(\theta, \omega')}^{+\infty} v dF(v | \omega') \right\rangle H(d\omega' | \omega) \\
&\leq \omega\theta + \beta \int_{\Omega} \delta^{\omega'} U(\omega') H(d\omega' | \omega) + \beta \int_{\Omega} (1 - \delta^{\omega'}) S_{L,r}(\theta, 1, \omega') H(d\omega' | \omega) \\
&\leq \omega\theta + \beta \int_{\Omega} S_{L,r}(\theta, 1, \omega') H(d\omega' | \omega).
\end{aligned}$$

This establishes that $S_{L,r}(\theta, 1, \omega) \leq \max_{\Omega} \omega\theta / (1 - \beta)$ for all ω , and so $\max_{\Omega} \omega\theta / (1 - \beta)$ is also a uniform upper bound for $S_{L,r}(\theta, L^*(\theta), \omega)$. This completes the proof of point 1 in the proposition.

We now go straight to point 3 before proving point 2. Consider the problem of a firm choosing W to best-respond to all other firms playing a RPE. By a simple improvement argument, $W \in [V^*(\underline{\theta}, \omega'), V^*(\bar{\theta}, \omega')]$. Since V^* is continuous and increasing, then offering any such best response W is equivalent to choosing a type τ to imitate such that $W = V^*(\tau, \omega')$. In any RPE, by Proposition 2, the best response by a firm θ of current size $L^*(\theta)$ is ‘truthful revelation’, $\tau^* = \theta$, which solves

$$S(\theta, L^*(\theta), \omega) = A(\theta, L^*(\theta), \omega) + \beta \int_{\Omega} \max_{\tau(\omega')} \langle S[\theta, \mathcal{L}(L^*(\theta), \omega', \tau(\omega')), \omega'] + B(L^*(\theta), \omega', \tau(\omega')) \rangle H(d\omega' | \omega)$$

where, with a slight abuse of notation:

$$\begin{aligned} \mathcal{L}(L, \omega', \tau) &= L \left(1 - \delta^{\omega'}\right) \left(1 - \lambda_1^{\omega'} \bar{F}(V^*(\tau, \omega') | \omega')\right) \\ &\quad + \lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \Lambda(\bar{\theta}) G(V^*(\tau, \omega') | \omega') \end{aligned}$$

and

$$\begin{aligned} B(L, \omega', \tau) &= L \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V^*(\tau, \omega')}^{+\infty} v dF(v | \omega') \\ &\quad - V^*(\tau, \omega') \left(\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \Lambda(\bar{\theta}) G(V^*(\tau, \omega') | \omega')\right) \end{aligned}$$

are continuous functions of L and τ . Using the RP property

$$\mathcal{L}(L^*(\theta), \omega', \tau) = L^*(\theta) \left(1 - \delta^{\omega'}\right) \left(1 - \lambda_1^{\omega'} \bar{F}(\tau)\right) + \lambda_0^{\omega'} (1 - \Lambda^*(\bar{\theta})) + \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \Lambda^*(\tau) \quad (32)$$

$$\begin{aligned} B(L^*(\theta), \omega', \tau) &= L^*(\theta) \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{\tau}^{\bar{\theta}} V^*(x, \omega') d\Gamma(x) \\ &\quad - V^*(\tau, \omega') \left(\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \Lambda^*(\tau)\right). \end{aligned}$$

Lemma 6 V^* is Lipschitz continuous, hence absolutely continuous and $V^*(\theta, \omega') = \int^{\theta} V^{*\prime}(x, \omega') dx$.

Proof. Fix θ and ω' . Optimality requires for all $h > 0$:

$$\begin{aligned}
& S[\theta, \mathcal{L}(L^*(\theta), \omega', \theta - h, \omega')] + L^*(\theta) \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{\theta-h}^{\bar{\theta}} V^*(x, \omega') d\Gamma(x) \\
& \quad - V^*(\theta - h, \omega') \left(\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda^*(\theta - h)\right) \\
& \leq S[\theta, \mathcal{L}(L^*(\theta), \omega', \theta), \omega'] + L^*(\theta) \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{\theta}^{\bar{\theta}} V^*(x, \omega') d\Gamma(x) \\
& \quad - V^*(\theta, \omega') \left(\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda^*(\theta)\right).
\end{aligned}$$

Rearranging:

$$\begin{aligned}
& [V^*(\theta, \omega') - V^*(\theta - h, \omega')] \cdot \left[\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda^*(\theta)\right] \\
& \leq S[\theta, \mathcal{L}(L^*(\theta), \omega', \theta), \omega'] - S[\theta, \mathcal{L}(L^*(\theta), \omega', \theta - h), \omega'] \\
& \quad - \lambda_1^{\omega'} (1 - \delta^{\omega'}) \left\{ L^*(\theta) \int_{\theta-h}^{\theta} V^*(x, \omega') d\Gamma(x) + V^*(\theta - h, \omega') \cdot [\Lambda^*(\theta) - \Lambda^*(\theta - h)] \right\} \\
& \leq S[\theta, \mathcal{L}(L^*(\theta), \omega', \theta), \omega'] - S[\theta, \mathcal{L}(L^*(\theta), \omega', \theta - h), \omega'] \\
& \quad - V^*(\theta - h, \omega') \cdot \lambda_1^{\omega'} (1 - \delta^{\omega'}) \{L^*(\theta) [\Gamma(\theta) - \Gamma(\theta - h)] + [\Lambda^*(\theta) - \Lambda^*(\theta - h)]\}, \}
\end{aligned} \tag{33}$$

where the second inequality is obtained by minorizing the integral term, remarking that V^* is increasing in θ . Now differentiability of S w.r.t. L and the definition (32) together imply that:¹⁷

$$\begin{aligned}
& S[\theta, \mathcal{L}(L^*(\theta), \theta, \omega'), \omega'] - S[\theta, \mathcal{L}(L^*(\theta), \theta - h, \omega'), \omega'] = S_L[\theta, \mathcal{L}(L^*(\theta), \theta, \omega'), \omega'] \\
& \quad \times \lambda_1^{\omega'} (1 - \delta^{\omega'}) \{L^*(\theta) [\Gamma(\theta) - \Gamma(\theta - h)] + [\Lambda^*(\theta) - \Lambda^*(\theta - h)]\} + o(h).
\end{aligned}$$

Substituting into (33), we obtain:

$$\begin{aligned}
& [V^*(\theta, \omega') - V^*(\theta - h, \omega')] \cdot \left[\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda^*(\theta)\right] \\
& \leq \{S_L[\theta, \mathcal{L}(L^*(\theta), \omega', \theta), \omega'] - V^*(\theta - h, \omega')\} \\
& \quad \times \lambda_1^{\omega'} (1 - \delta^{\omega'}) \{L^*(\theta) [\Gamma(\theta) - \Gamma(\theta - h)] + [\Lambda^*(\theta) - \Lambda^*(\theta - h)]\} + o(h)
\end{aligned}$$

Now dividing through by $h > 0$ and taking the limit superior:

$$0 \leq D^- V^*(\theta) \leq 2\lambda_1^{\omega'} (1 - \delta^{\omega'}) L^*(\theta) \gamma(\theta) \frac{S_L[\theta, \mathcal{L}(L^*(\theta), \omega', \theta), \omega'] - V^*(\theta, \omega')}{\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda^*(\theta)}$$

¹⁷The $o(h)$ term at the end comes from $o(\mathcal{L}(L^*(\theta), \omega', \theta) - \mathcal{L}(L^*(\theta), \omega', \theta - h)) = o(L^*(\theta) [\Gamma(\theta) - \Gamma(\theta - h)] + [\Lambda^*(\theta) - \Lambda^*(\theta - h)]) = o(h)$ by differentiability of Γ and Λ^* .

(the first inequality is a direct consequence of V^* being increasing). All terms in the r.h.s. are uniformly bounded above ($\gamma(\theta)$ by assumption and S_L by a property established earlier in this proof). So V^* is a continuous function with bounded (upper-left) Dini derivative, which is sufficient to ensure Lipschitz-continuity (see, e.g., Problem 20.c p112 in Royden, 1988). \square

We are finally in a position to prove claim 2 of the proposition, namely that in any RPE, $V^*(\theta)$ is continuously differentiable. Because V^* is increasing, we know that V_θ^* exists outside of a null set, say N_V . Therefore, for each $\theta \in [\underline{\theta}, \bar{\theta}] \setminus N_V$ we can take a derivative in the Bellman equation and write a NFOC:

$$V_\theta^*(\theta, \omega') = 2\lambda_1^{\omega'} (1 - \delta^{\omega'}) L^*(\theta) \gamma(\theta) \frac{S_L[\theta, \mathcal{L}(L^*(\theta), \omega', \theta), \omega'] - V^*(\theta, \omega')}{\lambda_0^{\omega'} (1 - \Lambda(\bar{\theta})) + \lambda_1^{\omega'} (1 - \delta^{\omega'}) \Lambda^*(\theta)} := q(\theta).$$

Notice that the RHS $q(\theta)$ is continuous in θ , where it exists, i.e. in the set $[\underline{\theta}, \bar{\theta}] \setminus N_V$ (recall that L^* is continuous by the assumption that L_0 is). Since $[\underline{\theta}, \bar{\theta}] \setminus N_V$ is the complement of a set of measure 0, it is dense in $[\underline{\theta}, \bar{\theta}]$. Therefore, for all $\theta \in N_V$, there exists a sequence $\{\theta_n\}$, $\theta_n \in [\underline{\theta}, \bar{\theta}] \setminus N_V$ such that $\theta_n \rightarrow \theta$. As $V_\theta^*(\theta_n, \omega')$ exists and equals $q(\theta_n)$ for all θ_n in this sequence, using the NFOC and continuity of q : $\lim_{n \rightarrow \infty} V_\theta^*(\theta_n, \omega') = \lim_{n \rightarrow \infty} q(\theta_n) = q(\theta)$. Let

$$\tilde{V}_\theta(\theta) := \begin{cases} V_\theta^*(\theta, \omega') & \theta \in [\underline{\theta}, \bar{\theta}] \setminus N_V \\ q(\theta) & \text{otherwise} \end{cases}$$

which, by the last argument, is continuous everywhere. Then,

$$V^*(\theta, \omega') = V^*(\underline{\theta}, \omega') + \int_{\underline{\theta}}^{\theta} V_\theta^*(x, \omega') dx = V^*(\underline{\theta}, \omega') + \int_{\underline{\theta}}^{\theta} \tilde{V}_\theta(x) dx$$

where the second equality follows from the fact that $V_\theta^*(\theta, \omega') \neq \tilde{V}_\theta(\theta)$ only on a null set. So $V^*(\theta, \omega')$ is the integral of a continuous function \tilde{V}_θ , hence it is differentiable with $V_\theta^*(\theta, \omega') = \tilde{V}_\theta(\theta)$ everywhere, and the FOC $V_\theta^*(\theta, \omega') = q(\theta)$ holds everywhere. Point 3 of the proposition is thus proven.

Finally on to point 2. Introducing a time index τ and using the notation

$$\begin{aligned} \Lambda_{\tau+1} &= \mathcal{T}(\omega_{\tau+1}, \Lambda_\tau) \\ \Delta(\theta, \omega_\tau) &= (1 - \delta^{\omega_\tau}) (1 - \lambda_1^{\omega_\tau} \bar{\Gamma}(\theta)) \\ \mathcal{U}(\theta, \omega_\tau, \Lambda_\tau) &= \mathbf{E}_{\omega_{\tau+1} | \omega_\tau} \left[\delta^{\omega_{\tau+1}} U(\omega_{\tau+1}, \Lambda_{\tau+1}) \right. \\ &\quad \left. + (1 - \delta^{\omega_{\tau+1}}) \lambda_1^{\omega_{\tau+1}} \int_{\theta}^{+\infty} V^*(x | \omega_{\tau+1}, \Lambda_{\tau+1}) d\Gamma(x) \right], \\ \text{and: } \mu(\theta, \omega_\tau, \Lambda_\tau) &= S_L(\theta, L_\tau^*(\theta), \omega_\tau, \Lambda_\tau), \end{aligned}$$

we can rewrite the Euler equation (31) as follows:

$$\mu(\theta, \omega_t, \Lambda_t) = \omega_t \theta + \beta \mathcal{U}(\theta, \omega_t, \Lambda_t) + \beta \mathbf{E}_{\omega_{t+1}|\omega_t} [\Delta(\theta, \omega_{t+1}) \mu(\theta, \omega_{t+1}, \Lambda_{t+1})]. \quad (34)$$

For any measurable function ϕ of ω and any t , define recursively the linear operator $X_0^t[\phi] = \phi$ and

$$X_s^t[\phi] = \mathbf{E}_{\omega_{t+s}|\omega_{t+s-1}} [\Delta(\theta, \omega_{t+s}) X_{s-1}^t[\phi]] \text{ for } s = 1, 2 \dots n-1.$$

After n forward substitutions, we can write (34) as

$$\mu(\theta, \omega_t, \Lambda_t) = \mu_n(\theta, \omega_t, \Lambda_t) + \beta^{n+1} X_{n+1}^t[\mu]$$

where

$$\mu_n(\theta, \omega_t, \Lambda_t) = \sum_{s=0}^n \beta^s \{ X_{s+1}^t[\omega\theta] + \beta X_{s+1}^t[\mathcal{U}(\theta, \omega, \Lambda)] \}.$$

Since $\mu > 0$ and $\Delta \in (0, 1)$ with probability 1

$$\begin{aligned} 0 < \mu(\theta, \omega_t, \Lambda_t) - \mu_n(\theta, \omega_t, \Lambda_t) &= |\mu(\theta, \omega_t, \Lambda_t) - \mu_n(\theta, \omega_t, \Lambda_t)| = \beta^{n+1} X_{n+1}^t[\mu] \\ &\leq \beta^{n+1} \mathbf{E}_{\omega_{t+1}|\omega_t} [\mathbf{E}_{\omega_{t+2}|\omega_{t+1}} [\dots \mathbf{E}_{\omega_{t+n}|\omega_{t+n-1}} [\mu(\theta, \omega_{t+n}, \Lambda_{t+n})]]] \\ &= \beta^{n+1} \mathbf{E}_{\omega_{t+n}|\omega_t} [\mu(\theta, \omega_{t+n}, \Lambda_{t+n})] \end{aligned}$$

Since a firm can always guarantee itself positive profits and employment by offering its workers the value of unemployment, then $L^*(\theta, \omega, \Lambda)$ is bounded away from 0 with probability one. So the TVC (22) implies that, as $n \rightarrow \infty$, the last term vanishes, thus μ_n converges pointwise (and indeed uniformly) to μ .

Next, taking derivatives

$$\begin{aligned} \frac{\partial \mu_n(\theta, \omega_t, \Lambda_t)}{\partial \theta} &= \omega_t + \frac{\lambda_1^\omega \gamma(\theta)}{1 - \lambda_1^\omega \bar{\Gamma}(\theta)} + \sum_{s=1}^n \beta^s X_s^t \left[\omega + \frac{\lambda_1^\omega \gamma(\theta)}{1 - \lambda_1^\omega \bar{\Gamma}(\theta)} \omega \theta \right] \\ &\quad + \sum_{s=2}^{n+1} \beta^s X_s^t \left[\frac{\lambda_1^\omega \gamma(\theta)}{1 - \lambda_1^\omega \bar{\Gamma}(\theta)} \mathcal{U}(\theta, \omega, \Lambda) - (1 - \delta^\omega) \lambda_1^\omega V^*(\theta | \omega, \Lambda) \gamma(\theta) \right] \end{aligned}$$

which is continuous in θ . As the arguments of the operator X_s^t in the last expression are continuous in ω, θ on the compact set $\Omega \times [\underline{\theta}, \bar{\theta}]$, the $X_s^t[\cdot]$ terms in the sums are bounded above and below uniformly with probability 1 by some upper bound $X' < \infty$ and lower bound $-X'$ for all $(\theta, \omega_t, \Lambda_t)$. Therefore, driven by discounting, the two sums converge as $n \rightarrow \infty$, and

$$\lim_{n \rightarrow \infty} \frac{\partial \mu_n(\theta, \omega_t, \Lambda_t)}{\partial \theta} = \frac{\partial \mu_\infty(\theta, \omega_t, \Lambda_t)}{\partial \theta}$$

exists for all $\theta, \omega_t, \Lambda_t$. Also

$$\begin{aligned} \frac{\partial \mu_n(\theta, \omega_t, \Lambda_t)}{\partial \theta} - \frac{\partial \mu_\infty(\theta, \omega_t, \Lambda_t)}{\partial \theta} &= - \sum_{s=n+1}^{\infty} \beta^s X_s^t \left[\omega + \frac{\lambda_1^\omega \gamma(\theta)}{1 - \lambda_1^\omega \bar{\Gamma}(\theta)} \omega \theta \right] \\ &\quad - \sum_{s=n+2}^{\infty} \beta^s X_s^t \left[\frac{\lambda_1^\omega \gamma(\theta)}{1 - \lambda_1^\omega \bar{\Gamma}(\theta)} \mathcal{U}(\theta, \omega, \Lambda) - (1 - \delta^\omega) \lambda_1^\omega V^*(\theta | \omega, \Lambda) \gamma(\theta) \right] \end{aligned}$$

So, for all $\theta, \omega_t, \Lambda_t$, $|X_s^t[\cdot]| < X'$ implies

$$\left| \frac{\partial \mu_n(\theta, \omega_t, \Lambda_t)}{\partial \theta} - \frac{\partial \mu_\infty(\theta, \omega_t, \Lambda_t)}{\partial \theta} \right| < X' \frac{\beta^{n+1} + \beta^{n+2}}{1 - \beta}$$

so that convergence of derivatives is uniform. By Theorem 7.17 in Rudin (1976), conclude that μ is continuously differentiable in θ with:

$$\frac{\partial \mu(\theta, \omega_t, \Lambda_t)}{\partial \theta} = \frac{\partial \mu_\infty(\theta, \omega_t, \Lambda_t)}{\partial \theta},$$

which completes the proof of the proposition. \square

C Proof of Proposition 4

We first show that the limit in (24) exists, is positive and uniformly bounded above. By assumption, $\omega\theta$ and b^ω are positive and uniformly bounded above by some $K < \infty$, therefore by the definition of \mathbf{T} , $\mathbf{T}(\frac{\omega\theta}{b^\omega})$ is also positive and uniformly bounded above by K , and by induction the same is true of $\mathbf{T}^j(\frac{\omega\theta}{b^\omega})$. Hence the sequence $\sum_{j=0}^{n-1} \beta^j \mathbf{T}^j(\frac{\omega\theta}{b^\omega})$ is increasing and uniformly bounded above by $K/(1 - \beta)$, so each of the two sums in this sequence must converge and the limit exists and is positive and bounded above by $K/(1 - \beta)$.

We next show that, if there exists a RPE, then it is given by (24). Suppose there exists a RPE $(\frac{\mu}{U})$. Then by definition of RPE $\mu \geq \mathbf{T}_V[\mu, U] \geq U \geq 0$ and $\mathbf{T}_V[\mu, U]$ is increasing in θ . Then, by inspection of \mathbf{T}_μ : $\mu_t(\bar{\theta} | \omega) = \omega \bar{\theta} + \beta \mathbf{T}_\mu(\frac{\mu}{U})(\bar{\theta} | \omega) \leq \omega \bar{\theta} + \beta \mathbf{E}_{\omega_{t+1} | \omega_t}[\mu_{t+1}(\bar{\theta} | \omega_{t+1})]$. Since $\mathbf{E}_{\omega_{t+1} | \omega_t}[\mu_{t+1}] \geq 0$, iterating \mathbf{T} forward and using the TVC:

$$0 \leq \beta^n \mathbf{T}_\mu^n[\mu, U](\bar{\theta} | \omega) \leq \beta^n \mathbf{E}_{\omega_{t+n} | \omega_t}[\mu_{t+n}(\bar{\theta} | \omega_{t+n})] \rightarrow 0,$$

so that $\mu(\bar{\theta} | \omega) \leq \omega \bar{\theta} / (1 - \beta) \leq K / (1 - \beta)$. By Proposition 3, in any RPE $\partial \mu / \partial \theta$ exists; the proof of Proposition 2 shows that S has increasing differences in (θ, L) , so that $\mu = S_L$ is increasing in θ . So for all θ , $0 \leq U \leq \mu(\theta | \omega) \leq K / (1 - \beta)$ and $(\frac{\mu}{U})$ is uniformly bounded above. By definition of a RPE, $(\frac{\mu}{U})$ must solve (23). Substituting forward in (23) and using $U \leq \mu(\theta | \omega) \leq K / (1 - \beta)$ we find $0 \leq \beta^n \mathbf{T}^n(\frac{\mu}{U}) \leq \beta^n K / (1 - \beta) \rightarrow 0$, so that $(\frac{\mu}{U}) = (\frac{\omega\theta}{b^\omega}) + \beta \mathbf{T}(\frac{\omega\theta}{b^\omega}) + \beta^2 \mathbf{T}^2(\frac{\mu}{U}) = \dots = (\frac{\mu^*}{U^*})$ as claimed.

We finally turn to existence. By construction, $(\frac{\mu^*}{U^*})$ solves (23). Moreover, since $(\frac{\mu^*}{U^*})$ is uniformly bounded above, it satisfies the TVC. So we only have to show that it satisfies $0 \leq U^* \leq \mathbf{T}_V[\mu^*, U^*]$ and $\mathbf{T}_V[\mu^*, U^*]$ increasing in θ . By definition of \mathbf{T}_V :

$$\begin{aligned} \frac{\partial \mathbf{T}_V[\mu^*, U^*]}{\partial \theta}(\theta | \omega) &= \frac{\frac{\partial Q_t}{\partial \theta}(\theta | \omega)}{Q_t(\theta | \omega)} (\mu^*(\theta | \omega) - \mathbf{T}_V[\mu^*, U^*](\theta | \omega)) \\ &= \frac{\frac{\partial Q_t}{\partial \theta}(\theta | \omega)}{Q_t(\theta | \omega)} \left(\mu^*(\theta | \omega) - U^*(\omega) - \int_{\underline{\theta}}^{\theta} (\mu^*(x | \omega) - U^*(\omega)) \frac{\frac{\partial Q_t}{\partial \theta}(x | \omega)}{Q_t(\theta | \omega)} dx \right) \end{aligned}$$

so it suffices to prove that $\mu^*(\underline{\theta} | \omega) \geq U(\omega)$ and μ^* is increasing in θ . Now consider the sequence of functions $\{\mu_n, U_n\}_{n \in \mathbb{N}}$ defined by $\mu_0(\theta | \omega) = \omega\theta$, $U_0(\omega) = b^\omega$, and $(\frac{\mu_{n+1}}{U_{n+1}}) = (\frac{\omega\theta}{b^\omega}) + \beta \mathbf{T}(\frac{\mu_n}{U_n})$. Clearly this sequence converges to $(\frac{\mu^*}{U^*})$. Suppose μ_n is increasing in θ and greater than U_n for some n . It is straightforward to see from the definition of \mathbf{T}_μ that μ_{n+1} is increasing in θ . Then:

$$\mu_{n+1}(\theta | \omega) - U_{n+1}(\omega) = (\omega\theta - b^\omega) + (\mathbf{T}_\mu[\mu_n, U_n](\theta | \omega) - \mathbf{T}_U[\mu_n, U_n](\omega)).$$

The condition $\omega\underline{\theta} \geq b^\omega$ in all states ensures that the first term is positive, while the conditions $(1 - \delta^\omega) \lambda_1^\omega \geq \lambda_0^\omega$ guarantees that the second terms is also positive. \square