Non-Stationary Search Equilibrium*

Giuseppe Moscarini†
Yale University
and
NBER

Fabien Postel-Vinay‡
University of Bristol
and
Paris School of Economics

November 2009

Abstract

We study a stochastic economy where both employed and unemployed workers search randomly for labor contracts posted by ex ante heterogeneous firms, while aggregate productivity is subject to shocks. A firm can commit to a (Markov) contract, which specifies a wage contingent on all payoff-relevant states, but must pay equally all of its workers, who have limited commitment and are free to quit at any time. Our exercise provides the first dynamic stochastic general equilibrium analysis of a popular class of search wage-posting models, drawing in part from the literature on recursive contracts under moral hazard. An equilibrium of the contract-posting game is Rank-Preserving [RP] if larger firms offer a larger value to their workers in all states of the world. We find two sufficient (but not necessary) conditions for every equilibrium to be RP: either firms only differ in their initial size, or they also differ in their fixed idiosyncratic productivity but more productive firms are also initially weakly larger. In the latter case, turnover is always efficient, as workers always move from less to more productive firms. In both cases, the ranking of firm sizes never changes on the RP equilibrium path, a property that has two useful implications. First, the stochastic dynamics of firm size provide an intuitive and robust explanation for the empirical finding that large employers in the US are more cyclically sensitive (Moscarini and Postel-Vinay, 2009). Second, RP equilibrium computation is fairly tractable, and we construct and simulate calibrated examples.

*This paper fully develops the theory that we first sketched in our companion paper "The Timing of Labor Market Expansions: New Facts and a New Hypothesis". We acknowledge useful comments to that paper and to the June 2008 draft of this paper from seminar and conference audiences at numerous venues. 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

We study the aggregate equilibrium dynamics of a frictional labor market where workers search randomly on and off the job for employment contracts posted by firms. Our exercise provides the first analysis of aggregate dynamics of a popular class of search wage-posting models, originating with Burdett and Mortensen (1998, henceforth BM). They provide the first successful formalization of the hypothesis that cross-sectional wage dispersion is largely a consequence of labor market frictions. In so doing the BM model has started what has now established itself as the most promising line of research in the analysis of wage inequality, as the vibrant and empirically very successful literature organized around that hypothesis continues to show.

That literature, however, is invariably cast in deterministic steady state. Ever since the inception 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 a continuum of firms of strategies that are all best responses to one another.

We find a way around this problem by considering a class of equilibria satisfying what we call the *Rank-Preserving property*, i.e. equilibria in which the workers' ranking of firms is time-invariant. We show that this class of equilibria is generic if all firms are equally productive and the environment is not subject to aggregate stochastic shocks. We further show that the same property holds in equilibrium when firms have heterogeneous productivity, where more productive firms offer a larger value and employ more workers at all points in time, if (but not only if) they have more employees to begin with. We view the fact that the workers' ranking of firms also reflects the hierarchy of productivity in a Rank-Preserving Equilibrium in the presence of productive heterogeneity across firms as a very appealing property of the model. It parallels a similar property of BM's static equilibrium, and in ensures constrained-efficient labor reallocation at all dates. Finally, we investigate existence of Rank-Preserving Equilibria in a stochastic environment.

Besides being of 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. Unlike the typical representative-agent model, the BM model makes predictions about the business cycle behavior of wage distributions, firm size distributions, or patterns of labor reallocation across firms. In Moscarini and Postel-Vinay (2009, MPV09), we quantitatively gauge those predictions against facts documented using various (often new) data sets. Moreover, by advocating the BM model as a potentially useful tool for the study of aggregate labor market dynamics, we hope to contribute to a synthesis between the BM 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.[1]

The rest of the paper is organized as follows. In Section 2 we lay out the basic environment. In Section 3 we describe and formally define an equilibrium of our model economy. We then introduce the notion of Rank Preserving Equilibria in Section 4, where we also state our main result about the generality of RPE and give a characterization of RPE. Then Section 5 shows how to practically simulate equilibrium paths for wages and employment in a RPE. Finally, Section 6 concludes.

2 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 with heterogeneous firm types. We present our model in discrete time, as it affords more clarity in the presentation of the contract posting problem under one-sided commitment as a recursive problem, following the seminal insights of Spear and Srivastava (1987).

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[2] Workers and firms are risk neutral, infinitely lived, and maximize payoffs discounted with factor $\beta \in (0, 1)$. Firms operate constant-return technologies with heterogeneous productivity levels $\omega\theta$, where ω is an aggregate component, evolving within some bounded set of values $\Omega \subset \mathbb{R}_+$ according to a discrete-time stationary

[1]Rudanko (2008) and Menzio and Shi (2008) formulate and solve wage contract-posting models with aggregate productivity shocks, where job search is directed, rather than random in the spirit of BM. This 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 idea and emergence of Rank-Preserving Equilibrium, while maintaining BM's assumption of random search common to the vast majority of the search literature. From a theoretical viewpoint, we see both programs as fruitful directions of exploration. From a quantitative viewpoint, the directed search approach is focused on the response of the job-finding rate to aggregate shocks. This approach does not generate a well-defined notion of employer size. Hence, it is silent on the wealth of new evidence on cyclical patterns of the size distribution and employer size/growth relationship that we offer in MPV09, and that we envision as central to our understanding of the propagation of aggregate shocks in labor markets.

[2]We 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 our companion paper Moscarini and Postel-Vinay (2008, MPV08). Finally, that the mass of firms and workers both have measure one is obviously innocuous and only there to simplify the notation.

first-order Markov process $H(d\omega' \mid \omega)$, and θ is a fixed, idiosyncratic heterogeneity component, distributed across firms $\theta \sim \Gamma(\cdot)$ with continuous density $\gamma = \Gamma'$ over $\left[\underline{\theta}, \overline{\theta}\right]$.

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 $\Gamma(\cdot)$.[3]. 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 $\delta^{\omega'}$. Note that all these transition probabilities, although exogenous, are allowed to depend on the aggregate state ω. Workers attach a common lifetime value of U to being unemployed.

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: a firm of type θ produces output and pays wages in state ω;

2. the flow benefit b^ω accrues to unemployed workers;

3. the new state ω' of aggregate labor productivity is realized;

4. employed workers can quit to unemployment;

5. jobs are destroyed exogenously with chance $\delta^{\omega'}$;

6. 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;

7. each previously unemployed worker receives an offer with probability $\lambda_0^{\omega'}$.

To close the model, we need to specify how are wages set. 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

[3]Later, we extend the model to allow for non-uniform sampling, in that different types-θ firms have different chances of being sampled by job searchers. This extension is theoretically straightforward and useful in quantitative applications

the set of feasible wage contracts to a non-steady-state environment.[4] Under commitment, such a wage function implies a value V for any worker to work for that firm, which is 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.

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 $\left[\underline{\theta}, \overline{\theta}\right]$, 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 $L_0(\theta) = \Lambda_0'(\theta) / \gamma(\theta)$, is a continuous function of θ.

3 Equilibrium

3.1 Definition

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

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

$$F(W \mid \zeta, V) := \int_{\underline{\theta}}^{\overline{\theta}} \mathbb{I}\{V(\theta, \zeta) \leq W\} \, d\Gamma(\theta) \tag{1}$$

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 $\overline{F} = 1 - F$ denote the survival function.

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

$$G(W \mid \zeta, \Lambda, V) := \frac{1}{\Lambda\left(\overline{\theta}\right)} \cdot \int_{\underline{\theta}}^{\overline{\theta}} \mathbb{I}\{V(\theta, \zeta) \leq W\} \, d\Lambda(\theta) \tag{2}$$

[4] We thus rule out, among other things, wage-tenure contracts (Stevens, 2004; Burdett and Coles, 2003), offer-matching or individual bargaining (Postel-Vinay and Robin, 2002; Dey and Flinn, 2005; Cahuc, Postel-Vinay and Robin, 2006), contracts conditioned on employment status (Carrillo-Tudela, 2009). 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. Shimer (2008) proposes an alternative formulation, which maintains BM's restriction of a constant posted wage, even out of steady state, and delivers a few of the same results.

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

Let primes denote next period values. Given a strategy $V \in \mathscr{V}_Z$ followed by all firms and the resulting F and Λ, an unemployed earns a value solving:

$$U(\zeta \mid V) = b^{\omega} + \beta \mathbf{E}_{\zeta' \mid \zeta} \left[\left(1 - \lambda_0^{\omega'}\right) U(\zeta' \mid V) + \lambda_0^{\omega'} \int \max \langle v, U(\zeta' \mid V) \rangle \, dF(v \mid \zeta', V) \right], \quad (3)$$

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.

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

$$\begin{aligned} L' = \mathscr{L}(\zeta, W \mid V) := {} & L\left(1 - \delta^{\omega}\right)\left(1 - \lambda_1^{\omega} \overline{F}(W \mid \zeta, V)\right) \\ & + \lambda_0^{\omega}\left[1 - \Lambda\left(\overline{\theta}\right)\right] \mathbb{I}\{W \geq U(\zeta \mid V)\} + \lambda_1^{\omega}\left(1 - \delta^{\omega}\right) \Lambda\left(\overline{\theta}\right) G(W \mid \zeta, V). \quad (4) \end{aligned}$$

After the new aggregate state ω' is drawn, of the measure L of workers currently employed by this firm, a fraction $\left(1 - \delta^{\omega'}\right)$ are not separated exogenously into unemployment. Of these survivors, a fraction $\lambda_1^{\omega'} \overline{F}(W \mid \zeta, V)$ quit because they draw from F an outside offer which gives them a value larger than W. The currently unemployed $1 - \Lambda\left(\overline{\theta}\right)$ find jobs with chance $\lambda_0^{\omega'}$, and accept an offer from a firm offering W 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 $\left(1 - \delta^{\omega'}\right) \Lambda\left(\overline{\theta}\right)$ 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 $G(W \mid \zeta, V)$), in which case they quit to this firm offering W.

A consistency condition requires the cumulated firm size to evolve as the sum of individual firm sizes on the equilibrium path. For any $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$:

$$\Lambda(\theta \mid \zeta, V)' = \int_{\underline{\theta}}^{\theta} \mathscr{L}(\zeta, V(x, \zeta) \mid V) \, d\Gamma(x) := \mathscr{T}(\zeta \mid V). \quad (5)$$

The support of Λ is contained in that of Γ, because there cannot be a firm θ with employees if there exists no such firm of type θ. By induction, starting from the initial distribution of employment and for every history of the game, Λ has a (possibly nil) Radon-Nikodym derivative $d\Lambda(\theta \mid \zeta, V) / d\Gamma(\theta)$ everywhere in θ, and the consistency condition requires this derivative to be $\mathscr{L}(\zeta, V(\theta, \zeta) \mid V)$. Therefore, F and G also exist at all nodes of the game when firms play the strategy V.

A value strategy $W \in \mathscr{V}_Z$ can also be represented by a wage strategy $w \in \mathscr{V}_Z$ such that the worker's Bellman equation is solved by W, in which case we say that w "implements"

W given that all other firms play V:

$$W(\theta,\zeta) = w(\theta,\zeta) + \beta \mathbf{E}_{\zeta'|\zeta}\Bigg[\delta^{\omega'} U(\zeta' \mid V) + \left(1-\delta^{\omega'}\right)\left(W(\theta,\zeta') + \lambda_1^{\omega'} \int_{W(\theta,\zeta')}^{+\infty} \left[v - W(\theta,\zeta')\right] dF(v \mid \zeta', V)\right)\Bigg] \quad (6)$$

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 follow a behavioral strategy V (a value policy) that must be a best-response against other firms at any node ζ of the game, including those reached with probability zero on the equilibrium path. For example, a firm may find itself losing more workers than predicted by current equilibrium play. This requires specifying a consistent belief assessment. In this sense, firms also act sequentially, optimizing at any game node, but subject to the constraint of delivering the value to the workers once hired. This constraint 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 $W = U$. The reputational underpinnings of the firm's commitment power have been widely explored in the wage-posting literature. As is standard, behavioral strategies in the extensive form dynamic game generate strategy profiles of the equivalent static, strategic form game where each firm chooses a map once and for all at time 0.

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. For the same reason, and because it is small, a firm takes its competitors' behavior (the distributions F and G) as given when choosing a strategy, because its own deviations cannot be detected and be subject to retaliation, so its actions cannot affect the distribution of offers in the economy. Each firm can only observe the public histories of $\{\omega, F, G, L, \Lambda, U\}$, which form a set Q. Hence $Z \subseteq Q$.

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 $F(\cdot \mid \zeta, V)$ is uniquely determined from (1), so it contains no independent information about ζ. Similarly, given ζ, $V(\cdot,\zeta)$ and $F(\cdot \mid \zeta, V)$, the value of unemployment $U(\zeta \mid V)$ is uniquely determined recursively forward by (3), so it contains no independent information about the past history ζ. The same is true, from (2), of $G(\cdot \mid \zeta, V)$, 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, 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 and covertly controlled by the firm and has a direct impact on the firm's continuation payoffs. Because the history of own size $\{L_s\}_{s=1}^{t}$ 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 $\zeta_t = \{\omega_1, \cdots, \omega_t, L_t\}$. 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, for example as a public randomization device, although they are no longer payoff-relevant given the Markov evolution of ω.

Definition 1 *A* SEQUENTIAL EQUILIBRIUM *of the contract posting game is a measurable function $V \in \mathscr{V}_Z$ of, and a set of consistent beliefs over 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 and given beliefs. Formally, at least one solution $w^\star \in \mathscr{V}_Z$ to* (6) *with $W = V$ also solves:*

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

where $\Lambda_t(\theta) = \int_{\underline{\theta}}^{\theta} \mathscr{L}(\zeta_{t-1}, V(x, \zeta_{t-1}) \mid V)\, d\Gamma(x)$, and F, G, U, W are defined by (1), (2), (3), (6) *with $\zeta = \zeta_t$.*

The equilibrium strategy V is a fixed point in the usual game-theoretic sense: 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 $W = V$, or $w = w^\star$.

3.2 Markov perfect equilibrium

Because the history of aggregate productivity is too large a space to be tractable, we look for equilibria in strategies that depend only on current values of payoff-relevant variables. From our discussion, it is clear that

$$\hat{\zeta} = \{\theta, L, \omega', \Lambda\} \tag{7}$$

is both the smallest and largest such state vector on which equilibrium strategies can depend. If all firms condition their current offers on these four objects in $\hat{\zeta}$, then from Definition 1 of equilibrium so should each firm in its best response. Let $\mathscr{V}_{\hat{Z}}$ be the space of measurable functions $\hat{Z} \to \mathbb{R}$. Then we focus on:

Definition 2 *A* Markov perfect equilibrium *of the contract posting game is a sequential equilibrium V in the set $\mathscr{V}_{\hat{Z}} \subset \mathscr{V}_Z$, a measurable function of $\hat{\zeta}$ defined in* (7)*.*

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, as we will see next.

From now on, we let the new value distributions $F(\cdot \mid \omega', \Lambda)$ and $G(\cdot \mid \omega', \Lambda)$, firm size $\mathscr{L}(L, \omega', \Lambda, W)$, employment distribution $\mathscr{T}(\omega', \Lambda)$, and value of unemployment $U(\omega', \Lambda)$ be defined as in (1) - (5), with the Markov state in (7) replacing ζ. Notice that only new firm size $\mathscr{L}$ 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, L.

4 Equilibrium characterization

4.1 The firm's contract-posting problem: recursive formulation

We look for a Markov perfect equilibrium of the contract-posting game. Suppose all other 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'. Then, by inspection of the firm's sequential profit maximization problem, these four objects are sufficient to pin down the firm's best response and evolve according to a Markov process. Therefore, 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. Therefore, we focus on the following recursive problem.

We fix the 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 $W(\omega') = U$. The firm solves:

$$\Pi\left(\theta, L, \omega, \Lambda, \overline{V}\right) = \sup_{w, W(\omega') \geq U(\omega', \mathscr{T}(\omega', \Lambda))} \Big\langle (\omega\theta - w) L + \beta \int_{\Omega} \Pi\left[\theta, \mathscr{L}\left(L, \omega', \Lambda, W(\omega')\right), \omega', \mathscr{T}(\omega', \Lambda), W(\omega')\right] H(d\omega' \mid \omega) \Big\rangle \quad (8)$$

subject to a Promise-Keeping (PK) constraint to deliver at least the promised $\overline{V}$, where the continuation value on the RHS comes from (6) after a small algebraic manipulation:

$$\overline{V} \leq w + \beta \cdot \int_{\Omega} \Bigg\{ \delta^{\omega'} U\left(\omega', \mathscr{T}(\omega', \Lambda)\right) + \left(1 - \delta^{\omega'}\right) \cdot \cdot \left[\left(1 - \lambda_1^{\omega'} \overline{F}\left(W(\omega') \mid \omega', \Lambda\right)\right) W(\omega') + \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF\left(v \mid \omega', \Lambda\right)\right] \Bigg\} H(d\omega' \mid \omega). \quad (9)$$

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, so that wage and continuation values deliver at least the current expected value $\overline{V}$ to the workers.

Notice that at time 0 the firm could extract full rents by offering $w = -\infty$, because it is "too late" for the initial workers to quit. To avoid this pathological outcome, we let the initial wage be chosen according to some bargaining procedure that splits rents from the contract and leaves the firm a non-negative cut, so the recursive formulation (8)-(9) only applies from period $t = 1$ on. Therefore, $\overline{V} \geq U(\omega, \Lambda)$ is always guaranteed, because $\overline{V}$ is the value promised a period before under the worker participation constraint $W(\omega') \geq U$.

4.2 An equivalent unconstrained recursive formulation

The constraints PC $W(\omega') \geq U(\omega', \mathscr{T}(\omega', \Lambda))$ and PK in (9) are appended to the firm's Bellman equation with their associated Lagrange multipliers, $m^{\omega'}$ for PC and π^{ω} for PK, respectively, to form a Lagrangian. Since the maximand is smooth in the wage w, we can take a derivative w.r.t. to w to see that optimality requires $-L + \pi^{\omega} \leq 0$. In other words, PK (9) must bind: if it did not, then $\pi^{\omega} = 0$, $-L < 0$, so the optimal wage would be as low as it can be, making (9) bind, a contradiction. So $\pi^{\omega} > 0$, and we can solve for the wage from (9) and replace it into the firm's Bellman equation. When we do this, we see that Π is differentiable w.r.t. the promised value, with $\Pi_{\overline{V}}\left(\theta, L, \omega, \Lambda, \overline{V}\right) = -L$.[5] This fact has two implications. First, firm profits are linear in promised value, and we can define the joint

[5]Except possibly for start-up firms that have $L = 0$ initially, but then for those the concept of a promised value inherited from the past is moot and the choice of w is irrelevant.

value of the firm-worker collective as a function independent of the value promised to the worker:

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

Second, we can replace this expression into the Bellman equation to obtain an equivalent DP problem in the surplus function S, rather than in profits Π, without the PK constraint:

$$\begin{aligned} S(\theta, L, \omega, \Lambda) = \omega\theta L + \beta \int_{\Omega} \Bigg\{ & \delta^{\omega'} U(\omega', \mathscr{T}(\omega', \Lambda)) L \\ & + \max_{W(\omega') \geq U(\omega', \mathscr{T}(\omega', \Lambda))} \Bigg\langle S(\theta, \mathscr{L}(L, \omega', \Lambda, W(\omega')), \omega', \mathscr{T}(\omega', \Lambda)) \\ & + L\left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF(v \mid \omega', \Lambda) \\ & - W(\omega') \left(\lambda_0^{\omega'} \left(1 - \Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'} \left(1 - \delta^{\omega'}\right) \Lambda\left(\overline{\theta}\right) G(W(\omega') \mid \omega', \Lambda)\right) \Bigg\rangle \Bigg\} H(d\omega' \mid \omega). \end{aligned} \tag{10}$$

An equilibrium is a solution V that coincides with the one followed by the other firms. To solve for equilibrium, we proceed as follows. We assume that an equilibrium strategy V, a value offered by each firm to workers which is also a best response to itself, exists, and we show which properties V must have to be an equilibrium. This allows us to restrict the set of possible equilibrium functions V, in particular, we prove that under certain sufficient conditions an equilibrium strategy must be strictly increasing in own productivity θ. In that smaller set, we will construct an equilibrium.

4.3 Properties of F and G in equilibrium

We begin by establishing that 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) *Any 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 all of their workers whenever they match, a positive measure, at a negligible cost. Because this argument invokes a possible upward deviation in value, it leaves open the possibility of an atom on the upper support of the value distribution. To exclude this last possibility, observe that either the upper bound value is below the PDV of productivity of the best firm—but then this firm can still deviate up and achieve higher profits—or it is equal to the productivity of the best firm—but then a deviation toward offering, e.g., $W = U$ in all states is profitable as it affords strictly positive profits.

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 $\overline{W} = \max \omega\theta / (1-\beta)$ is a natural upper bound to the offered value: the firm can always do weakly better by offering less than $\overline{W}$, as it can hope to make some profits. So the support of F is a convex and bounded subset of $\mathbb{R}_+$, which we can therefore take to be compact WLOG.

The properties stated in Proposition 1 will simplify our further characterization of equilibrium, to which we now turn.

4.4 Rank-Preserving Equilibria (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 $L^\star(\theta)$ 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 implemented 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:* $\theta \mapsto V(\theta, L^\star(\theta), \omega, \Lambda)$ *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 under the RP assumption: the proportion of firms that offer less than θ is simply that proportion of firms that are less productive than θ

$$F\left(V\left(\theta, L^\star(\theta), \omega', \Lambda\right) \mid \omega', \Lambda\right) \equiv \Gamma(\theta), \tag{11}$$

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:

$$\Lambda\left(\overline{\theta}\right) G\left(V\left(\theta, L^\star(\theta), \omega', \Lambda\right) \mid \omega', \Lambda\right) = \Lambda(\theta). \tag{12}$$

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. 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) *If Ω is finite and 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 equilibrium of the dynamic value-posting game 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.*

The proof is in Appendix A. Although that proof 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 in the Bellman equation of the modified but equivalent value-posting problem (10).[6] A more productive firm still wants and can afford to pay more, now in terms of values accruing to workers. If initially (or once) 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 values offered to workers remain ranked by firm productivity at all points in the future. This condition is only sufficient. We conjecture that it is 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 Perfect Equilibrium V exists, then V can be a best-response to itself if and only if it is increasing 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 Perfect Equilibrium of the contract-posting game.

4.5 Evolution of the firm size distribution in RPE

In a given aggregate state ω, firm sizes evolve following (4). In a RPE, in which firm size equals $L^\star(\theta) = \Lambda'(\theta)/\gamma(\theta)$, (4) reads as:

$$\begin{aligned} L^\star_{t+1}(\theta) &= L^\star_t(\theta)(1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right) + \lambda_0^\omega u_t + \lambda_1^\omega(1-\delta^\omega)\Lambda_t(\theta) \\ &= L^\star_t(\theta)(1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right) + \lambda_0^\omega u_t + \lambda_1^\omega(1-\delta^\omega)\int_{\underline{\theta}}^{\theta} L^\star_t(x)\,d\Gamma(x), \end{aligned} \tag{13}$$

where a discrete time index t was introduced for convenience, together with the notation $u_t := 1-\Lambda_t\left(\overline{\theta}\right)$ to designate the economy's unemployment rate. Equation (13) combines an ordinary differential equation and a first-order difference equation in Λ, a function of time and θ. We can solve it forward from any initial condition for any realization of the history of ω, independently of wages and offered values. Indeed multiplying through by $\gamma(\theta)$ in (13) and integrating with respect to θ yields:

$$\Lambda_{t+1}(\theta) = \lambda_0^\omega u_t\Gamma(\theta) + (1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)\Lambda_t(\theta).$$

[6] In a way similar to that in which Caputo (2003) appeals to single-crossing properties of the Hamiltonian in his analysis of comparative dynamics for infinite-horizon optimal control problems.

For any initial condition $\Lambda_0(\theta)$ at some (renormalized) initial date 0 such that the aggregate state remained at ω between 0 and t, the last law of motion is a first-order difference equation which solves as:

$$\Lambda_t(\theta) = \left[(1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)\right]^t \Lambda_0(\theta) + \lambda_0\Gamma(\theta)\sum_{s=1}^{t}\left[(1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)\right]^{s-1} u_{t-s}, \quad (14)$$

where the unemployment rate u_s solves the simple first-order, constant-coefficient difference equation $u_{s+1} = \delta^\omega(1-u_s) + (1-\lambda_0^\omega)u_s$ over $[0,t]$ with initial condition $u_0 = 1-\Lambda_0\left(\overline{\theta}\right)$. Next differentiating with respect to θ, one obtains a closed-form expression for the workforce of any type-θ firm:

$$\begin{aligned} L_t^\star(\theta) = \frac{\Lambda_t'(\theta)}{\gamma(\theta)} &= (1-\delta^\omega)^t\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)^{t-1}\left[\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)L_0^\star(\theta) + t\lambda_1^\omega\Lambda_0(\theta)\right] \\ &+ \lambda_0^\omega\left\{u_{t-1} + \sum_{s=2}^{t}(1-\delta^\omega)^{s-1}\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)^{s-2}\left[1-\lambda_1^\omega+\lambda_1^\omega s\Gamma(\theta)\right]u_{t-s}\right\}, \quad (15)\end{aligned}$$

where $L_0^\star(\theta)$ was the value of this solution under state ω' time of the last state switch from ω' to ω.

The steady-state versions of (14) and (15) (assuming the aggregate state forever stays at ω) are:

$$\begin{aligned} L_\infty^\star(\theta) &= \frac{\delta^\omega\lambda_0^\omega}{\delta^\omega+\lambda_0^\omega}\cdot\frac{1-(1-\delta^\omega)(1-\lambda_1^\omega)}{\left[1-(1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)\right]^2} \\ &\text{and} \\ \Lambda_\infty(\theta) &= \frac{\delta^\omega\lambda_0^\omega}{\delta^\omega+\lambda_0^\omega}\cdot\frac{\Gamma(\theta)}{1-(1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)}, \end{aligned} \quad (16)$$

which are the familiar steady-state expressions found in the BM model.[7]

Before going any further into characterizing Rank-Preserving Equilibria, we should notice that the analysis of firm size and employment dynamics carried out in this paragraph would apply to any job ladder model in which a similar concept of RPE can be defined. Indeed nothing in the dynamics of $L_t^\star$ or Λ_t 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.

[7]Incidentally, this is the point at which the necessity for sampling weights appears. Note from equation (16) that the steady-state size ratio of the largest to the smallest firm in the market in units of (non-normalized) employment is $L_\infty^\star\left(\overline{\theta}\right)/L_\infty^\star\left(\underline{\theta}\right) = (1-\lambda_1+\lambda_1/\delta)^2$, a ratio which is in the order of 25-30 given standard estimates of λ_1 and δ. Now of course the data counterpart of that size ratio is virtually infinite. It appears that the BM model requires a sampling distribution that is very heavily skewed toward high-productivity firms in order to replicate the observed distribution of firm sizes.

4.6 Properties of optimal contracts in RPE

Equation 15 combined with the assumption that initial firm size, $L_0^\star(\theta)$, is a continuous function of θ (see Section 2) ensures that $L_t^\star(\theta)$ 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 surplus and worker value functions 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* $L^\star(\theta)$*, i.e.* $S_L(\theta, L^\star(\theta), \omega, \Lambda)$ *exists for all* θ*. Moreover,* $S_L(\theta, L, \omega, \Lambda)$ *is continuous in* L *at* $L^\star(\theta)$*;*
2. $\theta \mapsto S_L(\theta, L^\star(\theta), \omega, \Lambda)$ *is continuously differentiable in* θ*;*
3. $\theta \mapsto V(\theta, L^\star(\theta), \omega, \Lambda)$ *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 $\theta \mapsto V(\theta, L^\star(\theta), \omega, \Lambda)$, 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 (11) and (12) w.r.t. θ to write:

$$f(V \mid \omega, \Lambda) \cdot \frac{dV}{d\theta} = \gamma(\theta) \qquad \text{and} \qquad g(V \mid \omega, \Lambda) \cdot \frac{dV}{d\theta} = L^\star(\theta)\gamma(\theta). \tag{17}$$

at $V = V(\theta, L_t^\star(\theta), \omega, \Lambda_t)$.

4.7 A RPE with differentiable contracts

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

$$\begin{aligned} \lambda_0^{\omega'}\left(1 - \Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'}\left(1 - \delta^{\omega'}\right)\Lambda\left(\overline{\theta}\right) G\left(W(\omega') \mid \omega', \Lambda\right) \\ = \left[S_L\left(\theta, \mathscr{L}\left(L, \omega', \Lambda, W(\omega')\right), \omega', \mathscr{T}(\omega', \Lambda)\right) - W(\omega')\right] \\ \times \left(1 - \delta^{\omega'}\right)\lambda_1^{\omega'}\left[Lf\left(W(\omega') \mid \omega', \Lambda\right) + \Lambda\left(\overline{\theta}\right) g\left(W(\omega') \mid \omega', \Lambda\right)\right] - m^{\omega'} \end{aligned} \tag{18}$$

with complementary slackness $m^{\omega'}\left[W(\omega') - U(\omega', \mathscr{T}(\omega', \Lambda))\right] = 0$. In a RPE, (18) is solved by $W = V(\theta, L^\star(\theta), \omega, \Lambda)$.

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

$$V_{t+1}(\theta \mid \omega) := V\left(\theta, L_t^\star(\theta), \omega, \Lambda_t\right)$$

and further define the costate variable:

$$\mu_{t+1}(\theta \mid \omega) := S_L\left(\theta, \mathscr{L}\left(L_t^\star(\theta), \omega, \Lambda_t, V_{t+1}(\theta \mid \omega)\right), \omega, \mathscr{T}(\omega, \Lambda_t)\right),$$

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 aggregate (uncontrolled) state variable Λ is subsumed into the time index in the above notation, which is licit as Λ evolves deterministically conditional on ω. Combining (18) and the various restrictions (11), (12), and (17) that hold in a RPE, we obtain the RPE version of the FOC:

$$\begin{aligned}\lambda_0^\omega u + & \lambda_1^\omega (1-\delta^\omega)\Lambda_t(\theta) \\ &= \lambda_1^\omega (1-\delta^\omega)\left[\mu_{t+1} - V_{t+1}\right]\left[L_t^\star(\theta) f\left(V_{t+1} \mid \omega, \Lambda_t\right) + (1-u_t) g\left(V_{t+1} \mid \omega, \Lambda_t\right)\right] - m_t^\omega \\ &= 2\lambda_1^\omega (1-\delta^\omega)\frac{L_t^\star(\theta)\gamma(\theta)}{dV_{t+1}/d\theta}\left(\mu_{t+1} - V_{t+1}\right) - m_t^\omega.\end{aligned}$$

Next, going back to the firm's problem (10), the Envelope condition w.r.t. firm size writes down as:

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

In a RPE, introducing again our time index t, this becomes:

$$\begin{aligned}\mu_t(\theta \mid \omega) = \omega\theta + \beta \int_\Omega \Bigg\{ & \delta^{\omega'} U\left(\omega', \mathscr{T}(\omega', \Lambda)\right) + \left(1-\delta^{\omega'}\right)\lambda_1^{\omega'} \int_\theta^{+\infty} V_{t+1}\left(x \mid \omega'\right) d\Gamma(x) \\ & + \mu_{t+1}\left(\theta \mid \omega'\right)\left(1-\delta^{\omega'}\right)\left(1 - \lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\Bigg\} H\left(d\omega' \mid \omega\right),\end{aligned} \tag{20}$$

Note that now the shadow marginal value μ only depends on the distribution of employment Λ through the cumulative distribution of employment across firms up to θ, $\Lambda(\theta)$ and the corresponding density $L^\star(\theta)\gamma(\theta)$. 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 and not the entire size distribution Λ.

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

$$\lim_{t\to\infty} E\left[\beta^t \mu_t(\theta \mid \omega) L_t^\star(\theta) \mid \mathbf{z}_0\right] = 0. \tag{21}$$

5 Practical implementation of RPE

5.1 A strategy to solve for for stochastic RPE

We now show in practice how to numerically "solve" for the RPE, by which we mean simulate the dynamic paths of the distributions of employment and wages across firms, given an initial state and a subsequent realization of a sequence of aggregate shocks. For the sake of illustration, we focus on the case where the aggregate state can take on two values, $\Omega = \{\omega, \omega'\}$, with conditional switching probabilities σ^{ω} and $\sigma^{\omega'}$. Generalization to any first-order Markov process over a finite set is conceptually trivial. With a two-state process for the aggregate productivity component, the Euler equation — or Envelope condition — (20) becomes:

$$\begin{aligned}\mu_t(\theta \mid \omega) = \omega\theta + \beta\sigma^{\omega}\Bigg\{&\delta^{\omega'} U\left(\omega', \mathscr{T}(\omega', \Lambda)\right) + \left(1 - \delta^{\omega'}\right)\lambda_1^{\omega'} \int_\theta^{+\infty} V_{t+1}(x \mid \omega')\, d\Gamma(x) \\ &+ \mu_{t+1}(\theta \mid \omega')\left(1 - \delta^{\omega'}\right)\left(1 - \lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\Bigg\} \\ + \beta(1 - \sigma^{\omega})\Bigg\{&\delta^{\omega} U\left(\omega, \mathscr{T}(\omega, \Lambda)\right) + (1 - \delta^{\omega})\lambda_1^{\omega} \int_\theta^{+\infty} V_{t+1}(x \mid \omega)\, d\Gamma(x) \\ &+ \mu_{t+1}(\theta \mid \omega)(1 - \delta^{\omega})\left(1 - \lambda_1^{\omega}\overline{\Gamma}(\theta)\right)\Bigg\}.\end{aligned}$$

Taking derivatives w.r. to θ on both sides:

$$\begin{aligned}\frac{\partial \mu_t}{\partial \theta}(\theta \mid \omega) &= \omega \\ &+ \beta\sigma^{\omega}\left(1 - \delta^{\omega'}\right)\left\{\lambda_1^{\omega'}\gamma(\theta)\,\pi_{t+1}(\theta \mid \omega') + \frac{\partial \mu_{t+1}}{\partial \theta}(\theta \mid \omega')\left(1 - \lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\right\} \\ &+ \beta(1 - \sigma^{\omega})(1 - \delta^{\omega})\left\{\lambda_1^{\omega}\gamma(\theta)\,\pi_{t+1}(\theta \mid \omega) + \frac{\partial \mu_{t+1}}{\partial \theta}(\theta \mid \omega)\left(1 - \lambda_1^{\omega}\overline{\Gamma}(\theta)\right)\right\},\end{aligned} \tag{22}$$

where $\pi_t(\theta \mid \omega) := \mu_t(\theta \mid \omega) - V_t(\theta \mid \omega)$ denotes the shadow value to the *firm* of the marginal worker. Together with the FOC for an interior solution of the promised value:

$$\frac{\partial \pi_t}{\partial \theta}(\theta \mid \omega) = \frac{\partial \mu_t}{\partial \theta}(\theta \mid \omega) - \frac{2\lambda_1^\omega \Lambda_t'(\theta)(1-\delta^\omega)}{\lambda_0^\omega u_t + \lambda_1^\omega (1-\delta^\omega)\Lambda_t(\theta)}\pi_t(\theta \mid \omega) \tag{23}$$

this gives a system of four PDEs in $\pi_t(\theta \mid \omega)$, $\partial\mu_t(\theta \mid \omega)/\partial\theta$, all functions of θ and t, a pair for each value of ω.

The main difficulty in solving this system lies in the dependence of $\partial\mu_t(\theta \mid \omega)/\partial\theta$ on $\partial\mu_{t+1}(\theta \mid \omega')/\partial\theta$ and $\pi_{t+1}(\theta \mid \omega')$, that is on the jump in the shadow marginal values of one worker, both to the firm (π) and to the collective (μ), caused by the possible occurrence of an aggregate state switch next period. To get around this problem, we can approximate that "jump term" by a known function J (e.g. polynomials) of the state variable $(\theta, L^\star(\theta), \omega, \Lambda(\theta))$ depending on a finite vector of unknown coefficients, $\mathbf{a}$, which can be determined iteratively by successive approximations.Thus the proposed simulation protocol is akin to a projection method to solve the system of PDEs that characterize equilibrium. Its specific feature is that projection is only used to approximate the jumps in π and $\partial\mu/\partial\theta$ caused by aggregate shocks, the rest of the system being solved "exactly". Practical details of the algorithm are given in Appendix C.

Simulation of the model then goes as follows: first, we pick an initial state of the economy $(\omega_0, \Lambda_0(\cdot), L_0^\star(\cdot))$ and simulate a path of ω. Second, given the simulated path of ω and the initial state of the economy, simulate the associated paths of $\Lambda(\cdot)$ and $L^\star(\cdot)$ as per equations (14) and (15). Third, given the previously simulated objects, we solve (22) and (23) subject to (21) using the algorithm sketched above and described more completely in Appendix C. Completion of those three steps produces a solution for $\{\Lambda, L^\star, \partial\mu_t/\partial\theta, \pi_t\}$ over some initially chosen time interval $t \in \{0, \cdots, T\}$ given any simulated sequence of aggregate states. Wages are finally retrieved from:[8]

$$\begin{aligned} w_t(\theta \mid \omega) = \omega\theta - \pi_t(\theta \mid \omega) + \beta\sigma^\omega\left(1-\delta^{\omega'}\right)\left(1-\lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\pi_{t+1}(\theta \mid \omega') \\ + \beta(1-\sigma^\omega)(1-\delta^\omega)\left(1-\lambda_1^\omega\overline{\Gamma}(\theta)\right)\pi_{t+1}(\theta \mid \omega). \end{aligned}$$

5.2 Simulation results

We now illustrate the quantitative properties of the model using the following calibration. First, all scalar parameters are given values as indicated in Table 1. Next, the sampling distribution of firm types is calibrated following the Bontemps, Robin and Van den Berg

[8]Note that this equation also features a "jump term" — i.e., $w_t(\theta \mid \omega)$ depends on future values of π in both aggregate states. We approximate that jump term in the same fashion as we do the jump term in the differentiated Euler equation (22).

(2000) estimation procedure in such a way that the predicted steady-state distribution in the high productivity state fits the business sector wage distribution observed from the CPS in 2006 (see our MPV08 companion paper for details).

	ω	λ_0^ω	λ_1^ω	δ^ω	σ^ω	β
Expansion	1.04	0.39	$0.25 \cdot \lambda_0^\omega$	.017	0.0083	0.9959
Contraction	1	0.33	$0.25 \cdot \lambda_0^\omega$	.019	0.0167	0.9959

Table 1: Model Calibration

The aggregate productivity shifter ω is normalized at one in the low aggregate state and assumed to be 4 percent higher in the high state, so that aggregate labor productivity will exhibit fluctuations of roughly $\pm 3\%$ around its trend (see Figure 4 below), which is in same order of magnitude as Robin's (2009) estimates. The high-state value of λ_0 is the monthly job finding rate from the Current Population Survey, averaged over NBER expansions, i.e. over periods between NBER troughs and the following NBER peak. The low-state value of λ_0 is defined in the same way, the average being taken over NBER contractions, i.e. over periods between NBER peaks and the following NBER trough.[9] The job loss rate δ is also constructed in a similar fashion, using averages of the monthly rate of job separation, excluding quits. The arrival rate of offers to employed job seekers, λ_1, is calibrated as a constant fraction of λ_0 which produces a share of job-to-job transitions in total separations of about 50%, as is typically observed (e.g. Hall, 2005). The aggregate productivity process is calibrated to reflect the average duration of booms and recessions in the US (10 and 4 years, respectively, according to the NBER). The discount factor β corresponds to an annual discount rate of 5 percent.

Finally, we abstract from entry or exit of firms over the cycle. In particular the lower bound of the support of firm-specific productivity components, $\underline{\theta}$, is normalized at 1 in both aggregate productivity states, and so is the minimum wage, $\underline{w}$. This implies that even the least productive firm in the economy remains viable in bad times, so that aggregate downturns do not cause firms to leave the market.

Figures 1-5 illustrate the output of a representative simulation. Aggregate productivity ω is in the high state at the initial date, while the economy's initial state is set to the low-ω steady state, so that the economy starts off in an expansionary phase, just out of a very (infinitely) long recession. The simulated series then cover a period of 30 years spanning three expansions and two recessions. On all plots except Figure 3 the cycle is materialized

[9] At the time of writing, the JOLTS data set covers the period 12/2000 to 06/2009, a period spanning one complete NBER contraction (March-November 2001), one complete NBER expansion (November 2001-December 2007), plus a few months in the previous expansion in the following recession.

by the dashed line which represents the unemployment rate (in deviations from its mean and rescaled for legibility).

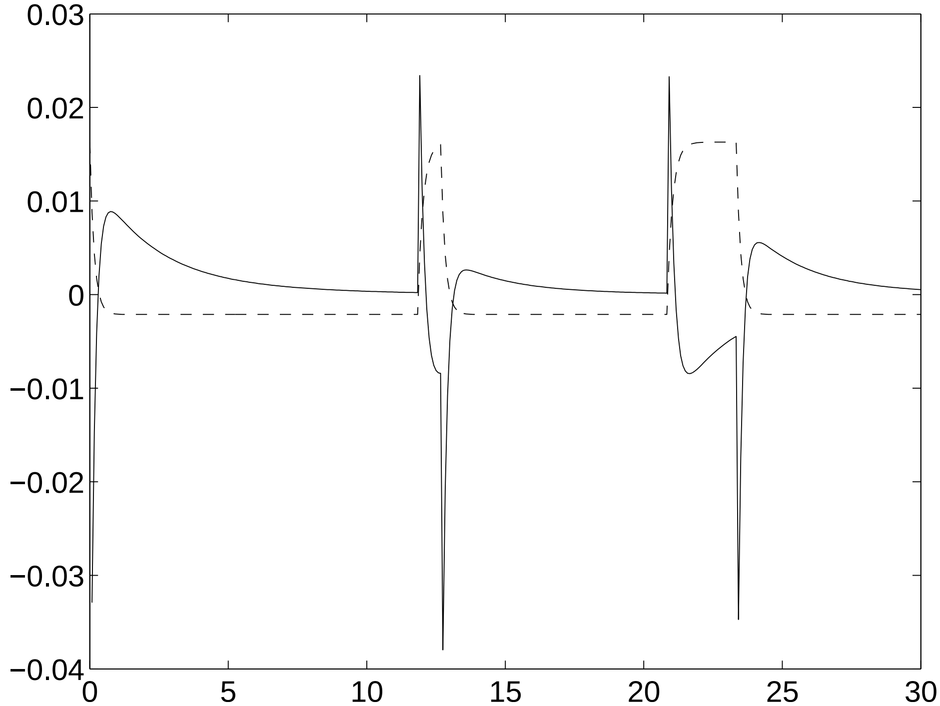

Fig. 1: Firm size growth differential

The results do not differ qualitatively from those of the deterministic transition dynamics analyzed in MPV08. Most remarkably, as documented in MPV08 and MPV09, the growth rate differential of employment at (initially) large minus small firms collapses upon a recession and rises slowly through an expansion, as shown on Figure 1 which plots the simulated differential in average growth rates across firm size classes.[10] The cyclical pattern of relative growth by firm size is naturally interpreted in our model as reflecting the slow upgrading of workers to better jobs through on the job search. The governing mechanism is the fol-

[10]Formally, the average growth rate differential is defined as:

$$\frac{\Lambda_{t+1}\left(\overline{\theta}\right)-\Lambda_{t+1}\left(\theta_{\ell}\right)}{\Lambda_{t}\left(\overline{\theta}\right)-\Lambda_{t}\left(\theta_{\ell}\right)}-\frac{\Lambda_{t+1}\left(\theta_{s}\right)}{\Lambda_{t}\left(\theta_{s}\right)},$$

where θ_ℓ and θ_s are thresholds defining the groups of "large" and "small" firms, respectively. Here the group of large firms is defined as the top quintile of the (high-) steady-state distribution of firm sizes among workers (i.e. all firms such as at least 80% of the employed workforce work at smaller firms), while small firms are the bottom 25 percent of that distribution (which comprises all firms such as at least 75% of the employed workforce work at larger firms). According to data on the distribution of establishment sizes from the County Business Patterns (CBP) those thresholds roughly correspond to establishments of over 500 and less than 20 employees, respectively, and are consistent with the thresholds used in our companion empirical paper, MPV09.

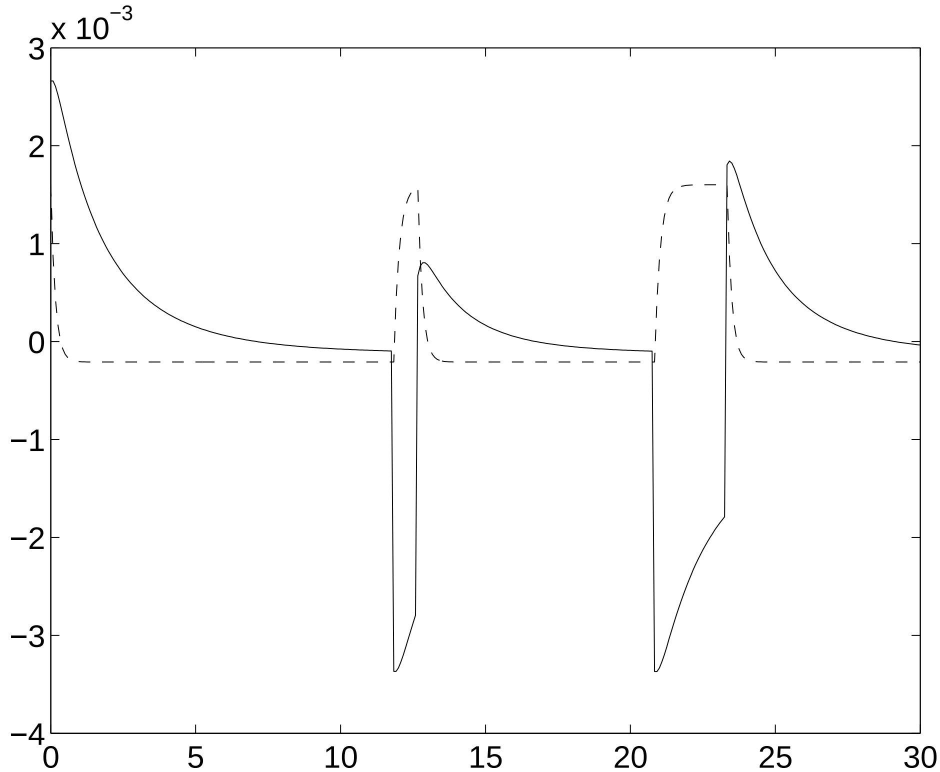

Fig. 2: Job-to-job transition rate

lowing. At the onset of an expansion, the many unemployed workers inherited from the previous recession are available for work at any (low) wage that makes work preferable to unemployment. As those unemployed job applicants are willing to accept any offer, the random search process allocates them into firms following the sampling distribution of firm types. However the magnitude of the job finding rate λ_0 is such that the pool of cheap, unemployed job seekers dries out quickly. To keep expanding, firms begin to raise wages to poach labor from their competitors. But as argued earlier, the BM single-crossing argument applies: more productive firms are both able and willing to raise wages further than their less productive competitors, as their opportunity cost of not producing is higher, especially so in an aggregate expansion. Thus workers gradually select themselves into more productive, better paying jobs — at the speed permitted by search frictions, which is determined by λ_1 — and more productive firms grow relatively faster and become relatively larger.[11] As a consequence, as the expansion progresses, the distribution of employment becomes increasingly skewed toward more productive, larger firms. Therefore, when the recession hits, while

[11] Indeed an immediate implication of the RP property of equilibrium is that a firm's rank in the distribution of firm sizes is the same as that firm's rank in the productivity distribution or in the distribution of offered worker values. As such our model provides a theoretical justification for the use of firm size as a proxy for firm productivity — as does any job-ladder model whose equilibrium has the RP property.

all firms have too many workers, larger firms have relatively more excess employment. The distribution of firm sizes must then ebb back toward smaller firms, so in net terms large firms shed proportionally more workers to reduce their share. Finally, note that the same intuition explains the procyclical rate of job-to-job transition predicted by our model (Figure 2).[12]

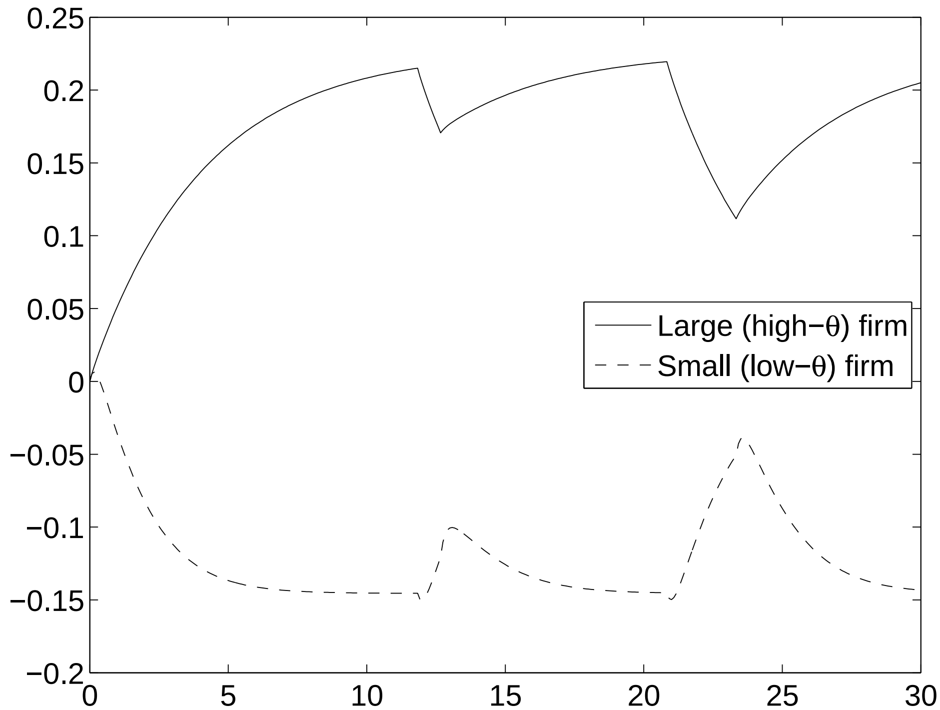

Fig. 3: Individual firm employment paths

To further comprehend the pattern of labor adjustments in the model, Figure 3 shows the firm-level employment paths (in relative deviation from high-ω steady-state values) at two different firms. The one labeled "large" (solid line on the graph) is the firm at the 90th percentile of the high-ω steady-state distribution of firm size among workers — i.e. the firm such that exactly 90% of the employed workforce work at smaller firms —, while the one labeled "small" (dashed line on the graph) is the firms at the 10th percentile in that distribution. Those are the median firms in the groups of small and large firms used to plot Figure 1. The graph shows that, while employment at the larger firm is clearly procyclical, the smaller firm actually swells in contractions and shrinks in expansions. Relative to expansions, contractions are times in which the job destruction rate δ is higher and the hiring

[12]We should raise the following caveat about the intuition just spelled out. Our maintained assumption of exogenous offer arrival and job destruction probabilities hides the restriction that, in a given aggregate state of the economy, λ_1 and δ are the same for all firms. If, for example, we allowed for endogenous vacancy posting, we might observe that large firms reduce their hiring effort (vacancies) by strictly less than small firms in a slump, which could attenuate or even overturn the recession pattern of Figure 1. This will have to be investigated in future research.

rate from unemployment λ_0 is lower, which unambiguously tends to make employment at all firm proyclical. But contractions are also times when the λ_1 is lower, which helps small firms and hurts large ones as it reduces the frequency at which the latter can raid the former for workers. Quantitatively the differential impact of λ_1 overcomes the combined impact of λ_0 and δ for small firms.

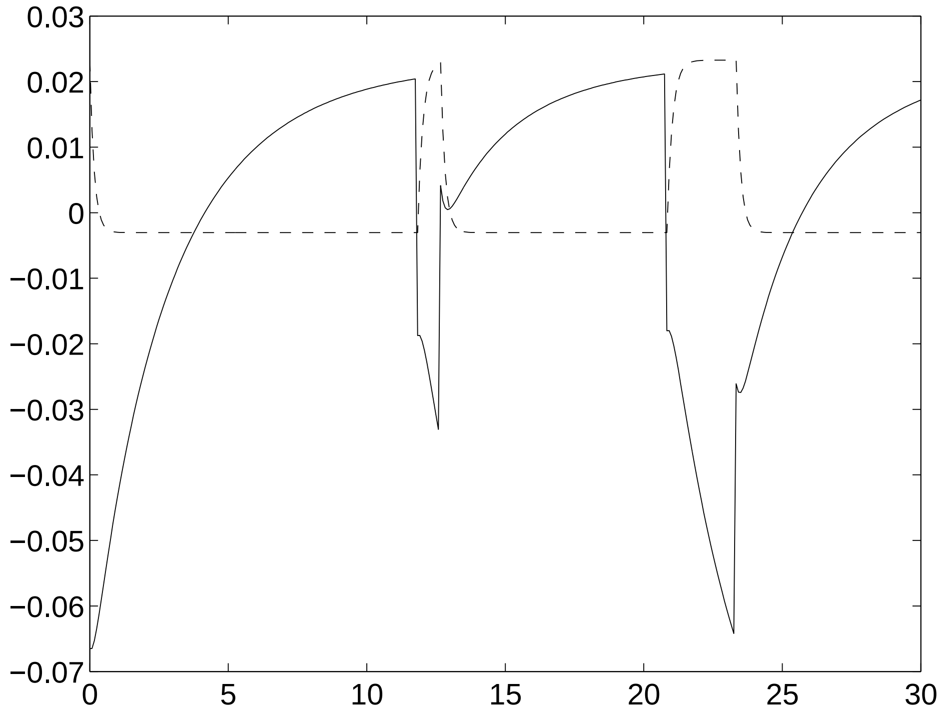

Fig. 4: Mean output per worker

The cyclical upgrading of workers also strongly propagates and amplifies the effects of the aggregate labor productivity shock on measured average labor productivity (Figure 4).[13] Labor being a quasi-fixed factor in this model, average output per worker jumps upon impact of an aggregate, proportional productivity shock by ± 4 percent, the calibrated magnitude of the aggregate shock (see Table 1). After that initial jump, output per worker continues to adjust in the direction of the shock in a smooth and quasi-monotonic fashion. This smooth adjustment is driven by the slow movements of workers up and down the job ladder: for example in an expansion, labor is slowly reallocated toward more productive firms, which

[13]Krause and Lubik (2006) describe a distantly related mechanism in a search-and-matching economy with wage bargaining. Firms create two types of jobs that only differ by creation costs, thus pay different wages. Unemployed workers choose which market to probe, and employed workers also choose the intensity of their search effort. In an aggregate expansion, employed workers increase their effort to upgrade to more abundant high-paying jobs. Unemployed face more congestion for high-wage jobs, so they revert to low-paying ones. Filling vacancies with unemployed workers becomes quite easy, thus profitable, raising the job-finding rate and making it as volatile as in the data.

increases average labor productivity beyond the initial increase in the aggregate productivity shifter through a composition effect. Moreover, because the driving force is the reallocation of workers into better jobs, the speed at which that composition effect drives average productivity up is limited by the speed at which employed workers are able to find better jobs, before they lose the ones they have. So the key parameter that determines the extent of propagation of aggregate shocks in this model is the ratio λ_1 δ, a standard measure of the extent of search frictions in wage-posting models.[14]

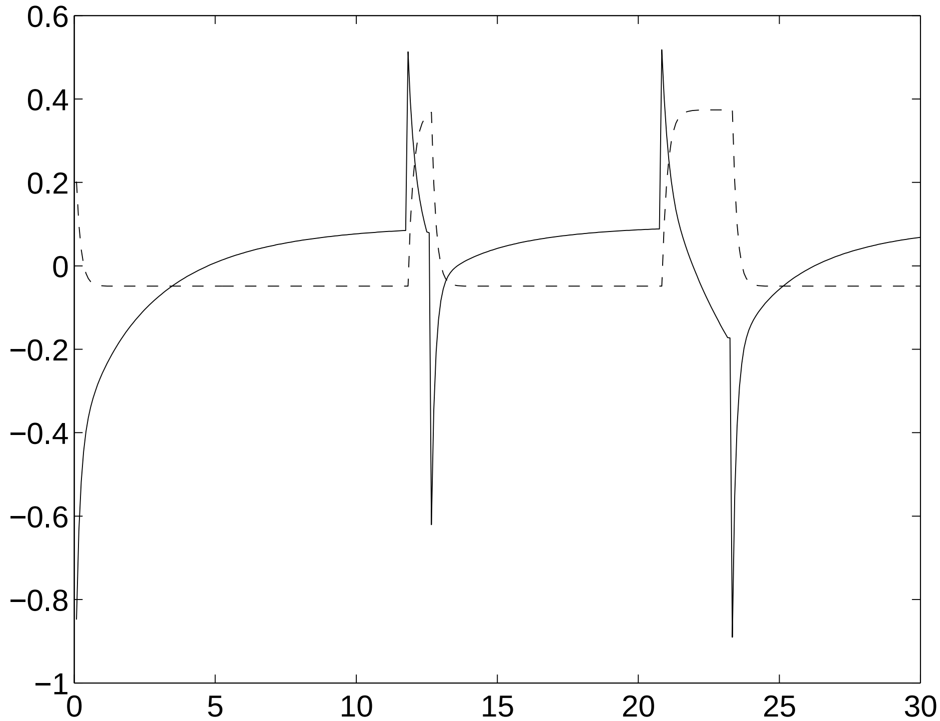

Fig. 5: Mean wage (in Log)

The model predicts procyclical wages in the sense that on average steady-state wages are lower in the low aggregate state than in the high aggregate state, although the wage jumps in a direction opposite to labor productivity when aggregate shocks hit. When a recession lasts long enough compared to the speed of adjustment toward steady state, as is the case in the second recession occurring at the beginning of year 21 in the simulation, this still produces an

[14]About Figure 4, the observant reader will have noticed a slight dip in average labor productivity at the beginning of some of the expansions, immediately after the initial jump directly caused by the shock. This dip is also due to a composition effect: as argued before, at early stages of an expansion, when the unemployed are many, new hires get allocated disproportionately into low-productivity firms, which tends to bring down aggregate productivity. If λ_0 (which governs the inflow of unemployed job applicants) is sufficiently large, that effect may initially dominate the positive effect of labor upgrading on mean output per worker. But then precisely because λ_0 is large, the pool of unemployed job applicants becomes depleted very quickly, so that the effect described in this footnote is short-lived.

impression of procyclical wages, however the short recession occurring at year 10 on Figure 5 suggests that it is not the always case. The forces combining into the observed dynamic path of wages are the following. First, the composition effect described in the previous paragraph as governing the smooth dynamics of productivity also affects wages: more productive firms pay higher wages at all dates and workers move up the productivity (or wage) ladder in expansions and down that ladder in recessions. Obviously, if this were the only mechanism at play, wages would unambiguously be procyclical. However, each firm-level wage also follows a dynamic path of its own. The fact mentioned in the discussion of Proposition 2 that larger larger firms have a motive to transfer higher values to their workers also applies across dates for a given firm: firms tend to post higher values whenever they grow in size, and vice versa. As a consequence, firm-level wages closely track firm-level employment, so that wages are countercyclical at firms toward the bottom of the productivity distribution, and procyclical toward the top of that distribution. When computing the aggregate mean wage, high-θ firms, which are larger, get more weight, so that their cyclical wage pattern dominates in the aggregate wage series.

6 Conclusion [in progress]

References

[1] Barlevy, G. (2008) "Identification of Search Models using Record Statistics", *Review of Economic Studies*, 75(1) 29-64.

[2] Bontemps, C., J.-M. Robin, and G. J. Van den Berg (2000) "Equilibrium Search with Continuous Productivity Dispersion: Theory and Non-Parametric Estimation", *International Economic Review*, 41 (2) 305-58.

[3] Burdett, K. and M. Coles (2003) "Equilibrium Wage-Tenure Contracts", *Econometrica*, 71(5), 1377-1404.

[4] Burdett, K. and D. T. Mortensen (1998) "Wage Differentials, Employer Size and Unemployment", *International Economic Review*, 39, 257-73.

[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] Corbæ, D., M. B. Stinchcombe, and J. Zeman (2009), *An Introduction to Mathematical Analysis for Economic Theory and Econometrics*, Princeton University Press.

[10] Dey, M. S. and C. J. Flinn (2005) "An Equilibrium Model of Health Insurance Provision and Wage Determination", *Econometrica*, 73(2), 571-627.

[11] Hall, R.E. (2005) "Job Loss, Job Finding, and Unemployment in the U.S. Economy over the Past Fifty Years", *in* M. Gertler and K. Rogoff, eds., *NBER Macroeconomics Annual*, 20, 1-137.

[12] Krause, M. and T. Lubik (2006), "The Cyclical Upgrading of Labor and On-the-job Search", *Labour Economics*, 13(4), 459-477.

[13] Menzio, G. and S. Shi (2008) "Efficient Search on the Job and the Business Cycle", PIER Working Paper 08-029.

[14] Mortensen, D. (2003) *Wage Dispersion: Why Are Similar Workers Paid Differently?*, MIT Press.

[15] Moscarini, G. (2004) "Job-to-Job Quits and Corporate Culture", Yale University *mimeo*.

[16] 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. [MPV08]

[17] Moscarini, G. and F. Postel-Vinay (2009) "Large Employers Are More Cyclically Sensitive", NBER Working Paper 14740. [MPV09]

[18] Postel-Vinay, F. and J.-M. Robin (2002) "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity", *Econometrica*, 70 (6), 2295-350.

[19] 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.

[20] Robin, J.-M. (2009) "Labour Market Dynamics with Sequential Auctions: On Countercyclical Wage Inequality and Unemployment Volatility", Walras-Bowley Lecture, North American Summer Meeting of the Econometric Society, Boston University, June 2009.

[21] Royden, H. L. (1988) *Real Analysis – Third Edition*, Prentice-Hall.

[22] Rudanko, L. (2008) "Aggregate and Idiosyncratic Risk in a Frictional Labor Market", mimeo Boston University.

[23] Rudin, W. (1976) *Principles of Mathematical Analysis – Third Edition*, McGraw-Hill.

[24] 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.

[25] Spear, S. and S. Srivastava, (1987), "On Repeated Moral Hazard With Discounting", *Review of Economic Studies*, 54, 599-617.

[26] Stevens, M. (2004) "Wage-Tenure Contracts in a Frictional Labour Market: Firms' Strategies for Recruitment and Retention", *Review of Economic Studies*, 71, 535-551.

[27] Thomas, J. and T. Worrall (1988) “Self-Enforcing Wage Contracts”, *Review of Economic Studies*, 55, 541-554.

Appendix

A Proof of Proposition 2

For convenience, we repeat the firm's DP problem (10):

$$S(\theta, L, \omega, \Lambda) = \omega\theta L + \beta \int_{\Omega} \Bigg\{ \delta^{\omega'} U\left(\omega', \mathscr{T}\left(\omega', \Lambda\right)\right) L$$

$$+ \sup_{W(\omega') \geq U(\omega', \mathscr{T}(\omega', \Lambda))} \Bigg\langle S\left(\theta, \mathscr{L}\left(L, \omega', \Lambda, W\left(\omega'\right)\right), \omega', \mathscr{T}\left(\omega', \Lambda\right)\right) + L\left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF\left(v \mid \omega', \Lambda\right)$$

$$- W\left(\omega'\right)\left(\lambda_0^{\omega'}\left(1 - \Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'}\left(1 - \delta^{\omega'}\right) \Lambda\left(\overline{\theta}\right) G\left(W\left(\omega'\right) \mid \omega', \Lambda\right)\right) \Bigg\rangle \Bigg\} H\left(d\omega' \mid \omega\right),$$

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:

$$A(\theta, L, \omega, \Lambda) := \omega\theta L + \beta \int_{\Omega} \delta^{\omega'} U\left(\omega', \mathscr{T}\left(\omega', \Lambda\right)\right) L H\left(d\omega' \mid \omega\right),$$

$$B\left(L, \omega', \Lambda; W\left(\omega'\right)\right) := L\left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{W(\omega')}^{+\infty} v dF\left(v \mid \omega', \Lambda\right)$$

$$- W\left(\omega'\right)\left(\lambda_0^{\omega'}\left(1 - \Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'}\left(1 - \delta^{\omega'}\right) \Lambda\left(\overline{\theta}\right) G\left(W\left(\omega'\right) \mid \omega', \Lambda\right)\right).$$

Our proof strategy is as follows. First, we show that in any equilibrium F and G are continuous and have a compact, convex support. Next, we define certain supermodularity properties SM of a value function that imply that the maximizer V in (10) is increasing in θ. Then, we fix an arbitrary Λ and show that the Bellman operator in (10) 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 (10) has a unique solution. Finally, *if* there exists a solution S to (10) when Λ is not fixed, then S must also solve the restricted problem (10) 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.

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

$$\mathbf{M}^{\mathbf{\Lambda}} \mathscr{S}(\theta, L, \omega) := A(\theta, L, \omega, \Lambda)$$

$$+ \beta \int_{\Omega} \max_{W(\omega')} \Big\langle \mathscr{S}\left[\theta, \mathscr{L}\left(L, \omega', \Lambda, W\left(\omega'\right)\right), \omega'\right] + B\left(L, \omega', \Lambda; W\left(\omega'\right)\right) \Big\rangle H\left(d\omega' \mid \omega\right). \quad (24)$$

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

1. *$\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ 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 $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$;*
3. *Any such measurable selection V is increasing in θ and L;*
4. *$\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ is increasing and convex in L and with increasing differences in (θ, L) over $(\underline{\theta}, \overline{\theta}) \times (0,1)$.*

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 a little bit.

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

To prove point 3, we first establish that the maximand in (24) 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. First note that, since $\mathscr{S}$ 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 L and $h > 0$ and define the difference in L of the maximand in (24):

$$\mathscr{D}(W) := \mathscr{S}(\theta, \mathscr{L}(L+h, W)) - \mathscr{S}(\theta, \mathscr{L}(L, W)) + h\left(1 - \delta^{\omega'}\right)\lambda_1^{\omega'} \int_W^{+\infty} v dF\left(v \mid \omega'\right).$$

First note that $\mathscr{D}$ is also a function of θ, which is trivially increasing in θ from the assumption that $\mathscr{S}$ has increasing differences in (θ, L). We thus keep the dependence of $\mathscr{D}$ on θ implicit, fix θ and focus on establishing that $\mathscr{D}(W)$ is increasing in W. We will do so by showing that the upper-right Dini derivative of $\mathscr{D}(W)$, which we denote as $D^+\mathscr{D}(W)$ and which exists everywhere (although possibly equalling $\pm\infty$), is everywhere positive. Take $x > 0$:

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

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 $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ has one-sided derivatives because it is convex. Majorizing the last integral:

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

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

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

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 $\mathscr{S}$ is convex in L by assumption, $\mathscr{S}_{L,r}$ is increasing in L. Combining all those properties, the latter inequality implies:

$$D^+\mathscr{D}(W) \geq \left[\mathscr{S}_{L,r}(\theta, \mathscr{L}(L,W)) - W\right] \cdot \left(1-\delta^{\omega'}\right)\lambda_1^{\omega'} h D_+F(W). \tag{25}$$

The only way the RHS in this last inequality can be negative is if $\mathscr{S}_{L,r}(\theta, \mathscr{L}(L,W)) - W < 0$. 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:

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

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

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

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

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

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

$$\mathscr{S}_{L,\ell}(\theta, \mathscr{L}(L,V)) - V \geq \frac{\lambda_0^{\omega'}\left(1-\Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda\left(\overline{\theta}\right)G(V \mid \omega')}{\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{LD_-F(V) + \Lambda\left(\overline{\theta}\right)D_-G(V)\right\}} \geq 0. \tag{26}$$

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

This, together with (25), shows that $D^{+}\mathscr{D}(V) \geq 0$ at all V which is an optimal selection, i.e. at all V in the support of F. To finally establish that $\mathscr{D}$ is increasing over the support of F, recall that, as F and G are continuous, so is $W \mapsto \mathscr{L}(L, W)$. Moreover, as $\mathscr{S}$ is convex in L (by assumption), it is continuous w.r.t. L. Thus by inspection, $\mathscr{D}$ is a continuous function of W. Continuity plus the fact that $D^{+}\mathscr{D}(V) \geq 0$ are sufficient to ensure that $\mathscr{D}$ 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 $(\theta_0, L_0) \in (\underline{\theta}, \overline{\theta}) \times (0,1)$ and $h > 0$ such that $(\theta_0 + h, L_0 + h)$ are still in $(\underline{\theta}, \overline{\theta}) \times (0,1)$. We first consider right-differentiability of $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ 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 $V\left(\theta_0, L_0^{+}, \omega'\right) := \lim_{h \to 0^{+}} V(\theta_0, L_0 + h, \omega')$ exists everywhere (and likewise for $V\left(\theta_0^{+}, L_0, \omega'\right)$). By point 3, $V\left(\theta_0, L_0^{+}, \omega'\right)$ is increasing in L_0. Then:

$$
\begin{aligned}
&\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}(\theta_0, L_0 + h) - \mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\left(\theta_0, L_0^{+}\right) = A(\theta_0, L_0 + h) - A(\theta_0, L_0) \\
&\quad + \beta \int_{\Omega} \Big\langle \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0 + h, V\left(\theta_0, L_0 + h, \omega'\right)\right)\right] - \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0, V\left(\theta_0, L_0^{+}, \omega'\right)\right)\right] \\
&\qquad + B\left(L_0 + h; V\left(\theta_0, L_0 + h, \omega'\right)\right) - B\left(L_0; V\left(\theta_0, L_0^{+}, \omega'\right)\right) \Big\rangle H\left(d\omega' \mid \omega\right) \\
&\geq A(\theta_0, L_0 + h) - A(\theta_0, L_0) \\
&\quad + \beta \int_{\Omega} \Big\langle \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0 + h, V\left(\theta_0, L_0^{+}, \omega'\right)\right)\right] - \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0, V\left(\theta_0, L_0^{+}, \omega'\right)\right)\right] \\
&\qquad + B\left(L_0 + h; V\left(\theta_0, L_0^{+}, \omega'\right)\right) - B\left(L_0; V\left(\theta_0, L_0^{+}, \omega'\right)\right) \Big\rangle H\left(d\omega' \mid \omega\right) \\
&= \left(\omega\theta_0 + \beta \int_{\Omega} \delta^{\omega'} U\left(\omega'\right) H\left(d\omega' \mid \omega\right)\right) \cdot h \\
&\quad + \beta \int_{\Omega} \Bigg\langle \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0 + h, V\left(\theta_0, L_0^{+}, \omega'\right)\right)\right] - \mathscr{S}\left[\theta_0, \mathscr{L}\left(L_0, V\left(\theta_0, L_0^{+}, \omega'\right)\right)\right] \\
&\qquad\qquad + h \cdot \left(1 - \delta^{\omega'}\right) \lambda_1^{\omega'} \int_{V\left(\theta_0, L_0^{+}, \omega'\right)}^{+\infty} v dF\left(v \mid \omega'\right) \Bigg\rangle H\left(d\omega' \mid \omega\right),
\end{aligned}
\tag{27}
$$

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

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

Now dividing through by h in (27) and (28), and invoking continuity w.r.t. V of $\mathscr{L}_L\left(L,V\right)=\left(1-\delta^{\omega'}\right)\left(1-\lambda_1^{\omega'}\overline{F}\left(V\right)\right)$ (by continuity of F), everywhere right-differentiability of $\mathscr{S}$ w.r.t. L (by convexity of $\mathscr{S}$), 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 $\frac{1}{h}\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\left(\theta_0, L_0+h\right)-\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\left(\theta_0, L_0^+\right)\right]$ exhibited in (27) and (28) both converge to the same limit as $h\to 0^+$, which, together with continuity of $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ in L at L_0 which implies $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\left(\theta_0, L_0^+\right)=\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\left(\theta_0, L_0\right)$, establishes right-differentiability of $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ w.r.t L with the following expression for $\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\right]_{L,r}\left(\theta, L\right)$

$$
\begin{aligned}
\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\right]_{L,r}\left(\theta, L\right)&=\omega\theta+\beta\int_\Omega\delta^{\omega'}U\left(\omega'\right)H\left(d\omega'\mid\omega\right)\\
&\quad+\beta\int_\Omega\Bigg\langle\mathscr{S}_{L,r}\left[\theta,\mathscr{L}\left(L, V\left(\theta, L^+,\omega'\right)\right)\right]\cdot\mathscr{L}_L\left(L, V\left(\theta, L^+,\omega'\right)\right)\\
&\qquad\qquad+\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{V\left(\theta, L^+,\omega'\right)}^{+\infty}v\,dF\left(v\mid\omega'\right)\Bigg\rangle H\left(d\omega'\mid\omega\right).
\end{aligned}
\tag{29}
$$

Straightforward inspection shows that $\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\right]_{L,r}\left(\theta, L\right)>0$, so that $\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}$ is increasing in L. We now show that $\left[\mathbf{M}^{\mathbf{\Lambda}}\mathscr{S}\right]_{L,r}\left(\theta, L\right)$ is increasing in L and θ. It is sufficient to show that the term under the $\int$ in (29) 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 $V_k=V\left(\theta, L_k^+,\omega'\right)$ for

$k = 1, 2$. Because V is increasing in L, $V_2 \geq V_1$. Then:

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

where the last equality stems from the definition of $\mathscr{L}_L$. Because $\mathscr{S}_{L,r}$ and $\mathscr{L}$ are both increasing in L, and because $\mathscr{L}$ is also increasing in V, the first term in the r.h.s. of the last equality above is positive. Finally, convexity of $\mathscr{S}$ combined with the first-order condition (26) implies that $\mathscr{S}_{L,r}\left[\theta, \mathscr{L}\left(L_2, V_2\right)\right] \geq \mathscr{S}_{L,\ell}\left[\theta, \mathscr{L}\left(L_2, V_2\right)\right] \geq V_2$, so that $\mathscr{S}_{L,r}\left[\theta, \mathscr{L}\left(L_2, V_2\right)\right] \geq v$ for all $v \leq V_2$, implying that the integral term is nonnegative. This shows that $\left[\mathbf{M}^{\boldsymbol{\Lambda}}\mathscr{S}\right]_{L,r}$ is (strictly) increasing in L. The proof that $\left[\mathbf{M}^{\boldsymbol{\Lambda}}\mathscr{S}\right]_{L,r}$ is strictly increasing in θ proceeds along similar lines (details available upon request). Thus $\mathbf{M}^{\boldsymbol{\Lambda}}\mathscr{S}$ 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). □

Now consider the set of functions defined over $\left[\underline{\theta}, \overline{\theta}\right] \times [0,1] \times \Omega$ that are continuous in (θ, L) and call it $C_{\left[\underline{\theta}, \overline{\theta}\right] \times [0,1] \times \Omega}$. 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 $C'_{\left[\underline{\theta}, \overline{\theta}\right] \times [0,1] \times \Omega} \subset C_{\left[\underline{\theta}, \overline{\theta}\right] \times [0,1] \times \Omega}$ 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 $C'_{\left[\underline{\theta}, \overline{\theta}\right] \times [0,1] \times \Omega}$ is closed in $C_{\left[\underline{\theta}, \overline{\theta}\right] \times [0,1] \times \Omega}$ under the sup norm.[16]

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

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

[16]While 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.

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. □

Lemma 3 *Let $X \subset \mathbb{R}^2$ be a convex set and $f_n : X \to \mathbb{R}$, $N \in \mathbb{N}$ 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. □

Corollary 1 *The set $C'_{[\underline{\theta}, \overline{\theta}] \times [0,1] \times \Omega}$ of functions defined over $[\underline{\theta}, \overline{\theta}] \times [0, 1] \times \Omega$ that are increasing and convex in L and have increasing differences in (θ, L) is a closed subset of $C_{[\underline{\theta}, \overline{\theta}] \times [0,1] \times \Omega}$ 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 $\mathbf{M}^\Lambda$ defined in (24) maps $C'_{[\underline{\theta}, \overline{\theta}] \times [0,1] \times \Omega}$ into itself and is a contraction of modulus β under the sup norm.*

Proof. That $\mathbf{M}^{\Lambda}$ maps $C'_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$ 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 (24) that $\mathbf{M}^{\Lambda}$ satisfies Blackwell's sufficient conditions with modulus β. □

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

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

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

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

If an equilibrium exists, then a firm has a best response and a value S which solves $S = \mathbf{M}S$. For every $\Lambda \in C_{[\underline{\theta},\overline{\theta}]}$, by definition of $\mathbf{M}$ and $\mathbf{M}_{\Lambda}$ this implies $S = \mathbf{M}_{\Lambda}S$. Since the fixed point of $\mathbf{M}_{\Lambda}$ is unique, if $S = \mathbf{M}S$ exists then for every fixed $\Lambda \in C_{[\underline{\theta},\overline{\theta}]}$ we have for all $(\theta, L, \omega) \in \left[\underline{\theta},\overline{\theta}\right] \times [0,1] \times \Omega$: $S\left(\theta, L, \omega, \Lambda\right) = \mathscr{S}^{\star}_{\Lambda}\left(\theta, L, \omega\right)$. Therefore, if the value function S and an equilibrium of the contract-posting game exist, then $S \in C'_{[\underline{\theta},\overline{\theta}]\times[0,1]\times\Omega}$: 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 $V\left(\theta, L, \omega, \Lambda\right)$ 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 $\mathscr{L}$ 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. □

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:

$$V^{\star}(\theta,\omega) := V(\theta, L^{\star}(\theta), \omega, \Lambda)$$

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

The main purpose of Proposition 3 is actually to establish claim 2, continuous differentiability of $V^{\star}$. Our proof strategy is as follows. We know from Proposition 2 that the optimal policy $V^{\star}$ 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 $V(\theta, L)$ in θ, both for fixed L and for $L = L^{\star}(\theta)$, and in L at $L = L^{\star}(\theta)$ 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 $L = L^{\star}(\theta)$; 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 $V^{\star}$ is Lipschitz continuous in θ. This implies that $V^{\star}$ 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. *$\theta \mapsto V(\theta, L^{\star}(\theta), \omega, \Lambda) = V^{\star}(\theta,\omega)$ is continuous;*
2. *$L \mapsto V(\theta, L, \omega, \Lambda)$ is continuous at $L = L^{\star}(\theta)$;*
3. *$\tau \mapsto V(\tau, L^{\star}(\theta), \omega, \Lambda)$ is continuous at $\tau = \theta$.*

Proof. $\theta \mapsto V^{\star}(\theta,\omega)$ is increasing by Proposition 2, so $V^{\star}$ can only have (countably many) jump discontinuities. But then a jump discontinuity in $V^{\star}$ 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 $V^{\star}$ (Lemma 5), $\exists\, \alpha > 0$: $\forall \eta \in (0, \alpha]$, $V^{\star}(\theta,\omega) \leq V^{\star}(\theta+\eta,\omega) \leq V^{\star}(\theta,\omega) + \varepsilon$. But then monotonicity of V in L and in θ (see Appendix A) further implies: $V^{\star}(\theta,\omega) \leq V(\theta, L^{\star}(\theta+\eta), \omega, \Lambda) \leq V^{\star}(\theta+\eta,\omega) \leq V^{\star}(\theta,\omega)+\varepsilon$, so that $\forall\, L \in [L^{\star}(\theta), L^{\star}(\theta+\eta)]$, $V(\theta, L, \omega, \Lambda) - V(\theta, L^{\star}(\theta), \omega, \Lambda) \leq$

ε, which establishes right-continuity of V in L at $L^\star(\theta)$. Left-continuity is established in the same way, and so is continuity of $\tau \mapsto V(\tau, L^\star(\theta), \omega, \Lambda)$ at $\tau = \theta$. $\square$

We now go on to establish point 1 of the proposition. 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 $S_{L,r}(\theta, L^+, \omega) = \lim_{h\to 0+} S_{L,r}(\theta, L+h, \omega)$, and symmetrically $S_{L,\ell}(\theta, L^-, \omega) = \lim_{h\to 0+} S_{L,\ell}(\theta, L-h, \omega)$. Now following exactly the same steps as in (27) and (28) (see the proof of Lemma 1 in Appendix A), only applied to S, we establish:

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

where the dependence of various functions on the aggregate state variable Λ was kept implicit to avoid notational overload. Next, the facts that V is increasing in L (see the proof of Proposition 2) and continuous in L at $L = L^\star(\theta)$ (from Lemma 5), combined with continuity of $\mathscr{L}$ and $\mathscr{L}_L$ w.r.t. V (by continuity of F), imply that $\mathscr{L}_L(L, V(\theta, L^+, \omega'), \omega') = \mathscr{L}_L(L, V(\theta, L, \omega'), \omega')$ and $S_{L,r}[\theta, \mathscr{L}(L, V(\theta, L^+, \omega'), \omega'), \omega'] = S_{L,r}\left[\theta, \mathscr{L}(L, V(\theta, L, \omega'), \omega')^+, \omega'\right]$ at $L = L^\star(\theta)$. As a further consequence:

$$\begin{aligned}S_{L,r}\left(\theta, L^\star(\theta)^+, \omega\right) &= \omega\theta + \beta \int_\Omega \delta^{\omega'} U\left(\omega'\right) H\left(d\omega' \mid \omega\right) \\ &\quad + \beta \int_\Omega \Big\langle S_{L,r}\left[\theta, \mathscr{L}\left(L, V^\star\left(\theta, \omega'\right), \omega'\right)^+, \omega'\right] \cdot \mathscr{L}_L\left(L, V^\star\left(\theta, \omega'\right), \omega'\right) \\ &\quad\quad + \left(1-\delta^{\omega'}\right)\lambda_1^{\omega'} \int_{V^\star(\theta, \omega')}^{+\infty} v dF\left(v \mid \omega'\right)\Big\rangle H\left(d\omega' \mid \omega\right). \quad (30)\end{aligned}$$

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

$$\begin{aligned}0 \le \mathscr{D}_{S_L}\left(\theta, L^\star(\theta), \omega\right) &= \beta \int_\Omega \mathscr{D}_{S_L}\left[\theta, \mathscr{L}\left(L, V^\star\left(\theta, \omega'\right), \omega'\right), \omega'\right] \cdot \mathscr{L}_L\left(L, V^\star\left(\theta, \omega'\right), \omega'\right) H\left(d\omega' \mid \omega\right) \\ &< \beta \int_\Omega \mathscr{D}_{S_L}\left[\theta, \mathscr{L}\left(L, V^\star\left(\theta, \omega'\right), \omega'\right), \omega'\right] H\left(d\omega' \mid \omega\right).\end{aligned}$$

At this point, if we can prove that $\mathscr{D}_{S_L}$ is uniformly bounded above by some $K > 0$, then iterating the last inequality will show that $0 \le \mathscr{D}_{S_L}(\theta, L^\star(\theta), \omega) < \beta^n K$ for all $n \in \mathbb{N}$, which implies that $\mathscr{D}_{S_L}(\theta, L^\star(\theta), \omega) = 0$ 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 $\mathscr{D}_{S_L}$ 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 $S_{L,r}(\theta, L^\star(\theta), \omega) \geq S_{L,\ell}(\theta, L^\star(\theta), \omega) \geq V^\star(\theta, \omega) \geq U(\omega)$ again invoking convexity in conjunction with the FOC (26):

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

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 $S_{L,r}(\theta, L^\star(\theta), \omega)$. 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, $W \in \left[V^\star\left(\underline{\theta}, \omega'\right), V^\star\left(\overline{\theta}, \omega'\right)\right]$. Since $V^\star$ is continuous and increasing, then offering any such best response W is equivalent to choosing a type τ to imitate such that $W = V^\star(\tau, \omega')$. In any RPE, by Proposition 2, the best response by a firm θ of current size $L^\star(\theta)$ is 'truthful revelation', $\tau^\star = \theta$, which solves

$$
\begin{aligned}
S(\theta, L^\star(\theta), \omega) = {} & A(\theta, L^\star(\theta), \omega) \\
& + \beta \int_\Omega \max_{\tau(\omega')} \left\langle S\left[\theta, \mathscr{L}\left(L^\star(\theta), \omega', \tau\left(\omega'\right)\right), \omega'\right] + B\left(L^\star(\theta), \omega', \tau\left(\omega'\right)\right) \right\rangle H\left(d\omega' \mid \omega\right)
\end{aligned}
$$

where, with a slight abuse of notation:

$$
\begin{aligned}
\mathscr{L}\left(L, \omega', \tau\right) = {} & L\left(1 - \delta^{\omega'}\right)\left(1 - \lambda_1^{\omega'} \overline{F}\left(V^\star\left(\tau, \omega'\right) \mid \omega'\right)\right) \\
& + \lambda_0^{\omega'}\left(1 - \Lambda\left(\overline{\theta}\right)\right) + \lambda_1^{\omega'}\left(1 - \delta^{\omega'}\right) \Lambda\left(\overline{\theta}\right) G\left(V^\star\left(\tau, \omega'\right) \mid \omega'\right)
\end{aligned}
$$

and

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

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

$$\mathscr{L}\left(L^{\star}(\theta),\omega',\tau\right)=L^{\star}(\theta)\left(1-\delta^{\omega'}\right)\left(1-\lambda_1^{\omega'}\overline{\Gamma}(\tau)\right)+\lambda_0^{\omega'}\left(1-\Lambda^{\star}\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}(\tau) \quad (31)$$

$$B\left(L^{\star}(\theta),\omega',\tau\right)=L^{\star}(\theta)\left(1-\delta^{\omega'}\right)\lambda_1^{\omega'}\int_{\tau}^{\overline{\theta}}V^{\star}\left(x,\omega'\right)d\Gamma(x) - V^{\star}\left(\tau,\omega'\right)\left(\lambda_0^{\omega'}\left(1-\Lambda\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}(\tau)\right).$$

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

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

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

Rearranging:

$$\begin{aligned}
&\left[V^{\star}\left(\theta,\omega'\right)-V^{\star}\left(\theta-h,\omega'\right)\right]\cdot\left[\lambda_0^{\omega'}\left(1-\Lambda\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}(\theta)\right]\\
&\le S\left[\theta,\mathscr{L}\left(L^{\star}(\theta),\theta,\omega'\right),\omega'\right]-S\left[\theta,\mathscr{L}\left(L^{\star}(\theta),\theta-h,\omega'\right),\omega'\right]\\
&-\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{L^{\star}(\theta)\int_{\theta-h}^{\theta}V^{\star}\left(x,\omega'\right)d\Gamma(x)+V^{\star}\left(\theta-h,\omega'\right)\cdot\left[\Lambda^{\star}(\theta)-\Lambda^{\star}(\theta-h)\right]\right\}\\
&\le S\left[\theta,\mathscr{L}\left(L^{\star}(\theta),\theta,\omega'\right),\omega'\right]-S\left[\theta,\mathscr{L}\left(L^{\star}(\theta),\theta-h,\omega'\right),\omega'\right]\\
&-V^{\star}\left(\theta-h,\omega'\right)\cdot\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{L^{\star}(\theta)\left[\Gamma(\theta)-\Gamma(\theta-h)\right]+\left[\Lambda^{\star}(\theta)-\Lambda^{\star}(\theta-h)\right],\right\}
\end{aligned} \quad (32)$$

where the second inequality is obtained by minorizing the integral term, remarking that $V^{\star}$ is increasing in θ. Now differentiability of S w.r.t. L and the definition (31) together imply that:[17]

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

Substituting into (32), we obtain:

$$\begin{aligned}
&\left[V^{\star}\left(\theta,\omega'\right)-V^{\star}\left(\theta-h,\omega'\right)\right]\cdot\left[\lambda_0^{\omega'}\left(1-\Lambda\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}(\theta)\right]\\
&\qquad\le\left\{S_L\left[\theta,\mathscr{L}\left(L^{\star}(\theta),\theta,\omega'\right),\omega'\right]-V^{\star}\left(\theta-h,\omega'\right)\right\}\\
&\qquad\times\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\left\{L^{\star}(\theta)\left[\Gamma(\theta)-\Gamma(\theta-h)\right]+\left[\Lambda^{\star}(\theta)-\Lambda^{\star}(\theta-h)\right]\right\}+o(h)
\end{aligned}$$

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

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

$$0 \leq D^{-}V^{\star}(\theta) \leq 2\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)L^{\star}(\theta)\gamma(\theta)\frac{S_L\left[\theta,\mathscr{L}\left(L^{\star}(\theta),\theta,\omega'\right),\omega'\right]-V^{\star}(\theta,\omega')}{\lambda_0^{\omega'}\left(1-\Lambda\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}(\theta)}$$

(the first inequality is a direct consequence of $V^{\star}$ 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 $V^{\star}$ 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). □

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

$$V_{\theta}^{\star}\left(\theta,\omega'\right) = 2\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)L^{\star}(\theta)\gamma(\theta)\frac{S_L\left[\theta,\mathscr{L}\left(L^{\star}(\theta),\theta,\omega'\right),\omega'\right]-V^{\star}(\theta,\omega')}{\lambda_0^{\omega'}\left(1-\Lambda\left(\overline{\theta}\right)\right)+\lambda_1^{\omega'}\left(1-\delta^{\omega'}\right)\Lambda^{\star}(\theta)} := q(\theta).$$

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

$$\tilde{V}_{\theta}(\theta) := \begin{cases} V_{\theta}^{\star}(\theta,\omega') & \theta \in \left[\underline{\theta},\overline{\theta}\right]\backslash N_V \\ q(\theta) & \text{otherwise} \end{cases}$$

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

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

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

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

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

we can rewrite the Euler equation (30) as follows:

$$\mu(\theta,\omega_t,\Lambda_t)=\omega_t\theta+\beta\mathscr{U}(\theta,\omega_t,\Lambda_t)+\beta\mathbf{E}_{\omega_{t+1}|\omega_t}\left[\Delta(\theta,\omega_{t+1})\,\mu(\theta,\omega_{t+1},\Lambda_{t+1})\right]. \tag{33}$$

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

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

After n forward substitutions, we can write (33) as

$$\mu(\theta,\omega_t,\Lambda_t)=\mu_n(\theta,\omega_t,\Lambda_t)+\beta^{n+1}X_1^{n+1,t}[\mu]$$

where

$$\mu_n(\theta,\omega_t,\Lambda_t)=\sum_{s=0}^{n}\beta^s\left\{X_1^{s+1,t}[\omega\theta]+\beta X_1^{s+1,t}[\mathscr{U}(\theta,\omega,\Lambda)]\right\}.$$

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

$$\begin{aligned}0<&\mu(\theta,\omega_t,\Lambda_t)-\mu_n(\theta,\omega_t,\Lambda_t)=|\mu(\theta,\omega_t,\Lambda_t)-\mu_n(\theta,\omega_t,\Lambda_t)|=\beta^{n+1}X_1^{n+1,t}[\mu]\\ \leq&\beta^{n+1}\mathbf{E}_{\omega_{t+1}|\omega_t}\left[\mathbf{E}_{\omega_{t+2}|\omega_{t+1}}\left[\cdots\mathbf{E}_{\omega_{t+n}|\omega_{t+n-1}}\left[\mu(\theta,\omega_{t+n},\Lambda_{t+n})\right]\right]\right]\\ &=\beta^{n+1}\mathbf{E}_{\omega_{t+n}|\omega_t}\left[\mu(\theta,\omega_{t+n},\Lambda_{t+n})\right]\end{aligned}$$

Since a firm can always guarantee itself positive profits and employment by offering its workers the value of unemployment, then $L^\star(\theta,\omega,\Lambda)$ is bounded away from 0 with probability one. So the TVC (21) implies that, as $n\to\infty$, the last term vanishes, thus μ_n converges pointwise (and indeed uniformly) to μ.

Next, taking derivatives

$$\begin{aligned}\frac{\partial\mu_n(\theta,\omega_t,\Lambda_t)}{\partial\theta}=&\,\omega_t+\frac{\lambda_1^\omega\gamma(\theta)}{1-\lambda_1^\omega\overline{\Gamma}(\theta)}+\sum_{s=1}^{n}\beta^sX_1^{s,t}\left[\omega+\frac{\lambda_1^\omega\gamma(\theta)}{1-\lambda_1^\omega\overline{\Gamma}(\theta)}\omega\theta\right]\\ &+\sum_{s=2}^{n+1}\beta^sX_1^{s,t}\left[\frac{\lambda_1^\omega\gamma(\theta)}{1-\lambda_1^\omega\overline{\Gamma}(\theta)}\mathscr{U}(\theta,\omega,\Lambda)-(1-\delta^\omega)\lambda_1^\omega V^\star(\theta\mid\omega,\Lambda)\gamma(\theta)\right]\end{aligned}$$

which is continuous in θ. As the arguments of the operator $X_1^{s,t}$ in the last expression are continuous in ω,θ on the compact set $\Omega\times\left[\underline{\theta},\overline{\theta}\right]$, the $X_1^{s,t}[\cdot]$ 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 $n\to\infty$, and

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

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

$$\begin{aligned}\frac{\partial\mu_n(\theta,\omega_t,\Lambda_t)}{\partial\theta}-\frac{\partial\mu_\infty(\theta,\omega_t,\Lambda_t)}{\partial\theta}=&-\sum_{s=n+1}^{\infty}\beta^sX_1^{s,t}\left[\omega+\frac{\lambda_1^\omega\gamma(\theta)}{1-\lambda_1^\omega\overline{\Gamma}(\theta)}\omega\theta\right]\\ &-\sum_{s=n+2}^{\infty}\beta^sX_1^{s,t}\left[\frac{\lambda_1^\omega\gamma(\theta)}{1-\lambda_1^\omega\overline{\Gamma}(\theta)}\mathscr{U}(\theta,\omega,\Lambda)-(1-\delta^\omega)\lambda_1^\omega V^\star(\theta\mid\omega,\Lambda)\gamma(\theta)\right]\end{aligned}$$

So, for all $\theta, \omega_t, \Lambda_t$, $\left|X_1^{s,t}[\cdot]\right| < X'$ 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. □

C Details of the simulation algorithm

This appendix explicates some details of the projection method that we use to solve for the shadow values μ and π in the PDE system (22,23) in each aggregate state ω. As explained in the main text, the idea is to approximate cross-state jumps in those shadow values by a known function J. Specifically, we rewrite (22) as:

$$\begin{aligned}\frac{\partial \mu_t}{\partial \theta}(\theta \mid \omega) = \omega + \beta(1-\delta^\omega)\left\{\lambda_1^\omega \gamma(\theta)\pi_{t+1}(\theta \mid \omega) + \frac{\partial \mu_{t+1}}{\partial \theta}(\theta \mid \omega)\left(1 - \lambda_1^\omega \overline{\Gamma}(\theta)\right)\right\} \\ + \beta\sigma^\omega\left\{\left(1-\delta^{\omega'}\right)\left[\lambda_1^{\omega'}\gamma(\theta)\pi\left(\theta \mid \omega'\right) + \frac{\partial \mu}{\partial \theta}\left(\theta \mid \omega'\right)\left(1 - \lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\right]\right. \\ \left. -(1-\delta^\omega)\left[\lambda_1^\omega \gamma(\theta)\pi(\theta \mid \omega) + \frac{\partial \mu}{\partial \theta}(\theta \mid \omega)\left(1 - \lambda_1^\omega \overline{\Gamma}(\theta)\right)\right]\right\}. \end{aligned} \quad (34)$$

The "jump term" that we need to approximate is the last term in curly brackets in the latter equation:

$$\begin{aligned}&\left(1-\delta^{\omega'}\right)\left[\lambda_1^{\omega'}\gamma(\theta)\pi\left(\theta \mid \omega'\right) + \frac{\partial \mu}{\partial \theta}\left(\theta \mid \omega'\right)\left(1 - \lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\right] \\ &\quad -(1-\delta^\omega)\left[\lambda_1^\omega \gamma(\theta)\pi(\theta \mid \omega) + \frac{\partial \mu}{\partial \theta}(\theta \mid \omega)\left(1 - \lambda_1^\omega \overline{\Gamma}(\theta)\right)\right] \\ &\qquad \simeq J\left(\theta, L^\star(\theta), \omega, \Lambda(\theta) \mid \mathbf{a}\right).\end{aligned}$$

Given a specific vector of coefficients $\mathbf{a}$, system (23,34), together with the transversality condition (21), becomes a pair of independent systems of PDEs, one for each aggregate state, which can be separately numerically solved over the infinite future for any initial value of $(\Lambda(\cdot), L^\star(\cdot))$ using the algorithm described in MPV08.

We thus proceed in the following steps:

0. Pick an initial state of the economy $(\omega_0, \Lambda_0(\cdot), L_0^\star(\cdot))$ and simulate a path of ω. Denote switching dates as $(s_1, s_2, \dots)$.

1. Fix a parameter **a**.

2. Given the choice of **a** made at step 1 and the implied J-function, solve (23,34,21) using the appropriate initial conditions. More specifically:

   (a) Solve (23,34,21) with initial condition $(\Lambda_0(\cdot), L_0^\star(\cdot))$ as if state ω_0 prevailed forever. This implies certain values for $\partial\mu_0(\theta,\omega_0)/\partial\theta$, $\partial\mu_{s_1}(\theta,\omega_0)/\partial\theta$, $\pi_0(\theta,\omega_0)$, $\pi_{s_1}(\theta,\omega_0)$ and $\left(\Lambda_{s_1}(\cdot), L_{s_1}^\star(\cdot)\right)$ at date s_1 when the next aggregate shock occurs.

   (b) Solve (23,34,21) with initial condition $\left(\Lambda_{s_1}(\cdot), L_{s_1}^\star(\cdot)\right)$ as if state ω_1 prevailed over $t \in [s_1, +\infty)$. This implies certain values for $\partial\mu_{s_1}(\theta,\omega_1)/\partial\theta$, $\partial\mu_{s_2}(\theta,\omega_1)/\partial\theta$, $\pi_{s_1}(\theta,\omega_1)$, $\pi_{s_2}(\theta,\omega_1)$ and $\left(\Lambda_{s_2}(\cdot), L_{s_2}^\star(\cdot)\right)$.

   (c) Solve (23,34,21) with initial condition $\left(\Lambda_{s_2}(\cdot), L_{s_2}^\star(\cdot)\right)$ as if state ω_2 prevailed over $t \in [s_2, +\infty)$, etc. That is, repeat step 2, mutatis mutandis, for the first K jumps in ω (in practice with a two-state process for ω, two jumps — one up, one down — are enough).

3. The simulations performed at stage 3 provide a vector of jumps in $\partial\mu/\partial\theta$:

$$\left(1-\delta^{\omega'}\right)\left[\lambda_1^{\omega'}\gamma(\theta)\,\pi_{s_k}\left(\theta \mid \omega'\right) + \frac{\partial\mu_{s_k}}{\partial\theta}\left(\theta \mid \omega'\right)\left(1-\lambda_1^{\omega'}\overline{\Gamma}(\theta)\right)\right]$$
$$-\left(1-\delta^{\omega}\right)\left[\lambda_1^{\omega}\gamma(\theta)\,\pi_{s_k}(\theta \mid \omega) + \frac{\partial\mu_{s_k}}{\partial\theta}(\theta \mid \omega)\left(1-\lambda_1^{\omega}\overline{\Gamma}(\theta)\right)\right],$$
$$k = 1, \cdots, K. \quad (35)$$

   Compare those with the jumps predicted from the initially chosen function $J(\cdot \mid \mathbf{a})$ and the simulated path of $(\omega, L^\star(\cdot), \Lambda(\cdot))$. If different, update **a** and start over at step 1. Exactly how **a** is updated depends on the chosen functional form for the approximate jump function J. In practice we use a projection on polynomials,[18] and the updated vector of coefficients **a** is obtained by regression of the "simulated jumps" in (35) on the elements of J.

[18] After many trials, a good compromise between accuracy and speed of convergence was found using projection on quartics in $\Lambda(\theta)$ and $\theta \times L(\theta)$. With that specification the root mean squared prediction error is in the order of 1/100th of a percent of the mean (absolute value) simulated jump.