Uncertainty and Unemployment*

Edouard Schaal†
New York University

August 2015

Abstract

This paper studies the impact of time-varying idiosyncratic risk at the establishment level on aggregate unemployment fluctuations and on the labor market over the period 1972-2009. I build a tractable search-and-matching model of the labor market with firm dynamics and heterogeneity in productivity and sizes, in which I introduce time-varying idiosyncratic volatility. The model features directed search and allows for endogenous separations, entry and exit of establishments, and job-to-job transitions. I show, first, that the model can replicate salient features of the behavior of firms at the microeconomic level. Second, I find that the introduction of time-varying idiosyncratic volatility improves the fit of search-and-matching models for a range of business cycle moments. In a series of counterfactual experiments, I then show that time-varying idiosyncratic risk is important to account for the magnitude of fluctuations in aggregate unemployment for past US recessions, including in particular the recessions of 1990-1991 and 2001. Though the model can account for about 40% of the total increase in unemployment for the 2007-2009 recession, uncertainty alone does not seem sufficient to explain the magnitude and persistence of unemployment observed during that period.

JEL Classification: E24, E32, D83

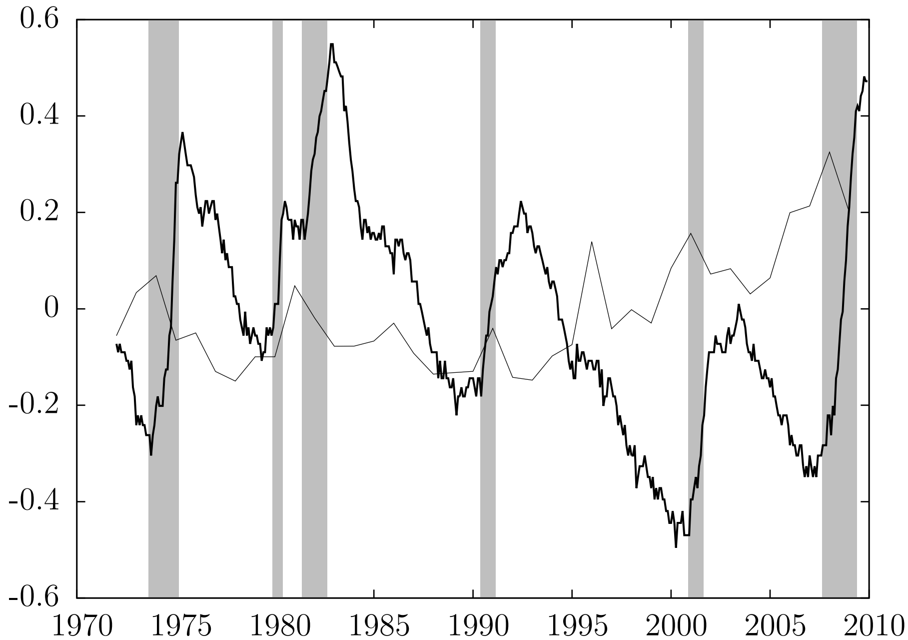
*I am greatly indebted to Nobuhiro Kiyotaki, Esteban Rossi-Hansberg and Oleg Itskhoki for their invaluable advice and support. I also wish to thank Andy Atkeson, Giuseppe Bertola, Nick Bloom, Pablo Fajgelbaum, Guido Menzio, Jean Marc Robin, Gilles Saint-Paul, Eric Smith, Mathieu Taschereau-Dumouchel, Gianluca Violante, seminar participants at several institutions, and three anonymous referees for their helpful suggestions. I gratefully acknowledge financial support from the Center for Economic Policy Studies at Princeton University. This paper previously circulated under the title "Uncertainty, Productivity and Unemployment in the Great Recession".

†Email: edouard.schaal@nyu.edu; Address: 19 W 4th St., 6Fl., New York, NY 10012.

1 Introduction

The recession that followed the 2007-2008 collapse of the financial markets resulted in one of the deepest downturns in post-war US labor markets. While GDP contracted by up to 6.8% in the fourth quarter of 2008, the unemployment rate grew from 5% in January 2008 to 10% in October 2009 according to the Bureau of Labor Statistics (BLS).

A large and growing body of literature has advanced the hypothesis that the heightened level of uncertainty over the period 2007-2012, as suggested by various measures at the macro[1] and micro[2] levels, may be partly responsible for the unusual magnitude and persistence of the slump. This paper examines to what extent fluctuations in uncertainty over the business cycle can shed light on the US labor market experience over various past recessions, including the Great Recession of 2007-2009.



Notes: Data are shown in log deviations from their long-run averages. The thick curve shows seasonally adjusted civilian unemployment from the BLS; the thin curve displays the interquartile range of establishment-level TFP shocks constructed by Bloom et al. (2014) from the Census of Manufactures and the Annual Survey of Manufactures. Shaded areas correspond to NBER recessions. See Appendix C for details.

Figure 1: Unemployment and establishment-level volatility in TFP

Uncertainty is a broad concept that encompasses notions as diverse as risk and ambiguity. This

[1] Some typical measures of aggregate uncertainty are the volatility of aggregate TFP as measured by a GARCH model, aggregate stock market volatility, survey-based measures of disagreement in forecasts or ex-ante forecast errors over aggregate variables such as output or inflation. Other more recent contributions include the measures proposed by Jurado et al. (2014) using common factor analysis or the news-based index of economic and policy uncertainty of Baker et al. (2014).

[2] Measures of micro-level risk suggested in the literature include the volatility of establishment-level TFP constructed by Bloom et al. (2014) using Census data, the cross-sectional dispersion of firm-level growth rates in sales from Compustat, the implied and realized stock market volatility as captured in the VIX/VXO series. Figure 3 in the Appendix compares these series.

paper focuses on a particular source of uncertainty, namely time-varying establishment-level volatility in TFP.[3] While being largely overlooked in labor market studies, fluctuations in micro-level volatility are large in the data. This suggests that volatility may be an important determinant of employment decisions and labor market reallocation. Figure 1 compares the evolution of the US unemployment rate to a measure of establishment-level volatility constructed from Census data by Bloom et al. (2014) over 1972-2009. Constructed using annual input-output data from the Census of Manufactures and the Annual Survey of Manufactures, this series presents the cross-sectional interquartile range (IQR) of innovations to establishment-level TFP, estimated from an AR(1) process.[4] Two important facts can be drawn from that figure. First, as it has been widely noted in previous literature using other measures of micro-level volatility,[5] idiosyncratic risk is countercyclical and rises during recessions. This particular measure peaks early as the economy enters a downturn and then declines relatively quickly as the recession unfolds. Second, fluctuations in idiosyncratic risk are large with an all-time high of 30% from its long-run average at the end of 2007.

How does uncertainty affect the level of economic activity? How does it contribute to aggregate unemployment fluctuations? Several channels have been put forward in the literature. The first one, on which a large part of the literature has focused, is the *real option* channel (Bernanke (1983); McDonald and Siegel (1986); Dixit and Pindyck (1994)). Firms usually face a substantial amount of uncertainty when making major investments decisions, such as buying new equipment, purchasing land and buildings, or expanding their workforce. These decisions frequently entail large sunk costs that are, at least partially, irreversible. The interaction of irreversibility and uncertainty creates an option value of waiting. In times of heightened uncertainty, firms have an incentive to postpone investment until conditions improve in order to avoid costly mistakes. A second important channel is the *risk premium* channel (Arellano et al. (2010); Gilchrist et al. (2010)). When uncertainty is high, risk premia rise: the cost of external financing increases and the ability of firms to undertake large investments or expand is reduced. A third channel is the *risk aversion* or *precautionary motive* channel. Because of risk aversion, investors and managers may turn away from risky, high return projects, potentially resulting in low growth and slow recovery. These precautionary motives may further negatively affect an economy subject to nominal rigidities as households may decide to cut on their consumption (Basu and Bundick (2012); Fernández-Villaverde et al. (2013)).

Employment decisions display several features likely to produce large real option effects: they typically involve important sunk costs (advertisement, search, screening and training costs); employment contracts are usually long-term relationships that cannot be easily reversed, both because of frictions and institutional reasons (labor contracts, dismissal costs, etc.). Because of these characteristic fea-

[3] In particular, I do not consider macro-level risk. A previous version of this paper experimented with the effects of time-varying volatility in aggregate TFP. However, because volatility in aggregate TFP is small in the data, its quantitative impact was negligible, while significantly increasing the computational cost. This finding is consistent with Leduc and Liu (2012) and, more recently, Backus et al. (2014).

[4] The estimation controls for time and plant-level fixed effects and 4-digit price deflators. See Appendix C for additional details and discussion.

[5] See Figure 3 in the Appendix for a comparison with other measures.

tures of labor markets, I focus my analysis on real option effects. In times of high uncertainty, hiring is risky because it is costly, and because productivity may revert quickly. As a result, the *option value of waiting* increases and firms should delay hiring. Hence, high uncertainty may induce a drop in the number of vacancies and in the job finding rate, ultimately resulting in a rise in unemployment.

This, however, only captures part of the story. First, not only hirings but also separations should be subject to an option value: with higher uncertainty, firms should become more reluctant to separate from their workforce, as it would be costly to search for new workers in the case of a rise in future productivity. The combined effect of lower hiring and separation flows on unemployment is thus ambiguous. Moreover, volatility shocks are known in the literature to produce countervailing effects that could affect the response of unemployment. For instance, volatility, by raising the actual dispersion across establishments, tends to increase the reallocation in what is commonly labeled as *realized volatility* effects: even though the option value of waiting rises with volatility, the shocks are larger and firms end up adjusting their workforce more often. To evaluate the impact of uncertainty on labor market fluctuations, I thus propose an equilibrium model that allows to disentangle and quantify these opposite forces, as well as understand the importance of general equilibrium effects and other characteristic features associated to time-varying risk.

The concept of establishment is often absent from search-and-matching models. In order to address the relevant aspects of the response of the labor market to establishment-level risk, I develop an equilibrium search-and-matching model with firm dynamics and heterogeneity in productivity and sizes. The concept of establishment is introduced through the assumption of decreasing returns to scale. The model allows for aggregate productivity shocks and time-varying volatility in idiosyncratic productivity. Despite being a large heterogeneous agent economy, the model retains its tractability, and dynamics can be easily computed. The model is estimated by simulated method of moments using a set of standard business cycle moments as well as targets drawn from labor market flow data. First, as a validation exercise, I show that the model is consistent with a range of cross-sectional and establishment-level facts. Second, I demonstrate that the introduction of time-varying idiosyncratic risk improves the general fit of search-and-matching models for a range of business cycle moments. Then, I analyze and decompose the response of the economy to aggregate productivity and idiosyncratic volatility shocks. A general lesson from my analysis is that search frictions do not lead to strong real option effects. This result stems from the fact that gross US labor market flows are large, implying that the estimated search costs are too low to create strong irreversibilities. My findings suggest, however, that volatility is still a major determinant of labor market flows through its impact on reallocation. For instance, I find that an increase in volatility leads to a large rise in unemployment, due to an increase in layoffs driven by realized volatility effects. This increase is accompanied by a modest rise in hiring, dampened by general equilibrium and real option effects, which turns out insufficient to compensate for the rise in layoffs. I finally run various counterfactual experiments to evaluate the ability of the model to replicate the US experience across all the recessions included in the 1972-2009 period. The model is quite successful in replicating output dynamics in general. In terms of unemployment, the model can account for about 80% of the 1973-1975, 1980-1982 and 2001 episodes, and virtually 100% of the 1990-1991 recession. Idiosyncratic volatility allows to explain up to 40% of the total increase in unemployment in the 2007-2009 recession, but a large fraction of the magnitude and persistence remains unexplained.

Beyond the analysis of the role of idiosyncratic risk on the labor market, this paper contributes to the search-and-matching literature by developing a model of firm dynamics and search frictions that is fully tractable under a rich structure of aggregate shocks. Dynamic models featuring heterogeneous firms usually raise a number of technical issues. In particular, one must keep track of the infinite-dimensional distribution of firms to solve the model. To address this issue, I use the directed search structure of Menzio and Shi (2010, 2011) in order to exploit the convenient property of *block recursivity*, which allows for an easy and complete characterization of the economy's out-of-steady-state dynamics. While they established this property in an environment with single-worker firms only, I show that block recursivity continues to hold with multiworker firms and decreasing returns to scale, under some additional conditions. In particular, a specific trick allows me to obtain this property despite the presence of two-sided heterogeneity. The model features realistic firm dynamics and a rich description of labor markets flows. In the model, heterogeneous firms can endogenously expand/contract and enter/exit over the business cycle. Workers search for new job opportunities both off and on the job. On-the-job search is especially important for quantitative applications to business cycles as it allows distinction between quits and layoffs, which have notably different cyclical properties. In section 3, I show that the model is able to reproduce a range of facts at the establishment and cross-sectional levels. First, it matches a number of features of the micro-level employment policies of establishments in terms of hires, layoffs and quits. It also does reasonably well in replicating the cross-sectional distribution of establishment growth rates as reported in Davis et al. (2006, 2011). Second, turning to the evolution of the cross section of firms over the business cycle, I find that the model is able to explain the finding that large establishments are more cyclically sensitive than small ones, as reported by Moscarini and Postel-Vinay (2012). Finally, I explore the wage predictions of the model in Appendix F and conclude that the model can generate a substantial wage dispersion in line with empirical estimates. The model can also produce a sizeable size-wage differential.

Related literature

This paper is related to several strands in the literature. It first relates to the growing literature on uncertainty-driven business cycles.[6] This paper studies the role of time-varying risk in interaction with search frictions in explaining labor market dynamics, unlike most of the literature with the exception

[6] See, for example, the time-varying volatility and real option value models of Bloom (2009), Bachmann and Bayer (2009), Bloom et al. (2014); the fiscal volatility paper of Fernández-Villaverde et al. (2013); the uncertainty and financial friction models of Christiano et al. (2009), Gilchrist et al. (2010), and Arellano et al. (2010); the New-Keynesian DSGE papers of Basu and Bundick (2012) and Leduc and Liu (2012); the uncertainty and ambiguity aversion paper of Bianchi et al. (2014); or the measurement papers of Bachmann et al. (2010), Baker et al. (2014), Jurado et al. (2014).

of Leduc and Liu (2012) and Lin (2014). The first paper adds search frictions to the New-Keynesian DSGE framework of Basu and Bundick (2012) and concludes that labor market imperfections provide strong amplification to uncertainty shocks. Lin (2014) builds on the more traditional RBC search-and-matching tradition of Merz (1995) and Andolfatto (1996) and finds that uncertainty shocks help explain jobless recovery episodes. While their DSGE frameworks allow for an easier comparison to standard RBC and New-Keynesian models as well as to examine, for instance, the role of uncertainty on inflation, the representative agent approach of these two papers restricts their analysis to macroeconomic uncertainty, whose size and impact are relatively small (see Leduc and Liu (2012)). This paper is more closely related to the firm dynamics and heterogeneous agent literature in the spirit of Hopenhayn (1992). This approach allows me, in particular, to examine the impact of the large empirical fluctuations in establishment-level risk and use micro-data to discipline the quantitative exercise. My model also relates to the rest of the literature that uses non-convex adjustment costs in labor to create a real option value (Bloom (2009); Bachmann and Bayer (2009); Bloom et al. (2014)). Search frictions manifest themselves, in my model, as an endogenous linear hiring cost at the firm level, a feature reminiscent of the kinked adjustment cost model of Bertola and Caballero (1990) and Hopenhayn and Rogerson (1993). This hiring cost creates a kink in the objective function, which leads to an inaction region able to generate the irreversibility essential to real option effects. The search approach of this paper opens up the possibility of using rich labor market flow data to discipline the size of this cost and the frequency of adjustment.

This paper also relates to the recent strand in the literature that has sought to introduce search-and-matching frictions to models of firm dynamics. Acemoglu and Hawkins (2010) and Elsby and Michaels (2013) extend the Mortensen and Pissarides (1994) model to firms with decreasing returns using the Stole and Zwiebel (1996) bargaining procedure. Acemoglu and Hawkins (2010) emphasize the time-consuming aspect of matching to generate persistence in unemployment. Elsby and Michaels (2013) show that the gap between average and marginal products of labor resulting from the decreasing returns allows a reasonable calibration of the model to generate large fluctuations in unemployment and vacancies. However, computing the out-of-steady-state dynamics in these models is difficult and requires the use of approximation methods. My paper explores another more tractable approach that exploits directed search and block recursivity. This tractability enables me to enrich the model further by adding job-to-job transitions and endogenous firm entry/exit, which play an important role in business cycles. Kaas and Kircher (2011) develop a model that applies a similar idea but differs in the techniques used. Addressing the question of efficiency of search models with large firms, they build a model in which firms offer long-term contracts and use a device similar to block recursivity for tractability. Block recursivity usually requires some indifference condition on either side of the labor market. By assuming that workers are homogeneous and cannot search on the job, Kaas and Kircher (2011) obtain this indifference condition on the worker side. As a result, firms are not indifferent between contracts, and their model can replicate the empirical fact that growing firms have higher job-filling rates. They cannot, however, address issues related to job-to-job transitions, which have very specific cyclical properties and account for the largest part of hires and separations in the data. In my model, there is heterogeneity on both sides of the market because workers with different contracts are allowed to search on the job and firms differ in size and productivity. Block recursivity still obtains, because firms, despite being heterogeneous, value workers in the same way, giving rise to an indifference condition on the firm side. As a consequence, workers in my model are not indifferent between contracts, and the model can replicate some new features of the data, in particular the optimal firm policy in terms of quits and layoffs (figure 5), as evidenced in Davis et al. (2011), and can study the dynamics of job-to-job transitions over the business cycle.

The paper is structured as follows. Section 2 introduces the model and presents results on the existence and efficiency of the equilibrium. In section 3, I calibrate the model and evaluate the performance of the model using some establishment-level and cross-sectional facts. Section 4 analyzes and discusses the impact of aggregate productivity and idiosyncratic volatility shocks, before evaluating the ability of the model to account for the US labor market experience over the 1972-2009 period. Section 5 concludes.

2 Model

In order to study the role of time-varying firm-level volatility in explaining fluctuations in unemployment over the business cycle, I build a dynamic search model with i) heterogeneous firms that operate a decreasing returns to scale technology, ii) idiosyncratic risk with time-varying volatility, iii) aggregate fluctuations in productivity, and iv) endogenous separations and on-the-job search to allow for the most complete description of the labor market. The assumption of decreasing returns is crucial, in particular, to provide a well-defined notion of firm size, which enables me to study the dynamics of employment in the cross section of firms in response to aggregate productivity and volatility shocks. The model builds on the directed search framework of Menzio and Shi (2010) in order to exploit the property of block recursivity, defined below.

2.1 Population and technology

Time is discrete. The economy is populated by a continuum of measure 1 of equally productive, infinitely-lived workers and an endogenous measure of firms with free entry. Firms and workers are risk neutral and share the same discount rate β. Firms all produce an identical homogeneous good. The aggregate state of nature is described by a variable s that takes a finite number of value in $\mathcal{S}$ and follows a Markov process with transition matrix $\pi_s(s'|s)$. Aggregate productivity depends on the state of nature and is given by $y(s)$. Firms differ in their idiosyncratic productivity z, independent across firms, that lies in the finite set $\mathcal{Z}$ and follows a Markov process $\pi_z(z'|z,s)$, which I allow to depend on the aggregate state of nature. A firm with a measure n of workers operates the production technology,

$$\mathrm{e}^{y(s)+z}F(n),$$

where F is a strictly increasing, concave production function with $F(0) = 0$. Upon entry, firms must pay a sunk entry cost k_e. Finally, I follow Hopenhayn (1992) in assuming that firms must pay a fixed operating cost $k_f > 0$ every period in order to use the production technology. This operating cost is crucial to generate endogenous exits in the model. It can be interpreted in two ways: either as the fixed cost of using some resources, or similarly as the value of some outside option for firms.

The aggregate state space in this economy comprises the current aggregate state of nature s and should, in principle, include the distribution of employment across firms. Fortunately, the aggregate state space reduces to the variable s under the property of block recursivity. I assume below that this property holds and derive conditions in section 2.9 under which such an equilibrium exists.

2.2 Labor market

Search is directed on the worker and firm sides. Firms announce contracts to attract workers. Since utility is transferable between firms and workers, a sufficient statistics for each contract is the utility x that it delivers to the worker upon matching. Firms offering identical contracts compete on the same market segment, and we therefore describe the labor market as a continuum of submarkets indexed by the utility $x \in [\underline{x}, \overline{x}]$ that firms promise to workers. Firms must pay a cost c for each vacancy that they post. Workers can direct their search and choose in which submarket to look for a job. Match creation on each market segment is governed by a standard matching function with constant returns to scale. Denote $\theta(s, x)$ the vacancy-unemployment ratio or tightness of submarket x in state s. On a submarket with tightness θ, workers find jobs with probability $p(\theta)$, while firms find candidates with probability $q(\theta) = p(\theta)/\theta$. As is common in the literature, we assume that p is increasing, while q is decreasing, and that $p\,(0) = 0$, $q\,(0) = 1$. Workers and firms must therefore solve a trade-off between the level of utility of a given contract and the corresponding probability of being matched. Search takes time and I assume that firms and workers can only visit one submarket at a time.

Employed workers are allowed to search on the job, but are less efficient in doing so than unemployed workers. Denoting λ as the relative search efficiency of the employed compared to the unemployed, the job-finding probability of employed workers is $\lambda p(\theta)$. The equilibrium tightness can be written as $\theta\,(s, x) = \nu/\mu$, where ν stands for the number of vacancies posted on submarket x and μ the corresponding efficiency-weighted number of searching workers.[7]

The amount of vacancies ν that a firm posts is not restricted to be discrete and should be interpreted as a mass. As a result, a law of large numbers applies and firms do not face uncertainty about the number of workers that they recruit. In particular, a firm that posts ν vacancies meets exactly a measure $\nu q\,(\theta)$ of workers.

[7] In particular, $\mu = \mu_u + \lambda\mu_e$, where μ_u is the number of unemployed workers and μ_e the corresponding number of employed workers searching on that market.

2.3 Contracting and timing

Contracts specify various elements relevant to the firm and its workers. To simplify the exposition, let us assume for now that contracts are complete, state-contingent, and that there is full commitment from both the worker and firm sides. A contract specifies $\{w_{t+j}, \tau_{t+j}, x_{t+j}, d_{t+j}\}_{j=0}^{\infty}$, where w is a wage, τ a layoff probability, x the submarket where the worker searches while employed[8] and d an exit dummy. Each element at time $t+j$ is contingent on the entire history of shocks (s^{t+j}, z^{t+j}). I maintain the assumptions of completeness and full commitment throughout this section, but show, however, in section E.2 of the appendix that completeness and commitment from the worker side can be relaxed along some dimensions.

The contracts offered by firms are large objects, but can be written in their recursive form. As a convention, the contracts are rewritten every period after matching occurs and when production takes place (stage B on Figure 2 below). At this stage, the firm starts with some utility W, promised in the past to its incumbent workers or new recruits. A recursive contract $\omega = \{w, \tau, x, d, W'\}$ for the current period specifies the current wage w and next period's quantities $\{\tau(s', z'), x(s', z'), d(s', z'), W'(s', z')\}$, contingent on next period' state, where W' is some future promised utility. Because of commitment, contract ω is required to deliver at least the promised utility W to the worker.

The timing is illustrated in Figure 2. At the beginning of period t, the aggregate state of nature s is realized. Firms then decide whether to enter or not. Immediately after, incumbent and entering firms learn their idiosyncratic productivity z and decide whether to exit $(d = 1)$ or stay. In the following stage, separations occur at probability τ. Search and matching follows with new and incumbent firms on one side and unemployed/employed workers on the other side. Production takes place in the final stage of the period.

2.4 Worker's problem

As a convention, the following value functions are expressed at stage B of the period, when production takes place. We write the value of unemployment as follows:

$$\mathbf{U}(s) = \max_{x_u(s')} b + \beta \mathbb{E}\left[p\left(\theta(s', x_u(s'))\right) x_u(s') + \Big(1 - p\left(\theta\left(s', x_u(s')\right)\right)\Big)\mathbf{U}(s')\right]. \tag{1}$$

When unemployed, workers enjoy a utility b from home production or leisure. In the following period, they choose a market segment, $x_u(s')$, for their job search. In doing so, they must solve a trade-off between the offered utility, x_u, and the likeliness to get a job, $p(\theta(s', x_u))$, which also depends on

[8] The fact that the contract specifies x, the submarket on which the worker should be looking for a job, is a consequence of the assumption of completeness. Part E.2 in the appendix shows that this feature can be relaxed as the firm can write an incentive compatible contract that makes the worker search on the optimal market segment.

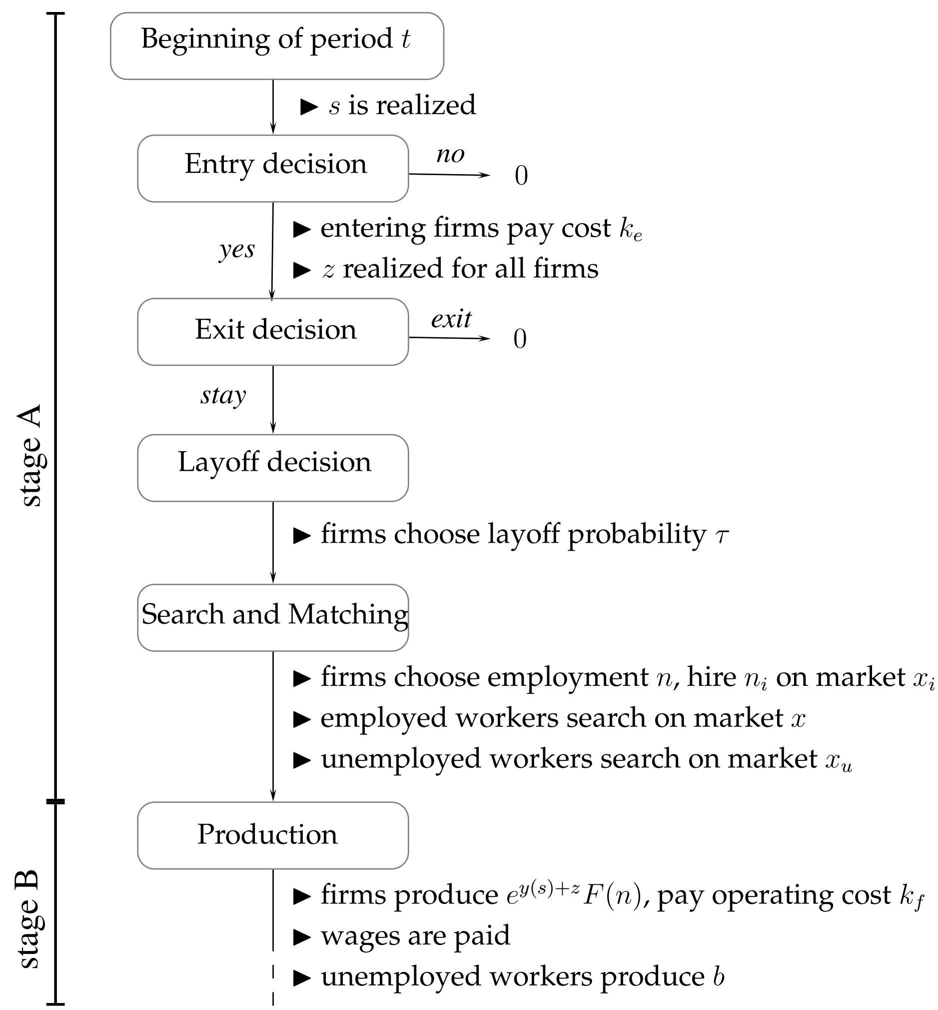


Figure 2: Timing

the aggregate state of the economy. When successful, they enjoy the promised utility x_u, but remain unemployed otherwise.

In the case of a worker employed in a firm with productivity z under the contingent contract $\omega = \{w, \tau(s', z'), x(s', z'), d(s', z'), W'(s', z')\}$, the value can be written:

$$\begin{aligned}\mathbf{W}(s, z, \omega) = w + \beta\mathbb{E}\Big[&\big(d + (1-d)\,\tau\big)\mathbf{U}(s') + (1-d)\,(1-\tau)\lambda p\left(\theta\left(s', x'\right)\right) x \\ &+ (1-d)\,(1-\tau)\left(1 - \lambda p\left(\theta\left(s', x\right)\right)\right)W'\left(s', z'\right)\Big]\end{aligned} \tag{2}$$

The worker first receives the wage w. The following period may then lead to three different outcomes, which correspond to three terms in brackets: i) in the case of exit, d, or layoff, τ, the worker goes back to unemployment with value $\mathbf{U}(s')$; ii) he finds a job in a different firm under a contract with value x at probability $\lambda p(\theta(s', x))$; or iii) he stays in the firm and receives the promised utility $W'(s', z')$ in the following period. Notice that laid-off workers have to spend one period unemployed before looking for a job.

2.5 Firm's problem

Consider the problem of a firm at the production stage with a measure n of employed workers. Workers within the same firm may differ in their levels of promised utility. Each worker is identified by an index $j \in [0, n]$ and a corresponding level of promised utility $W(j)$.

The problem of a firm consists of choosing a list of contracts for its incumbent workers,

$$\omega(j) = \left\{w(j), \tau\left(s', z'; j\right), x\left(s', z'; j\right), W'\left(s', z'; j\right), d\left(s', z'\right)\right\}, \ \forall j \in [0, n].$$

In addition, the firm must decide on a submarket for its hiring in the next period $x_i(s', z')$ and choose a number of workers to hire $n_i(s', z')$. We may describe the problem faced by firms as follows:

$$\begin{aligned}\mathbf{J}\left(s, z, n, \{W(j)\}_{j\in[0,n]}\right) = \max_{\substack{n_i(s', z'),\, x_i(s', z'), \\ \{\omega(j)\}_{j\in[0,n]}}} \quad & e^{y(s)+z}F(n) - k_f - \int_0^n w(j)\,\mathrm{dj} \\ &+ \beta\mathbb{E}\left\{-n_i\frac{c}{q\Big(\theta\big(s', x_i\big)\Big)} + \mathbf{J}\left(s', z', n', \left\{\hat{W}\left(s', z'; j'\right)\right\}_{j'\in[0,n']}\right)\right\}^+\end{aligned} \tag{3}$$

subject to the laws of motion, for all (s', z'):

$$n'(s',z') = \int_0^n \left(1-\tau(s',z';j)\right)\Big(1-\lambda p\Big(\theta\big(s',x(s',z';j)\big)\Big)\Big)\mathrm{d}j + n_i(s',z') \tag{4}$$

$$\hat{W}(s',z';j') = \begin{cases} W'(s',z';j) \text{ for } j' \in [0,n'-n_i'] \text{ and } j' = \Phi(s',z';j) \\ x_i(s',z') \text{ for } j' \in [n'-n_i',n'], \end{cases} \tag{5}$$

where $\Phi(s',z';j) = \int_0^j (1-\tau(s',z';k))\Big(1-\lambda p(\theta(s',x(s',z';k)))\Big)\mathrm{d}k$.

In the current period, the firm earns revenue from production, $e^{y(s)+z}F(n)$, minus the fixed operating cost k_f and wage bill $\int_0^n w(j)\,\mathrm{d}j$. In the following period, the firm chooses whether to exit or not. The $\{\cdot\}^+$ notation, standing for $\max(\cdot,0)$, captures this decision, which we summarize in the dummy $d(s',z') \in \{0,1\}$ ($d=1$ for exit). Following this decision, the firm then chooses a number of workers to hire $n_i(s',z')$ and a submarket $x_i(s',z')$ for their recruitment. Because each vacancy has a probability $q(\theta(s',x_i))$ to be filled, the total vacancy cost incurred for these new recruits is $n_i c/q(\theta(s',x_i))$.

Constraint (4) is the law of motion of total employment. Employment $n'(s',z')$ in the next period is the sum of the new hires $n_i(s',z')$ with the remaining workers after the departure of those laid off with probability $\tau(s',z';j)$ and of those moving to other jobs with probability $\lambda p(\theta(s',x(s',z';j)))$. Equation (5) keeps track of the promised utilities across workers. Because the measure of workers within the firm evolves over time, I use the mapping Φ to reindex the job stayers and make sure that a worker with an original index $j \in [0,n]$ is assigned a new index $\Phi(s',z';j) \in [0,n'-n_i]$ in the next period. New hires with promised utility, $x_i(s',z')$, are assigned an index in $[n'-n_i,n']$.

In addition to these constraints, because of commitment, the firm is subject to the following *promise-keeping* constraint:

$$\forall j \in [0,n], \quad W(j) \quad \leq \quad \mathbf{W}\Big(s,z,\omega(j)\Big). \tag{6}$$

Constraint (6) checks that the contract $\omega(j)$, assigned to worker j, delivers at least the promised lifetime utility $W(j)$. Note finally that there is no non-negativity constraint on the firm's value, implying that firms have deep pockets and no limited liability.

2.6 Joint surplus maximization

The structure of the economy allows us to greatly simplify the firm's problem. The completeness of contracts, the commitment assumption and the transferability of utility guarantee that optimal policies always maximize the joint surplus of a firm and its workers. The model can thus be solved in two stages: a first stage in which we maximize the surplus and a second step in which we can design the contracts that implement the allocation.

Define the joint surplus maximization problem for a firm and its current workers by the following

Bellman equation:

$$\mathbf{V}(s,z,n) = \max_{\substack{d(s',z'),\, n_i(s',z'),\, x_i(s',z'),\\ \{\tau(s',z';j),\, x(s',z';j)\}_{j\in[0,n]}}} \mathrm{e}^{y(s)+z}F(n) - k_f + \beta\mathbb{E}\Bigg\{ nd\mathbf{U}(s')$$
$$+ (1-d)\left[\mathbf{U}(s')\int_0^n \tau \mathrm{dj} + \int_0^n (1-\tau)\,\lambda p\left(\theta\left(s',x\right)\right) x \mathrm{dj}\right.$$
$$\left.\left. - \left(\frac{c}{q\left(\theta\left(s',x_i\right)\right)} + x_i\right) n_i + \mathbf{V}\left(s',z',n'\right)\right]\right\} \tag{7}$$

subject to

$$n'\left(s',z'\right) = \int_0^n \left(1-\tau\left(s',z';j\right)\right)\Big(1-\lambda p\big(\theta\left(s',x\left(s',z';j\right)\right)\big)\Big)\mathrm{d}j + n_i\left(s',z'\right), \quad \forall\left(s',z'\right).$$

The surplus maximization problem characterizes the optimal allocation of physical resources within a firm: the optimal amount of layoffs, job-to-job transitions, new hires and the decision whether to exit or not. Since utility is transferable, transfers between the firm and its workers leave the surplus unchanged. Elements of the contracts that describe the way profits are split, such as wages and continuation utilities, thus disappear in the surplus maximization problem. In particular, the distribution of promised utilities, $\{W(j)\}_{j\in[0,n]}$, is not part of the state space and only the size of employment at the production stage n matters.

The first element in the surplus maximization problem is production followed by the payment of the operating cost k_f. In the next period, the firm chooses whether to exit or not, a decision captured by the exit dummy $d(s',z')$. If a firm chooses to exit, all the workers return to unemployment while the firm's value is set to zero, yielding a total utility of $n\mathbf{U}(s')$. If it chooses not to exit, the firm may then proceed with its layoffs. The total mass of layoffs is $\int_0^n \tau(s',z';j)\,dj$, which provides a total expected utility of $\mathbf{U}(s')\int_0^n \tau dj$ to the worker-firm group. During the search-and-matching stage, some workers move to other jobs with value $x(s',z';j)$ and contribute the amount $\int_0^n (1-\tau(s',z';j))\,\lambda p(\theta(s',x(s',z';j)))\,x(s',z';j)\,\mathrm{dj}$ to the total surplus. Simultaneously, the firm proceeds with its hiring. For each new hire on the labor market segment $x_i(s',z')$, the firm incurs a total vacancy cost of $c/q(\theta(s',x_i(s',z')))$ and must offer the lifetime utility-wage $x_i(s',z')$ to its new recruits, which appears as a cost to the current worker-firm group.

The following proposition formally establishes the equivalence between the firm's problem and the joint surplus maximization.

Proposition 1. *The firm's problem and joint surplus maximization are equivalent in the following sense:*

(i) The surplus and firm's profit verify

$$\mathbf{V}(s,z,n) = \mathbf{J}(s,z,n,\{W(j)\}_{j\in[0,n]}) + \int_0^n W(j)\,dj,$$

(ii) Given any profit maximizing policy $\left\{\{\omega(j)\}_{j\in[0,n]}, d(s',z'), n_i(s',z'), x_i(s',z')\right\}$*, the firm policy* $\left\{\{\tau(s',z';j), x(s',z';j)\}_{j\in[0,n]}, d(s',z'), n_i(s',z'), x_i(s',z')\right\}$ *maximizes the joint surplus,*

(iii) Conversely, if $\left\{\{\tau(s',z';j), x(s',z';j)\}_{j\in[0,n]}, d, n_i, x_i\right\}$ *maximizes the joint surplus, there exists a set of contracts* $\{\omega(j)\}_{j\in[0,n]}$ *with wages and continuation utilities* $\{w(j), W'(s',z';j)\}_{j\in[0,n]}$ *such that the policy* $\left\{\{\omega(j)\}_{j\in[0,n]}, d(s',z'), n_i(s',z'), x_i(s',z')\right\}$ *maximizes profits.*

Proposition 1 tells us that it is possible to find the optimal allocation of physical resources $\left\{\{\tau(s',z';j), x(s',z';j)\}_{j\in[0,n]},\ d(s',z'),\ n_i(s',z'),\ x_i(s',z')\right\}$ first and solve for the contracts that implement that allocation later, in a second stage. This proposition establishes, in particular, that such contracts always exist and are, in fact, easy to construct once the allocation is known. This result is of particular interest in practice since equation (7) is a Bellman equation that can easily be solved with standard numerical methods.

The fact that one can maximize the joint surplus regardless of the distribution of promised utilities is an important result, which stems from the transferability of utility, the compleness of contracts and the assumption of commitment. Transferability ensures that a firm and its workers evaluate the benefits from their actions using a common utility scale and agree on a definition of a joint surplus. Completeness guarantees that there always exists sophisticated enough schemes of transfers, through wages and continuation utilities, that can implement the surplus maximizing allocation by suitably redistributing the benefits between the firm and its workers for any initial distribution of promises. Finally, commitment ensures that firms cannot extract a larger part of the surplus by reneging on their promises, for instance by laying off workers with high utility-wages. Since promises have to be fulfilled in all circumstances, including upon separation, tweaking the allocation away from the surplus maximizing allocation cannot deliver higher profits: it is optimal for firms to maximize the physical size of the "cake", i.e., the surplus, pay the workers their dues and enjoy the remaining profits, which are then maximized. The distribution of promises is thus irrelevant to the determination of the physical allocation of resources, but only matters for the way the surplus is split between the firm and its workers, in particular for wages.

Notice, finally, from equation (7), that since all the contracting aspects have disappeared, the surplus maximization problem is purely forward looking and the firm's current state (s,z) has no impact on the optimal policy $\left\{\{\tau,x\}_{j\in[0,n]},\ d,\ n_i,\ x_i\right\}$ chosen by a firm in state (s,z,n). As a result, while the equilibrium policy should in principle depend on the entire state space $(s',z';s,z,n)$, it solely depends on the firm's state at the beginning of next period $(s',z';n)$. This result is assumed throughout the rest of the paper.

2.7 Free entry

Every period after the aggregate shock s is realized, potential entrants decide whether or not to enter. Upon entry, firms must pay an entry cost k_e, after which they draw their idiosyncratic productivity z from some distribution g_z. Depending on the outcome, firms may decide to exit or stay, in which case they can start hiring and producing as any normal firm.

We define the problem faced by an entering firm of type z as follows:

$$\mathbf{J}_e(s,z) = \max_{n_e(s,z),x_e(s,z)} \left[\mathbf{V}(s,z,n_e) - n_e x_e - n_e \frac{c}{q(\theta(s,x_e))} \right]^+ . \tag{8}$$

Having drawn the idiosyncratic productivity z, the potential entrant first decides whether or not to exit, a decision captured by the notation $\{\cdot\}^+$ and summarized in the dummy $d_e(s,z)$. If it stays, the firm chooses a measure of workers to hire, $n_e(s,z)$, and a market for recruitment, $x_e(s,z)$, in order to maximize its profits net of the total vacancy cost $n_e c/q(\theta(s,x_e))$. Using proposition 1, these profits can be written as the joint surplus $\mathbf{V}(s,z,n_e)$ minus the total utility $n_e x_e$ that the firm must deliver to its new recruits.[9]

An important feature of this economy is that the submarket x_e, in which workers are hired, solely appears through the term $\frac{c}{q(\theta(s,x_e))} + x_e$, which we can describe as a *hiring cost per worker*, common to both entering and incumbent firms. The first term, $c/q(\theta(s,x_e))$, captures the total vacancy cost of hiring exactly one worker. The second term, x_e, is the utility-wage that firms offer to their new recruits. As a result, the decision of entering firms can be decomposed as a two-stage problem: a first stage, during which firms choose where to search for their workers; a second stage, in which firms decide on the number of workers to recruit. In the first stage, firms choose the submarkets that minimize hiring costs per worker. Define the minimal hiring cost κ as[10]

$$\kappa(s) = \min_{\underline{x}\le x\le \overline{x}} \left[x + \frac{c}{q(\theta(s,x))} \right]. \tag{9}$$

Optimal entry further imposes the requirement that only the submarkets that minimize this hiring cost are open in equilibrium, which we summarize in the following complementary slackness condition:

$$\forall x, \quad \theta(s,x) \left[x + \frac{c}{q(\theta(s,x))} - \kappa(s) \right] = 0. \tag{10}$$

Equation (10) expresses that submarkets either minimize the hiring cost, $\kappa(s) = x + c/q(\theta(s,x))$, or remain unvisited, $\theta(s,x) = 0$. In equilibrium, active submarkets will have the same hiring cost $\kappa(s)$ and firms will be indifferent between them. Notice that equation (10) provides us with an expression

[9]The ex-post profits after entry for a firm of type z coincide with the surplus net of the promised utility x_e, $\mathbf{J}\left(s,z,n_e,\{x_e\}_{j\in[0,n_e]}\right) - n_e \frac{c}{q(\theta(s,x_e))} = \mathbf{V}(s,z,n_e) - n_e x_e - n_e \frac{c}{q(\theta(s,x_e))}$.

[10]Note in particular that the cost minimization problem is the same across firms. This property is key to obtain the indifference condition on the firm side required for block recursivity, as we discuss in section 2.10.

for the equilibrium market tightness on every active market,

$$\theta(s,x) = q^{-1}\left(\frac{c}{\kappa(s) - x}\right). \tag{11}$$

The job filling probability q being a decreasing function, this expression tells us, in particular, that equilibrium tightness must decrease with the level of utility promised to workers as these offers succeed in attracting more workers, while firms refrain from posting such expensive contracts. The probability of finding a job for workers thus declines with the attractiveness of the offer.

We may now describe the full free-entry condition in this economy. Firms enter the economy as long as expected profits exceed the entry cost k_e, driving these profits down to k_e. Therefore, expected surplus from entering must be equal to k_e in equilibrium:

$$k_e = \sum_{z \in \mathcal{Z}} \mathbf{J}_e(s,z)\, g_z(z), \ \forall s. \tag{12}$$

Note that the free-entry condition is crucial to guarantee the existence of a block-recursive equilibrium. Section 2.10 discusses this property in more details and explain why it obtains in this setup.

2.8 Unemployment and firm distribution dynamics

Using the optimal decision of firms, we may now describe the evolution of employment over time. Let u be the unemployment rate and $g(z,n)$ the distribution of employment across firms in stage B of the current period, when production takes place.

Given a current aggregate state (s,g), the evolution of the unemployment rate is governed by the following equation:

$$\begin{aligned} u' = {} & \Big(1 - p\left(\theta\left(s', x_u\left(s'\right)\right)\right)\Big)u \\ & + \sum_{z,z',n} n\left[d\left(s',z';n\right) + \left(1 - d\left(s',z';n\right)\right)\tau\left(s',z';n\right)\right]\pi_z\left(z' \mid z,s\right) g(z,n). \end{aligned} \tag{13}$$

Equation (13) states that unemployment at the start of the next period corresponds to the fraction $1-p\left(\theta\left(s', x_u\left(s'\right)\right)\right)$ of unemployed workers that do not find a job next period in addition to the workers that lose their jobs because of exits, d, or layoffs, τ.

The dynamics of the distribution of employment across firms can be described by

$$\begin{aligned} g'(z',n') = {} & \sum_{z,n} \mathbb{I}\left\{n'\left(s',z';n\right) = n'\right\}\left(1 - d\left(s',z';n\right)\right)\pi_z(z' \mid z,s) g(z,n) \\ & + m_e\left(s',g\right)\mathbb{I}\left\{n_e\left(s',z'\right) = n'\right\}\left(1 - d_e\left(s',z'\right)\right) g_z(z'). \end{aligned} \tag{14}$$

where $\mathbb{1}\{\cdot\}$ denotes an indicator function. Equation (14) defines the number of firms with individual state (z', n') in the next period as the sum of the surviving incumbents and entering firms that end up in this state. The term m_e is the endogenous measure of new entrants, which depends on the aggregate state of nature s' and distribution g. It is defined as the number of entering firms required to reach the equilibrium market tightness on every market segment. Fortunately, because firms are indifferent between the various submarkets, these equilibrium conditions can be summarized by a unique aggregate condition which states that the total number of jobs found by workers has to equal the number of jobs created by firms. More formally, m_e is implicitly defined by the equation

$$JF_{\text{total workers}}\left(s', g\right) = JC_{\text{total incumbents}}\left(s', g\right) + m_e\left(s', g\right) JC_{\text{entrant}}\left(s'\right), \tag{15}$$

where JF holds for the number of jobs found by workers across all submarkets and JC for the number of jobs created by firms. In particular, JC_{entrant} is the number of jobs created by a measure one of entrants. An explicit formula for m_e is derived in appendix B.

2.9 Existence and efficiency

We may now define a block-recursive equilibrium in this economy. For this purpose, I proceed in a constructive way and introduce the notion of a quasi-equilibrium as a candidate equilibrium. I define a quasi-equilibrium as a block-recursive solution to both the workers' problems (1)-(2) and firms' problem (3), which further satisfies the free-entry condition (12). Unfortunately, without further restrictions on the parameters, the labor market equilibrium condition as described by equation (15) may imply negative entry in some cases. Under such circumstances, the assumption of free-entry is not valid and block-recursivity does not obtain. For a quasi-equilibrium to be a well defined block-recursive equilibrium, one must verify that entry is non-negative in every possible state of the world.

Definition 1. Define the following concepts:

(i) A quasi-equilibrium of this economy is a) a set of value functions $\mathbf{U}(s)$, $\mathbf{W}(s, z, \omega)$, $\mathbf{J}\left(s, z, n, \{W(j)\}_{j\in[0,n]}\right)$, $\mathbf{V}(s, z, n)$ and $\mathbf{J}_e(s, z)$, b) a decision rule for unemployed workers $\{x_u(s')\}$, for entering firms $\{d_e(s, z), n_e(s, z)\}$ and for incumbent firms $\left\{d(s', z'; n), n_i(s', z'; n), x_i(s', z'; n), \{\omega(j; s, z, n)\}_{j\in[0,n]}\right\}$, c) a hiring cost $\kappa(s)$ and corresponding labor market tightness $\theta(s, x)$ such that equations (1)-(12) are satisfied;

(ii) A block-recursive equilibrium is a quasi-equilibrium such that entry is non-negative in any state of the world.

Proposition 2 below establishes existence and efficiency results.

Proposition 2. *(i) Under weak regularity conditions stated in Appendix G.2, a quasi-equilibrium always exists;*

(ii) A block-recursive equilibrium, when it exists, is efficient.

First, proposition 2 shows that a quasi-equilibrium always exists under some weak regularity conditions. In particular, it uses Schauder's fixed point theorem to show the existence of a joint solution to the surplus maximization, free-entry and unemployed workers' problems. The existence of a solution to the firm's profit maximization problem and corresponding contracts ensues from proposition 1. Unfortunately, the existence of a full block-recursive equilibrium, namely a quasi-equilibrium with positive entry, cannot be easily proved. The key issue is that the measure of entrants m_e, implicitly defined in (15), depends on the infinite dimensional distribution g and cannot be put into a recursive form. Although one can derive sufficient conditions on parameters to guarantee that entry is always non-negative, it is easier to check this condition numerically in practice. Proposition 2 thus provides us with a constructive way to solve for block-recursive equilibria: 1) solve for a quasi-equilibrium in a first stage, and 2) check that the obtained policy functions imply positive entry in every state of the world. Note, in addition, that the non-negative entry condition is a weak restriction for empirically relevant cases as data from the US and other developed economies always display positive entry, even in the midst of large recessions.

Turning to (ii), this proposition establishes the efficiency of block recursive equilibria. It guarantees, in particular, that a quasi-equilibrium with positive entry, once found, maximizes welfare and must be, as such, unique in a payoff-equivalent sense. This extends standard results in competitive search models with single-worker firms. This model thus offers an efficient benchmark in which unemployment is efficient and there is no mispricing, nor inefficient separations.

Section E in the Appendix characterizes additional properties of the optimal contracts and provides an alternative version of the model relaxing commitment and completeness of the contracts.

2.10 Block recursivity

In this section, I explain the intuition behind the property of block recursivity as it appears in the literature and discuss why it obtains in this setup. Readers interested in the quantitative exercise may directly skip to section 3.

In search-and-matching models with sufficient heterogeneity, the distribution of workers across firms is in general required for agents to forecast wages and the labor market tightness. This feature is problematic when the distribution is an infinite-dimensional object, which standard numerical techniques cannot handle. To address that problem, Menzio and Shi (2011) introduced the concept of block recursivity using from key insights. The first insight is the use of directed search instead of random search. In random search, wages are negotiated and depend on workers' and firms' outside options, which usually depend on the distribution of workers across firms in equilibrium. In a directed search setup, firms and workers do not need to forecast wages (or contracts) because wages are choice variables, which, as such, do not depend on who they meet: firms choose the wage that they offer; workers choose where to apply. In a directed search environment, the only remaining channel through which the distribution of employment may affect the equilibrium is the vacancy-unemployment ratio or market tightness. The second key insight, based on a clever use of the free-entry condition, comes into play at this stage. This condition equalizes the cost of opening a vacancy to the value of a job. This value depends on the probability at which a job is created—a function of the market tightness—and on the surplus of the match, but does not directly depend on the distribution of employment in the economy. Since the cost of opening a vacancy is a constant in these models, the free-entry condition pins down the value of the market tightness as a function of the value of a new job. Likewise, the value of a new job is not directly affected by the employment distribution, but only indirectly through the expectation of future market tightness. But because the free-entry condition pins down future market tightness independently from the distribution of firms, it is then possible to construct a full equilibrium in which *neither* the value functions, *nor* the market tightness depend on the employment distribution across firms: the equilibrium is *block recursive*.

Unfortunately, this reasoning does not easily apply to a setup with too much heterogeneity. The free-entry condition only pins down the equilibrium market tightness on a single market: the one chosen by entering firms. To characterize the tightness on the other submarkets, homogeneity is often assumed on either side of the labor market. With homogeneous workers or firms, an indifference condition arises that can be used to ensure that the free-entry condition pins down the tightness on *every* active submarket in the economy. In the environment proposed in this paper, there is heterogeneity on both sides of the market. Firms differ in productivity and sizes. Workers differ in their employment statuses—employed or unemployed— and in their current utility levels depending on whether they work in high-paying jobs or not. A contribution of this paper is to show that block recursivity may still obtain in the presence of two-sided heterogeneity and proves the existence of such equilibria. The "trick" relies on two assumptions: the transferability of utility—which guarantees that all contracts are viewed in an identical way by agents—and the fact that firms hire a continuum of workers. Under these two conditions, the decision over the market for hiring can be summarized by the minimization of the cost $\kappa(s)$. Therefore, despite heterogeneity on the firm side, firms are effectively indifferent across submarkets because they face the same hiring cost $\kappa(s)$. Even though firms differ in productivity and sizes, they all seek to minimize this cost and thus post their offers on the same markets. As a consequence, indifference on the firm side in combination with the free-entry condition allows to characterize the equilibrium tightness of every active submarket and generalizes the block recursive property to the whole economy.

3 Business cycle and Establishment-level Properties

In this section, I calibrate the model and evaluate its predictions at various levels of aggregation. Starting at the aggregate level, I present some standard business cycle statistics from model simulations and compare them to the same model with aggregate productivity shocks only. I show that the presence time-varying idiosyncratic volatility generally leads to more realistic fluctuations in unemployment and other variables. Turning to the establishment-level implications of the model, I discuss some of its properties in terms of growth rates, and show that it can replicate salient features of the employment behavior of firms.

3.1 Calibration

Functional forms and stochastic processes

I parameterize the model as follows. The production function is the concave function $F(n) = An^{\alpha}$, where α governs the amount of diminishing returns in the economy. Since time is discrete, I must choose a job-finding probability function bounded between 0 and 1, which rules out Cobb-Douglas matching functions. Following Menzio and Shi (2010), I pick the CES contact rate functions

$$p(\theta) \equiv \theta(1+\theta^{\gamma})^{-1/\gamma}, \quad q(\theta) \equiv p(\theta)/\theta = (1+\theta^{\gamma})^{-1/\gamma}.$$

In addition to providing a good fit to the data, these functions satisfy all the regularity conditions required for the existence of an equilibrium stated in Appendix G.2. To parameterize them, I estimate function p by non-linear least squares using the job-finding rate series constructed by Shimer (2007)[11] and a measure of market tightness θ. I construct the latter using data on vacancies from the Job Openings and Labor Turnover Survey (JOLTS) and the Conference Board's Help Wanted Index with unemployment data from the Bureau of Labor Statistics (BLS).[12] The regression yields $\gamma = 1.60$ with a good fit ($R^2 = 0.90$).

The aggregate and idiosyncratic productivity shocks follow the AR(1) processes

$$\begin{aligned} y_t &= \rho_y y_{t-1} + \sigma_y \sqrt{1-\rho_y^2}\varepsilon_{y,t}, \quad \varepsilon_{y,t} \sim \mathcal{N}(0,1) \\ z_t &= \rho_z z_{t-1} + v_{t-1}\sqrt{1-\rho_z^2}\varepsilon_{z,t}, \quad \varepsilon_{z,t} \sim \mathcal{N}(0,1), \end{aligned}$$

where v_t denotes the time-varying volatility of idiosyncratic productivity. I assume that its log follows the AR(1) process with mean $\log \overline{v}$:

$$\log v_t = (1-\rho_v)\log \overline{v} + \rho_v \log v_{t-1} + \sigma_v \sqrt{1-\rho_v^2}\varepsilon_{v,t}, \quad \varepsilon_{v,t} \sim \mathcal{N}(0,1),$$

which ensures that idiosyncratic volatility remains positive. In the data, idiosyncratic volatility is countercyclical. I therefore allow the innovations $\varepsilon_{y,t}$ and $\varepsilon_{v,t}$ to be correlated and denote $\sigma_{yv} = \operatorname{cor}(\varepsilon_{y,t}, \varepsilon_{v,t})$. Innovations to z_t are independent across agents. The aggregate state of nature is

[11] These data were constructed by Robert Shimer. For additional details, please see Shimer (2007) and his webpage, http://sites.google.com/site/robertshimer/research/flows.

[12] Because the direct vacancy measure by JOLTS is available only since 2001, I use the Conference Board's Help Wanted Index to complete the measure from 1951Q1 to 2000Q4 and adjust its level to match the JOLTS stock of vacancies in 2001Q1. Unemployment comes from the monthly seasonally adjusted unemployment rate constructed by the BLS. Data are averaged over three-month periods and detrended using an HP filter with parameter 1600.

$s_t = (y_t, v_t)$.

Calibration strategy

The model is estimated using a method of simulated moments. For the largest part, I follow the search-and-matching literature in choosing the moments to target. The chosen calibration strategy targets mostly aggregate labor market flows as in Shimer (2005). It is conservative in the sense that such a strategy usually leads to the unemployment volatility puzzle.

The time period is set to one month. I set the discount rate β to 0.996 so that the annual interest rate is about 5%. I set the decreasing returns to scale parameter $\alpha = 0.85$ in the middle of the range of empirical estimates in the literature (Basu and Fernald (1995); Basu (1996); Basu and Kimball (1997)).[13] Without firm-level panel data, I do not have observations on the idiosyncratic productivity process of firms. I thus follow the investment literature and set $\rho_z = (0.95)^{\frac{1}{3}}$ in order to match an approximate quarterly autocorrelation of 0.95 as in Khan and Thomas (2008) and Bloom et al. (2014).

The parameters left to estimate are the following: the productivity parameters $(\rho_y, \sigma_y, \overline{v}, \rho_v, \sigma_v, \sigma_{yv})$, the home production b, the vacancy posting cost c, the entry cost k_e, the fixed operating cost k_f and the relative search efficiency of employed workers compared to unemployed ones λ.

To discipline the choice of the aggregate productivity parameters (ρ_y, σ_y), I target the autocorrelation and standard deviation of log-detrended output, using seasonally adjusted quarterly real GDP from the Bureau of Economic Analysis.[14] Regarding the idiosyncratic productivity parameters $(\overline{v}, \rho_v, \sigma_v, \sigma_{yv})$, I select moments from the establishment-level volatility series constructed by Bloom et al. (2014). I target, in particular, the average interquartile range (IQR) of innovations to idiosyncratic TFP, its autocorrelation, standard deviation and correlation with aggregate output. To inform the estimation of the labor market parameters (c, b, λ), I include in my moments the following historical averages of the monthly transition rates: an Unemployment-Employment (UE) rate of 45%, an Employment-Unemployment (EU) rate of 2.6% according to Shimer (2005), and an Employment-Employment (EE) rate of 2.9% following estimates by Nagypál (2007). To discipline the entry cost k_e, I target an average fraction of jobs created by opening establishments of 21%, according to the Business Employment Dynamics (BED) over the period 1992Q3-2009Q4. Finally, because the operating cost k_f governs the rate of exit in the economy and the degree of dynamic selection, I target an average establishment size of 15.6, as in the 2002 Economic Census.

The parameters are jointly estimated using a search algorithm in the parameter space that minimizes the distance between the empirical and simulated moments, with weights chosen to yield relative errors of the same amplitude for each moment. Section D in the Appendix describes the numerical

[13] I choose to match the total decreasing returns at the firm level because I am interested in explaining firm dynamics, despite the absence of capital in the model. A previous version of this paper targeted a wage share of 0.66, with little difference on the final results.

[14] During the estimation procedure, all time series are computed in log deviations from an HP-trend with parameter 1600 for quarterly data and 100 for annual data.

implementation. Table 1 summarizes the parameter values that result from the calibration. Table 2 shows the fit of the model with the targeted moments. The fit is, overall, quite satisfactory. Note that the autocorrelation of output produced by the model is slightly below its empirical counterpart, because the less persistent volatility shocks introduce extra variation in output.

Targeting an annual interquartile range of 0.393, the long-run standard deviation $\overline{v}$ of idiosyncratic volatility is estimated to be 0.533, about ten times as large as that of output, a result in line with Bloom et al. (2014). The standard deviation of the volatility process is large, $\sigma_v = 13.2\%$, but still relatively low compared to the 30% increase in the IQR observed from 2004 to 2008. The labor market parameters can be interpreted as follows. The estimated home production b represents about 63% of the average output per person in the economy, consistent with the 71% found by Hall and Milgrom (2008). The vacancy cost c is about 33% of the average quarterly compensation of workers, which is about twice as much as the 14% estimated in Silva and Toledo (2009). There is, unfortunately, no widely accepted empirical estimate for the entry and operating costs. Comparing their values to the average output produced by a single firm in a month, my estimated costs represent about 38% for the entry cost k_e and about 5% for the operating cost k_f.

3.2 Business cycle statistics

To evaluate the performance of the model at the aggregate level, I simulate it for a large number of periods and compute some business cycle moments. In particular, I calculate the standard deviation and contemporaneous correlation with output of several variables. These variables include unemployment, total vacancies and various labor market flows such as total hirings, quits and layoffs. In order to understand the contribution of idiosyncratic volatility shocks, I further compute the same moments in a version of my model with aggregate productivity shocks only. For this purpose, I recalibrate the model using the same targets except those related to the time-varying volatility, namely $\rho\left[IQR\left(e_{z,t}\right)\right]$, $\sigma\left[IQR\left(e_{z,t}\right)\right]$ and $corr\left[IQR\left(e_{z,t}\right), Y_t\right]$.

The results are presented in Table 3. A first striking result is that the model proposed in this paper explains about 50% of the volatility in unemployment with aggregate productivity shocks only (column 3). This finding suggests that the introduction of heterogeneous multiworker firms and the presence of a slow-moving distribution of employment across establishments adds amplification to search-and-matching models, which are known to produce little volatility in aggregate unemployment when calibrated to match moments as those chosen in my estimation.[15]

Most importantly, column 5 shows that the addition of stochastic idiosyncratic volatility makes substantial progress in explaining the volatility of labor market variables. With these additional shocks, the model accounts for 75% of the total volatility in unemployment and approximately doubles

[15] To emphasize this point, Table 4 in the Appendix evaluates the same business cycle moments to a standard Diamond-Mortensen-Pissarides model, calibrated along the strategy described in Shimer (2005). The calibration is identical to the one in the original article except that I target the autocorrelation (0.839) and standard deviation of output (0.017) instead of output per person to harmonize it with my estimation.

the volatility of other variables, improving the general fit of the model with the exception of an excessive volatility in layoffs. Turning to comovements, aggregate productivity shocks being the only source of business cycle fluctuations, column 4 displays in general excessively high contemporaneous correlations with output. The introduction of time-varying volatility in column 6 breaks this result and helps the model produce correlations more in line with the data, as evidenced by quits and layoffs. The correlation of vacancies and hirings with output are, however, too weak compared to the data because they tend to rise with volatility shocks, as we will see in the next section. The cyclicality of each variable is, however, correctly predicted by both versions of the model.

Overall, the introduction of heterogeneous multiworker firms allows the model to have more realistic predictions than a typical search-and-matching model. In addition, volatility shocks generate larger fluctuations in unemployment and other labor market variables, offering a plausible mechanism to account for the volatility unexplained by standard models.

3.3 Establishment-level properties

Because I do not target any establishment-level or cross-sectional moment other than the interquartile range of idiosyncratic productivity, I now examine several implications of the model in the cross-section of establishments as a validation exercise. For that purpose, it is convenient to introduce the following measure of establishment growth rates as used by Davis et al. (1996). Denoting $n_{i,t}$ as the total employment of establishment i at date t, define growth rate $g_{i,t}$ as

$$g_{i,t} = \frac{n_{i,t} - n_{i,t-1}}{\frac{1}{2}(n_{i,t} + n_{i,t-1})}.$$

This measure takes the ratio of net employment growth to the average size of the establishment between periods $t-1$ and t. This measure is convenient in that it can account for the entry and exit and treats them in a symmetric fashion. A growth rate of 2 means entry, while -2 stands for exit.

Growth rate distribution

Davis et al. (2011)[16] report the quarterly employment growth rate distribution of establishments using data from the BED dataset in 2008. I simulate the model for a large number of periods, aggregate the data over three-month periods and compare the empirical and simulated growth rate distributions.

Figure 4 displays the two distributions. Given that the only cross-sectional moment in the estimation was the IQR of idiosyncratic productivity shocks, the model generated distribution displays a reasonable fit to its empirical counterparts. Yet, the fit is imperfect and the reason is worth highlighting for future extensions. On the positive side, both present a large peak at 0 (16% in the data, 21% in the model), which indicates that a substantial number of establishments do not adjust their

[16] I would like to thank Steven Davis, Jason Faberman and John Haltiwanger for allowing me to use their tabulations from the BED dataset.

employment at all in a quarter. The model can replicate this feature because the search frictions manifest themselves as a kink in the firm's problem, thereby producing a region of inaction for firms. On the negative side, the distribution generated by the model is more left-skewed than its empirical counterpart. This result stems, in the model, from the endogenous exit and dynamic selection of firms. Since mostly unproductive firms exit, large productive firms tend to be overrepresented in the sample of surviving firms. At the same time, these productive firms have a stronger tendency to contract over time because of mean reversion in their fundamentals. The combination of these two facts explains why the simulated distribution is asymmetric. A possible way to improve this dimension would be to introduce permanent productivity differences across firms, so that transitory productivity shocks would have a lower impact on exits and firm sizes.

Finally, regarding the cyclicality of firms' growth rates, Moscarini and Postel-Vinay (2012) have documented that large firms are more cyclically sensitive than small ones. The model is able to replicate this feature as it produces a positive correlation of 0.28 between output and the differential growth rate between the 50% largest and smallest firms in the economy, a result robust to various definitions of size.[17]

Employment policy

Empirical evidence shows that firms with different growth rates have different hiring, layoff and quit rates. The composition of hirings against separations and the balance between layoffs and quits present some important regularities at the establishment level. Davis et al. (2011) show with some empirical exercises that capturing these regularities may be important to improve the time-series predictions of search models. Being one of the few models in the literature with multiworkers firms and a meaningful distinction between quits and layoffs, I examine my model's predictions along this dimension.

Figure 5 displays the empirical and simulated employment-weighted levels of hirings, quits and layoffs as a function of establishment growth. To produce this graph, I simulate the model for a large number of periods and compute the corresponding series by aggregating over three-month periods. Quite surprisingly, without targeting any of these observations in the cross-section, the model can replicate a number of qualitative and quantitative features of hiring, quit and layoff rates at the establishment level. In particular, it is able to match the change in the composition of quits versus layoffs for contracting firms. Establishments that contract by a small amount tend to favor quits over layoffs, as they internalize the fact that workers can be directly employed without experiencing

[17] The reason behind this result is the different cyclical sensitivity of the hiring cost and the value of unemployment. Being determined by the free entry condition, the hiring cost absorbs much of the variation in the value of firms that results from aggregate shocks, and is highly procyclical. On the opposite, the value of unemployment is rather sluggish because it includes the value of home production b, which remains constant over time. As a result, the firing threshold shifts out more in recessions than the hiring threshold. Because of mean reversion in productivity, large firms have a larger tendency to lie in a region close to the highly cyclical separation threshold, rather than small firms which are closer to the hiring threshold, explaining the differential cyclical sensitivity across firm sizes.

unemployment. However, the job-to-job transition technology becomes congested at some point, and firms that contract by a significant amount use layoffs more intensively.[18] The key qualitative feature that the model misses is *churning*: expanding establishments in the data separate from a sizeable fraction of their workforces; contracting establishments on the other hand hire a positive amount of workers. The model is able to generate churning to some extent through time aggregation, as evidenced in the non-zero amount of quits for expanding establishments, but too much churning is suboptimal in the model since workers are homogeneous. Accounting for the observed level of churning in the data would likely require adding worker heterogeneity in productivity to the model.

4 Understanding the Forces at Work

With time-varying idiosyncratic volatility and multiworker firms heterogeneous in productivities and sizes, this paper introduces two important dimensions to standard search-and-matching models. Before running the final counterfactual experiments, I pause in this section to describe the workings of the model in details and explore how each of these dimensions affect the labor market.

I first describe the equilibrium and, in particular, how search frictions affect the employment decision of firms as a function of their productivities and sizes. The optimal policy takes the form of various action thresholds—or *triggers*—in the spirit of the kinked adjustment cost literature. I then examine the impact of aggregate productivity and idiosyncratic volatility shocks. The response of the economy to these shocks hides a rich variety of effects which I decompose between first moment, general equilibrium, option value and realized volatility effects.

4.1 Equilibrium Description

Labor market equilibrium

The labor market is organized in a continuum of submarkets, indexed by the contracts that firms offer. We take a closer look in this section at how firms and workers allocate themselves across these submarkets in equilibrium.

Figure 6 depicts the different labor market segments on the axis $[\underline{x}, \overline{x}]$ with the equilibrium market tightness $\theta(s, x)$. An important feature is that the market tightness decreases with the value of the contract. To maximize profits, firms prefer offering low utility contracts and post more vacancies in markets with low x. However, as these markets become more crowded, the job filling probability declines and the cost of searching rises. As a result, some firms find it profitable to raise their offers, trading off lower profits from higher utility-wages for a greater probability of filling the job, until they become effectively indifferent across markets. The equilibrium tightness, captured in equation (11), is

[18]Note that the levels at which the quit rate settles for contracting establishments differ in the model and the data. This result is an artefact of the estimation. Because we are ultimately interested in the aggregate predictions of the model, the estimation targets the aggregate Employment-to-Employment (EE) rate in the data. However, since expanding establishments in the model do not use quits, the estimation compensates with larger rates for contracting ones.

consequently a decreasing function of x with the implication that the job filling probability for firms rises with the value of the contract, while the job finding probability for workers declines.

While recruiting firms are indifferent across the various submarkets, workers are not. Having different outside options, unemployed and employed workers search on different market segments as illustrated on the graph: unemployed workers tend to apply to low-paid jobs, while employed workers apply to higher-paid jobs, as they are more willing to tolerate low job finding probabilities.

Employment policy at the establisment level

Establishments can use various margins—hires, quits, layoffs, or exit—to adjust employment over the cycle. I examine in this section how the decision of firms to use these margins varies as a function of their individual characteristics (z, n) at the beginning of a period.

Figure 7 displays the optimal policy of firms, as it appears in my baseline calibration. As one would naturally expect, hirings take place at small productive firms, whose marginal value of adding jobs is high, while separations—quits and layoffs—occur at large unproductive firms. Interestingly, because search frictions show up in the surplus (7) as a linear hiring cost, $\kappa(s) = c/q(s, x_i) + x_i$, a wedge appears in the adjustment cost faced by firms at $n' = n$. More specifically, laying a worker off earns a value of $\mathbf{U}(s)$ to the worker-firm group, while hiring incurs the cost $\kappa(s)$, strictly greater than the value of unemployment in equilibrium.[19] Arising from this kink in adjustment costs, a band of inaction emerges between two thresholds, a hiring and a separation thresholds, which play the role of triggers in the firm's employment strategy. Whenever a firm falls in the hiring region, in the lower right area, its optimal strategy consists of hiring workers up until it reaches the hiring threshold—a point at which the marginal value of adding jobs equals the hiring cost. Symmetrically, whenever a firm finds itself in the separation region, its optimal sequence of actions is to separate from workers, using a mix of quits and layoffs, until it reaches the separation threshold, at which the marginal value of employment equals the marginal value of quitting.[20] The presence of an inaction region implies the existence of a non-negligible mass of firms not adjusting employment within a period, a fact well supported in the data as evidenced by Davis et al. (1996).

Exits take place at small unproductive firms. Indeed, the operating cost k_f being fixed, the decision to exit mostly affects small firms with low productivity, as their current production and expected future surpluses fall short of the total operating costs. This feature is consistent with empirical observations, as evidenced in Evans (1987).

[19]Because the market $x_u(s) > \mathbf{U}(s)$ where unemployed workers search is active in equilibrium, we know that $\kappa(s) = c/q(\theta(s, x_u(s))) + x_u(s) > \mathbf{U}(s)$. Similarly, a wedge appears between the value of quitting and the cost of hiring, as the best contract offered in equilibrium is $\hat{x}(s) = \kappa(s) - c < \kappa(s)$.

[20]Notice here that the hiring and separation thresholds both play the role of *triggers* and *return points,* in accordance to the kinked adjustment cost literature (see Bertola and Caballero (1990)). Despite important similarities, this strategy is different from the Ss-type policies that arise in the fixed adjustment cost literature, where the trigger and return point differ (see Khan and Thomas (2008)).

4.2 Productivity shocks

Aggregate productivity shocks are the common source of business cycle fluctuations in the search-and-matching literature. In this section, I analyze the impact of negative aggregate productivity shocks at the macroeconomic level and use the model to study at a deeper level how these shocks affect firms in the cross-section. The response of the economy reflects the combination of various effects from partial to general equilibrium, that affect entrants and incumbents in remarkably different ways.

Employment policy

Figure 8 illustrates how a negative one standard deviation productivity shock affects the employment strategy of firms in partial equilibrium (upper panel), holding the tightness and the value of unemployment fixed, and in general equilibrium (lower panel). The black continuous lines depict the hiring, separation and exit thresholds before the shock, when aggregate productivity and volatility are set to their means; the dashed blue lines describe how these thresholds are affected when the shock hits.

How a negative productivity shock affects the employment strategy of firms in partial equilibrium is straightforward. When productivity declines, the marginal value of employment decreases. As a result, expanding firms cut on hiring and grow less—the hiring threshold shifts down—while separations rise as the value of jobs in other less productive firms fall below the separation threshold—the separation region widens. An increase in exits is simultaneously observed as the decline in production makes firms at the margin of profitability unable to cover the costs of operation, causing a rightward shift of the exit threshold.

This picture is, however, incomplete without considering general equilibrium effects. As firms cut on hiring and larger flows of workers enter unemployment, the tightness falls on active segments of the labor market. Consequently, the job finding rate of workers dips, causing a fall in the value of unemployment as job prospects deteriorate. At the same time, as the degree of competition on the labor market diminishes, the cost of hiring drops and firms find it easier to hire workers. Resulting from these two general equilibrium effects, relatively low productivity firms have weaker incentives to separate, while productive one are encouraged to hire more. Which of these first moment or general equilibrium effects dominate is, in principle, ambiguous. The total effect of productivity shocks in the baseline calibration is displayed in the lower panel of Figure 8. The general equilibrium effects seem to dominate as the hiring and separation thresholds shift leftward, while the opposite forces affecting the exit threshold exactly cancel out.[21]

[21] The key to understand why general equilibrium effects are so strong in this economy lies in the free entry condition. The requirement that the value of firms remains constantly equal to the entry cost necessitates a strong reactivity of general equilibrium objects. For instance, the fall in the value of entering firms must be largely compensated by a decline in the hiring cost and tightness, ultimately benefiting incumbent firms to the point that they would rather grow than contract in response to a fall in productivity. Entrants, on the other hand, take the largest hit, and entry falls significantly.

Impulse responses

Figure 9 displays the impulse responses of various variables after the economy is hit by a negative 1% transitory shock to aggregate productivity. As one would expect, output and output-per-worker drop on impact and recover slowly, closely tracking the recovery in productivity. Total vacancies and hirings decrease, largely driven by a fall in entry that dominates the mild increase in hiring by incumbents. Consistent with Figure 8, total separations decrease, masking a fundamentally different impact on quits and layoffs. As entry falls and unemployed workers start flooding the labor markets, the probability of finding a job decreases for workers, making job-to-job (quits) transition less appealing. As a result, contracting firms reduce their use of quits, but intensify layoffs. The joint increase in layoffs with a reduction in hiring results in an overall rise in unemployment of about 3%, confirming our previous finding that the addition of firm heterogeneity to search-and-matching models may provide some amplification to aggregate productivity shocks. Turning to exits, even though the exit threshold is unaffected by aggregate productivity, total exits decline, reflecting mostly a lower aggregate rate of employment. Figure 10 breaks down the response of the economy according to partial equilibrium, entry and general equilibrium.

4.3 Volatility shocks

We are now ready to address the main question that motivated this study: what is the impact of uncertainty or volatility shocks on the economy as a whole, the labor market and the cross-section of firms? Volatility shocks produce a variety of effects that are, in general, difficult to disentangle, including real option effects, Oi-Hartman-Abel effects, realized volatility and general equilibrium effects.

Employment policy

Let us first examine how an increase in idiosyncratic volatility affects the optimal employment policy of firms as a function of their individual characteristics. Using the same convention as in the previous section, Figure 11 presents the impact of a positive one standard deviation shock to volatility v in both partial (upper panel) and general equilibrium (lower panel).

The partial equilibrium figure allows us to isolate the real option effects. Consistent with previous literature, an increase in volatility raises the option value of waiting, and firms have stronger incentives to delay decisions that involve irreversibilities. Because search incurs sunk costs, the decision to hire a worker is partially irreversible and firms have a tendency to defer recruitment to future periods. Likewise, firms may prefer to delay laying off workers, in order to avoid repaying the search costs if conditions were to improve. Consequently, the hiring and separation regions shrink, leading to a widening of the inaction band. For the same reason, exits, being fully irreversible, subside substantially and the exit threshold falls back left.

The upper panel of Figure 11 reveals an important finding: search frictions alone do not seem large enough to generate strong option value effects. Despite being qualitatively consistent with the uncertainty literature, the *wait-and-see* effects, visible in the widening of the inaction band, are surprisingly small. This finding stems from the fact that labor market mobility in the US is high. For instance, Davis et al. (2013) estimate the job filling rate probability to be 5.2% per day, about 80% per month, while the average job finding probability per worker is about 45% per month. Because these numbers are high, the degree of irreversibility of a hire or a layoff cannot be too large. Hence, any model calibrated to match average labor market flows in the US would have difficulty generating strong option value effects, unless additional costs or heteregeneity among workers were considered.[22] One should not, however, jump too quickly to the conclusion that volatility is unimportant to explain the dynamics of the labor market. As we will see in the next section, time-varying volatility will prove to be important to explain several episodes in the data, mostly through its impact on reallocation.

The lower panel of Figure 11 incorporates the general equilibrium dimension. The movements in the various thresholds, in this case, are mostly due to an effect commonly called the *Oi-Hartman-Abel effect*.[23] Because of the Oi-Hartman-Abel effect and an embedded real option value, idiosyncratic volatility shocks increase the value of firms, causing a large flow of firms to enter the economy. Consequently, the labor market tightness rises, the cost of hiring shoots up, making firms hire less and pushing the hiring threshold further down. Simultaneously, a higher tightness leading to a greater job finding probability, the value of quitting and the value of unemployment rise, leading firms to separate more and the separation region to expand, effectively overriding the option value effect. Finally, the exit region widens in comparison to the partial equilibrium case, as the greater hiring costs reduce the expectation of future surpluses.

Impulse responses

Figure 12 displays the aggregate impulse responses of several variables to a transitory +5% idiosyncratic volatility shock. The response of the economy reflects the combination of various components: i) partial equilibrium (isolating real option effects), ii) general equilibrium, iii) entry, and iv) realized volatility. The term *realized volatility* designates the fact that dispersion across firms actually increases once the volatility shock is realized. As a result, even though the inaction band may widen, firms may hit the action thresholds more often and become more active in response to an increase in uncertainty. The model predicts that this effect is strong, and I thus attempt to quantify the relative importance of each component.

Figure 13 presents four different simulations under the same transitory shock to volatility, in which

---

[22] To explore the robustness of this claim, I run the same exercise in partial equilibrium by increasing the vacancy posting cost c and lowering the efficiency of the matching function m, such that $p(\theta) = \theta q(\theta) = m \cdot \theta (1 + \theta^{\gamma})^{-\frac{1}{\gamma}}$. The option value only becomes sizeable for extreme value of c and m that would be difficult to reconcile with the data.

[23] This effect, described in Oi (1961); Hartman (1972); Abel (1983), is well known in the uncertainty literature. Because firms increase employment when idiosyncratic productivity rises, while they reduce employment when productivity is low, the value of a firm is in general a convex function of productivity. As a result, a mean-preserving spread of idiosyncratic productivity tends to increase the value of firms.

I shut down the contribution of each component one by one. The black continuous line presents the response of the full economy. The blue dotted line attempts to shut down the realized volatility effect by imposing that the actual volatility remains constant in the simulation, even though firms' beliefs are hit by the uncertainty shock. The green dashed line additionally controls for the entry of firms by keeping the entry process constant to its steady-state level. Finally, the red dash-dotted line presents the partial equilibrium response of firms, holding realized volatility and entry constant, in an attempt to isolate firms from changes in general equilibrium objects.

The red dash-dotted line shows the importance of the real option effects. Consistent with our findings from Figure 11, hirings, separations and exits drop on impact because of an increase in the option value of waiting. Firms sensibly turn away from layoffs and substitute with quits to the point that quits rise, while layoffs plunge. Interestingly, hirings and quits end up rising to a point above their initial levels; an effect largely driven by the fact that more firms survive with the decline in exits. Adding the general equilibrium effects, the green dashed line reflects our findings from Figure 11: hirings fall even deeper, while the higher job finding probability encourages lower layoffs, as firms substitute towards additional quits. The blue line, which allows for endogenous entry, shows a similar pattern to the green dashed line except that hirings pick up immediately because of the surge in entry caused by the Oi-Hartman-Abel effect, discouraging future entrants and lowering hirings in the subsequent periods. The difference between the Full model (black continuous line) and the blue dotted lines captures the effect of realized volatility. As the figure illustrates, the realized volatility effects are extremely large and dominate all the previously mentionned effects. Because they are hit by more dispersed shocks, firms hit their action thresholds more often and pure volatility shocks result in more turnover across firms: hirings, quits, layoffs and exits all rise.

With our decomposition of labor market flows, we may now return to the aggregate impulse responses of Figure 12 to analyze the overall contribution of volatility to output and unemployment. Because of the Oi-Hartman-Abel effect, total output and output-per-person, aggregated over the cross-section of firms, rise as volatility increases. Vacancies increase, mirroring the evolution of total hires. Unemployment rises quite substantially because unemployment inflows (layoffs) dominate the outflows (hires from unemployment). Indeed, even though hirings increase in response to a volatility shock, a large part of that increase is accounted by job-to-job transitions, as workers reallocate from low to high productivity firms. Therefore, unemployment surges unambiguously reflecting the fact that a greater number of firms receive bad shocks.

5 Counterfactual Exercises

After our detailed analysis on the impact of productivity and idiosyncratic volatility on the labor market, we are now prepared to conduct our main quantitative exercise. We ask, in this section, whether the model can account for the US labor market experience over the period 1972-2009 and how much variation can be attributed to fluctuations in productivity and idiosyncratic volatility.

5.1 Description

I jointly estimate two series of shocks for aggregate productivity $\{y_t\}_{1972:1}^{2009:12}$ and idiosyncratic volatility $\{v_t\}_{1972:1}^{2009:12}$ by matching two natural empirical counterparts: i) the quarterly output-per-person series from the BLS, and ii) the annual cross-sectional IQR of innovations to idiosyncratic TFP from the Census. Since these two series are endogenous in the model, I use a procedure of search in the space of productivity and volatility shocks, which minimizes the distance between the empirical and simulated series. In both cases, the simulated series are computed following the same steps as in the data.[24]

Instead of using a standard HP filter to detrend the data in this exercise, I use the band-pass filter developed by Christiano and Fitzgerald (2003) and restrict my attention to fluctuations in the range of 6 to 32 quarters, as is commonly done in the business cycle literature. I adopt this method in order to remove high-frequency noise components from the empirical series.[25] To illustrate the difference between the two detrending approaches, Figure 14 presents the output-per-person series in panel (a) and the IQR series in panel (b), detrended using both methods. As the figure shows, the two series are very close, but the HP-detrended series display more high frequency variations, which turn out to be difficult to match with the model without extremely volatile, negatively autocorrelated shocks that cause a spurious amount of reallocation in the labor market.

I perform two counterfactual experiments. In the first experiment, I use the full model calibrated as in part 3 with both aggregate productivity and idiosyncratic volatility shocks. In the second experiment, in order to isolate the contribution of volatility, I run the same exercise in the version of the model with productivity shocks only and fit the output-per-person series alone, while volatility is kept constant to its mean. Figure 14 shows the fit with the empirical series on panel (a) and (b). As the figure illustrates, the fit of the simulated series with their empirical counterparts is almost perfect. Panel (c) and (d) report the imputed aggregate productivity and idiosyncratic volatility shock series.

5.2 Results

I now analyze the ability of the model to account for the various NBER-dated recessions over the period 1972-2009.[26] Figure 15 and 16 report output and unemployment in the data and in the model across the five episodes. Because the labor market flow data from the Job Openings and Labor Turnover Survey (JOLTS) is unavailable before 2001, the recession of 2007-2009 is the unique episode entirely covered by the dataset. Figure 17 displays the fit of the model for the various labor market flows provided by the JOLTS during this episode. The series are presented in log deviation from the

[24]Since the IQR measure controls for selection and productivity shocks have an impact on the selection of firms, the IQR responds to productivity. This effect is, however, small and the two series of shocks are well identified.

[25]Since an *iid* process has a flat frequency decomposition, note also that the band-pass filter reduces the incidence of *iid* measurement errors.

[26]These recessionary episodes are 1973Q4-1975Q1, 1980Q1-Q3, 1981Q3-1982Q4, 1990Q3-1991Q1, 2001Q1-Q4 and 2007Q4-2009Q2. I group together the recessions of 1980Q1-Q3 and 1981Q3-1982Q4 as a single recessionary episode because the two events are too close in time to allow the identification of separate turning points for peak-to-trough analysis.

peak (or trough for countercyclical variables) preceding the recession. Peaks (troughs) are identified as the local maxima (minima) that precede recessions.[27] Peak-to-trough measures in both variables are detailed on Table 5 for the various episodes.

The model is quite successful at explaining fluctuations in output with either version of the model. Early recessions in particular, including the recession of 1990-1991, display very little difference between the versions with and without volatility shocks. Productivity thus appears to be the main force driving variation in output. During the more recent recessions, the presence of volatility shocks help explain an additional 0.5% to 1% decline in peak-to-trough measures. The recession of 2001 displays the largest discrepancy with the data, but the overall fit is nonetheless very satisfactory.

The unemployment series suggest a more important role for volatility. The model with productivity shocks explains in general between 40% to 60% of the total increase in unemployment in the early recessions of 1972 to 1991. The contribution of productivity to unemployment variation then falls to about 20-30% in the last two recessions. The introduction of volatility shocks, as we know, contributes to unemployment through a combination of various effects, including real option effects, but mostly by intensifying the reallocation of labor across firms. The simulated series confirm the importance of volatility shocks to explain variation in unemployment. The full model explains between 70% and 80% of the total increase in unemployment during the recessions of 1973-1975 and 1980-1982. It captures reasonably well the rise and subsequent fall in unemployment during 1990-1991, though the reversal in employment takes place earlier in the model than in the data. Volatility appears to have played a major role during the recession of 2001 as it explains about 50% of the total increase in unemployment, while only 30% is attributable to productivity. The model, however, cannot justify the slow decline in unemployment that took place after 2003.

Surprisingly, despite a large peak in volatility in 2007, the presence of volatility shocks only increase the explanatory power of the model from 20% to 40% of the total rise in unemployment. Part of the reason stems from the fact that volatility rose slowly from 2005 to 2007 in the Census data, as shown in Figure 1. As a result, the reallocation of labor in the model occurs progressively during that period and only few workers have to experience unemployment before finding new jobs. The labor market flow series of Figure 17, however, provide encouraging support that volatility is essential to understand the US labor market experience. As the figure illustrates, the full model does a very satisfactory job at explaining the evolution of hirings, layoffs and quits during the 2007-2009 period, and clearly outperforms the model with productivity shocks only. In particular, the model with volatility shocks accounts for more than 80% of the total peak-to-trough variations in these variables. It misses, however, the evolution of total vacancies which do not fall as much as in the data. Since the model captures most of the fall in hirings, the discrepancy between the model and the data must result from a larger decline in the model's vacancy yield and an insufficient decrease in the labor market tightness. While my findings suggest that time-varying volatility has played an important role in the

[27] I also impose that a peak (trough) must be preceded by at least three quarters of consecutive growth (decline) to avoid selecting blips in the data.

2007-2009 recession, it also shows that volatility alone does not seem sufficient to account for the total variation in unemployment without additional ingredients such as, for instance, financial frictions.

To summarize our conclusions, we have seen that the model with both productivity and volatility shocks can quite sastisfactorily account for the joint evolution of output and unemployment across various past episodes. Time-varying volatility, mostly through its impact on the reallocation of labor, appears to be an important driver of labor market flows and contributes to offer a more complete view of the labor market. We found, however, only partial support for the role of productivity and volatility in the 2007-2009 recession, as the combination of both shocks explains at most 40% of the total rise in unemployment.

5.3 Labor wedge

Several authors have reported large movements in the labor wedge over past recessions, including in particular the recession of 2007-2009. I conclude this section with an exploration of the model's predictions for the labor wedge, namely the ratio of the marginal rate of substitution of consumption for leisure and the marginal product of labor. Following Chari et al. (2007), I define the labor wedge as the implicit labor tax,

$$1 - \tau_{l,t} = -\frac{u_H}{u_C}\left(C_t, H_t\right)/F_H\left(K_t, H_t\right) = \frac{C_t}{Y_t} H_t^{1+\nu},$$

assuming $u\left(C, H\right) = \log C - \xi \frac{H^{1+\nu}}{1+\nu}$ and $F\left(K, H\right) = K^\alpha H^{1-\alpha}$, where C holds for total consumption, H hours and K capital.[28] I first compute the response of the labor wedge to aggregate productivity and volatility shocks and report the results on Figure 18 in the Appendix. Interestingly, both the negative productivity and positive volatility shocks lead to a decline in the labor wedge, equivalent to an implicit increase in a tax on labor income. Interestingly however, while the drop in the labor wedge is almost negligible in the case of a productivity shock, it is quite large in the case of a +5% volatility shock, which represents about a third of its standard deviation. The decline is due to the fact that volatility shocks imply an increase in unemployment, mostly through intensified reallocation of labor, as well as an increase in output through the Oi-Hartman-Abel effect, which both push the labor wedge down. Productivity shocks on the other hand do not produce such a large decline in the labor wedge because the drop in employment is largely compensated by a fall in output.

Going back to the counterfactual exercises, I compute the labor wedge both in the model and in the data and report the peak-to-trough measures in Table 5. As the table shows, recessions are usually followed by a worsening of the labor wedge, implying that the implicit tax on labor rises in the aftermath of a recession. As was pointed out before, the last recession appears as the worst episode with a fall of about 7% in the labor wedge under the chosen specification. On the other hand, the

[28]See appendix C for more details on the parametrization and data sources. Consumption is defined by the resource constraint in the model as total output net of costs (vacancy, entry and operating costs). Since there is no intensive margin of labor, I use total employment instead of hours in the model.

recession of 1990-1991 appears as the mildest episode. Table 5 shows that the full model with volatility shocks is in general more successful at explaining movements in the labor wedge than the version with productivity shocks only, as one could have expected from the previous paragraph. The model does reasonably well for the recessions of 1980-1982 to 2001, during which uncertainty seems to have played a larger role, but only explains a fraction of the decline in the wedge for the last recession, mirroring the fact that uncertainty explains a limited part of unemployment during this episode. Surprisingly, the model fails at replicating the deterioration of the labor wedge during the recession of 1973-1975. This result is, however, driven by the fact that the model overpredicts the fall in output during this recession, limiting the decline in the wedge, and because the imputed volatility shock during this episode is rather small as Figure 14 shows.

6 Conclusion

In this paper, I have developed a dynamic search-and-matching model of the labor market with firm dynamics and heterogeneity in productivity and sizes. The model is based on directed search and allows for endogeneous separations, on-the-job search and endogenous entry and exit of firms. Despite the amount heterogeneity, the model is highly tractable and can accomodate a variety of aggregate shocks, thanks to the property of block recursivity, which I exploit to analyze the out-of-steady-state dynamics of the model.

After showing that the model can replicate salient features of firm behavior at the establishment level, I use this framework to analyze the role of time-varying idiosyncratic risk on aggregate unemployment fluctations and on the labor market. I show that the response of the economy to productivity and volatility shocks is complex and hides a variety of effects. The response of the economy to volatility shocks, in particular, is the combination of various effects ranging from real option, Oi-Hartman-Abel to general equilibrium effects. My findings suggest, however, that the real option effects are mild and dominated by realized volatility effects. In other words, volatility shocks intensify the reallocation process, inducing larger gross labor market flows and higher unemployment.

In a series of counterfactual experiments, I examine the ability of the model to account for the US labor market experience during past historical episodes. Feeding the model with a series of shocks that match the productivity and volatility data, I show that the model offers a quite satisfactory account of various past recessions. Time-varying volatility appears as an important driver of labor market fluctuations, in particular for the recessions of 1990-1991, 2001 and 2007-2009. The success is, however, only partial for the last recession, as the joint combination of productivity and volatility explains at most 40% of the observed increase in unemployment.

The model is quite flexible and could be used in a variety of setups with aggregate shocks or transitional dynamics. For instance, because it allows for decreasing returns, a possible extension would be to introduce monopolistic competition and study the model's dynamic implications for international trade. Applications to markets other than the labor market may also provide interesting insights. For instance, Boualam (2014) proposes an application to the banking industry and studies the dynamics of the credit market. Other extensions, such as the introduction of concave utility or skill heterogeneity among workers, also seem promising.

Regarding the role of uncertainty, the model has the—perhaps surprising—implication that real option effects are weak. This result stems from the fact the employment decisions can be easily reversed when search frictions are the only costs associated to the reallocation of labor. This conclusion may, however, change with the introduction of additional sunk costs, such as job-specific human capital investments. Uncertainty may also affect employment through other channels. For instance, adding stronger discount factor effects could attenuate the Oi-Hartman-Abel effects and lower the response in entry and hiring. Other sources of uncertainty not considered in this paper could also reveal important, for instance policy uncertainty as studied in Baker et al. (2014) and Fernández-Villaverde et al. (2013). Financial frictions, in interaction with uncertainty shocks, could also improve the response of the model during the recession of 2007-2009.

References

Abel, A. B. (1983): "Optimal investment under uncertainty," *The American Economic Review*, 228–233.

Acemoglu, D. and W. Hawkins (2010): "Wages and Employment Persistence with Multi-worker Firms," Working paper, MIT and University of Rochester.

Andolfatto, D. (1996): "Business Cycles and Labor-Market Search," *The American Economic Review*, 86, 112–132.

Arellano, C., Y. Bai, and P. Kehoe (2010): "Financial Markets and Fluctuations in Uncertainty," Working paper, Federal Reserve Bank of Minneapolis and NBER; Arizona State University; Princeton University, University of Minnesota and Federal Reserve Bank of Minneapolis.

Azariadis, C. (1975): "Implicit contracts and underemployment equilibria," *The Journal of Political Economy*, 1183–1202.

Bachmann, R. and C. Bayer (2009): "Firm-Specific Productivity Risk over the Business Cycle: Facts and Aggregate Implications," Working paper, University of Michigan; IGIER, Universitá Bocconi.

Bachmann, R., S. Elstner, and E. Sims (2010): "Uncertainty and Economic Activity: Evidence from Business Survey Data," Working paper, University of Michigan; University of Notre Dame; Ifo Institute for Economic Research.

Backus, D., A. Ferrière, and S. Zin (2014): "Risk and Ambiguity in Models of Business Cycles," Working Paper 20319, National Bureau of Economic Research.

Baily, M. N. (1974): "Wages and employment under uncertain demand," *The Review of Economic Studies*, 37–50.

Baker, S. R., N. Bloom, and S. J. Davis (2014): "Measuring Economic Policy Uncertainty," *Working paper*.

Basu, S. (1996): "Procyclical Productivity: Increasing Returns or Cyclical Utilization?" *Quarterly Journal of Economics*, 111, 719–751.

Basu, S. and B. Bundick (2012): "Uncertainty shocks in a model of effective demand," Working paper, National Bureau of Economic Research.

Basu, S. and J. G. Fernald (1995): "Are Apparent Productive Spillovers a Figment of Specification Error?" *Journal of Monetary Economics*, 36, 165–188.

Basu, S. and M. S. Kimball (1997): "Cyclical Productivity with Unobserved Input Variation," Working Paper 5915, National Bureau of Economic Research.

Beaudry, P. and J. DiNardo (1991): "The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro data," *Journal of Political Economy*, 665–688.

Bernanke, B. S. (1983): "Irreversibility, Uncertainty, and Cyclical Investment," *Quarterly Journal of Economics*, 98, 85–106.

Bertola, G. and R. J. Caballero (1990): "Kinked adjustment costs and aggregate dynamics," 237–296.

Bianchi, F., C. Ilut, and M. Schneider (2014): "Uncertainty Shocks, Asset Supply amd Pricing over the Business Cycle," Working paper.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623–685.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. J. Terry (2014): "Really Uncertain Business Cycles," Working paper, Stanford University.

Boualam, Y. (2014): "Bank Dynamics and Credit Market Competition," Working paper, University of Pennsylvania.

Brown, C. and J. Medoff (1989): "The Employer Size-Wage Effect," *Journal of Political Economy*, 97, 1027–1059.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2007): "Business Cylce Accounting," *Econometrica*, 75, pp. 781–836.

Christiano, L., R. Motto, and M. Rostagno (2009): "Financial Factors in Economic Fluctuations," Working paper, Northwestern University and European Central Bank.

Christiano, L. J. and T. J. Fitzgerald (2003): "The Band Pass Filter," *international economic review*, 44, 435–465.

Davis, S. J., R. J. Faberman, and J. Haltiwanger (2006): "The Flow Approach to Labor Markets: New Data Sources and Micro-Macro Links," *Journal of Economic Perspectives*, 20, 3–26.

——— (2011): "Labor Market Flows in the Cross Section and Over Time," *Journal of Monetary Economics*, In Press, Corrected Proof.

Davis, S. J., R. J. Faberman, and J. C. Haltiwanger (2013): "The Establishment-Level Behavior of Vacancies and Hiring*." *Quarterly Journal of Economics*, 128.

Davis, S. J., J. C. Haltiwanger, and S. Schuh (1996): *Job Creation and Destruction*, MIT Press.

Dixit, A. K. and R. S. Pindyck (1994): *Investment under uncertainty*, Princeton university press.

Elsby, M. W. and R. Michaels (2013): "Marginal Jobs, Heterogeneous Firms, and Unemployment Flows," *American Economic Journal: Macroeconomics*, 5, 1–48.

Evans, D. S. (1987): "The Relationship Between Firm Growth, Size, and Age: Estimates for 100 Manufacturing Industries," *Journal of Industrial Economics*, 35, 567–581.

Fernández-Villaverde, J., P. A. Guerrón-Quintana, K. Kuester, and J. Rubio-Ramírez (2013): "Fiscal volatility shocks and economic activity," Working paper, National Bureau of Economic Research.

Gilchrist, S., J. Sim, and E. Zakrajšek (2010): "Uncertainty, Financial Frictions, and Investment Dynamics," Working paper, Boston University and Federal Reserve Board.

Guvenen, F., S. Ozkan, and J. Song (2014): "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 122, 621–660.

Haefke, C., M. Sonntag, and T. van Rens (2013): "Wage rigidity and job creation," *Journal of Monetary Economics*, 60, 887 – 899.

Hagedorn, M. and I. Manovskii (2013): "Job selection and wages over the business cycle," *The American Economic Review*, 103, 771–803.

Hall, R. E. and P. R. Milgrom (2008): "The Limited Influence of Unemployment on the Wage Bargain," *American Economic Review*, 98, 1653–1674.

Hartman, R. (1972): "The effects of price and cost uncertainty on investment," *Journal of economic theory*, 5, 258–266.

Hopenhayn, H. and R. Rogerson (1993): "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 101, 915–38.

Hopenhayn, H. A. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica: Journal of the Econometric Society*, 1127–1150.

Hornstein, A., P. Krusell, and G. Violante (2007): "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," Working paper, 13674, National Bureau of Economic Research.

Jurado, K., S. Ludvigson, and S. Ng (2014): "Measuring Uncertainty," *Working paper*.

Kaas, L. and P. Kircher (2011): "Efficient Firm Dynamics in a Frictional Labor Market," Working paper, University of Konstanz and University of Pennsylvania.

Khan, A. and J. K. Thomas (2008): "Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics," *Econometrica*, 76, 395–436.

Leduc, S. and Z. Liu (2012): "Uncertainty shocks are aggregate demand shocks," *Federal Reserve Bank of San Francisco Working Paper*, 10.

Lin, T.-T. T. (2014): "The Role of Uncertainty in Jobless Recoveries," .

McDonald, R. and D. Siegel (1986): "The Value of Waiting to Invest," *Quarterly Journal of Economics*, 101, 707–728.

Menzio, G. and S. Shi (2010): "Block Recursive Equilibria for Stochastic Models of Search on the Job," *Journal of Economic Theory*, 145, 1453–1494.

——— (2011): "Efficient Search on the Job and the Business Cycle," *Journal of Political Economy, 2011, 119 (3), 468-510.*, 119 (3), 468–510.

Merz, M. (1995): "Search in the labor market and the real business cycle," *Journal of Monetary Economics*, 36, 269–300.

Mortensen, D. T. and C. A. Pissarides (1994): "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61, 397–415.

Moscarini, G. and F. Postel-Vinay (2012): "The contribution of large and small employers to job creation in times of high and low unemployment," *The American Economic Review*, 102, 2509–2539.

Nagypál, E. (2007): "Labor-Market Fluctuations and On-the-Job Search," *Manuscript, Northwestern University.*

Oi, W. Y. (1961): "The desirability of price instability under perfect competition," *Econometrica: journal of the Econometric Society*, 58–64.

Prescott, E., S. Cociuba, and A. Ueberfeldt (2011): "U.S. Hours and Productivity Behavior Using CPS Hours Worked Data: 1947-III to 2011-IV," .

Schmieder, J. F. (2009): "Labor Costs and the Evolution of New Establishments," *manuscript.*

Shimer, R. (2005): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, 95, 25–49.

——— (2007): "Reassessing the Ins and Outs of Unemployment," Working paper, 13421, National Bureau of Economic Research.

Silva, J. I. and M. Toledo (2009): "Labor turnover costs and the cyclical behavior of vacancies and unemployment," *Macroeconomic Dynamics*, 13, 76–96.

Stole, L. A. and J. Zwiebel (1996): "Intra-Firm Bargaining under Non-Binding Contracts," *Review of Economic Studies*, 63, 375–410.

A Tables and Figures

Table 1: Estimated parameters

Parameter	Value	Description
Calibrated:		
A	1	Technology parameter
β	0.996	Monthly discount factor
γ	1.60	Job-finding probability parameter
α	0.85	Decreasing returns to scale coefficient
ρ_z	$0.95^{\frac{1}{3}}$	Autocorrelation of idiosyncratic productivity z
Estimated:		
ρ_y	0.990	Autocorrelation of aggregate productivity y
σ_y	0.042	Standard deviation of aggregate productivity y
$\overline{v}$	0.533	Standard deviation of idiosyncratic productivity z
ρ_v	0.979	Autocorrelation of volatility process v
σ_v	0.132	Standard deviation of volatility process v
ρ_{yv}	-0.400	Correlation between $\varepsilon_{y,t}$ and $\varepsilon_{v,t}$
b	1.403	Home production
c	1.789	Vacancy posting cost
λ	0.366	Relative search efficiency of employees
k_e	14.21	Entry cost
k_f	1.956	Operating cost

Table 2: Targeted moments

Moment	Empirical value	Simulated
$\rho[Y_t]$	0.839	0.780
$\sigma[Y_t]$	0.017	0.016
$IQR(e_{z,t})$	0.393	0.384
$\rho\left[IQR\left(e_{z,t}\right)\right]$	0.760	0.763
$\sigma\left[IQR\left(e_{z,t}\right)\right]$	0.049	0.049
$corr\left[IQR\left(e_{z,t}\right), Y_t\right])$	-0.092	-0.106
UE rate	0.450	0.435
EU rate	0.026	0.026
EE rate	0.029	0.029
Average establishment size	15.6	15.2
Entry / Total job creation	0.21	0.27

Notes: UE, EU and EE are monthly transition rates. The notation ρ stands for autocorrelation and σ for standard deviation. Y_t denotes output. The autocorrelation and standard deviation of log-detrended output are quarterly. $IQR\left(e_{z,t}\right)$ denotes the interquartile range of annual innovations to idiosyncratic productivity.

Table 3: Business cycle statistics

	Data		Model (y only)		Model (y+v)	
	(1)	(2)	(3)	(4)	(5)	(6)
	Std Dev.	cor(Y,x)	Std Dev.	cor(Y,x)	Std Dev.	cor(Y,x)
Y	0.017	1	0.017	1	0.017	1
Y/L	0.012	0.590	0.014	0.998	0.012	0.923
U	0.121	-0.859	0.062	-0.960	0.089	-0.722
V	0.138	0.702	0.037	0.687	0.053	0.267
Hirings	0.058	0.677	0.035	0.571	0.049	0.202
Quits	0.102	0.720	0.064	0.855	0.071	0.648
Layoffs	0.059	-0.462	0.045	-0.981	0.086	-0.600

Notes: Time series are aggregated to a quarterly frequency and presented in log-deviation from an HP trend with parameter 1600. Y is output, Y/L output per person, U unemployment, V vacancies. Quits are identified as job-to-job transitions in the model. See appendix C for data sources.

Table 4: Comparison with standard Diamond-Mortensen-Pissarides model

	Data		Shimer (2005)		Model (y only)	
	Std Dev.	cor(Y,x)	Std Dev.	cor(Y,x)	Std Dev.	cor(Y,x)
Y	0.017	1	0.017	1	0.016	1
Y/L	0.012	0.590	0.017	1	0.014	0.998
U	0.121	-0.859	0.007	-0.982	0.062	-0.960
V	0.138	0.702	0.021	0.993	0.037	0.687
Hirings	0.058	0.677	0.003	0.448	0.035	0.571
Quits	0.102	0.720	-	-	0.064	0.855
Layoffs	0.059	-0.462	0.001	0.931	0.045	-0.981

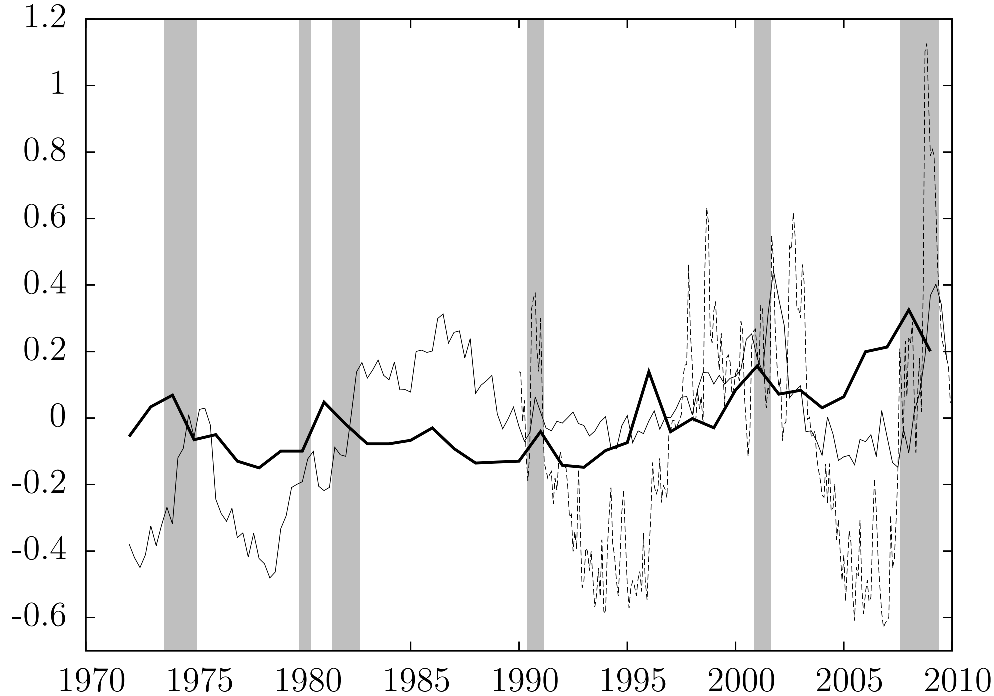
Notes: Time series are presented in logs. Quarterly time series detrended using an HP filter with parameter 1600. Y is output, Y/L output per person, U unemployment, V vacancies. Quits are identified as job-to-job transitions in the model. See appendix C for data sources. I compare simulated moments from a standard DMP model calibrated as in Shimer (2005) to my model with productivity shocks only.

Table 5: Peak-trough variations across various recessions

	1973-1975	1980-1982	1990-1991	2001	2007-2009
Output					
Data	-0.082	-0.069	-0.019	-0.038	-0.048
Model (y+v)	-0.096	-0.064	-0.025	-0.040	-0.042
Model (y only)	-0.089	-0.058	-0.019	-0.030	-0.038
Unemployment					
Data	0.490	0.441	0.124	0.328	0.521
Model (y+v)	0.403	0.319	0.168	0.267	0.212
Model (y only)	0.325	0.214	0.061	0.101	0.121
Labor wedge					
Data	-0.059	-0.046	-0.015	-0.056	-0.069
Model (y+v)	-0.016	-0.039	-0.020	-0.030	-0.021
Model (y only)	-0.023	-0.016	-0.005	-0.009	-0.010

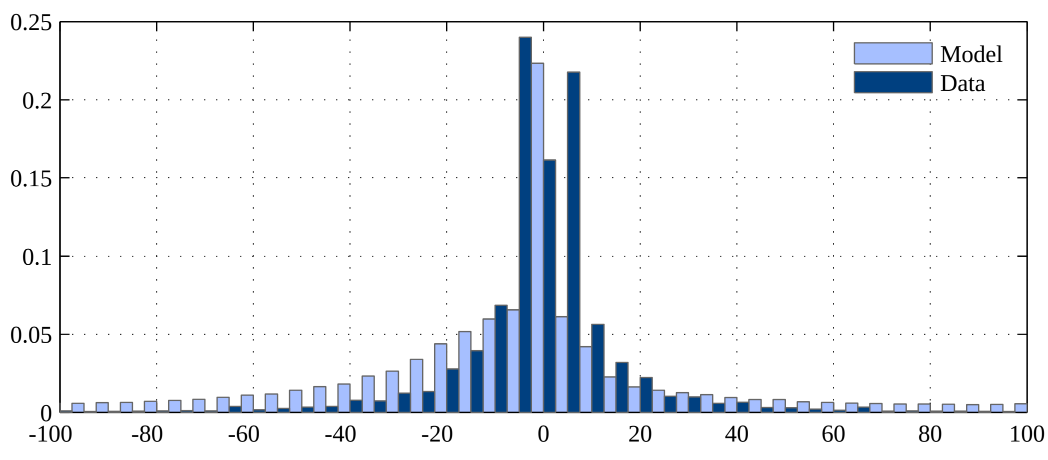
Notes: The peak-trough measures are computed in log deviation. The time series are detrended using a band pass filter for fluctuations from 6 to 32 quarters. Peaks (troughs) are identified as the first local maximum (minimum) preceding the recessionary period which follows at least three quarters of growth (decline). Simulated data is aggregated at the quarterly level.

Figure 3: Various measures of micro-level risk



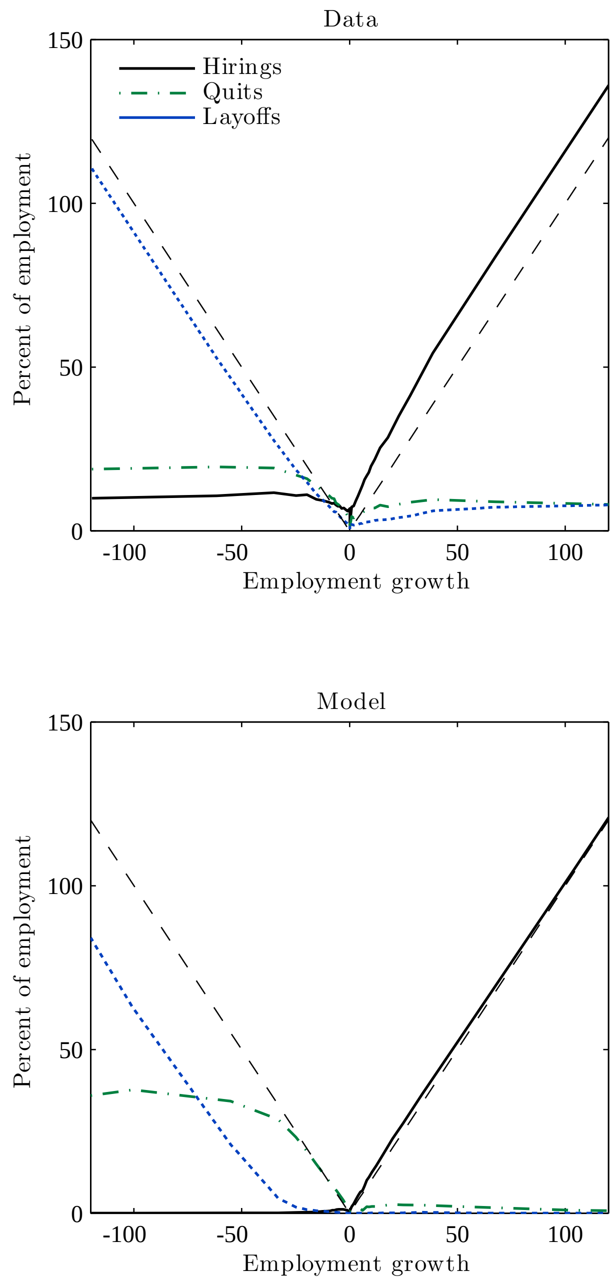
Notes: Data are shown in log deviations from their long-run averages. The thick curve shows the idiosyncratic risk measure from Census data constructed by Bloom et al. (2014); the thin curve shows the cross-sectional dispersion of annual sales growth from Compustat; the dashed line represents the VIX measure constructed by the CBOE. Shaded areas correspond to NBER recessions. See Appendix C for details.

Figure 4: Distribution of quarterly establishment growth rates



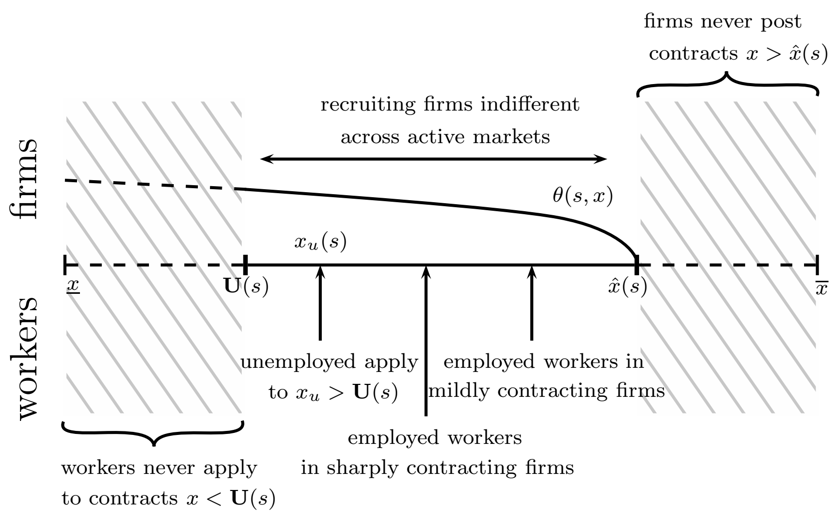
Notes: Quarterly data from 2008 tabulated from the BED dataset by Davis et al. (2011). Simulated distribution aggregated over a three-month interval.

Figure 5: Empirical and simulated employment policies as a function of growth



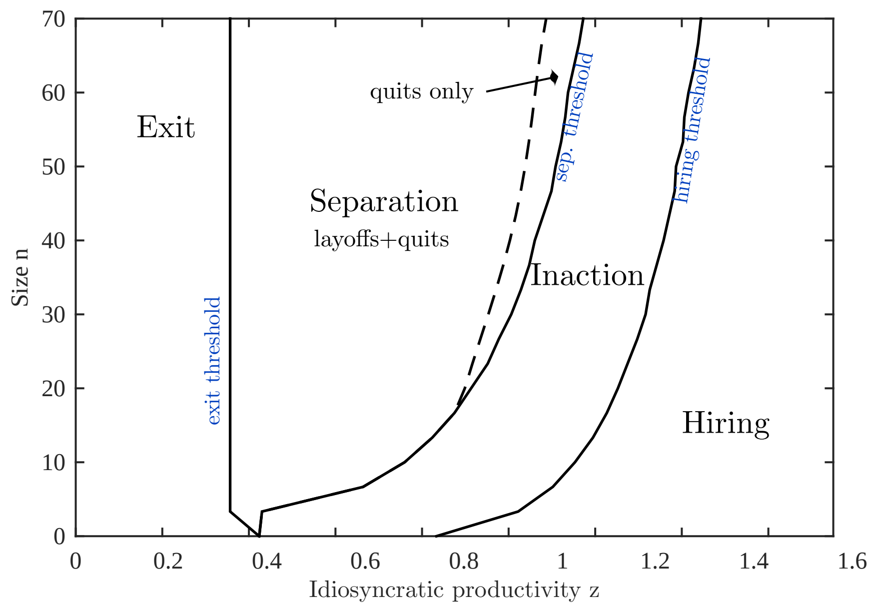
Notes: Tabulations from the BED dataset by Davis et al. (2011). Simulations aggregated over a three-month period. The averages are employment weighted. The dashed lines are the -45° and $+45^\circ$ lines to show the minimal level of separations and hirings needed to achieve the corresponding growth rate.

Figure 6: Description of labor market equilibrium



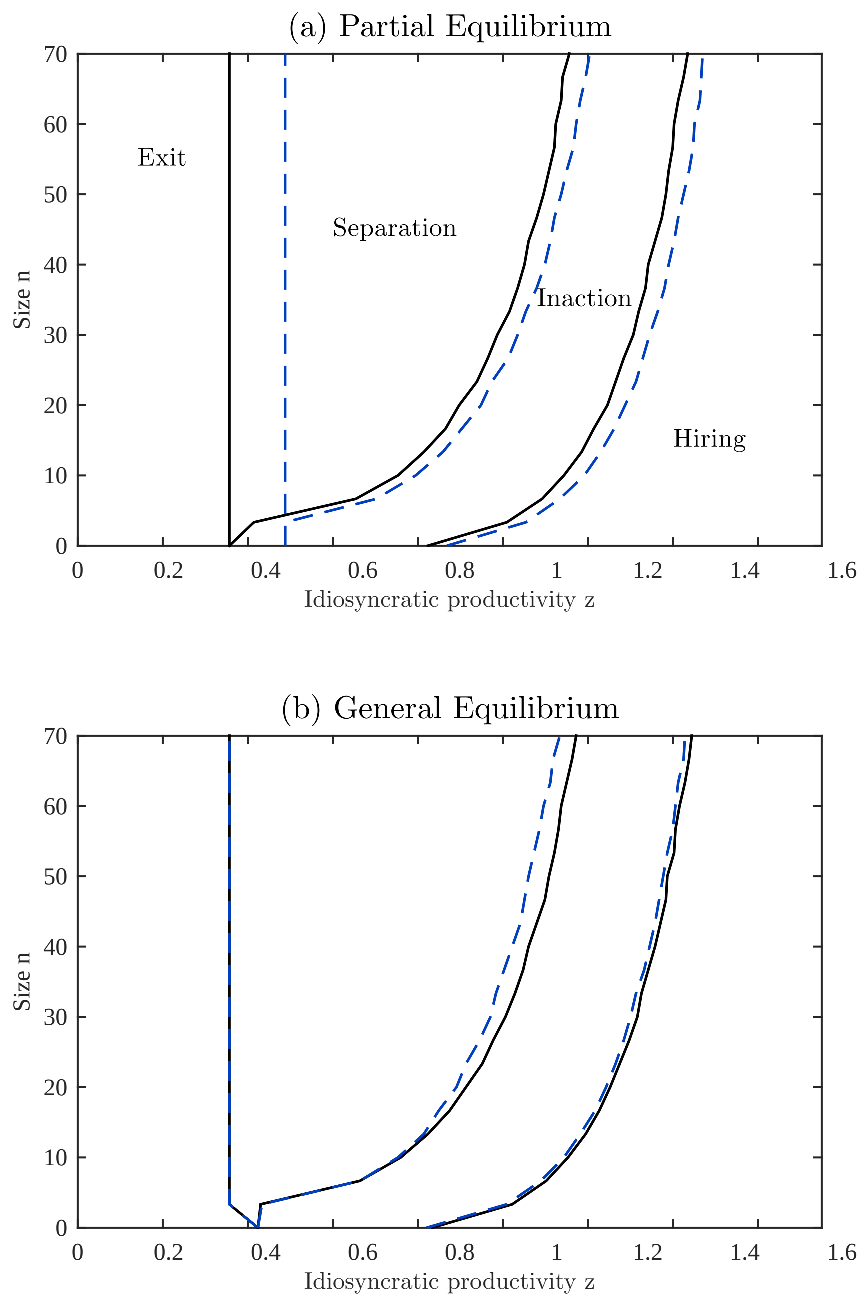
Notes: 1) The equilibrium market tightness $\theta(s,x)$ decreases with the value of the contract x. Because offering lower-paying contracts yields higher profits to firms, more vacancies are posted for low-x markets until the job filling probability drops sufficiently so that recruiting firms are effectively indifferent between active markets. 2) Workers are not indifferent between markets, because they have different outside options. Having the lowest outside option, i.e., unemployment, unemployed workers are less willing to tolerate low job finding probability and apply to markets with a low wage-utility x but high job finding probability. Because of efficiency, the relevant concept of outside option for employed workers is the shadow value of maintaining employment. It is thus possible to rank where employed workers apply for jobs: workers in sharply contracting firms have a lower outside option and apply to lower paid jobs than workers in mildly contracting firms. 3) Markets such that $x < \mathbf{U}(s)$ are inactive in equilibrium because unemployed workers never apply to jobs with a value below that of unemployment. Similarly, firms never post vacancies in markets with $x > \hat{x}(s)$, a point at which tightness is 0 and the job filling probability is 1, because offering higher-paying contracts cannot increase the job filling probability further.

Figure 7: Firm's action thresholds in the space of (z, n)



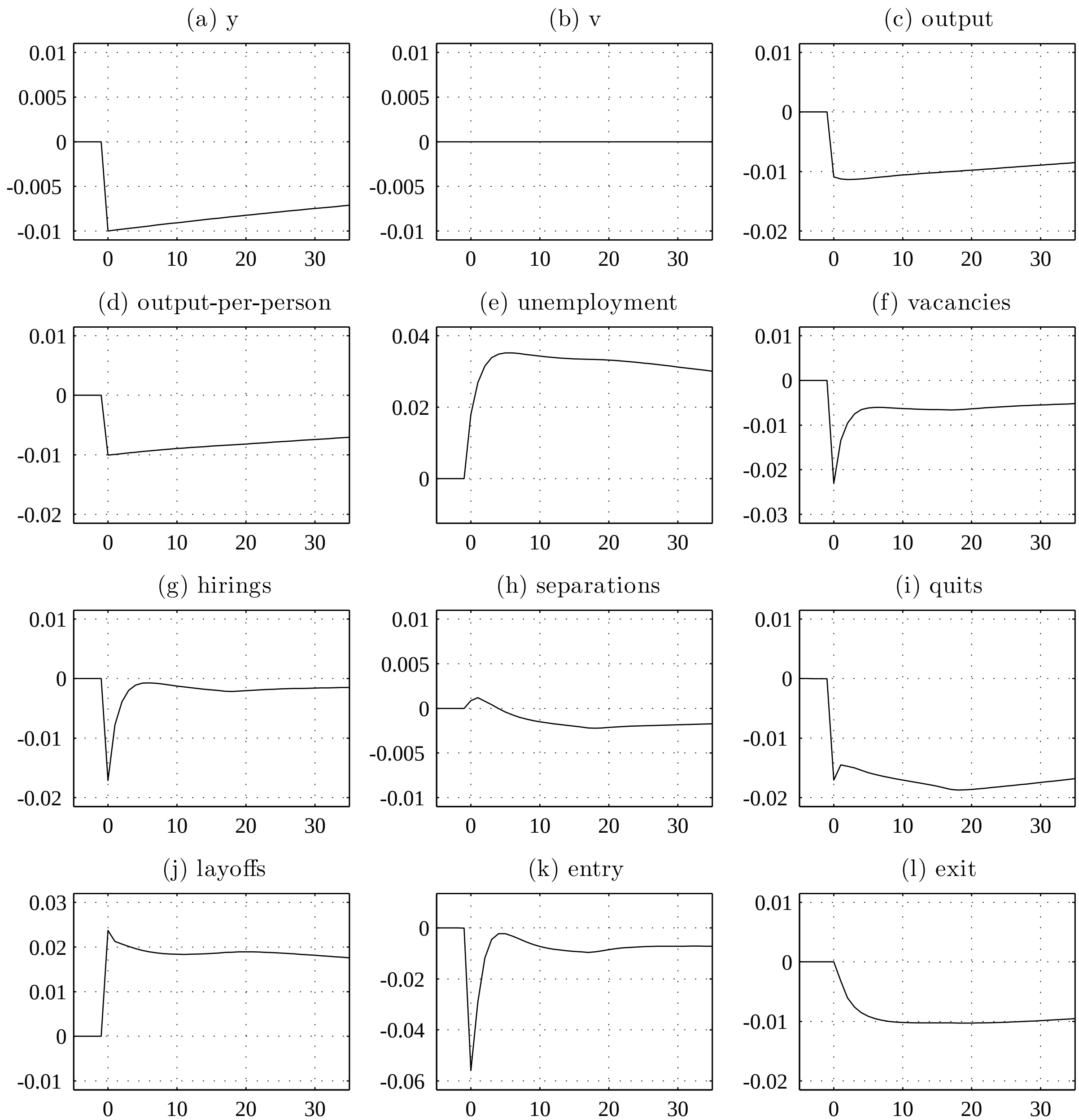
Notes: The optimal policies depicted on this figure correspond to the baseline calibration, holding the aggregate productivity y and volatility v to their mean values. Several points are worth noticing. 1) The areas corresponding to the different margins of adjustment are distinct and do not overlap, with the exception that firms separating from some of their workers tend to use, in general, a mix of quits and layoffs. However, hires and separations never occur at the same time because it is more costly for firms to hire new workers than retain the current workforce. 2) There exists a narrow band between the dashed line on the figure and the separation threshold, where firms exclusively separate from their workforce using quits. This feature is due to the fact that workers are strictly better off switching jobs directly, instead of going through a painful spell of unemployment. Firms successfully internalize this fact and send their workers looking for jobs outside before laying them off. However, the job-to-job transition technology is limited and quickly crowds out, so that firms willing to separate from a larger fraction of their workforce also use layoffs.

Figure 8: Firm's optimal policy after a negative one standard deviation shock to y



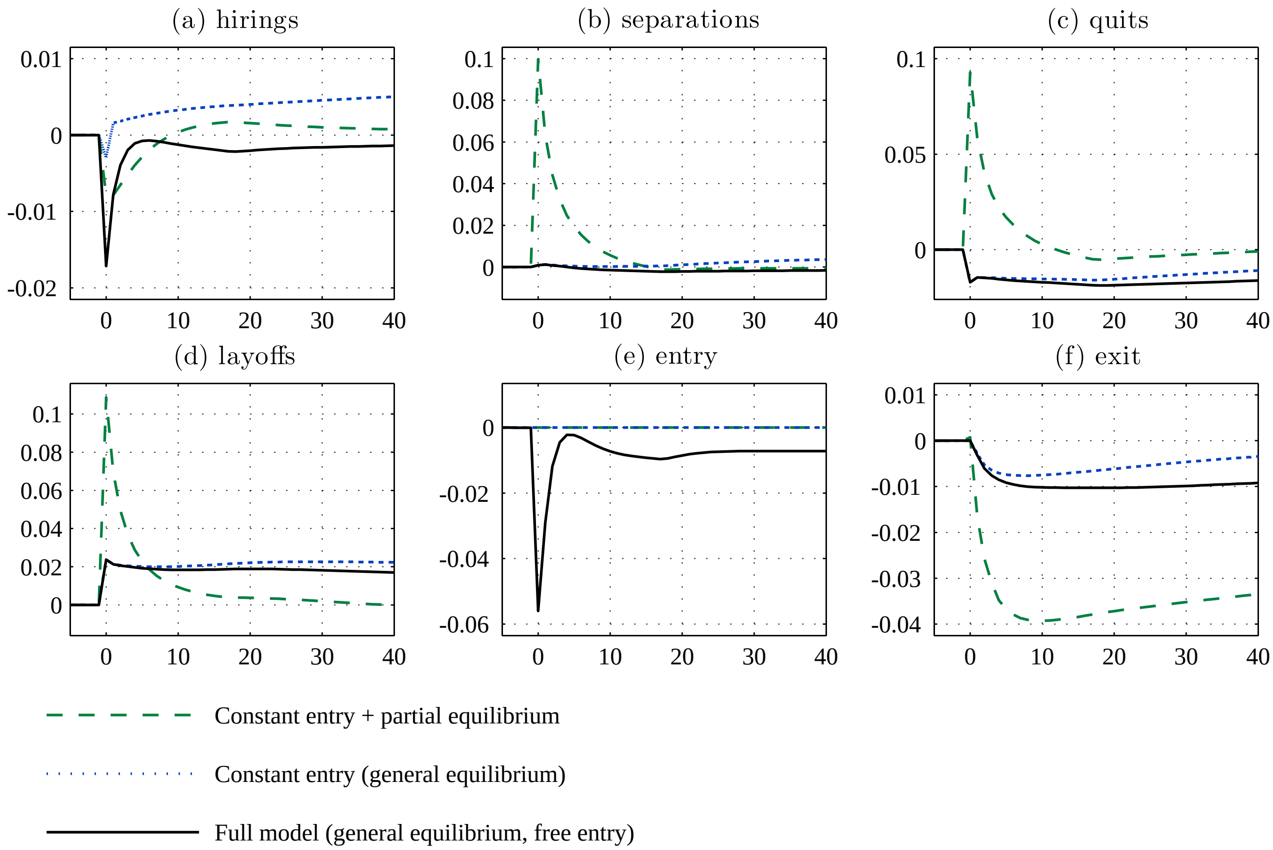
Notes: The black continuous line corresponds to the firm's optimal policy before the shock, and the dashed blue line is after the shock. The general equilibrium panel corresponds to the full model. The partial equilibrium is computed holding the hiring cost and the value of unemployment constant after the shock.

Figure 9: Response to a -1% transitory shock to aggregate productivity y



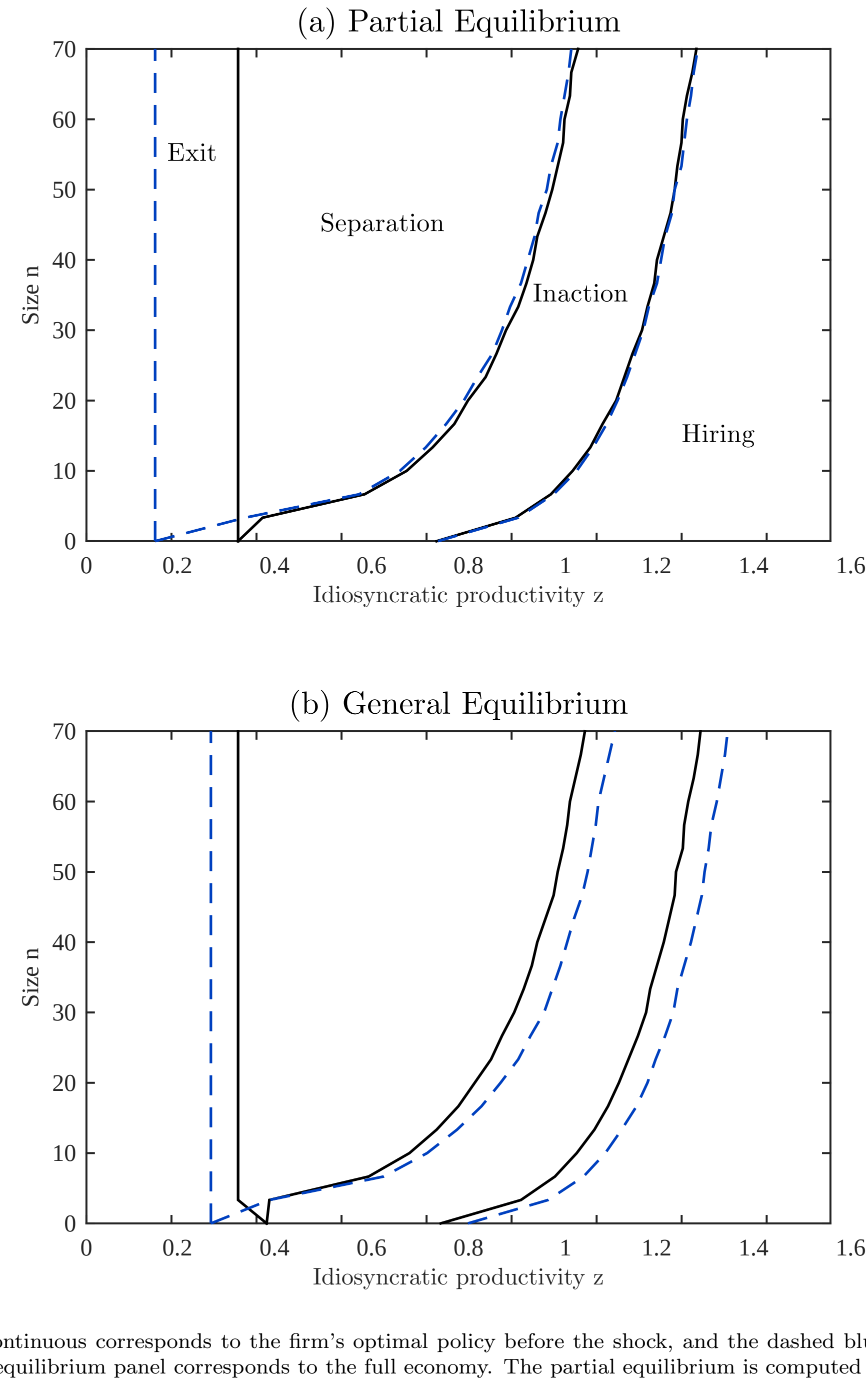
Notes: Series presented in log deviation from their steady state values when aggregate productivity and volatility are set to their means. The time period is a month and the shock hits at time $t = 0$. Separation is the sum of quits and layoffs. Entry and exit are expressed in total employment.

Figure 10: Breakdown of response to -1% transitory shock to aggregate productivity y



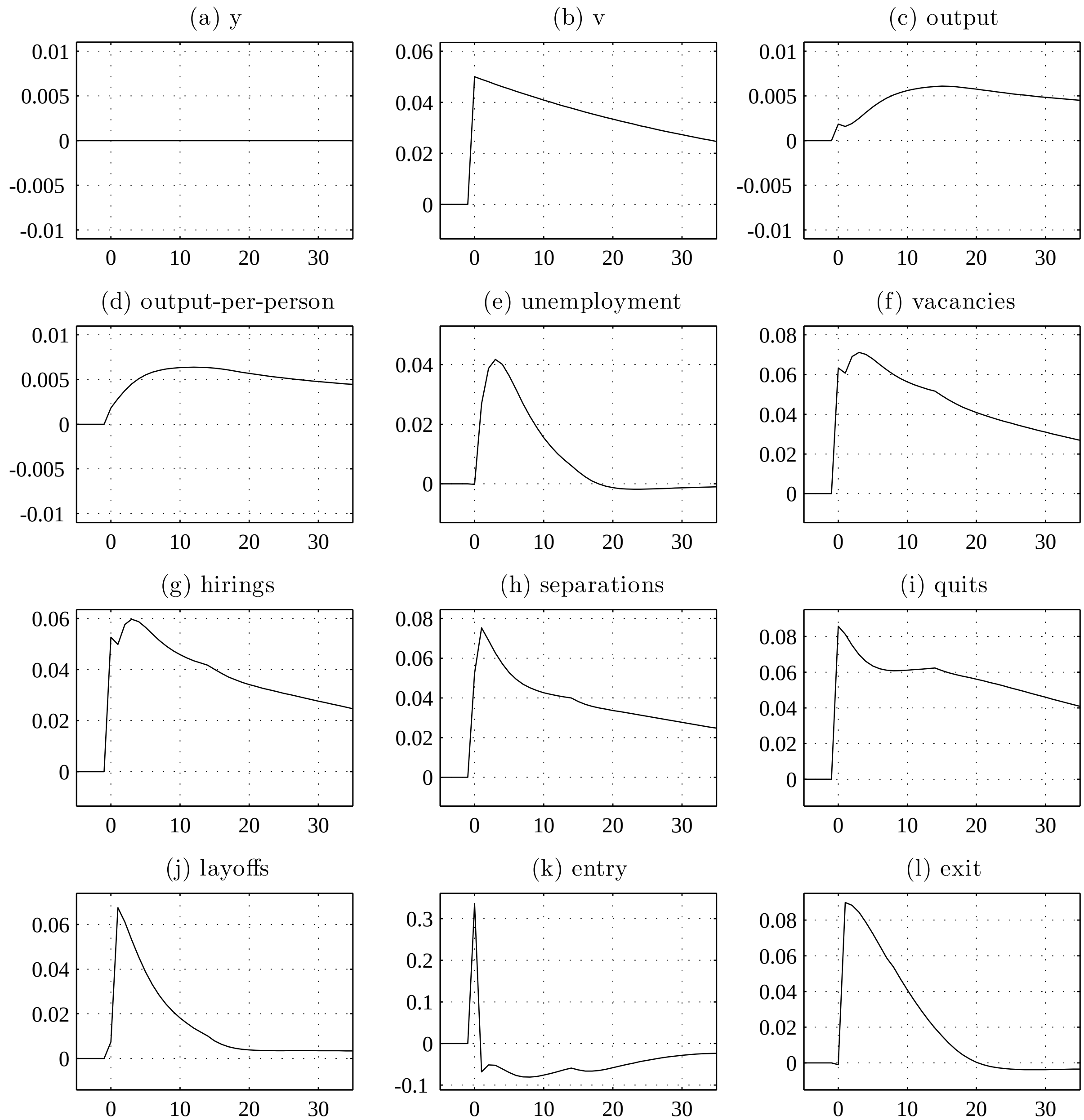
Notes: The black continuous line corresponds to the full economy; the blue dotted line to an economy with constant entry set to its steady state value; the green dashed line to an economy with constant entry and partial equilibrium. The series are presented in log deviation from the steady state when aggregate productivity and volatility are set to their means. The time period is a month and the shock hits at time $t = 0$. The shock is identical to that in Figure 9.

Figure 11: Firm's optimal policy after a positive one standard deviation shock to v



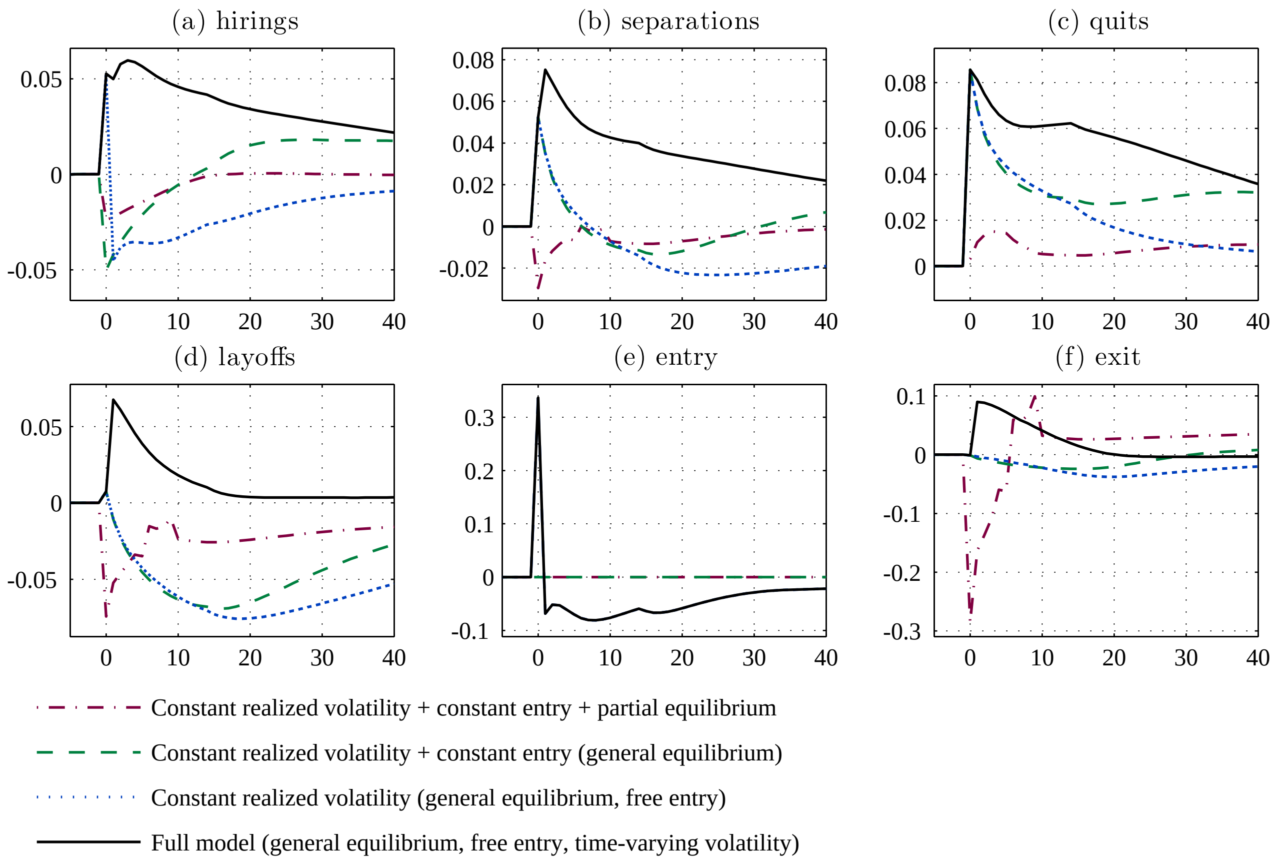
Notes: The black continuous corresponds to the firm's optimal policy before the shock, and the dashed blue line is after the shock. The general equilibrium panel corresponds to the full economy. The partial equilibrium is computed holding the hiring cost and the value of unemployment constant after the shock.

Figure 12: Response to +5% transitory shock to idiosyncratic volatility v



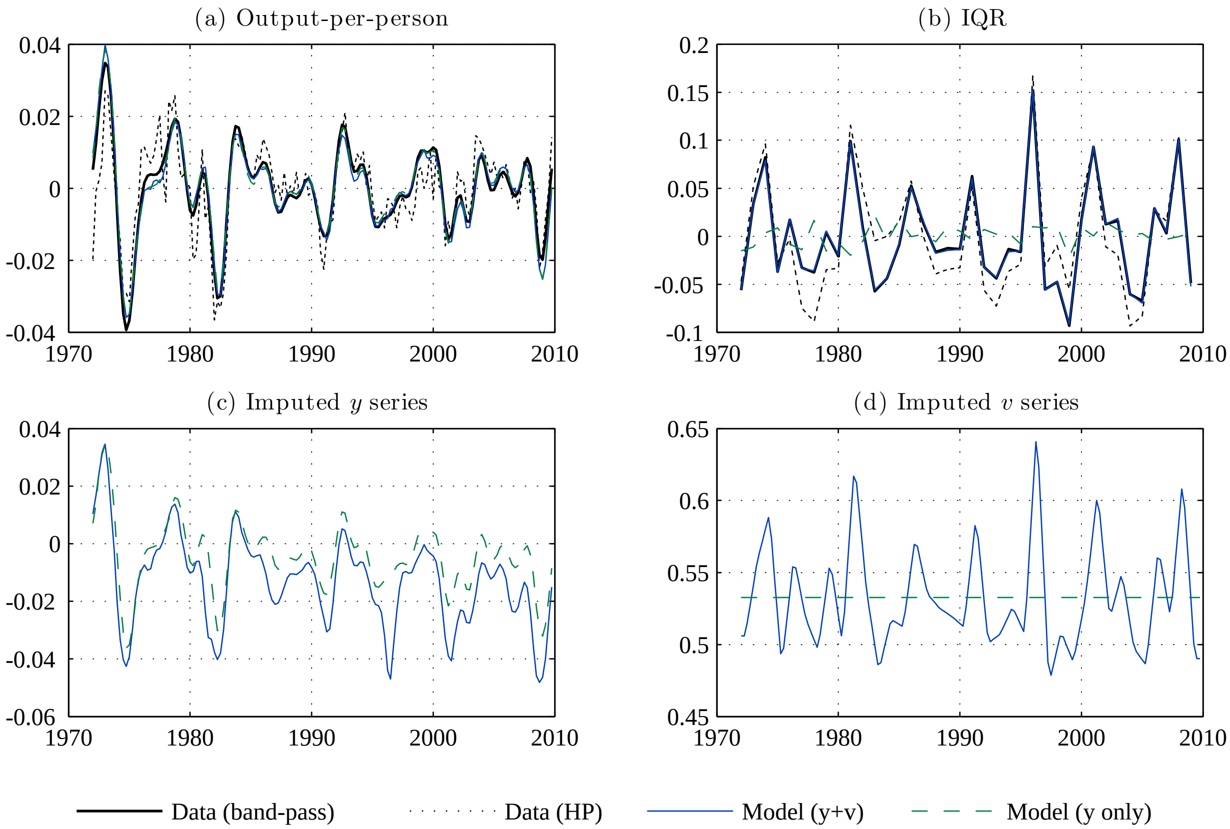
Notes: Series presented in log deviation from their steady state values when innovations to aggregate shocks are set to 0 for a long time. The time period is a month and the shock hits at time $t = 0$. Separation is the sum of quits and layoffs. Entry and exit are expressed in total employment.

Figure 13: Breakdown of response to +5% transitory shock to volatility v



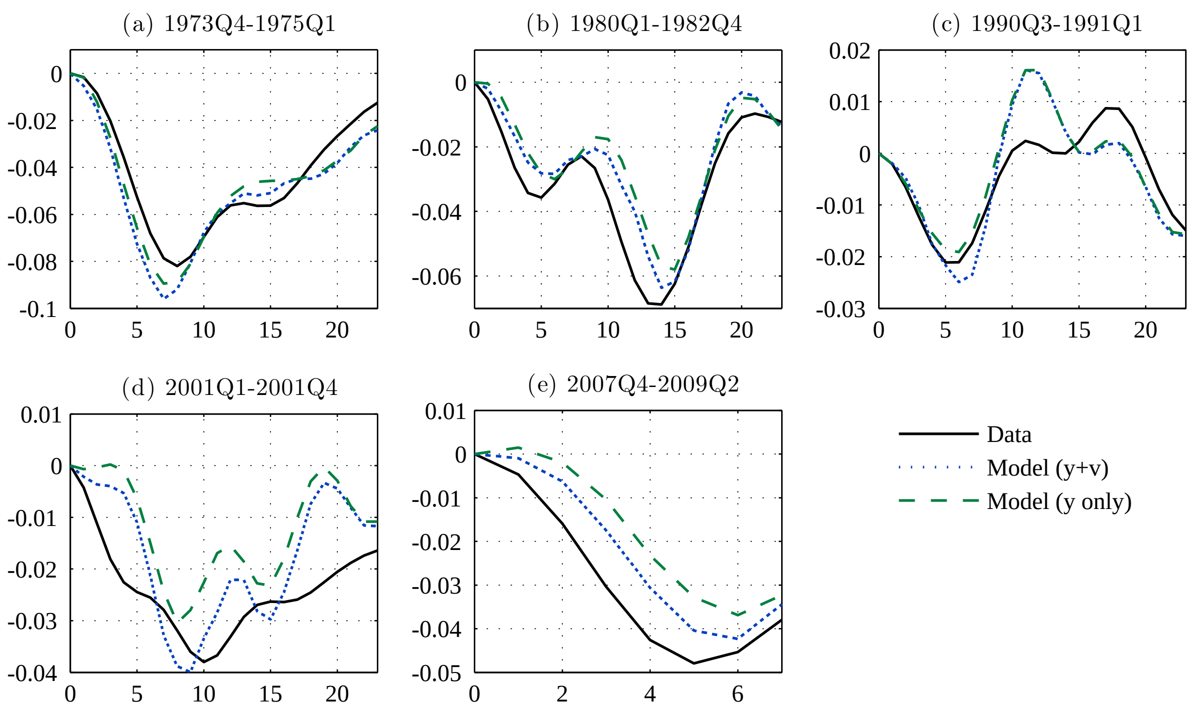
Notes: The black continuous line corresponds to the full economy; the blue dotted line to an economy with constant realized volatility fed but the same entry process as the full economy; the green dashed line to an economy with constant realized volatility and entry set to its steady state value; the red dash-dotted line to an economy with constant realized volatility, entry set to its steady state value and partial equilibrium. Series presented in log deviation from the steady state when innovations to aggregate shocks are set to 0 for a long time. The time period is a month and the shock hits at time $t = 0$. The shock is identical to that in Figure 12.

Figure 14: Fit and imputed shocks for the counterfactual exercise



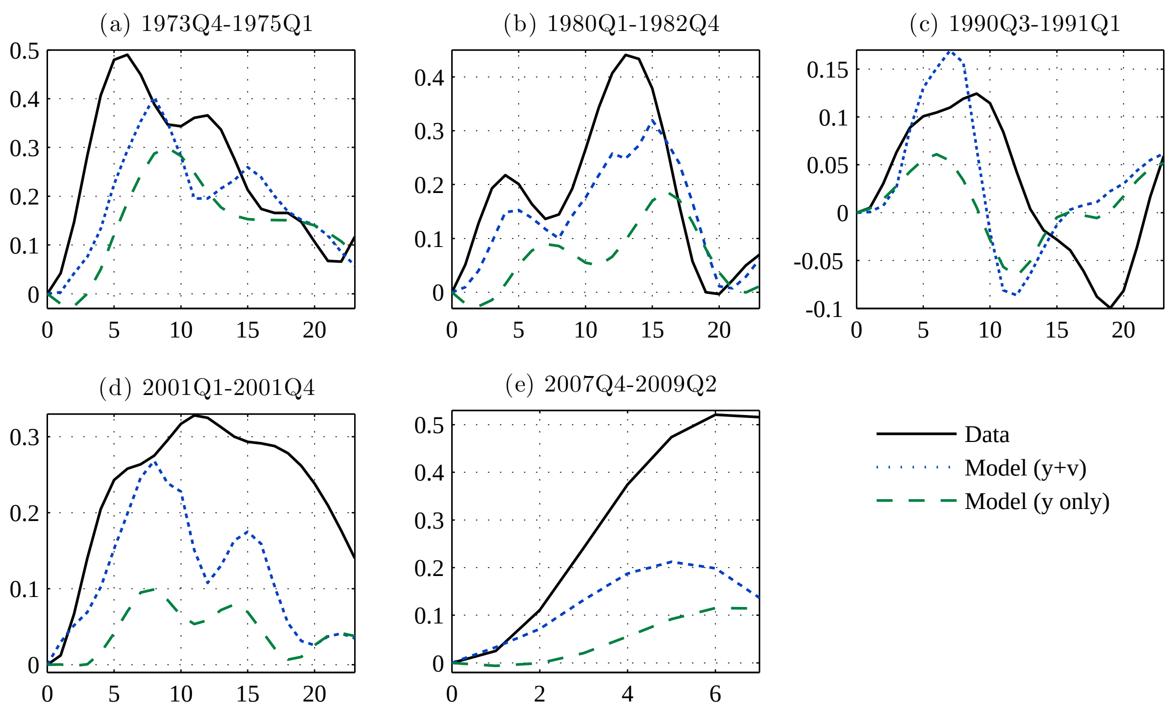
Notes: The black continuous line presents the data detrended using a band-pass filter (6 to 32 quarters for quarterly data, 2 to 8 years for annual data), the dotted black line is the data detrended using an HP filter (smoothing parameter 1600 quarterly, 100 annual), the blue continuous line corresponds to the model with productivity and volatility shocks, the green dashed line is the model with productivity shocks only. Note that the large peak in the IQR series in 1996 is an artefact of the Census data due to the change from SIC87 to NAICS classification in 1997, which biases the measure upward by more than 5% as reported in Bloom et al. (2014).

Figure 15: Counterfactual time series for output



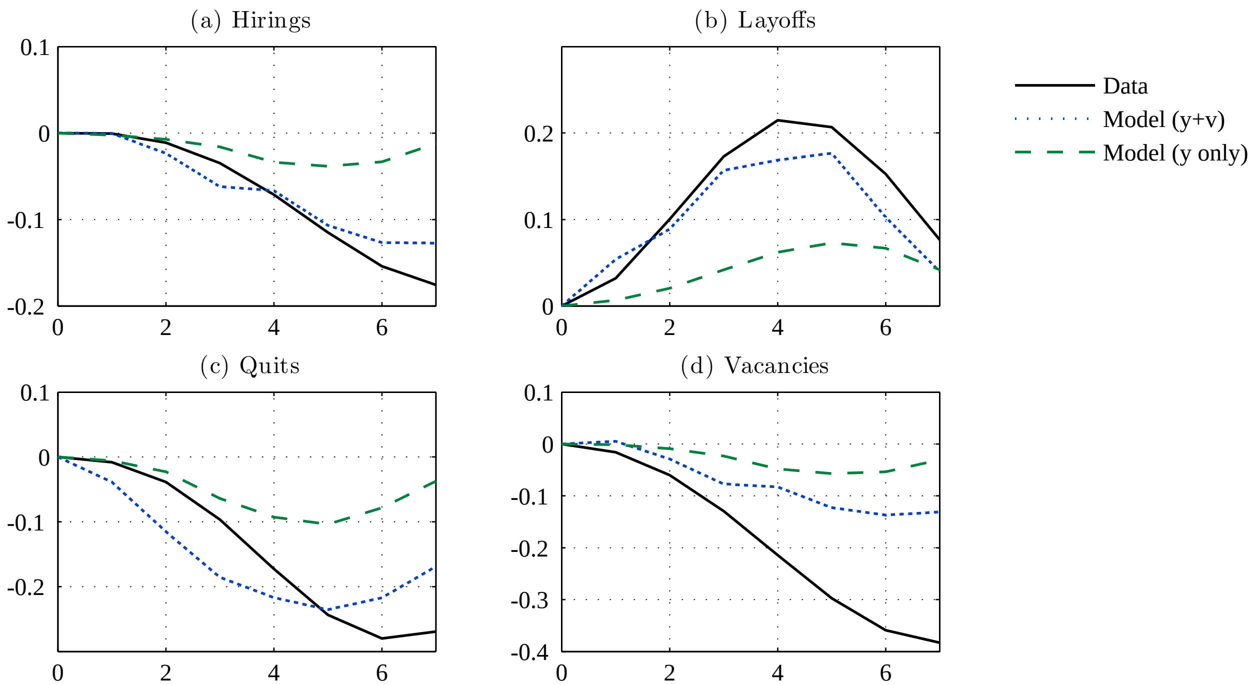
Notes: The black continuous line presents the data, the blue dotted line is the model with both aggregate productivity and volatility shocks, the green dashed line is the model with productivity shocks only and constant volatility. Responses shown in log deviation from the peak preceding recession. The aggregate productivity shock and volatility shock series are estimated to match the empirical output per person series and the IQR series from the Census.

Figure 16: Counterfactual time series for unemployment



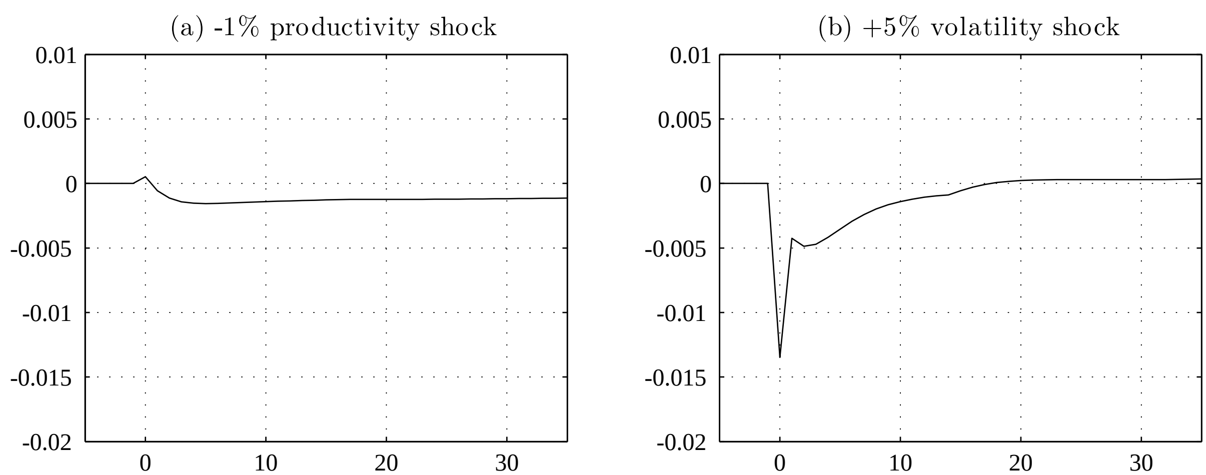
Notes: The black continuous line presents the data, the blue dotted line is the model with both aggregate productivity and volatility shocks, the green dashed line is the model with productivity shocks only and constant volatility. Responses shown in log deviation from the peak preceding recession. The aggregate productivity shock and volatility shock series are estimated to match the empirical output per person series and the IQR series from the Census.

Figure 17: Counterfactual time series for labor market flows in the 2007-2009 recession



Notes: The black continuous line presents the labor market flow data from the JOLTS dataset, the blue dotted line is the model with both aggregate productivity and volatility shocks, the green dashed line is the model with productivity shocks only and constant volatility. The aggregate productivity shock and volatility shock series are estimated to match the empirical output per person series and the IQR series from the Census. The procyclical variables (hirings, quits and vacancies) are shown in log deviation from the peak preceding the recession, the countercyclical variables (layoffs) from the preceding trough.

Figure 18: Response of the labor wedge to aggregate productivity and volatility shocks



Notes: The labor wedge is the ratio between the marginal rate of substitution between consumption and leisure to the marginal productivity of labor. Series presented in log deviation from their steady state values when aggregate productivity and volatility are set to their means. The time period is a month and the shock hits at time $t = 0$.

B Computing the Measure of Entrants

This section explains how to compute the measure of entering firms in every period. The number of entering firms is implicitly determined by the equilibrium conditions on each labor market segment. More specifically, recall that the equilibrium market tightness on a given submarket x is such that

$$\mu\left(s', g, x\right)\theta\left(s', x\right) = \nu\left(s', g, x\right), \quad \forall x,$$

where $\nu\left(s', g, x\right)$ is the measure of vacancies posted on that submarket and $\mu\left(s', g, x\right)$ the efficiency-weighted measure of searching workers. Multiplying both sides by $q\left(\theta\left(s', x\right)\right)$ and using the identity $p\left(\theta\right) = \theta q\left(\theta\right)$, this condition is equivalent to

$$JF\left(s', g, x\right) \equiv \mu\left(s', g, x\right) p\left(\theta\left(s', x\right)\right) = \nu\left(s', g, x\right) q\left(\theta\left(s', x\right)\right) \equiv JC\left(s', g, x\right), \quad \forall x, \tag{16}$$

where $JF\left(s', g, x\right)$ is the total number of jobs found by workers on submarket x and $JC\left(s', g, x\right)$ is the total number of jobs created by firms on the same submarket. Since firms are indifferent between the various submarkets, the continuum of equilibrium conditions (16) can be summarized by a unique aggregate conditions which guarantees that the total number of jobs found by workers across the various submarkets is equal to the total number of jobs created

$$JF_{\text{total workers}}\left(s', g\right) \equiv \int_{\underline{x}}^{\overline{x}} JF\left(s', g, x\right) dx = \int_{\underline{x}}^{\overline{x}} JC\left(s', g, x\right) dx \equiv JC_{\text{total firms}}\left(s', g\right).$$

To compute the number of entrants m_e', calculate the total number of jobs found by workers in the economy for a given period,

$$\begin{aligned} JF_{\text{total workers}}\left(s', g\right) = p\Big(\theta\left(s', x_u'\left(s'\right)\right)\Big) u \\ + \sum_{z, z', n} \pi_z\left(z' \mid z, s\right) g(z, n)\left(1 - d'\left(s', z'; n\right)\right) \int n\left(1 - \tau'\left(s', z'; j, n\right)\right) \lambda p\left(\theta\left(s', x'\left(s', z'; j, n\right)\right)\right) dj, \end{aligned}$$

which includes the number of successful hires from unemployment and the number of successful job-to-job transitions. Then, compute the total number of jobs created by incumbent firms,

$$JC_{\text{total incumbents}}\left(s', g\right) = \sum_{z, z', n} \pi_z\left(z' \mid z, s\right) g(z, n)\left(1 - d'\left(s', z'; n\right)\right) n_i\left(s', z'; n\right),$$

and the number of jobs created by a measure one of entrants,

$$JC_{\text{entrant}}\left(s'\right) = \sum_{z'} g_z(z')\left(1 - d_e\left(s', z'\right)\right) n_e\left(s', z'\right).$$

The measure of entrants may finally be computed using our aggregate condition:

$$JF_{\text{total workers}}(s',g) = JC_{\text{total firms}}(s',g) = JC_{\text{total incumbents}}(s',g) + m_e(s',g)\, JC_{\text{entrant}}(s').$$

C Data Description

This section details the construction and sources of the empirical time series used throughout the paper.

C.1 Measures of micro-level risk

Establishment-level volatility of TFP

The establishment-level volatility of TFP is taken from Bloom et al. (2014) constructed using data from the Census of Manufactures and the Annual Survey of Manufactures by the Census Bureau. This dataset contains output and inputs data for more than 50,000 establishments. Frequency is annual and the dataset covers the period 1972 to 2009. Establishments with less than 25 years of data are excluded. Establishment-level TFP $\hat{z}_{j,t}$ is calculated using a standard approach, controlling for demand side effects with 4-digit industry price deflators. TFP shocks are then estimated using:

$$\log(\hat{z}_{j,t}) = \rho \log(\hat{z}_{j,t-1}) + \mu_j + \lambda_t + e_{j,t},$$

where μ_j is an establishment fixed effect and λ_t a year fixed effect. The base measure for micro-level risk is then defined as the cross-sectional interquartile range of the residual $e_{j,t}$. See Bloom et al. (2014) for additional details on the construction of this measure.

A potential concern is whether the variation in the cross-sectional dispersion captured by this measure should be interpreted as time-varying volatility in TFP. This measure controls for i) demand side effects using price deflators, ii) unobservable heterogeneity using establishment-level fixed effects, and iii) selection by choosing only establishments with 25+ years of data. One remaining concern lies in the possibility that unobservable heterogeneity could lead to differences in cyclical sensitivity across firms. In that case, an increase in cross-sectional dispersion could simply reflect the heterogeneous response of firms to a first-moment shock. This effect is, however, difficult to control for and, despite this caveat, the proposed measure is arguably the best that can be constructed with available data and I therefore use it throughout the paper as my benchmark idiosyncratic volatility measure.

Alternative measures of micro-level risk

The Compustat sales growth dispersion measure is constructed using quarterly sales (SALEQ) in dollars for active US firms over the period 1972Q1-2009Q4. I keep firms that have 100+ observations. Annual sales growth is computed according to $g_{i,t} = \frac{s_{i,t}-s_{i,t-4}}{1/2(s_{i,t}+s_{i,t-4})}$. The growth measures are detrended with time-industry dummies (2-digit NAICS). The micro-level risk measure derived from this series is the cross-sectional interquartile range of detrended $\hat{g}_{i,t}$.

The VIX measure is the monthly average of the implied volatility (new method) of stock market returns constructed by the CBOE over 1990-2009.

C.2 Other series

- Output is taken from the NIPA tables constructed by the Bureau of Economic Analysis. I use quarterly GDP in 2005 dollars from 1972Q1 to 2009Q4.
- Productivity Y/L is seasonally adjusted real average output per person in the non-farm sector over the period 1972Q1-2009Q4 from the Bureau of Labor Statistics.
- Unemployment is the seasonally adjusted monthly unemployment rate constructed by the BLS from the Current Population Survey over the period January 1972-December 2009 (for people aged 16 and over). Similarly, I use the total civilian labor force for people aged at least 16 from the BLS over the same period. The series are averaged over quarters.
- Vacancy is the quarterly average of the monthly vacancy measure from the Job Openings and Labor Turnover Survey. Since the measure is available only since 2001, I use the Conference Board's Help Wanted Index to complete the measure from 1972Q1 to 2000Q4 and adjust its level to match the JOLTS stock of vacancies in 2001Q1.
- Historical UE and EU monthly transition rates are taken from Shimer (2007) over the period 1972Q1-2007Q1. For later periods, I use the monthly series on labor force status flows from the Current Population Survey constructed by the BLS over February 1990 to March 2010.
- EE is constructed by taking the ratio of quits from JOLTS over employment $(1-U)$ from January 2001 to December 2009.
- Labor market flows for hiring, quits and layoffs are quarterly sums of the JOLTS measures from January 2001 to December 2009. The series are normalized by total labor force.
- The empirical labor wedge was constructed using quarterly, seasonally adjusted, chained 2009 dollars "Real Personal Consumption Expenditure" from Fred (PCECC96), total hours worked tabulated from the CPS by Prescott et al. (2011) normalized by total population aged 16-64 from the BLS, and the output measure from the NIPA described above. The wedge was computed following Chari et al. (2007) with the expression

$$1-\tau_{l,t} = -\frac{u_H/u_C}{F_H\left(K_t, H_t\right)} = \frac{C_t H_t^{1+\nu}}{Y_t}$$

  derived under the assumptions of $u(C,H) = \log C - \xi \frac{H^{1+\nu}}{1+\nu}$ and $F\left(K,H\right) = K^\alpha H^{1-\alpha}$ with the normalization $\xi = 1$ and assuming $\nu = 0.25$, which implies a Frisch elasticity of 4, a value within the range of standard macro estimates.

D Numerical Implementation - ONLINE APPENDIX

This section describes the implementation of the model that I use for the quantitative exercises.

D.1 Description of the problem

Under the stochastic processes chosen in section 3, the aggregate state of nature is $s = (y, v)$. Since all the contracting aspects are absent from the joint surplus maximization problem, it is more convenient to solve for the surplus (7) at the beginning of a period in stage A instead of stage B. Define the surplus $\mathbf{V}^A$ in stage A as follows:

$$\mathbf{V}^A(y, v, z, n) = \max_{\substack{n_i, x_i, \tau \\ x, d \in \{0,1\}}} n\mathbf{U}(y, v)d + (1-d)\Big\{n\tau\mathbf{U}(y, v) + n(1-\tau)\lambda p(\theta(y, v, x))x$$
$$-\kappa(y, v)n_i + \mathrm{e}^{y+z}F(n') - k_f + \beta\mathbb{E}\mathbf{V}^A\left(y', v', z', n'\right)\Big\} \tag{17}$$

subject to

$$n' = n(1-\tau)\Big(1 - \lambda p\big(\theta(y, v, x)\big)\Big) + n_i,$$

where n denotes the employment level reached at the end of the previous period.[29] Note that I have used the properties from proposition (3) that $x(j)$ is uniform across workers, $x(j) = x, \forall j$, and that the distribution of layoffs across workers is undetermined to impose symmetry in the layoffs rates, $\tau(j) = \tau, \forall j$. Notice also that I have used the definition of (9) to substitute for the hiring cost $\kappa(y, v)$. The hiring costs is implicitly defined by the free-entry problem of (12),

$$k_e = \sum_{z\in\mathcal{Z}} g_z(z)\Bigg\{\max_{\substack{n_e(y,v,z),\, x_e(y,v,z), \\ d_e(y,v,z) \in \{0,1\}}} (1 - d_e(y, v, z))\left[e^{y+z}F(n_e(y, v, z)) - k_f\right.$$
$$\left.-\kappa(y, v)n_e(y, v, z) + \beta\mathbb{E}\mathbf{V}^A\left(y', v', z', n_e(y, v, z)\right)\right]\Bigg\}, \quad \forall (y, v). \tag{18}$$

The equilibrium market tightness implied by the free-entry condition is defined by combining equations

[29] Under this notation, surplus at stage A of a period ($\mathbf{V}^A$) and surplus at stage B ($\mathbf{V}$) are related in the following way:

$$\mathbf{V}(y, v, z, n) = e^{y(s)+z}F(n) - k_f + \beta\mathbb{E}\mathbf{V}^A(y', v', z', n),$$

and

$$\mathbf{V}^A(y, v, z, n) = \max\, n\mathbf{U}(y, v)d + (1-d)\{n\tau\mathbf{U}(y, v) + n(1-\tau)\lambda p(\theta(y, v, x))x - \kappa(y, v)n_i + \mathbf{V}(y, v, z, n')\}.$$

(10) and (11), so that

$$\theta(y,v,x) = \begin{cases} q^{-1}\left(\frac{c}{\kappa(y,v)-x}\right) & \text{for active markets } (x \le \kappa(y,v) - c) \\ 0 & \text{for inactive markets } (x > \kappa(y,v) - c) \end{cases}$$

Finally, the value of unemployment is defined by (1),

$$\mathbf{U}(y,v) = \max_{x_u(s')} b + \beta\mathbb{E}\left[p\left(\theta(s', x_u(s'))\right) x_u\left(s'\right) + \Big(1 - p\left(\theta\left(s', x_u\left(s'\right)\right)\right)\Big)\mathbf{U}(y',v')\right]. \tag{19}$$

D.2 Algorithm

The problems of (17), (18) and (19) define three nested fixed point problems that we must solve to find a quasi-equilibrium. I describe below the algorithm that I use to solve for them. The value functions are computed on a $ny \times nv \times nz \times nn$ grid (`ny=21; nv=15; nz=15; nn=30` in my baseline calibration).

1. Set $k = 0$. Guess a value function $V^{(0)}(y,v,z,n)$;
2. Using the free-entry condition, solve numerically for $\kappa^{(k)}(y,v)$ such that

$$\begin{aligned} k_e &= \sum_z g_z(z)\Bigg[\max_{n_e(y,v,z)} e^{y+z}F\left(n_e(y,v,z)\right) - k_f - \kappa(y,v)\,n_e(z) \\ &\quad + \beta\mathbb{E}V^{(k)}\left(y',v',z',n_e(y,v,z)\right)\Bigg]^+,\ \forall (y,v); \end{aligned}$$

   The RHS of this equation being monotonic in κ, I use a quick bisection method for that step. Save the decision rules $n_e(y,v,z)$ and $d_e(y,v,z)$. Using this new value of $\kappa^{(k)}(y,v)$, compute the equilibrium market tightness from (10) and (11):

$$\theta^{(k)}(y,v,x) = \begin{cases} q^{-1}\left(\frac{c}{\kappa^{(k)}(y,v)-x}\right) & \text{for } x \le \kappa^{(k)}(y,v) - c \\ 0 & \text{for } x > \kappa^{(k)}(y,v) - c \end{cases}$$

3. By value function iteration, find the fixed point of the mapping,

$$U^{(k)}(y,v) = \max_{x'_u(y',v')} b + \beta\mathbb{E}\left[p(\theta(y',v',x_u))x_u + \Big(1 - p\left(\theta\left(y',v',x_u\right)\right)\Big)U^{(k)}(y',v')\right],$$

   and save the corresponding decision rule $x_u(y',v')$.

4. Compute one iteration of the mapping:

$$
\begin{aligned}
V^{(k+1)}(y,v,z,n) &= \max_{\tau,x,n_i} \Big\{ n\tau U^{(k)}(y,v) + n\left(1-\tau\right)\lambda p\left(\theta^{(k)}\left(y,v,x\right)\right)x \\
&\quad -\kappa^{(k)} n_i + \mathrm{e}^{y+z}F(n') - k_f + \beta\mathbb{E}V^{(k)}\left(y',v',z',n'\right)\Big\}^+ \\
\text{s.t.} &\quad n' = n(1-\tau)\Big(1-\lambda p\big(\theta^{(k)}\left(y,v,x\right)\big)\Big) + n_i
\end{aligned}
$$

and save the corresponding decision rules $n'\left(y,v,z,n\right)$, $n_i\left(y,v,z,n\right)$, $x\left(y,v,z,n\right)$, $\tau\left(y,v,z,n\right)$ and $d\left(y,v,z,n\right)$.

5. Stop if $||V^{(k+1)} - V^{(k)}|| \leq \varepsilon$. Otherwise, go back to step 2 with $k \leftarrow k+1$.

D.3 Additional remarks

A number of remarks are in order:

- For the distribution of entrants g_z, I pick the stationary distribution of z when volatility v is held constant, equal to its mean $\bar{v}$;
- The choice over x_u and x has to be computed very precisely:
  - in step 3, I use the first order condition of the maximization problem and solve for the value of $x_u\left(y',v'\right)$ using a bisection algorithm;
  - in step 4, to simplify the maximization over (x,τ,n_i), I proceed in two steps:
    * for all pairs (n,n') on the $nn \times nn$ grid, compute $r = \frac{n'}{n}$. If $r<1$, solve the subproblem

$$
\begin{aligned}
\omega\left(y,v,r\right) &= \max_{x,\tau} \tau U^{(k)}\left(y,v\right) + \left(1-\tau\right)\lambda p\left(\theta^{(k)}\left(y,v,x\right)\right)x \\
\text{s.t.} &\quad \left(1-\tau\right)\left(1-\lambda p\left(\theta^{(k)}\left(y,v,x\right)\right)\right) = r,
\end{aligned}
$$

      which yields the optimal mix of layoffs/quits for a given (n,n'). Save the decision rules $x\left(y,v,r\right)$, $\tau\left(y,v,r\right)$ and the value $\omega\left(y,v,r\right)$. If $r \geq 1$, set $\tau\left(y,v,r\right) = 0$, $x\left(y,v,r\right) = \kappa^{(k)} - c$ and $\omega\left(y,v,r\right) = 0$. This problem can be solved quite accurately using its first order conditions;
    * using this optimal mix, the maximization of step 4 can be turned into the simple one-dimensional maximization problem:

$$
\begin{aligned}
V^{(k+1)}(y,v,z,n) &= \max_{n'} \Big\{ \mathrm{e}^{y+z}F(n') - k_f - \kappa^{(k)}\left(n'-n\right)^+ \\
&\quad + n\omega\left(y,v,\frac{n'}{n}\right) + \beta\mathbb{E}V^{(k)}\left(y',v',z',n'\right)\Big\}^+ .
\end{aligned}
$$

      This procedure provides a very accurate and smooth solution. Because of the reduction of the state-space, it also runs very quickly.

- I use two cubic splines in step 4 to smooth the choice of $n'(y,v,z,n)$ over $[0,n]$ and $[n,\bar{n}]$;
- The whole algorithm takes about 5 minutes to converge for the baseline calibration on my Dell Precision T7600 equipped with a NVidia Tesla C2075 GPU.

D.4 Computing wages

Section E.2 proposes a version of the model without commitment on the worker side in which wages are uniquely determined. This subsection describes how one can easily compute wages from the surplus maximizing allocation. In what follows, it is convenient to use the timing introduced in subsection D.1, expressing value functions and policies at the beginning of a period (stage A).

We start by solving for the incentive constraint (21) described in E.2. For every state (y,v,z,n), compute the promised utility $W'(y,v,z,n)$ such that

$$x(y,v,z,n) = \underset{x}{\operatorname{argmax}}\ p(\theta(y,v,x))\left(x - W'(y,v,z,n)\right).$$

Because of the monotonicity of the problem, this can be done efficiently using a bisection method.

It is then useful to write the utility of a worker employed by a firm (z,n) at the beginning of a period (stage A). Define

$$\begin{aligned}\mathbf{W}^A(y,v,z,n) \;=\; & d(y,v,z,n)\,\mathbf{U}(y,v) + (1-d(y,v,z,n))\Big[\tau(y,v,z,n)\,\mathbf{U}(y,v) \\ & + (1-\tau(y,v,z,n))\,\lambda p(\theta(y,v,x(y,v,z,n)))\,x(y,v,z,n) \\ & + (1-\tau(y,v,z,n))\left[1-\lambda p(\theta(y,v,x(y,v,z,n)))\right]W'(y,v,z,n),\end{aligned}$$

where $W'(y,v,z,n)$ is the promised utility at the end of the period. It is now easy to solve for wages. We can use the promise-keeping constraint (6) to derive their wages:

$$w^{incumbent}(y,v,z,n) = W'(y,v,z,n) - \beta\mathbb{E}\left[\mathbf{W}^A\left(y',v',z',n'(y,v,z,n)\right)\right].$$

Similarly, one can derive the wage of workers hired from unemployment with promised utility $x_u(y,v)$:

$$w^{unemp}(y,v,z,n) = x_u(y,v) - \beta\mathbb{E}\left[\mathbf{W}^A\left(y',v',z',n'(y,v,z,n)\right)\right].$$

Finally, a worker successfully moving from a firm with state $(\tilde{z},\tilde{n})$ to a firm with state (z,n), hired with promised utility $x(y,v,\tilde{z},\tilde{n})$ receives the wage

$$w^{j2j}(y,v,z,n;\tilde{z},\tilde{n}) = x(y,v,\tilde{z},\tilde{n}) - \beta\mathbb{E}\left[\mathbf{W}^A\left(y',v',z',n'(y,v,z,n)\right)\right].$$

E Additional Theoretical Results - ONLINE APPENDIX

E.1 Properties of the optimal contracts

This section characterizes various properties of the equilibrium contracts and, in particular, how different elements of the contracts (layoff probability τ, market for on-the-job search x, etc.) vary across workers within a single firm.

Proposition 3. *Under the conditions of proposition 2, in a quasi-equilibrium with surplus maximizing policy* $\{\{\tau_j, x_j\}_{j\in[0,n]}, d, n_i, x_i\}$ *the following is true:*

(i) If workers can commit, wages are not uniquely determined. In particular, the transformation $\left\{w_j + a\Delta, \tau_j, x_j, W'_j - \Delta, d\right\}$ *leaves worker* j *and the firm indifferent, with* $a = \beta\mathbb{E}(1-d)(1-\tau_j)\left(1-\lambda p\left(\theta\left(s', x_j\right)\right)\right)]$ *and* $\Delta \in \mathbb{R}$*;*

(ii) The market for on-the-job search x *is identical for all workers in the same firm;*

(iii) Only the total number of layoffs $\int \tau_j \mathrm{d}j$ *is uniquely determined; the distribution of layoffs* $\{\tau_j\}_{j\in[0,n]}$ *over workers is not.*

Proposition 3 first establishes that wages, w, and continuation values, W', are not unique. There are two reasons behind this result: i) workers and firms are risk neutral and ii) there is commitment from both workers and firms. Under these two conditions, the timing of wages is irrelevant. Only the total discounted value of future wages upon hiring is determined in equilibrium. This result shows the flexibility of the setup proposed in this paper as it can accomodate various profiles of wages over the life-cycle. I propose one particular way to determine wages in section E.2 by relaxing the commitment assumption on the worker side. In that case, the incentive problem uniquely pins down wages and I explore the quantitative properties of that particular assumption in section F.

Second, this proposition shows that all workers within a firm search on the same labor market segment. This result is due to the strict concavity of the search problem. Finally, as was suggested in proposition 1, the distribution of layoff probabilities across workers of a given firm is not uniquely determined. As is evident from the definition of the joint surplus, any permutation or convex combination of these probabilities across workers leaves the surplus unchanged. However, the total number of layoffs at the firm level is uniquely determined.

E.2 Relaxing commitment and completeness

I present in this section an extension of the model in which I relax the assumption of commitment on the worker side and the completeness of contracts. These assumptions may seem, indeed, somewhat unrealistic. First, I show in this subsection that commitment on the worker side is not required because firms have enough instruments to write incentive-compatible contracts that implement the efficient allocation. Second, I prove that firms may write down contracts that only specify $\{w, \tau(s', z'), d(s', z'), W'(s', z')\}$. In particular, this means that firms do not have to specify the labor market segment $x(s', z')$ in which their workers should be searching on the job—arguably the most unrealistic feature of the form of contracts assumed so far. Under the incentive-compatible contracts, firms can balance the current wage vs. continuation utility in such a way that workers choose to search in the optimal submarket.

Notice, however, that commitment on the firm side cannot be relaxed without losing block recursivity. Indeed, as discussed in the main text, it is key for block recursivity to obtain that firms stick to the contracts they advertise. Without commitment, firms would pay wages to workers that make them indifferent between a new job and their current situation. In particular, a firm would need to know the distribution of workers across firms before making its hiring decision, thereby breaking our main tractability results.

If we relax the assumption of commitment on the worker side, two additional constraints arise in the design of the contract. When workers are employed, the firm is worried about two things: 1) either the worker does not want to stay in the firm and decides to return to unemployment at the time when separations take place, or 2) the worker would like to search on a different submarket than the one specified in the contract. When designing a contract $(w, \tau(s', z'), x(s', z'), W'(s', z'), d(s', z'))$, the firm must take into consideration a participation constraint,

$$\lambda p(\theta(s', x))x + (1 - \lambda p(\theta(s', x)))W'(s', z') \geq \mathbf{U}(s'), \forall s' \tag{20}$$

which makes sure that the worker does not prefer to return to unemployment, and we have the following incentive constraint,

$$\begin{aligned} x(s', z') &= \underset{\tilde{x}}{\operatorname{argmax}}\ \lambda p(\theta(s', \tilde{x}))\tilde{x} + \big(1 - \lambda p(\theta(s', \tilde{x}))\big)W'(s', z') \\ \Leftrightarrow x(s', z') &= \underset{\tilde{x}}{\operatorname{argmax}}\ p(\theta(s', \tilde{x}))(\tilde{x} - W'(s', z')), \end{aligned} \tag{21}$$

which verifies that the submarket x specified in the contract coincides with the one chosen by the worker. I now show that for any given contract (w, τ, x, W', d), there is a unique equivalent contract with wage w_{IC} and future utility W'_{IC} that satisfies the above incentive and participation constraints and delivers the same promised utility to the worker.

Proposition 4. *For any optimal contract $\omega = \{w, \tau, x, W', d\}$, there exists a unique equivalent incentive-compatible contract $\omega_{IC} = \{w_{IC}, \tau_{IC}, W'_{IC}, d_{IC}\}$ such that $\forall (s', z')$:*

1. $\tau_{IC}(s', z') = \tau(s', z')$ *and* $d_{IC}(s', z') = d(s', z')$,
2. $\lambda p(\theta(s', x(s', z')))x(s', z') + (1 - \lambda p(\theta(s', x'(s', z'))))W'_{IC}(s', z') \geq \mathbf{U}(s')$,
3. $x(s', z') = \underset{\tilde{x}}{argmax}\ p(\theta(s', \tilde{x}))(\tilde{x} - W'_{IC}(s', z'))$,
4. $\mathbf{W}(s, z, \omega) = \mathbf{W}(s, z, \omega_{IC})$.

Proposition 4 tells us that the allocation that maximizes the worker-firm joint surplus can be implemented by an incentive-compatible contract. In particular, the layoff and exit probabilities are the same: $\tau_{IC} = \tau$, $d_{IC} = d$, and the submarket x chosen by the worker coincides with the efficient one. The wage and future utility (w_{IC}, W'_{IC}) are the only elements that adjust to ensure that the two additional constraints (21) and (20) are satisfied. In addition to being more realistic than complete contracts with full commitment, these contracts offer the advantage of uniquely pinning down wages. They thus offer an alternative to other wage determination procedures. Appendix D.4 presents to numerically implement this procedure. Appendix F shows that the wages this procedure implies match a number of empirical facts, such as a realistic wage dispersion and size-wage differential.

F Wage Predictions - ONLINE APPENDIX

The use of optimal dynamic contracts in search models provides an alternative to the standard assumptions of Nash or Stole and Zwiebel bargaining. However, as shown in proposition 3, wages are not uniquely pinned down if workers can commit to stay in the firm and search on the optimal labor market while employed. In section E.2 of the Appendix, I show how relaxing this commitment assumption yields a unique characterization of wages and contracts, as employers have to design contracts that give the right incentives for workers to stay/leave the firm and apply to the right labor market. Under this specification, wages could in principle vary substantially across workers belonging to the same firm. I explore in this section the quantitative implications of this wage setting mechanism. Because of a rich incentive structure, the model is able to predict an important wage dispersion for observationally equivalent workers and accounts for larger fraction of the empirical variation than standard search model. It also predicts a quantitatively accurate size-wage differential.

F.1 Wage dispersion and elasticity

Hornstein et al. (2007) report that standard calibrations of search-and-matching models without on-the-job search cannot generate much dispersion in wages. In their basic calibration of a standard random search model, they obtain a mean-min ratio of 1.036 for wages, while their preferred empirical estimate is about 1.70 with a corresponding coefficient of variation of only 1/12th of the variation in the data. Using wage data from the 1990 Census with different sets of controls, they estimate an empirical coefficient of variation of residual wages ranging from 0.35 to 0.49. I estimate the same dispersion measure in my model by simulating over a large number of periods and obtain an average coefficient of variation of 0.22, which explains between 45% and 63% of the observed residual dispersion in wages, outperforming standard search-and-matching models.

Regarding the evolution of wages over the business cycle, the average wage appears highly procyclical. The elasticity of wages with respect to productivity (output per person) is close to 1 in my model, slightly higher than the elasticity of wages for new hires of 0.79 estimated in Haefke et al. (2013) using CPS data. However, without any explicit mechanism for wage stickiness, the model is unable to replicate the elasticity for all the workers in the CPS, estimated at 0.24 by the same authors. An interesting extension would be to introduce risk aversion for workers. Combined with the dynamic contracting framework of the model, this extension would connect search theory to the implicit contract literature and provide us with a theory of endogenous stickiness, in which case this dimension could be significantly improved.

Turning to earnings risk over the business cycle, Guvenen et al. (2014) report, using administrative data, that the distribution of transitory shocks to log earnings are negatively skewed with a skewness ranging between -0.08 and -0.23. Computing annual growth im log earnings in my model, I find an average skewness of -0.04 that can fall as low as -0.26 over long simulations. However, the model is unable to produce the same cyclicality of earnings risk described by the same authors. They find substantial evidence of countercyclical risk in the left-tail of earnings shocks. My model, however, predicts a very mild time-variation in earnings risk (2.8% standard deviation in the dispersion of transitory log earnings shocks). Consistent with their findings, the right-tail risk, measured by the difference between the 90th percentile (P90) and the 50th percentile (P50) in log earnings growth, is procyclical. However, the left-tail risk, measured by P50 - P10 (10th percentile), is not countercyclical, as the authors show, but procylical in my model. The reason behind this failure appears to stem from the feature of the model, shared by most search models, that the value of earnings by unemployed workers, b in my notation, is constant over the cycle. As a result, workers in the model face strong procyclical upside risk due to the many opportunities to climb the job ladder in good times, but face little downside risk in recessions as the value of unemployment bounds earnings losses from below.

F.2 Size-wage differential

A common finding in the literature is that firm size can explain part of the variation in wages. Brown and Medoff (1989) report that, in a variety of datasets, a substantial size-wage differential remains despite various controls for labor quality and institutions: employees working at large firms earn higher wages than employess at small firms. To investigate whether the model can reproduce this finding, I compute the wages in every establishment at the aggregate steady state. I then run the following regression,

$$\log(wage) = \alpha + \beta \log(employment) + \varepsilon,$$

and evaluate by how much the wage of a worker varies with establishment size. I obtain a coefficient $\beta = 0.008$, about half of the estimate of 0.014 reported in that paper. Interestingly, this size-wage differential can be explained by a mechanism due to search frictions quite different from standard explanations based on labor quality or institutions. The mechanism at work in the model is due to the way firms deal with worker incentives. In this economy, firms that want to expand prefer to retain their current workers in order to save on hiring costs. To do so, they must promise them higher continuation utility. Therefore, all other things being equal, firms that grow tend to offer higher wages on average than firms that shrink. Turning back to firm size, large firms are those that received high idiosyncratic shocks and have grown in the recent past. As a result, they inherit high-paying contracts from the previous periods and tend to pay high wages. This mechanism emphasizes establishment growth as a key determinant for wages. Schmieder (2009) finds supporting evidence in German matched employer-employee data that fast growing establishment offer higher wages.

F.3 Relationship to implicit contract literature

The contracting framework used in this paper is reminiscent of the implicit contract literature initiated by Baily (1974) and Azariadis (1975). These articles considered the optimal contractual arrangement between risk-neutral firms and risk-averse workers and determined conditions under which the optimal contract insulated workers from aggregate labor market conditions by offering rigid wages. The question whether wages are set by spot markets or implicit contracts inspired a large empirical literature led by Beaudry and DiNardo (1991) that derived simple testable implications of both theories and applied them on US panel data. In particular, Beaudry and DiNardo (1991) showed that wages determined on spot markets should solely adjust to current labor market conditions, while wages determined by implicit contracts should display history dependence. Using the aggregate unemployment rate as a proxy for labor market conditions, the authors designed a simple empirical test by running panel regressions of log wages on current unemployment (spot market model), unemployment at the start of the job (contract model with low mobility) and the minimum unemployment rate since the start of the job (contract model with high mobility) in addition to a vector of individual characteristics. Their results showed a greater dependence of wages on past rather than current unemployment rates, offering support to the contracting approach.

These results were later criticized by Hagedorn and Manovskii (2013) who argued that such dependence of wages on past unemployment rates could be driven by selection and was consistent with a search model where wages depended solely on current labor market conditions. They showed in particular that past unemployment rates were a proxies for match quality and that using better measures of match quality virtually eradicated the dependence on past unemployment rates.

In this paper, wages are determined through long-term contracts. However, several features distinguish this framework from the implicit contract literature. First, workers are risk neutral, so that the motive for firms to insure their workers against income risk is absent. Second, the frictions faced at the contracting stage are different: under lack of commitment from workers, as considered in sections E.2 and F, firms use wages to incentivize workers to stay or direct their search on the job to some specific market segments. In the resulting incentive-compatible contract, wages are uniquely determined and solely depend on a firm's state at the beginning of a period $(s', z'; n)$. Wages are, in particular, independent from past unemployment rates. In that sense, this paper is closer to the search model of Hagedorn and Manovskii (2013) in which wages only depend on current conditions, but in which the dynamic matching of workers with firms over the business cycle leads to a dynamic selection of jobs consistent with the above results.

Simulating a population of workers from my model for a large number of periods, I first replicate the results from Beaudry and DiNardo (1991) by running the same regressions on simulated wages in Table 6. Consistent with their findings, I find a large negative, significant impact of current and past

Table 6: Results from simulated wage regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Contemporaneous	-4.332***			-3.535***	0.010		
unemployment rate	(0.013)			(0.029)	(0.040)		
Unemployment		-3.917***		-0.063***		-0.020	
at start of job		(0.013)		(0.033)		(0.024)	
Minimum unemployment			-4.285***	-0.888***			0.013
since start of job			(0.014)	(0.042)			(0.031)
Aggregate productivity y_t					1.171***	0.957***	0.986***
					(0.008)	(0.026)	(0.026)
Volatility v_t					0.241***	0.201***	0.211***
					(0.003)	(0.005)	(0.006)
Output-per-person						0.078***	0.071***
						(0.009)	(0.009)
Job tenure							0.002***
							(2.9e-5)
Constant	0.769***	0.753***	0.755***	0.772***	0.445***	0.294***	0.285***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.018)	(0.018)

Notes: Standard errors are in parentheses. The dependent variable is the logarithm of monthly wages simulated from a population of 1000 workers for 1200 periods (100 years). Output-per-person is the aggregate output divided by employment in a given period. The job tenure variable is the number of months less than a year that a worker has spent in the same job.

unemployment rates on wages in columns 1 to 3. Testing the three specifications at the same time in column 4, current, initial and minimum unemployment rates all preserve their negative, highly significant impact. However, consistent with the findings of Hagedorn and Manovskii (2013), this dependence is largely driven by spurious correlations and selection, to the extent that past unemployment rates correlate with the distribution of existing jobs. Controlling for the aggregate state of the economy as captured by the two shocks y_t and v_t, column 5 shows that the dependence on the current unemployment rate vanishes. Next, controlling for a measure of productivity for current existing jobs in column 6, output-per-person in the present case, cancels out the dependence on the unemployment rate at the start of the job.[30] Similarly, my findings suggest that the minimum unemployment rate also proxies for match quality: a low minimum unemployment rate, distinct from the current rate, proxies for a long tenure in a given job. Long tenures indicate good matches and higher wages. Adding a control for job tenure in column 7 eradicates the dependence on the minimum unemployment rate.

As a conclusion, this model is able to replicate the observation of history dependence of wages from Beaudry and DiNardo (1991), but this dependence is driven by the dynamic selection of jobs, consistent with the recent findings of Hagedorn and Manovskii (2013).

[30] Aggregate conditions in the past, as measured by the unemployment rate at the start of the job, have an impact on the current distribution of jobs through the type and employment of firms that entered/exited in the past. My result suggests that the initial unemployment rate proxies for the general productivity of matches in the pool of of existing jobs.

G Proofs - ONLINE APPENDIX

G.1 Proofs of part 2.6

Proof of proposition 1. Let me first introduce some notation. For a generic firm policy $\gamma = \left\{ \{\omega(j)\}_{j\in[0,n]}, d(s',z'), n_i(s',z'), x_i(s',z') \right\}$, define $\tilde{\mathbf{J}}(s,z,n,\gamma)$ the value of a firm evaluated at that policy in the current period:

$$
\begin{aligned}
\tilde{\mathbf{J}}(s,z,n,\gamma) \quad &= \quad \mathrm{e}^{y(s)+z}F(n) - k_f - \int_0^n w(j)\,\mathrm{dj} \\
&\quad + \beta\mathbb{E}\left\{(1-d)\left(-n_i\frac{c}{q(\theta(s',x_i))} + \mathbf{J}\left(s',z',n',\left\{\hat{W}'(s',z';j')\right\}_{j'\in[0,n']}\right)\right)\right\},
\end{aligned}
$$

subject to (4) and (5). Define the corresponding surplus:

$$
\begin{aligned}
&\tilde{\mathbf{V}}(s,z,n,\gamma) \equiv \mathbf{J}(s,z,n,\gamma) + \int_0^n \mathbf{W}(s,z,\omega(j))\mathrm{d}j \qquad (22)\\
&= \mathrm{e}^{y(s)+z}F(n) - k_f + \beta\mathbb{E}\Bigg\{nd\mathbf{U}(s') + (1-d)\Bigg[\mathbf{U}(s')\int_0^n \tau\mathrm{d}j + \int_0^n (1-\tau)\lambda p(\theta(s',x))\,x\mathrm{d}j \\
&\quad - n_i\frac{c}{q(\theta(s',x_i))} + \mathbf{J}\left(s',z',n',\left\{\hat{W}'(s',z';j')\right\}_{j'\in[0,n']}\right) + \int (1-\tau)(1-\lambda p(\theta(s',x)))\,W\mathrm{d}j\Bigg]\Bigg\}.
\end{aligned}
$$

Under this notation, for any optimal policy γ^*, we have $\mathbf{J}(s,z,n,\{W(j)\}_{j\in[0,n]}) = \tilde{\mathbf{J}}(s,z,n,\gamma^*)$. The proof proceeds in the following steps: a) I show that the promise keeping constraint for incumbent workers (6) must bind for any optimal policy γ^*, b) I show the equivalence between the maximization of $\tilde{\mathbf{J}}$ and $\tilde{\mathbf{V}}$, c) I show how the maximization of $\tilde{\mathbf{V}}$ can be equivalently written under the form of equation (7).

a) We can write the firm's problem as

$$
\begin{aligned}
\mathbf{J}(s,z,n,\{W(j)\}_{j\in[0,n]}) \quad &= \quad \max_{\gamma} \tilde{\mathbf{J}}(s,z,n,\gamma) \\
\text{subject to} \qquad &\quad (4),\ (5) \text{ and} \\
&\quad W(j) \leq \mathbf{W}\Big(s,z;\omega(j)\Big),\ \forall j\in[0,n].
\end{aligned}
$$

The wage $w(j)$ only appears linearly in the term $\int_0^n w(j)\mathrm{d}j$ and in the promise keeping constraint. In particular, it does not affect the incentive structure of the problem. It is therefore optimal to offer the lowest possible wage, so that the promise keeping constraint binds with equality. For a given policy $\gamma = \left\{ \{w,\tau,x,W'\}_{j\in[0,n]}, d, n_i, x_i \right\}$, the optimal wage $w(j)$ is such that $W(j) = \mathbf{W}(s,z;\omega(j))$, i.e.,

$$
\begin{aligned}
w(j) \quad = \quad & W(j) - \beta\mathbb{E}\Big[\big(d(s',z') + \big(1-d(s',z')\big)\tau(s',z';j)\big)\mathbf{U}(s') \\
& + \big(1-d(s',z')\big)(1-\tau(s',z';j))\lambda p\big(\theta\big(s',x(s',z';j)\big)\big)\,x(s',z';j) \\
& + \big(1-d(s',z')\big)\big(1-\tau(s',z';j)\big)\big(1-\lambda p\big(\theta\big(s',x(s',z';j)\big)\big)\big)W'(s',z';j)\Big], \forall j. \qquad (23)
\end{aligned}
$$

The firm's problem is thus equivalent to

$$\begin{aligned}\mathbf{J}(s,z,n,\{W(j)\}_{j\in[0,n]}) &= \max_{\substack{\{\omega(j)\}_{j\in[0,n]},\, d(s',z'),\\ n_i(s',z'),\, x_i(s',z')}} \tilde{\mathbf{J}}\left(s,z,n,\left\{\{\omega(j)\}_{j\in[0,n]},d,x_i,n_i\right\}\right)\\ \text{subject to} &\quad (4),\ (5)\text{ and }(23).\end{aligned}$$

b) Let me now define the surplus maximization problem

$$\begin{aligned}\mathbf{V}(s,z,n) &= \max_{\substack{\gamma=\left\{\{\omega(j)\}_{j\in[0,n]},\right.\\ \left. d,n_i,x_i\right\}}} \tilde{\mathbf{V}}(s,z,n,\gamma)\\ &\quad \text{subject to (4) and (5).}\end{aligned}$$

The surplus is invariant with the wage, so for any decision rules $\{\{\tau,x,W'\}_{j\in[0,n]},d,\ n_i,\ x_i\}$, it is always possible to set the wage $w(j)$ according to (23). In that case, from the definition of the surplus (22), :

$$\tilde{\mathbf{V}}(s,z,n,\gamma)=\tilde{\mathbf{J}}(s,z,n,\gamma)+\int_0^n W(j)dj.$$

In this equation, $\int_0^n W(j)dj$ is a predetermined constant. Therefore, it is absolutely equivalent to maximize the left hand side under constraints (4) and (5), as to maximize the right hand side under the same constraints with the addition of (23), which corresponds to the firm's problem according to step (i). We therefore conclude that

$$\mathbf{V}(s,z,n)=\mathbf{J}\left(s,z,n,\{W(j)\}_{j\in[0,n]}\right)+\int_0^n W(j)\,dj.$$

Any policy that solves the firm's problem must maximize the joint surplus. On the other hand, for any policy $\gamma=\left\{\{\tau,x,W'\}_{j\in[0,n]},d,n_i,x_i\right\}$ that maximizes the joint surplus, there exists a wage (set according to (23)) that maximizes the firm's profits.

c) Because of the above equivalence, we may now write the surplus maximization problem as

$$\begin{aligned}
\mathbf{V}(s,z,n) = & \max_{\substack{\{\tau,x,W'\}_{j\in[0,n]}, \\ d, n_i, x_i}} \mathrm{e}^{y(s)+z}F(n) - k_f + \beta\mathbb{E}\Bigg\{ n\mathbf{U}(s')d + (1-d)\Bigg[\mathbf{U}(s')\int_0^n \tau \mathrm{d}j \\
& + \int (1-\tau)\lambda p\left(\theta\left(s',x\right)\right)x\mathrm{d}j - n_i\frac{c}{q\left(\theta\left(s',x_i\right)\right)} \\
& + \underbrace{\mathbf{J}(s',z',n',\left\{\hat{W}'\left(s',z';j'\right)\right\}_{j'\in[0,n']}) + \int (1-\tau)\left(1-\lambda p\left(\theta\left(s',x\right)\right)\right)W'\mathrm{d}j\Big)}_{=\mathbf{V}(s',z',n')-n_i x_i}\Bigg]\Bigg\}. \\
= & \max_{\substack{\{\tau,x,W'\}_{j\in[0,n]}, \\ d, n_i, x_i}} \mathrm{e}^{y(s)+z}F(n) - k_f + \beta\mathbb{E}\Bigg\{ n\mathbf{U}(s')d + (1-d)\Bigg[\mathbf{U}(s')\int_0^n \tau \mathrm{d}j \\
& + \int (1-\tau)\lambda p\left(\theta\left(s',x\right)\right)x\mathrm{d}j - n_i\left(\frac{c}{q\left(\theta\left(s',x_i\right)\right)} + x_i\right) + \mathbf{V}(s',z',n')\Bigg\},
\end{aligned} \tag{24}$$

subject to (4).

This expression shows that the distribution of continuation utilities, $\{W'\left(s',z';j\right)\}_{j\in[0,n]}$, is irrelevant for the joint surplus. The joint surplus maximization problem may be equivalently written as

$$\begin{aligned}
\mathbf{V}(s,z,n) = & \max_{\substack{d\left(s',z'\right), n_i\left(s',z'\right), x_i\left(s',z'\right), \\ \{\tau\left(s',z';j\right), x\left(s',z';j\right)\}_{j\in[0,n]}}} \mathrm{e}^{y(s)+z}F(n) - k_f + \beta\mathbb{E}\Bigg\{ nd\mathbf{U}(s') + (1-d)\Bigg[\mathbf{U}(s')\int_0^n \tau \mathrm{dj} \\
& + \int_0^n (1-\tau)\lambda p\left(\theta\left(s',x\right)\right)x\mathrm{dj} - \left(\frac{c}{q\left(\theta\left(s',x_i\right)\right)} + x_i\right)n_i + \mathbf{V}\left(s',z',n'\right)\Bigg]\Bigg\}
\end{aligned}$$

subject to (4) which is the definition of the surplus in equation (7). Because the joint surplus does not depend on the distribution of contracts, we can conclude in particular that any combination of wages and continuation utilities, $\{w(j), W'(s'z';j)\}$ that satisfy (6) with equality implement the allocation and maximize profits. In practice, any profile of future promised utilities $\{W'\left(s',z';j\right)\}_{j\in[0,n]}$ may be implemented as long as wages are set according to (23). There is thus a multiplicity of contracts that implement the allocation and these contracts can easily be solved. □

G.2 Proofs of part 2.9

This section demonstrates all the proofs of existence, efficiency and uniqueness.

Proposition 2.(i): Existence

Let me first introduce a number of assumptions and definitions required to show the existence of a solution to the free-entry condition and joint surplus maximization problem. Denote $\mathcal{Z} = \{\underline{z} < \ldots < \overline{z}\}$, $\overline{y} = \max_{s\in\mathcal{S}} y\left(s\right)$ and $\underline{y} = \min_{s\in\mathcal{S}} y\left(s\right)$.

Assumption 1. *F is bi-Lipschitz continuous, i.e., there exists $(\underline{\alpha}_F, \overline{\alpha}_F)$ such that*

$$\forall (n_1, n_2), \quad \underline{\alpha}_F |n_2 - n_1| \leq |F(n_2) - F(n_1)| \leq \overline{\alpha}_F |n_2 - n_1|.$$

Assumption 2. *(i) p, q are twice continuously differentiable; (ii) p is strictly increasing and strictly concave; q is strictly decreasing and strictly convex; (iii) $p(0) = 0$, $q(0) = 1$, (iv) $p \circ q^{-1}$ is strictly concave.*[31]

To prove the existence of a solution to the free-entry problem, I make one additional assumption about the distribution of idiosyncratic productivity of entrants. Denote $g'_z(s, z)$ the cross-sectional distribution of z' one period after entry in state s, i.e., $g_{z'}(s, z') = \sum_{z \in \mathcal{Z}} g_z(z) \pi_z(z'|s, z)$.

Assumption 3. *For all $s \in \mathcal{S}$, the distribution g_z first order stochastically dominates $g_{z'}(s, z')$.*

Assumption 3 is an assumption on the productivity process which guarantees that entrants are (weakly) more productive on average than incumbents. This is a key condition to ensure that a non-zero measure of entrants hire a strictly positive number of workers upon entry, so that the free-entry condition may effectively pin down the value of κ in equilibrium.

To proceed with the proof of proposition 2.(i), I show that there exists a common solution to the joint surplus maximization, free-entry condition and unemployed workers' problem. This establishes the behavior of variables $(\{\tau(s', z'; j), x(s', z'; j)\}_{j=0}^{n}, d(s', z'), n_i(s', z'), x_i(s', z'))$ without the need to describe the set of contracts that implement the efficient allocation. Contracts may then be solved following the proof of proposition 1 or using the refinement of subsection E.2 in this appendix when the assumption of commitment on the worker side is relaxed. Let us first define the set where our optimal surplus $\mathbf{V}$ lies and introduce our last assumption on parameters. Let $\overline{n}$ be an arbitrary upper bound on employment chosen sufficiently large so that it does not constrain the equilibrium.

Definition 2. Let $\mathcal{V}$ be the set of value functions $V : (s; z, n) \in \mathcal{S} \times \mathcal{Z} \times [0, \overline{n}] \longrightarrow \mathbb{R}$ (i) strictly increasing in n, (ii) satisfying $\forall s, \sum_z g_z(z) [V(s, z, 0)]^+ \leq \beta k_e$, (iii) bounded in $[\underline{V}, \overline{V}]$, (iv) bi-Lipschitz continuous in n such that

$$\forall V \in \mathcal{V}, \forall (s, z), \forall n^{(1)} \geq n^{(2)}, \quad \underline{\alpha}_V (n^{(2)} - n^{(1)}) \leq V(s, z, n^{(2)}) - V(s, z, n^{(1)}) \leq \overline{\alpha}_V (n^{(2)} - n^{(1)}),$$

with

$$\begin{aligned}
\underline{\alpha}_V &= e^{\underline{y}+\underline{z}} \underline{\alpha}_F + \beta(1-\beta)^{-1} b > 0, \\
\overline{\alpha}_V &= (1-\beta)^{-1} \left(\mathrm{e}^{\overline{y}+\overline{z}} \overline{\alpha}_F + \beta \left(\lambda \overline{x} + (1-\beta)^{-1} (b + \beta \overline{x}) \right) \right) \\
\underline{V} &= -k_f, \\
\overline{V} &= (1-\beta)^{-1} [\mathrm{e}^{\overline{y}+\overline{z}} F(\overline{n}) - k_f + \beta \overline{n} \left(\lambda \overline{x} + (1-\beta)^{-1} (b + \beta \overline{x}) \right)].
\end{aligned}$$

[31] (iv) is a regularity condition ensuring that workers' problem is well defined and concave.

Assumption 4. *Assume* $\overline{n} > \underline{\alpha}_V^{-1}(k_e + k_f)$.

Assumption 4 is a sufficient condition on parameters that guarantees that there is always a solution to the free-entry problem. We can now establish the existence of a solution to the free-entry problem.

Lemma 1. *Under Assumptions 1-4, for* $V \in \mathcal{V}$, $s \in \mathcal{S}$, *the free-entry problem* (9)-(12) *admits a solution. There exists a unique hiring cost per worker* $\kappa(s)$, *an optimal level of hiring for entering firms* $n_e^V(s,z)$ *and exit decision* $d_e^V(s,z)$ *such that*

1. *Submarket* x *is active* $\Rightarrow \theta^V(s,x) > 0 \Rightarrow c/q(\theta(s,x)) + x = \kappa^V(s)$,
2. *For all* $s \in \mathcal{S}$,

$$k_e \;=\; \max_{n_e(s,z)} \sum_z g_z(z) \left[V(s,z,n_e(s,z)) - \kappa^V(s)\, n_e(s,z)\right]^+ ,$$

3. $\theta^V(s,x) = \begin{cases} q^{-1}\left(\frac{c}{\kappa^V(s)-x}\right), & \textit{for } \underline{x} \le x \le \kappa^V(s) - c, \\ 0, & \textit{for } x \ge \kappa^V(s) - c. \end{cases}$

Proof. For $V \in \mathcal{V}$, $s \in \mathcal{S}$ and $\kappa \in \mathbb{R}$, let us define the following auxiliary function

$$\psi^{s,V}(\kappa) = \max_{0 \le n_e^{s,V}(z) \le \overline{n}} \sum_z g_z(z) \left[V\left(s,z,n_e^{s,V}(z)\right) - \kappa n_e^{s,V}(z)\right]^+ .$$

The objective of this proof is to show that, for all $s \in \mathcal{S}$, there exists a unique $\kappa^V(s)$ such that $k_e = \psi^{s,V}(\kappa(s))$. Because V is continuous in $n \in [0,\overline{n}]$ and z has a finite support, $\psi^{s,V}$ is a well-defined function for $\kappa \in \mathbb{R}$. The Theorem of the Maximum tells us that ψ^V is a continuous function of κ. Notice that V being increasing in n, $\psi^{s,V}(0) = \sum_z g_z(z)\left[V(s,z,\overline{n})\right]^+$. Also, since V is bi-Lipschitz continuous with parameters $(\underline{\alpha}_V, \overline{\alpha}_V)$, for $\kappa \ge \overline{\alpha}_V$, the maximum is reached at $n_e = 0$ and $\psi^{s,V}(\kappa) = \sum_z g_z(z)\left[V(s,z,0)\right]^+$. Let us show that $\psi^{s,V}$ is a decreasing function of κ. Take $\kappa_1 < \kappa_2$ and the corresponding $n_{e,i}^{s,V}(z)$, $i = 1,2$, that solve the maximization problem. Denote $\mathcal{Z}_i^{s,V} = \{z \in \mathcal{Z} | V^{s,V}(s,z,n_{e,i}^{s,V}(z)) - \kappa_i n_{e,i}^{s,V}(z) \ge 0\}$. Then we have

$$\begin{aligned}
\psi^{s,V}(\kappa_1) - \psi^{s,V}(\kappa_2) &= \sum_z g_z(z)\left[V(s,z,n_{e,1}^{s,V}(z)) - \kappa_1 n_{e,1}^{s,V}(z)\right]^+ \\
&\quad - \sum_z g_z(z)\left[V(s,z,n_{e,2}^{s,V}(z)) - \kappa_2 n_{e,2}^{s,V}(z)\right]^+ \\
&\ge \sum_{z \in \mathcal{Z}_2^{s,V}} g_z(z)\left[V(s,z,n_{e,2}^{s,V}(z)) - \kappa_1 n_{e,2}^{s,V}(z)\right] \\
&\quad - \sum_{z \in \mathcal{Z}_2^{s,V}} g_z(z)\left[V(s,z,n_{e,2}^{s,V}(z)) - \kappa_2 n_{e,2}^{s,V}(z)\right] \\
&\ge (\kappa_2 - \kappa_1) \sum_{z \in \mathcal{Z}_2^{s,V}} g_z(z)\, n_{e,2}^{s,V}(z) .
\end{aligned}$$

Symmetrically, we can establish that $\psi^{s,V}(\kappa_1) - \psi^{s,V}(\kappa_2) \leq (\kappa_2 - \kappa_1) \sum_{z \in \mathcal{Z}_1^{s,V}} g_z(z)\, n_{e,1}^{s,V}(z)$. Thus ψ^V is decreasing. But this also tells us that if we denote $\overline{\kappa}$ the smallest κ such that $\psi^{s,V}(\kappa) = \sum_z g_z(z) \left[\varphi^{s,V}(z,0)\right]^+$ (i.e., for which $n_e = 0$ is optimal for all z), then we have that ψ^V strictly decreases on $[0, \overline{\kappa}]$ from $\sum_z g_z(z) \left[V(s,z,\overline{n})\right]^+$ to $\sum_z g_z(z) \left[V(s,z,0)\right]^+$ and remains constant thereafter.

If $\sum_z g_z(z) \left[V(s,z,0)\right]^+ < k_e < \sum_z g_z(z) \left[V(s,z,\overline{n})\right]^+$, the Intermediate Value Theorem tells us that there exists a unique $\kappa^V(s)$ such that $\psi^{s,V}\left(\kappa^V(s)\right) = k_e$. This establishes the existence of a solution to the free-entry problem. Part (1) of the proposition ensues:

$$\theta^V(s,x) > 0 \Leftrightarrow c/q(\theta(s,x)) + x = \kappa^V(s).$$

Also, we have (2): there exists a $n_e^V(s,z) \geq 0$ chosen by entering firms so that

$$k_e = \sum_z g_z(z) \left[V\left(s,z,n_e^V(s,z)\right) - \kappa n_e^V(s,z)\right]^+$$

and a corresponding exit decision $d_e(s,z)$.

To conclude, we only need to check that

$$\sum_z g_z(z) \left[V(s,z,0)\right]^+ < k_e < \sum_z g_z(z) \left[V(s,z,\overline{n})\right]^+.$$

The left-hand side is guaranteed by the fact that $V \in \mathcal{V}$. The right-hand side is guaranteed by assumption 4, as we have $\sum_z g_z(z) \left[V(s,z,\overline{n})\right]^+ \geq \sum_z g_z(z) V(s,z,\overline{n}) \geq \sum_z g_z(z) \left(V(s,z,0) + \underline{\alpha}_\varphi \overline{n}\right) \geq -k_f + \underline{\alpha}_V \overline{n} > k_e$, because of assumption 4.

(3) The complementary slackness condition (10) implies that either

$$\theta(s,x) = 0 \quad \text{or} \quad c/q(\theta(s,x)) + x = \kappa^V(s).$$

For $x > \kappa^V(s) - c$, the second expression admits no solution, as the probability q must remain below 1. So θ must be 0 in this region. For $x \leq \kappa^V(s) - c$, it admits the unique solution $q^{-1}\left(\frac{c}{\kappa^V(s) - x}\right)$. In this region: $c/q(0) + x < \kappa^V$, so $\psi^{s,V}(c/q(0) + x) > k_e$. $\theta(s,x)$ cannot be 0 otherwise it would violate the free-entry condition (12). To summarize our results:

$$\theta^V(s,x) = \begin{cases} q^{-1}\left(\frac{c}{\kappa^V(s,x) - x}\right), & \text{for } \underline{x} \leq x \leq \kappa^V(s) - c, \\ 0, & \text{for } x \geq \kappa^V(s) - c. \end{cases}$$

□

We now prove the main proposition that establishes the existence of a quasi-equilibrium.

Proposition 5. *Under Assumptions 1-4, there exists a block-recursive solution to equations (1)-(12),* *i.e., the mapping $T : \mathcal{V} \longrightarrow \mathcal{V}$ such that*

$$TV(s,z,n) = \max_{\substack{d(s',z'),n_i(s',z'),x_i(s',z'),\\ \{\tau(s',z';j),x(s',z';j)\}_{j\in[0,n]}}} \mathrm{e}^{y(s)+z}F(n) - k_f + \beta\mathbb{E}\Bigg\{nd\mathbf{U}^V(s') + (1-d)\Bigg[\mathbf{U}^V(s')\int_0^n \tau \mathrm{dj}$$
$$+\int_0^n (1-\tau)\lambda p\left(\theta^V(s',x)\right)x\mathrm{dj} - \kappa^V(s')n_i + V(s',z',n')\Bigg]\Bigg\}$$

with $n' = \int(1-\tau(j))(1-\lambda p\left(\theta^V(s,x(j))\right)x(j)\mathrm{d}j + n_i$, (θ^V,κ^V) *solution to the free-entry problem* (9)-(12) *and* $\mathbf{U}^V$ *solution to* (1) *admits a fixed point.*

Proof of proposition 5. To prove the existence, I will proceed in four steps: (1) establish existence, uniqueness and boundedness of $U^V(s)$ given some $V \in \mathcal{V}$, (2) show that T is a well-defined mapping from $\mathcal{V}$ to $\mathcal{V}$, (3) T is a continuous mapping, (4) $T(\mathcal{V})$ is an equicontinuous family. Since $\mathcal{V}$ is closed, bounded and convex, using Schauder's Fixed Point Theorem as stated in Stokey and Lucas, Theorem 17.4 p.520, this will establish the existence of a solution $\mathbf{V}$ in $\mathcal{V}$ to Bellman equation (7).

Step 1. For $V \in \mathcal{V}$, lemma 1 gives the existence and uniqueness of functions κ^V, n_e^V, d_e^V and θ^V. We are going to show that the following mapping M_V that defines U^V is a contraction from the space of functions $U : \mathcal{S} \longrightarrow \mathbb{R}$, bounded between some $\underline{U}$ and $\overline{U}$, to be defined later:

$$M^V U(s) = \max_{x_u(s')} b + \beta\mathbb{E}\left\{p\left(\theta^V\left(s',x_u\right)\right)x_u + \left(1-p\left(\theta^V\left(s',x_u\right)\right)\right)U\left(s'\right)\right\}.$$

Applying Blackwell's sufficient conditions for a contraction mapping, check *discounting*: for $a \geq 0$,

$$\begin{aligned} M^V(U+a) &= \max_{x_u(s')} b + \beta\mathbb{E}\left\{p\left(\theta^V\left(s',x_u\right)\right)x_u + \left(1-p\left(\theta^V\left(s',x_u\right)\right)\right)\left(U\left(s'\right)+a\right)\right\} \\ &\leq M^V U + \beta a. \end{aligned}$$

Check *monotonicity*: for $U_1 \leq U_2$, and corresponding optimal choices $x_u^{(i)}$, for $i = 1,2$,

$$\begin{aligned} &M^V(U_2) - M^V(U_1) \\ &\geq \left(1-p\left(\theta^V\left(s,x_u^{(2)}\right)\right)\right)\beta\mathbb{E}\left(U_2\left(s'\right) - U_1\left(s'\right)\right) \geq 0. \end{aligned}$$

It is easy to show now that if $\underline{U} \leq U \leq \overline{U}$, then

$$b + \beta\underline{U} \leq M^V U \leq b + \beta\left(\overline{x} + \overline{U}\right).$$

The unique fixed point of M^V is therefore bounded between $\underline{U} = (1-\beta)^{-1}b$ and $\overline{U} = (1-\beta)^{-1}(b+\beta\overline{x})$.

Step 2. Let us now check that T is a well-defined mapping from $\mathcal{V}$ to $\mathcal{V}$. For what follows, it is useful to denote some policy $\gamma = \{\{\tau(s',z';j), x(s',z';j)\}_{j\in[0,n]}, d(s',z'), n_i(s',z'), x_i(s',z')\}$, and define

$$\Phi^V(s,z,n,\gamma) = \mathrm{e}^{y(s)+z}F(n) - k_f + \beta\mathbb{E}\Bigg\{nd\mathbf{U}^V(s') + (1-d)\Bigg[\mathbf{U}^V(s')\int_0^n \tau \mathrm{dj} \\ + \int_0^n (1-\tau)\lambda p\left(\theta^V\left(s',x\right)\right)x\mathrm{dj} - \kappa^V\left(s'\right)n_i + V\left(s',z',n'\right)\Bigg]\Bigg\}.$$

Φ^V denotes the current joint surplus evaluated at some arbitrary policy γ.

(i) If $V \in \mathcal{V}$, then TV is strictly increasing in n. Take $n^{(1)} < n^{(2)}$ and the corresponding optimal policies $\gamma^{(1)}$ and $\gamma^{(2)}$.

$$\begin{aligned} TV(s,z,n^{(2)}) - TV(s,z,n^{(1)}) &= \Phi(s,z,n^{(2)},\gamma^{(2)}) - \Phi(s,z,n^{(1)},\gamma^{(1)}) \\ &\geq \Phi(s,z,n^{(2)},\tilde{\gamma}) - \Phi(s,z,n^{(1)},\gamma^{(1)}) \end{aligned}$$

with a suboptimal policy $\tilde{\gamma} = \left\{\{\tilde{\tau}(s',z';j), \tilde{x}(s',z';j)\}_{j\in[0,n^{(2)}]}, \tilde{d}, \tilde{n}_i, \tilde{x}_i\right\}$ such that $\tilde{x}(j) = x(j)^{(1)}$, $\tilde{d} = d^{(1)}$, $\tilde{n}_i = n_i^{(1)}$, $\tilde{x}_i = x_i^{(1)}$, and $\tilde{\tau}(j) = \tau(j)^{(1)}$ for $j \in [0, n^{(1)}]$ and 1 for $j \in [n^{(1)}, n^{(2)}]$. In that case, we have $\tilde{n} = n^{(1)}$, and many terms cancel to yield the desired result that TV is strictly increasing in n.

$$\begin{aligned} &TV(s,z,n^{(2)}) - TV(s,z,n^{(1)}) \geq \Phi(s,z,n^{(2)},\tilde{\gamma}) - \Phi(s,z,n^{(1)},\gamma_1) \\ &\geq e^{y(s)+z}\left(F\left(n^{(2)}\right) - F\left(n^{(1)}\right)\right) + \beta\mathbb{E}\left[\left(n^{(2)} - n^{(1)}\right)\mathbf{U}^V(s')\right] > 0. \end{aligned}$$

(ii) If $V \in \mathcal{V}$, then $\forall s \in \mathcal{S}, \sum_{z\in\mathcal{Z}} g_z(z)\, TV(s,z,0)^+ \leq \beta k_e$. Recall that

$$TV(s,z,0) = \max_{d,n_i,x_i} -k_f + \beta\mathbb{E}\Bigg\{(1-d)\left[-\kappa^V(s)n_i + \beta\mathbb{E}V(s',z',n_i)\Big)\right]\Bigg\}.$$

Since κ^V is the solution to the free-entry condition and because of Assumption 3, we have

$$TV(s,z,0) \leq -k_f + \beta k_e.$$

Since $TV(s,z,0) \geq -k_f$, we have

$$\begin{aligned} TV(s,z,0)^+ &= \max\ \{TV(s,z,0),0\} \leq \max\ \{TV(s,z,0) + k_f, 0\} \\ &\leq TV(s,z,0) + k_f \leq \beta k_e, \end{aligned}$$

and therefore $\sum_{z\in\mathcal{Z}} g_z(z)\, TV(s,z,0)^+ \leq \beta k_e$.

(iii) If $V \in \mathcal{V}$, then TV is bounded in $[\underline{V}, \overline{V}]$ with $\underline{V} = 0$ and $\overline{V} = (1-\beta)^{-1}[\mathrm{e}^{\overline{y}+\overline{z}}F(\overline{n}) - k_f +$

$\beta\overline{n}\left(\lambda\overline{x}+(1-\beta)^{-1}\left(b+\beta\overline{x}\right)\right)]$:

$$TV(s,z,n)\leq \mathrm{e}^{\overline{y}+\overline{z}}F(\overline{n})-k_f+\beta\left(\overline{n}\overline{U}+\overline{n}\lambda\overline{x}+\overline{V}\right)\leq\overline{V}.$$

Now, for the lower bound:

$$\begin{aligned}TV(s,z,n)&\geq\Phi(s,z,n,\tilde{\gamma})\\&\geq e^{\underline{y}+\underline{z}}F\left(n\right)-k_f+\beta n\underline{U}\geq -k_f=\underline{V}\end{aligned}$$

with suboptimal policy $\tilde{\gamma}$ such that $\tilde{d}=1$.

(iv) If $V\in\mathcal{V}$, then

$$\forall(s,z),\forall n_2\geq n_1,\quad \underline{\alpha}_V(n_2-n_1)\leq TV(s,z,n_2)-TV(s,z,n_1)\leq\overline{\alpha}_V(n_2-n_1).$$

Take $n_2\geq n_1$ and corresponding optimal policies γ_i, $i=1,2$. Choose a suboptimal policy $\tilde{\gamma}$ such that $\tilde{d}=d_2$, $\tilde{x}(s',z';j)=x_2(s',z';j)$, $\tilde{n_i}=n_{i2}$, $\tilde{x}_i=x_{i2}$, $\tilde{\tau}(s',z';j)=\tau_2(s',z';j)$ for $j\in[0,n_1]$:

$$\begin{aligned}&TV(s,z,n_2)-TV(s,z,n_1)=\Phi(s,z,n_2,\gamma_2)-\Phi(s,z,n_1,\gamma_1)\\&\leq\Phi(s,z,n_2,\gamma_2)-\Phi(s,z,n_1,\tilde{\gamma})\\&\leq \mathrm{e}^{y+z}(F(n_2)-F(n_1))+\beta\mathbb{E}\Bigg\{(n_2-n_1)d_2\mathbf{U}^V(s')+(1-d_2)\Bigg(\mathbf{U}^V(s')\int_{n_1}^{n_2}\tau_2dj+\\&+\int_{n_1}^{n_2}(1-\tau_2)\lambda p^V(x_2)x_2\mathrm{d}j+V(s,z,n_2')-V(s,z,\tilde{n}_1')\Bigg)\Bigg\}\\&\leq\left[\mathrm{e}^{\overline{y}+\overline{z}}\overline{\alpha}_F+\beta\left(\overline{U}++\lambda\overline{x}+\overline{\alpha}_V\right)\right](n_2-n_1)=\overline{\alpha}_V(n_2-n_1).\end{aligned}$$

Proceed similarly for the other side and choose a policy $\tilde{\gamma}$ such that $\tilde{d}=d_1$, $\tilde{x}(s',z';j)=x_1(s',z';j)$, $\tilde{n_i}=n_{i1}$, $\tilde{x_i}=x_{i1}$, $\tilde{\tau}(s',z';j)=\tau_1(s',z';j)$ for $j\in[0,n_1]$ and 1 for $j\in[n_1,n_2]$:

$$\begin{aligned}&TV(s,z,n_2)-TV(s,z,n_1)=\Phi(s,z,n_2,\gamma_2)-\Phi(s,z,n_1,\gamma_1)\\&\geq\Phi(s,z,n_2,\tilde{\gamma})-\Phi(s,z,n_1,\gamma_1)\\&\geq \mathrm{e}^{y+z}(F(n_2)-F(n_1))+\beta\mathbb{E}\Bigg\{(n_2-n_1)d_1\mathbf{U}^V(s')+(1-d_1)\left(n_2-n_1\right)\mathbf{U}^V(s')\Bigg\}\\&\geq\left[\mathrm{e}^{\underline{y}+\underline{z}}\underline{\alpha}_F+\beta\underline{U}\right](n_2-n_1)=\underline{\alpha}_V(n_2-n_1).\end{aligned}$$

Therefore, TV is bi-Lipschitz continuous with the desired coefficients.

Step 3. We are now going to show that $T:\mathcal{V}\longrightarrow\mathcal{V}$ is a continuous mapping. Denote by $||\cdot||$ the infinite norm, i.e., $||V||=\sup_{(s,z,n)\in\mathcal{S}\times\mathcal{Z}\times[0,\overline{n}]}V(s,z,n)$. Take $V_1,V_2\in\mathcal{V}$. For (s,z,n) fixed, denote by γ_k, $k = 1, 2$, the corresponding optimal policies. Denote $\tilde{\gamma}$ the policy exactly equal to γ_1 except that $\tilde{x}(s', z'; j)$ is chosen such that $p\left(\theta^{V_1}(s', x_1')\right) = p\left(\theta^{V_2}(s', \tilde{x}')\right)$. This means in particular that $\tilde{x}(s', z'; j) = x_1(s', z'; j) + \kappa^{V_2}(s') - \kappa^{V_1}(s'), \forall (s', z')$.

$$\begin{aligned}
&TV_1(s,z,n) - TV_2(s,z,n) = \Phi^{V_1}(s,z,n,\gamma_1) - \Phi^{V_2}(s,z,n,\gamma_2) \\
&\le \Phi^{V_1}(s,z,n,\gamma_1) - \Phi^{V_2}(s,z,n,\tilde{\gamma}) \\
&\le \beta\mathbb{E}\Bigg\{ d_1 n \left(\mathbf{U}^{V_1}(s') - \mathbf{U}^{V_2}(s')\right)) + (1-d_1)\Bigg((\mathbf{U}^{V_1}(s') - \mathbf{U}^{V_2}(s')) \int \tau_1 \mathrm{d}j - (\kappa^{V_1}(s') - \kappa^{V_2}(s')) n_{i1} \\
&+ \int (1-\tau_1)\lambda p\left(\theta^{V_1}\left(s', x_1\right)\right)(x_1 - \tilde{x})\,\mathrm{d}j + V_1(s', z', n_1') - V_2(s', z', n_1')\Bigg)\Bigg\} \\
&\le \beta\Big[\overline{n}||\mathbf{U}^{V_1} - \mathbf{U}^{V_2}|| + \overline{n}||\kappa^{V_1} - \kappa^{V_2}|| + \overline{n}\lambda||\kappa^{V_1} - \kappa^{V_2}|| + ||V_1 - V_2||\Big].
\end{aligned}$$

According to lemma 2 below, we can control each term:

$$TV_1(y,s,z,n) - TV_2(y,s,z,n) \le \beta\left[\overline{n}\alpha_U + \overline{n}\,(1+\lambda)\,\alpha_\kappa + 1\right]||V_1 - V_2||,$$

which can be made arbitrarily small as $||V_1 - V_2||$ gets smaller. Therefore, T is a continuous mapping.

Lemma 2. *If $V_1, V_2 \in \mathcal{V}$, then*
[(i)]

1. $||\kappa^{V_1} - \kappa^{V_2}|| \le \alpha_\kappa ||V_1 - V_2||$, *with* $\alpha_\kappa = \frac{\beta}{n_{min}}$,
2. $||\theta^{V_1} - \theta^{V_2}|| \le \alpha_\theta ||V_1 - V_2||$, *with* $\alpha_\theta = \frac{\beta}{c|q'(\theta_{max})|n_{min}}$,
3. $||\mathbf{U}^{V_1} - \mathbf{U}^{V_2}|| \le \alpha_U ||V_1 - V_2||$, *with* $\alpha_U = (1-\beta)^{-1}\beta\alpha_\kappa$.

Proof. To prove the lemma, we first need to establish the following two results. Let us prove that there exists $\theta_{max} > 0$ such that

$$\forall V \in \mathcal{V}, \theta^V(\cdot) \le \theta_{max},$$

and there exists $n_{min} > 0$ such that

$$\forall V \in \mathcal{V}, \sum_z g_z(z)\, n_e^V(s,z) \ge n_{min}.$$

The first result can be established by the fact that $\kappa_V \le \overline{\alpha}_V$ as we showed in lemma 1. Then for some $x \in [\underline{x}, \overline{x}]$:

$$c/q(\theta^V(s,x)) + x \le \overline{\alpha}_V \Rightarrow q(\theta^V(s,x)) \ge c(\overline{\alpha}_V - x)^{-1} \Rightarrow \theta^V(s,x) \le q^{-1}[c(\overline{\alpha}_V - \underline{x})^{-1}].$$

Setting $\theta_{max} = q^{-1}[c(\overline{\alpha}_V - \underline{x})^{-1}]$ yields the desired result.

Now, for the second result, remember the free-entry condition:

$$k_e = \sum g_z(z)\,[V(s,z,n_e^{s,V}(z)) - \kappa^V(s)\,n_e^{s,V}(z)]^+.$$

Then, using the fact that V is bi-Lipschitz:

$$\begin{aligned} k_e &\leq \sum g_z(z)\,[V(s,z,n_e^{s,V}(z))]^+ \leq \sum g_z(z)\,[V(s,z,0) + \overline{\alpha}_V n_e^{s,V}(z)]^+ \\ &\leq \overline{\alpha}_V \sum g_z(z)\,n_e^{s,V}(z) + \sum g_z(z)\left[-k_f + \beta\mathbb{E}V(s',z',0)\right]^+. \end{aligned}$$

Since $\sum g_z(z)\,[-k_f + \beta\mathbb{E}V(s',z',0)]^+ \leq \beta k_e$ as we argued before, we have

$$\mathrm{E}_{g_z} n_e^V \geq \overline{\alpha}_V^{-1}(1-\beta)k_e \equiv n_{min}.$$

(i) The free-entry condition gives us for $i = 1,2$:

$$k_e = \sum_z g_z(z)\left[V\left(s,z,n_e^{s,V_i}(z)\right) - \kappa^{V_i}(s)\,n_e^{s,V_i}(z)\right]^+.$$

Denote $\mathcal{Z}_i = \{z \in \mathcal{Z} | V(s,z,n_e^{s,V_i}(z)) - \kappa^{V_i} n_e^{s,V_i}(z) \geq 0\}, i = 1,2$. Substracting both:

$$\begin{aligned} 0 &= \sum_z g_z(z)\left[V\left(s,z,n_e^{s,V_1}(z)\right) - \kappa^{V_1}(s)\,n_e^{s,V_1}(z)\right]^+ - \sum_z g_z(z)\left[V\left(s,z,n_e^{s,V_2}(z)\right) - \kappa^{V_2}(s)\,n_e^{s,V_2}(z)\right]^+ \\ &\geq \sum_{z\in\mathcal{Z}_2} g_z(z)\left[\left(\kappa^{V_2}(s) - \kappa^{V_1}(s)\right)n_e^{s,V_2}(z) + \beta\mathbb{E}\left[V_1\left(s,z,n_e^{s,V_2}(z)\right) - V_2\left(s,z,n_e^{s,V_2}(z)\right)\right]\right] \end{aligned}$$

which yields

$$\kappa^{V_2}(s) - \kappa^{V_1}(s) \leq \frac{\beta}{n_{min}}||V_1 - V_2||.$$

Symmetrically, establish that $\kappa^{V_1}(s) - \kappa^{V_2}(s) \leq \frac{\beta}{n_{min}}||V_1 - V_2||$, which establishes the desired result for $\alpha_\kappa = \beta/n_{min}$.

(ii) Pick an $s \in \mathcal{S}$ and $x \in [\underline{x}, \overline{x}]$, consider the case in which submarket x is open under value functions V_1 and V_2. In that case, we know that:

$$\kappa^{V_i}(s) = \frac{c}{q\left(\theta^{V_i}(s,x)\right)} + x,$$

therefore

$$\frac{c}{q\left(\theta^{V_1}(s,x)\right)} - \frac{c}{q\left(\theta^{V_2}(s,x)\right)} = \kappa^{V_1}(s) - \kappa^{V_2}(s),$$

so that

$$q\left(\theta^{V_2}(s,x)\right) - q\left(\theta^{V_1}(s,x)\right) = c^{-1} q\left(\theta^{V_1}(s,x)\right) q\left(\theta^{V_2}(s,x)\right)\left(\kappa^{V_1}(s) - \kappa^{V_2}(s)\right),$$

and we can easily conclude that

$$\left|\theta^{V_2}(s,x)-\theta^{V_1}(s,x)\right| \leq \frac{1}{c\left|q'(\theta_{max})\right|}\left(\kappa^{V_1}(s)-\kappa^{V_2}(s)\right) \leq \frac{\beta}{c\left|q'(\theta_{max})\right| n_{min}}||V_1-V_2||.$$

Now consider the case in which submarket x is active under value V_2, but not under value V_1. We have:

$$\begin{aligned}\kappa^{V_2}(s) &= \frac{c}{q\left(\theta^{V_2}(s,x)\right)}+x,\\ \kappa^{V_1}(s) &\leq \frac{c}{q\left(\theta^{V_1}(s,x)\right)}+x,\end{aligned}$$

but also $\theta^{V_1}(s,x)=0$ and $\theta^{V_2}(s,x)>0$ from the complementary slackness condition. We can still derive the inequality:

$$\frac{c}{q\left(\theta^{V_2}(s,x)\right)}-\frac{c}{q\left(\theta^{V_1}(s,x)\right)} \leq \kappa^{V_2}(s)-\kappa^{V_1}(s),$$

so that:

$$0 \leq \theta^{V_2}(s,x)-\theta^{V_1}(s,x) \leq \frac{\beta}{c\left|q'(\theta_{max})\right| n_{min}}||V_1-V_2||.$$

Finally, the case in which submarket x is closed for both value functions is trivial, $\theta^{V_1}(s,x)=\theta^{V_2}(s,x)=0$.

(iii) Fix s. Denote by $x_{uk}, k=1,2$ the corresponding optimal choices for unemployed workers. Pick the suboptimal policy $\tilde{x}_u(s')$ such that $p\left(\theta^{V_1}(s',x_{u1}(s'))\right)=p\left(\theta^{V_2}(s',\tilde{x}_u(s'))\right)$, i.e., $\tilde{x}_u(s')=x_{u1}(s')+\kappa^{V_2}(s')-\kappa^{V_1}(s'), \forall(s',z')$

$$\begin{aligned}&U^{V_1}(s)-U^{V_2}(s)=\beta\mathbb{E}\left[p\left(\theta^{V_1}\left(s',x_{u1}(s)\right)\right)x_{u1}\left(s'\right)+\left(1-p\left(\theta^{V_1}\left(s',x_{u1}(s)\right)\right)\right)U^{V_1}\left(s'\right)\right]\\ &\quad-\beta\mathbb{E}\left[p\left(\theta^{V_2}\left(s',x_{u2}(s)\right)\right)x_{u2}\left(s'\right)+\left(1-p\left(\theta^{V_1}\left(s',x_{u2}(s)\right)\right)\right)U^{V_2}\left(s'\right)\right]\\ &\leq\beta\mathbb{E}\left[p\left(\theta^{V_1}\left(s',x_{u1}\left(s'\right)\right)\right)\left(x_{u1}\left(s'\right)-\tilde{x}_u\left(s'\right)\right)+\left(1-p\left(\theta^{V_1}\left(s',x_{u1}\left(s'\right)\right)\right)\right)\left(U^{V_1}\left(s'\right)-U^{V_2}\left(s'\right)\right)\right]\\ &\leq\beta\mathbb{E}\left[p\left(\theta^{V_1}\left(s',x_{u1}\left(s'\right)\right)\right)\left(\kappa^{V_1}\left(s'\right)-\kappa^{V_2}\left(s'\right)\right)+\left(1-p\left(\theta^{V_1}\left(s',x_{u1}\left(s'\right)\right)\right)\right)\left(U^{V_1}\left(s'\right)-U^{V_2}\left(s'\right)\right)\right]\\ &\leq\beta\alpha_\kappa||V_1-V_2||+\beta||U^{V_1}-U^{V_2}||\end{aligned}$$

We can now conclude that

$$||U^{V_1}-U^{V_2}|| \leq (1-\beta)^{-1}\alpha_\kappa||V_1-V_2||. \quad \square$$

Step 4. We can now proceed to the last step of the proof of proposition 5. We must show that the family $T(\mathcal{V})$ is equicontinuous, i.e., $\forall\varepsilon>0$, there exists $\delta>0$ such that for $\xi_k=(s_k,z_k,n_k)$,$k=1,2$,

$$||\xi_1-\xi_2||<\delta \Rightarrow |TV(\xi_1)-TV(\xi_2)|<\varepsilon, \forall V\in\mathcal{V}.$$

Fix $\varepsilon > 0$ and denote

$$\begin{cases} \eta_s = \min\limits_{s_1 \neq s_2 \in \mathcal{S}} |s_1 - s_2| \\ \eta_z = \min\limits_{z_1 \neq z_2 \in \mathcal{Z}} |z_1 - z_2|. \end{cases}$$

Choose $\delta < \min(\eta_s, \eta_z, \varepsilon/\overline{\alpha}_V)$. Take (ξ_1, ξ_2) such that $||\xi_1 - \xi_2|| < \delta$. Therefore, $s_1 = s_2$ and $z_1 = z_2$. Take $V \in \mathcal{V}$. Using the fact that V is bi-Lipschitz:

$$|TV(\xi_1) - TV(\xi_2)| \leq \overline{\alpha}_V |n_1 - n_2| \leq \overline{\alpha}_V ||\xi_1 - \xi_2|| < \varepsilon.$$

Conclusion: $T(\mathcal{V})$ is equicontinuous. Schauder's Fixed Point Theorem applies and tells us that there exists a fixed point $\mathbf{V}$ to the mapping T. All other equilibrium objects $\mathbf{U}$, $\mathbf{W}$, $\mathbf{J}$, θ, κ and optimal policy functions are then well defined. This achieves the proof of proposition 5 which corresponds to proposition 2.(i) in the text. □

Proposition 2.(ii): Efficiency

Proof. To study efficiency, I now introduce the planning problem of this economy. I proceed in four steps. First, I define the planning problem. In step 2, I simplify one important constraint in the planner's problem and provide an equivalent formulation. In step 3, I show that the planner's problem is a well-defined pseudo-concave problem subject to quasiconcave constraints, so that the first order conditions of the Lagrangian problem are sufficient for optimality. Finally, I show in step 4 that a block-recursive allocation, when it exists, satisfies the first-order conditions of the planner's problem and is therefore efficient.

Step 1. Using the same convention as in part 2.8, I denote u_t and $g_t(z_t, n_t)$ the unemployment rate and distribution of firms at stage B of period t when production takes place. For notational simplicity, I also introduce distribution $g_t^A(z_t, n_{t-1})$ which is the distribution of firms at the beginning of the period in stage A. The two distributions are related in the following way:

$$\begin{aligned} g_t^A(z_t, n_{t-1}) &= \sum_{z_{t-1}} \pi_z(z_t \mid s_{t-1}, z_{t-1}) \, g_{t-1}(z_{t-1}, n_{t-1}) \\ g_t(z_t, n_t) &= \sum_{n_{t-1}} \mathbb{I}\left\{n'(s_t, z_t; n_{t-1}) = n_t\right\} g_t^A(z_t, n_{t-1}) \\ &\quad + m_{e,t} \mathbb{I}\left\{n_e(s_t, z_t) = n_t\right\} g_z(z_t). \end{aligned}$$

Since the planner can freely allocate workers between firms without respect to any promised utility, the only relevant information concerning each labor market segment is its tightness. Let us therefore label each submarket by its tightness, θ, instead of its corresponding contract, x. Denote by $(\theta_x, \theta_i, \theta_u)$ the markets chosen respectively by firms for on-the-job search, for hirings, and the one chosen by unemployed workers to search. Furthermore, all workers are identical in the eyes of the planner. Given the strict concavity of the problem in $\theta_{x,t}$, I focus directly on allocations in which $\theta_{x,t}$ is the same across workers within a same firm. Similarly, as proposition 3 will make it clear, only the total number of layoffs at the firm level is determined in equilibrium, whereas the exact distribution of layoffs across workers in the same firm is not. I thus focus directly on allocations in which τ_t is the same across workers, so that the total number of layoffs is $n_{t-1}\tau_t$. It should be understood that transformations of τ_t across workers that leave the total number of layoffs unchanged are also solutions of the planning problem. All decisions at time t depend implicitly on the entire history of past aggregate shocks $s^t = \{s_t, s_{t-1}, ...\}$. The planner's objective is to maximize the total welfare in the economy,

$$\max_{\substack{u_t, g^A_{t+1}, \theta_{u,t}, d_t, n_t, \tau_t, \\ n_{i,t}, \theta_{i,t}, d_{e,t}, n_{e,t}, \theta_{e,t}}}$$
$$\mathbb{E}\sum_t \beta^t \Bigg\{ u_t b + \sum_{z_t, n_{t-1}} g^A_t(z_t, n_{t-1})\left(1 - d_t\left(z_t, n_{t-1}\right)\right) \times ...$$
$$... \times \left(e^{y(s_t)+z_t} F\left(n_t\left(z_t, n_{t-1}\right)\right) - k_f - \frac{c}{q\left(\theta_{i,t}\left(z_t, n_{t-1}\right)\right)} n_{i,t}\left(z_t, n_{t-1}\right)\right)$$
$$+ m_{e,t}\left[-k_e + \sum_{z_t} g_z\left(z_t\right)\left(1 - d_{e,t}\left(z_t\right)\right) \times ... \right.$$
$$\left. ... \times \left(e^{y(s_t)+z_t} F\left(n_{e,t}\left(z_t\right)\right) - k_f - \frac{c}{q\left(\theta_{e,t}\left(z_t\right)\right)} n_{e,t}\left(z_t\right)\right)\right]\Bigg\}, \tag{25}$$

which is the discounted sum of production net of operating cost k_f and vacancy posting cost c over all existing firms, minus total entry costs for new firms $m_{e,t}$ every period, plus home production b of unemployed agents. The planner is subject to the laws of motion of the unemployment rate, u_t, the level employment for every firm, and the distribution of firms, g_t.

$$u_t = \Big(1 - p\left(\theta_{u,t}\right)\Big) u_{t-1} + ...$$
$$+ \sum_{z_t, n_{t-1}} n_{t-1}\left[d_t\left(z_t, n_{t-1}\right) + \left(1 - d_t\left(z_t, n_{t-1}\right)\right)\tau_t\left(z_t, n_{t-1}\right)\right] g^A_t(z_t, n_{t-1}), \tag{26}$$

$$n_t\left(z_t, n_{t-1}\right) = n_{t-1}\left(1 - \tau_t\left(z_t, n_{t-1}\right)\right)\left(1 - \lambda p(\theta_{x,t})\right) + n_{i,t}\left(z_t, n_{t-1}\right), \quad \forall (z_t, n_{t-1}) \tag{27}$$

$$g^A_{t+1}(z_{t+1}, n_t) = \sum_{(z_t, n_{t-1}) \mid n_t(z_t, n_{t-1}) = n_t} \left(1 - d_t\left(z_t, n_{t-1}\right)\right)\pi_z(z_{t+1} \mid s_t, z_t) g^A_t(z_t, n_{t-1})$$
$$... + m_{e,t} \sum_{z_t \mid n_{e,t}(z_t) = n_t} \left(1 - d_{e,t}\left(z_t\right)\right)\pi_z(z_{t+1} \mid s_t, z_t) g_z(z_t), \quad \forall (z_t, n_{t-1}) \tag{28}$$

In addition, the planner is subject to two additional types of constraints: a non-negativity constraint for entry, and a constraint verifying that each labor market segment is in equilibrium, i.e., that the number of workers finding a job is equal to the number of successful job openings on a given submarket.

More precisely, in every period, the planner is subject to:

$$m_{e,t} \geq 0, \tag{29}$$

$$JF_t^w(\theta) + JF_t^u(\theta) = JC_t^f(\theta) + JC_t^e(\theta), \quad \forall \theta, \tag{30}$$

where $JF_t^W(\theta)$ is the total number of jobs found by incumbent workers, equal to the number of successful job-to-job transitions,

$$JF_t^W(\theta) = \sum_{\substack{(z_t, n_{t-1})| \\ \theta_{x,t}(z_t, n_{t-1}) = \theta}} g_t^A(z_t, n_{t-1})(1 - d_t(z_t, n_{t-1})) n_{t-1}(1 - \tau_t(z_t, n_{t-1})) \lambda p(\theta_{x,t}(z_t, n_{t-1})),$$

$JF_t^U(\theta)$ is the number of jobs found for unemployed, equal to the number of successful unemployed candidates,

$$JF_t^u(\theta) = \mathbb{1}(\theta_{u,t} = \theta) p(\theta_{u,t}) u_{t-1},$$

$JC_t^f(\theta)$ is the number of jobs created by incumbent firms on market θ,

$$JC_t^f(\theta) = \sum_{(z_t, n_{t-1})|\theta_{i,t}(z_t, n_{t-1}) = \theta} g_t^A(z_t, n_{t-1})(1 - d_t(z_t, n_{t-1})) n_{i,t}(z_t, n_{t-1}),$$

and $JC_t^e(\theta)$ that of entering firms,

$$JC_t^e(\theta) = m_{e,t} \sum_{z|\theta_{e,t}(z) = \theta} g_z(z)(1 - d_{e,t}(z)) n_{e,t}(z).$$

As a summary, the planner's problem is to maximize (25) subject to constraints (26)-(30).

Step 2. The constraints defined in (30) are difficult to handle in practice. We now provide an easier equivalent formulation of the problem. Notice first that under constraint (30), we have the following equality:

$$\begin{aligned} & \sum_{z_t, n_{t-1}} g_t^A(z_t, n_{t-1})(1 - d_t) \frac{c}{q(\theta_{i,t})} n_{i,t} + m_{e,t} \sum_{z_t} g_z(z_t)(1 - d_{e,t}) \frac{c}{q(\theta_{e,t})} n_{e,t} \\ = \quad & c \sum_{z_t, n_{t-1}} g_t^A(z_t, n_{t-1})(1 - d_t)(1 - \tau_t) \lambda n_{t-1} \theta_{x,t} + \theta_{u,t} u_{t-1}, \end{aligned} \tag{31}$$

where I have used the identity $p(\theta) = \theta q(\theta)$. Substituting (31) into the objective function (25), we notice that the markets for hiring $\theta_{i,t}$ and $\theta_{e,t}$ do not affect the objective function:

$$\max_{\substack{u_t, g^A_{t+1}, \theta_{u,t}, d_t, n_t,\\ \tau_t, n_{i,t}, d_{e,t}, n_{e,t},}} \mathbb{E}\sum_t \beta^t \Bigg\{ u_t b + \sum_{z_t, n_{t-1}} g^A_t(z_t, n_{t-1})(1-d_t)\left(e^{y(s_t)+z_t} F(n_t) - k_f\right)$$
$$+ m_{e,t}\left[-k_e + \sum_{z_t} g_z(z_t)(1-d_{e,t})\left(e^{y(s_t)+z_t} F(n_{e,t}) - k_f\right)\right]$$
$$- c\left(\theta_{u,t} u_{t-1} + \sum_{z_t, n_{t-1}} g_t(z_t, n_{t-1}) n_{t-1}\lambda(1-d_t)(1-\tau_t)\theta_{x,t}\right)\Bigg\}. \tag{32}$$

This means that, as long as constraint (30) is satisfied, variables $(\theta_{i,t}, \theta_{e,t})$ leave aggregate welfare unchanged. This result echoes our finding in the competitive equilibrium that firms are indifferent between markets. What this means is that we can replace constraint (30) by an easier one. Summing over all the submarkets, constraint (30) gives an expression for the measure of entrants:

$$m_{e,t} = \left(\sum_{z_t} g_z(z_t)(1-d_{e,t}) n_{e,t}\right)^{-1}\left[\sum_{z_t, n_{t-1}} g_t(z_t, n_{t-1})(1-d_t)\left(\lambda n_{t-1} p(\theta_{x,t}) - n_{i,t}\right) + p(\theta_{u,t}) u_{t-1}\right] \tag{33}$$

Because $\theta_{i,t}$ and $\theta_{e,t}$ do not affect welfare under constraint (30), it is equivalent to maximize (25) under constraint (30) as to maximize (32) under constraint (33). Indeed, as long as (33) is satisfied, we can always arbitrarily distribute incumbent and entering firms across markets so that (30) is satisfied for all active submarket.

Step 3. We now show that the planner's problem is a well-behaved pseudo-concave problem. To show this, I rewrite the maximization of (32) under constraints (26), (27), (28) and (33) in such a way that the objective function is pseudoconcave and all constraints are quasiconcave. In that purpose, it is useful to write the summation over distribution $g^A_t(z_t, n_{t-1})$ as a summation over firms' indices, so that we can ignore the law of motion of g^A_t. Every firm is indexed by the period it was born, t_0, and an firm-specific index, j, among that cohort. Let me also introduce the variables $\xi_{u,t} = p(\theta_{u,t})$ and $\xi^{(t_0,j)}_{x,t} = p\left(\theta^{(t_0,j)}_{u,t}\right)$ which are useful to turn the problem concave along some dimensions. The planning problem may be equivalently written:

$$\max_{\substack{u_t, \theta_{u,t}, \xi_{u,t} h_t, v_t,\\ \left\{d^{(t_0,j)}_t, n^{(t_0,j)}_t, \tau^{(t_0,j)}_t, n^{(t_0,j)}_{i,t}, \theta^{(t_0,j)}_{x,t}, \xi^{(t_0,j)}_{x,t}\right\}_{(t,t_0,j)}}} \tag{34}$$
$$\mathbb{E}\sum_t \beta^t\left[\sum_{t_0=-\infty}^{t}\int\prod_{l=t_0}^{t}\left(1-d^{(t_0,j)}_l\right)\left(e^{y(s_t)+z^{(t_0,j)}_t} F\left(n^{(t_0,j)}_t\right) - k_f\right)dj + u_t b - c v_t - m_{e,t} k_e\right]$$

subject to

$$n_{t-1}^{(t_0,j)}\left(1-\tau_t^{(t_0,j)}\right)\left(1-\lambda\xi_{x,t}^{(t_0,j)}\right)+n_{i,t}^{(t_0,j)}-n_t^{(t_0,j)}=0, \tag{35}$$

$$\sum_{t_0=-\infty}^{t}\int\left[\prod_{l=t_0}^{t-1}\left(1-d_l^{(t_0,j)}\right)\right]\left(1-d_t^{(t_0,j)}\right)\left(1-\tau_t^{(t_0,j)}\right)\lambda n_{t-1}^{(t_0,j)}\theta_{x,t}^{(t_0,j)}dj$$
$$...+\theta_{u,t}u_{t-1}-v_t=0, \tag{36}$$

$$\left(1-\xi_{u,t}\right)u_{t-1}+\sum_{t_0=-\infty}^{t}\int\left[\prod_{l=t_0}^{t-1}\left(1-d_l^{(t_0,j)}\right)\right]n_{t-1}^{(t_0,j)}\left(d_t^{(t_0j)}+\left(1-d_t^{(t_0,j)}\right)\tau_t^{(t_0,j)}\right)-u_t=0, \tag{37}$$

$$\sum_{t_0=-\infty}^{t}\int\left[\prod_{l=t_0}^{t-1}\left(1-d_l^{(t_0,j)}\right)\right]\left\{\left(1-d_t^{(t_0,j)}\right)\left[n_{t-1}^{(j)}\left(1-\tau_t^{(t_0,j)}\right)\lambda\xi_{x,t}^{(t_0,j)}-n_{i,t}^{(t_0,j)}\right]dj\right.$$
$$...+\xi_{u,t}u_{t-1}=0, \tag{38}$$

$$\int\left(1-d_t^{(t,j)}\right)dj-m_{e,t}=0, \tag{39}$$

$$p\left(\theta_{u,t}\right)-\xi_{u,t}=0 \text{ and } p\left(\theta_{x,t}^{(t_0,j)}\right)-\xi_{x,t}^{(t_0,j)}=0. \tag{40}$$

The objective function is concave and non-stationary. It is therefore pseudoconcave. The constraints are all sums of linear and positive cross-product terms and are therefore quasiconcave. We may then conclude that the first-order conditions of the Lagrangian problem are sufficient to guarantee optimality.

Step 4. I will now show that a block-recursive equilibrium solves the planner's first order conditions. For that purpose, let us write the Lagrangian of version (25) of the planner's problem, summing over firms' indices. Write μ_t the Lagrange multiplier on constraint (26) and $\eta_t(\theta)$ the one for each submarket equilibrium (30).

$$\mathcal{L}=\mathbb{E}\sum_t\beta^t\left\{\sum_{t_0=-\infty}^{t}\int\left[\prod_{l=t_0}^{t-1}\left(1-d_l^{(t_0,j)}\right)\right]\left[\left(1-d_t^{(t_0,j)}\right)\left(e^{y(s_t)+z_t(t_0,j)}F\left(n_t^{(t_0,j)}\right)-k_f\right.\right.\right.$$
$$...-\frac{c}{q\left(\theta_{i,t}^{(t_0,j)}\right)}n_{i,t}^{(t_0,j)}-\eta_t\left(\theta_{i,t}^{(t_0,j)}\right)n_{i,t}^{(t_0,j)}+\eta_t\left(\theta_{x,t}^{(t_0,j)}\right)n_{t-1}^{(t_0,j)}\left(1-\tau_t^{(t_0,j)}\right)\lambda p\left(\theta_{x,t}^{(t_0,j)}\right)\Big)$$
$$...+\mu_t n_{t-1}^{(t_0,j)}\left(d_t^{(t_0,j)}+\left(1-d_t^{(t_0,j)}\right)\tau_t^{(t_0,j)}\right)\Big]$$
$$\left....-m_{e,t}k_e+u_tb-\mu_t\big(u_t-u_{t-1}\left(1-p(\theta_{u,t})\right)\big)+\eta_t\left(\theta_{u,t}\right)u_{t-1}p\left(\theta_{u,t}\right)\right\}, \tag{41}$$

where constraint (27) is implicitly substituted. To complete the proof, I am now going to show that a block-recursive competitive equilibrium (with non-negative entry) satisfies the first-order conditions of the planner. Pick a block-recursive equilibrium by $\{\mathbf{V},\ \mathbf{U},\ \kappa^*(s),\ \theta^*(s,x)\}$. Guess the following Lagrange multipliers:

$$\mu_t(s^t) = \mathbf{U}(s_t)$$
$$\eta_t(s^t,\theta) = x(s_t,\theta) \ \text{ s.t } x(s_t,\theta) = \theta^{*-1}(s_t,\theta).$$

In particular, notice that the Lagrange multipliers only depend on the current aggregate state of the economy, s_t, and not on its entire history. One may worry here about the invertibility of the equilibrium function θ^*, but we know thanks to lemma 1 that there always exists a corresponding promised utility x for all values of θ in $[0,\infty)$ given by $x = \kappa(s) - c/q(\theta)$.[32] Given this guess, we can now recognize that the planner's objective is to sum the joint-surplus $\mathbf{V}$ of incumbent and entering firms and the utility of unemployed workers $\mathbf{U}$. Each of these problems can be solved independently and we know that the policies obtained in the competitive equilibrium maximize each of them. To see this, let us have a look at the parts of the Lagrangian corresponding to a single existing firm given our choice of Lagrange multipliers:

$$\max_{\{\tau_t,\theta_{x,t},d_t,n_{i,t},\theta_{it}\}_t}$$
$$\mathbb{E}\sum_t \beta^t \left[\prod_{l=-\infty}^{t-1}(1-d_l)\right]\Big[(1-d_t)\Big(e^{y(s_t)+z_t}F(n_t) - k_f - \left(\frac{c}{q(\theta_{i,t})} + x(s_t,\theta_{i,t})\right)n_{i,t}$$
$$... + n_{t-1}(1-\tau_t)\lambda p(\theta_{x,t})\, x(s_t,\theta_{x,t})\Big) + n_{t-1}(d_t + (1-d_t)\tau_t)\mathbf{U}(s_t)\Big],$$

which is the sequential formulation of the surplus maximization problem in the competitive equilibrium. Turning to firms entering at date t:

$$\max_{\left\{\tau_{t'},\theta_{x,t'},d_{t'},n_{i,t'},\theta_{i,t'}\right\}_{t'\geq t}} m_{e,t}\Big\{-k_e + \mathbb{E}\sum g_z(z_t)\times ...$$
$$\sum_{t'=t}^{\infty}\beta^{t'-t}\left[\prod_{l=t}^{t'-1}(1-d_l)\right]\Big[(1-d_{t'})\Big(e^{y(s_{t'})+z_{t'}}F(n_{t'}) - k_f - \left(\frac{c}{q(\theta_{i,t'})} + x(s_{t'},\theta_{i,t'})\right)n_{i,t}$$
$$+ n_{t'-1}(1-\tau_{t'})\lambda p(\theta_{x,t'})x(s_{t'},\theta_{x,t'})\Big) + n_{t'-1}(d_{t'} + (1-d_{t'})\tau_{t'})\mathbf{U}(s_{t'})\Big]\Big\}.$$

This is the sequential formulation of the free-entry problem solved in the competitive equilibrium. The planner increases the number of entrants $m_{e,t}$ as long as the expected surplus from entering is equal to the entry cost k_e. Now, let us examine the part of the Lagrangian related to unemployed

[32] The bounds $[\underline{x},\overline{x}]$ are chosen so that the optimal x lies in the interior, so that we are not constraining the equilibrium.

workers:

$$\max_{\{\theta_{u,t},u_t\}_t} \sum_t \beta^t \Big[u_t b - \mathbf{U}(s_t)\big(u_t - u_{t-1}(1-p(\theta_{ut}))\big) + u_{t-1}p(\theta_{u,t})x(s_t,\theta_{u,t})\Big]$$

The first-order conditions with respect to u_{t+1} and θ_{ut} are equal to

$$\begin{aligned}
&[u_t] \quad b - \mathbf{U}(s_t) + \beta \mathbb{E}\left[(1-p(\theta_{ut+1}))\,\mathbf{U}(s_{t+1}) + p(\theta_{ut+1})\,x(s_{t+1},\theta_{ut+1})\right] = 0\\
&[\theta_{u,t}] \quad -u_{t-1}p'(\theta_{ut})\,\mathbf{U}(s_t) + u_{t-1}p'(\theta_{ut})\,x(s_t,\theta_{ut}) + u_{t-1}p(\theta_{u,t})\,x_\theta(s_t,\theta_{u,t}) = 0.
\end{aligned}$$

We recognize in the first equation the Bellman equation faced by unemployed workers and, in the second equation, the first-order condition corresponding to their problem. Therefore, the policies obtained from the competitive equilibrium maximize the planner's problem given our choice of Lagrange multipliers. The first-order conditions are thus satisfied. Block-recursive equilibria are thus efficient. □

G.3 Proofs of part E

(ii) First, recall that

$$p(\theta(s,x)) = p \circ q^{-1}\left(\frac{c}{\kappa(s) - x}\right).$$

Under assumption 2, $p(\theta(s,x))$ is a strictly decreasing, strictly concave function of $x \in [\underline{x}, \kappa(s) - c]$.

Proof of proposition 3. *(i)* Pick a contract $\omega = \{w,\tau,x,W',d\}$ that implement the firm's optimal policy. Consider now the modified contract $\tilde{\omega} = \{w + a\Delta, \tau, x, W' - \Delta, d\}$ where $a = \beta\mathbb{E}\left[(1-d)(1-\tau)\left(1-\lambda p(\theta(s'\right.\right.$
The worker's utility under this new contract is

$$\begin{aligned}
\mathbf{W}(s,z,\tilde{\omega}) = w + a\Delta + \beta\mathbb{E}\Big[&(d + (1-d)\,\tau)\,\mathbf{U}(s') + (1-d)(1-\tau)\,\lambda p\left(\theta\left(s',x\right)\right)x\\
&... + (1-d)(1-\tau)\Big(1-\lambda p\left(\theta\left(s',x\right)\right)\Big)\left(W'-\Delta\right)\Big]\\
= \mathbf{W}(s,z,\omega)&
\end{aligned}$$

The worker's utility is unchanged. His promise-keeping constraint is thus still satisfied. Turning to the firm's profits:

$$\begin{aligned}
\mathbf{J}\left(s,z,n,\{W(j)\}_{j\in[0,n]}\right) &= e^{y(s)+z}F(n) - k_f - \int_0^n w(j)dj\\
&\quad + \beta\mathbb{E}\left[(1-d')\left(-n_i'\frac{c}{q\left(\theta\left(s',x_i'\right)\right)} + \mathbf{J}\left(s',z',n',\left\{\hat{W}'\right\}\right)\right)\right]\\
&= e^{y(s)+z}F(n) - k_f - \int_0^n w(j)dj + \beta\mathbb{E}\Big[(1-d')\Big(\mathbf{V}\left(s',z',n'\right)\\
&\quad - \int_0^n (1-\tau)\left(1-\lambda p\left(\theta\left(s',x\right)\right)\right)W'dj - n_i'\left(c/q\left(\theta\left(s',x_i\right)\right) + x_i\right)\Big)\Big].
\end{aligned}$$

Under the new contract $\tilde{\omega}$, we have

$$\begin{aligned} & -\int_0^n \tilde{w}(j)dj + \beta\mathbb{E}\int_0^n (1-d)(1-\tau)\left(1-\lambda p\left(\theta\left(s',x\right)\right)\right)\tilde{W}'\left(s',z';j\right)dj \\ = & -\int_0^n w(j)dj + \beta\mathbb{E}\int_0^n (1-d)(1-\tau)\left(1-\lambda p\left(\theta\left(s',x\right)\right)\right)W'\left(s',z';j\right)dj, \end{aligned}$$

so the firm's profit is unchanged. The new contract leaves the firm and workers indifferent and implements the firm's optimal policy as well.

(ii) It is useful to rewrite the surplus maximization problem as a two-step problem

$$\begin{aligned} \mathbf{V}(s,z,n) &= \max_{d,n} e^{y(s)+z}F(n) - k_f + \beta\mathbb{E}\Big\{dn\mathbf{U}\left(s'\right) + (1-d)\Big[v\left(s',z',n,n'\right) \\ &\quad +\mathbf{V}\left(s',z',n'\right)\Big]\Big\} \end{aligned}$$

with

$$\begin{aligned} v\left(s,z,n,n'\right) &= \max_{n_i,x_i,\{\tau(j),x(j)\}} \mathbf{U}(s)\int_0^n \tau(j)\mathrm{d}j + \int_0^n (1-\tau(j))\,\lambda p\left(\theta\left(s,x(j)\right)\right)x(j)\mathrm{d}j \\ &\quad -\left(\frac{c}{q\left(\theta\left(s,x_i\right)\right)} + x_i\right)n_i \\ \text{subject to} \quad & n' = \int_0^n (1-\tau)\left(1-\lambda p\left(\theta\left(s,x\right)\right)\right)dj + n_i. \end{aligned}$$

First, it is easy to show that if $n' \geq n$, then it is optimal to set $n_i = n' - n$, $\tau = 0$ and $x = \kappa(s) - c$ so that $p(\theta(s,x)) = 0$. Indeed, since it is costly to hire workers, it is never optimal to layoff or let any worker leave for another firm if it wants to expand. Let us now focus on the case in which $n' < n$. Again, it is easy to show in this case that $n_i = 0$. However, the firm must solve a trade-off between layoffs and job-to-job transitions which we can write as

$$\begin{aligned} \max_{\{\tau(j),x(j)\}} \quad & \mathbf{U}(s)\int_0^n \tau(j)\mathrm{d}j + \int_0^n (1-\tau(j))\,\lambda p\left(\theta\left(s,x(j)\right)\right)x(j)\mathrm{d}j \\ \text{subject to} \quad & n' = \int_0^n (1-\tau)\left(1-\lambda p\left(\theta\left(s,x\right)\right)\right)dj. \end{aligned}$$

Proceeding with the change of variables $\theta(j) = q^{-1}\left(c/\left(\kappa(s) - x(j)\right)\right)$, the problem becomes strictly concave in $\theta(j)$. Taking the first order conditions with respect to $x(j)$,

$$(\kappa(s) + \mu)\,p'\left(\theta(j)\right) = c,$$

where μ is the Lagrange multiplier on the constraint. We thus conclude that $\theta(j)$ and thus $x(j)$ are identical across workers within a given firm.

(iii) Imposing that $x(j) = x, \quad \forall j \in [0,n]$, it is trivial to see that any permutation of the τ's between workers or any transformation that leave the total mass of layoff unchanged does not affect the objective function. The total number of layoff though is uniquely determined using the constraint: $\int_0^n \tau(s', z'; j)dj = n - (1 - \lambda p(\theta))^{-1} n'$. $\square$

Proof of proposition 4. I will prove the result in two steps. I will first show that if the firm can choose any continuing utility $W'(s', z')$, it is possible to find a schedule $W'(s', z'; x')$ that makes the worker choose x exactly. We will then show that this continuing utility must satisfy the participation constraint, i.e., $\lambda p(\theta(s', x))x + (1 - \lambda p(\theta(s', x)))W'(s', z'; x') \geq \mathbf{U}(s')$.

Step 1. Fix (s', z'). Recall that workers solve the problem[33]

$$x = \underset{\tilde{x} \in [\underline{x}, \kappa(s') - c]}{\operatorname{argmax}} \; p\left(\theta(s', \tilde{x})\right)\left(\tilde{x} - W'\left(s', z'\right)\right).$$

Define

$$\tilde{D}(x, W') = p(\theta(s', x))(x - W') \text{ and } \begin{cases} D(s', W') = \max\limits_{x \in [\underline{x}, \kappa(s) - c]} \tilde{D}(x, W') \\ C(s', W') = \underset{\tilde{x} \in [\underline{x}, \kappa(s') - c]}{\operatorname{argmax}} \tilde{D}(x, W') \end{cases}$$

$\tilde{D}$ is a continuous function of x and W'. It reaches a non-negative maximum in x on $[W', \kappa(s') - c]$. Assumption 2 guarantees that $\tilde{D}$ is strictly concave in x on $[W', \kappa(s') - c]$. The Theorem of the Maximum tells us therefore that $D(W')$ and $C(W')$ are continuous functions of W'. Thus, p being strictly positive over $[\underline{x}, \kappa(s') - c)$, D is strictly decreasing on $[-\infty, \kappa(s') - c]$. Therefore, C is strictly increasing on $[-\infty, \kappa(s') - c]$, as can be seen from the following: take $W_1 < W_2 \leq \kappa(s') - c$. Denote $x_k = C(W_k), k = 1, 2$. Then the following is true:

$$p\left(\theta\left(s', x_1\right)\right)\left(x_1 - W'_1\right) - p\left(\theta\left(s', x_2\right)\right)\left(x_2 - W'_2\right) < p\left(\theta\left(s', x_1\right)\right)\left(W'_2 - W'_1\right),$$

and

$$p\left(\theta\left(s', x_1\right)\right)\left(x_1 - W'_1\right) - p\left(\theta\left(s', x_2\right)\right)\left(x_2 - W'_2\right) > p\left(\theta\left(s', x_2\right)\right)\left(W'_2 - W'_1\right).$$

Therefore, $\theta(s', x_1) > \theta(s', x_2)$, and since in equilibrium $\theta(s', x) = q^{-1}(c/(\kappa(s') - x))$ is decreasing in x, we have:$x_2 > x_1$ and C is strictly increasing.

Now, let us show that C reaches $\underline{x}$ and $\kappa(s) - c$. For $W' = \kappa(s) - c$, function $\tilde{D}$ trivially reaches its maximum at $x = W' = \kappa(s') - c$. Does it reach $\underline{x}$? Rewrite the maximization problem of the worker over θ:

$$\begin{aligned} \tilde{D} &= \max_{\theta \in [0, \theta(s, \underline{x})]} p(\theta)\left(\tilde{x}(\theta) - W'\right) \\ &= \max_{\theta \in [0, \theta(s, \underline{x})]} p(\theta)\left(\kappa\left(s'\right) - W'\right) - c\theta, \end{aligned}$$

[33] Remember that $x = \kappa(s) - c$ is the highest active submarket in equilibrium. It satisfies $\theta(s, x) = 0$.

where I have used the equilibrium relationship: $\kappa(s) = x + c/q(\theta(s,x))$. This is a well defined strictly concave maximization problem and its derivative with respect to θ is

$$p'(\theta)(\kappa(s') - W') - c,$$

so that $\theta = (p')^{-1}(c/(\kappa(s') - W'))$. Therefore, setting W' to equal $\kappa(s') - c/p'(\theta(s', \underline{x}))$, the optimum is reached at $\theta(s', \underline{x})$ and the worker chooses to search in submarket $\underline{x}$. $C(W')$ is thus a continuous strictly increasing function that reaches $\underline{x}$ and $\kappa(s') - c$. By the Intermediate Value Theorem, for any $x \in [\underline{x}, \kappa(s') - c]$, there exists a unique $W'_{IC}(x')$ such that $\max_{\tilde{x}} \tilde{D}(\tilde{x}, W'_{IC}(x))$ is reached at x exactly. In other words, there exists a unique continuation utility $W'_{IC} \in [-\infty, \kappa(s') - c]$ that makes the worker choose exactly x. To finish this first step, we must choose the rest of the contract. Set $\tau_{IC} = \tau$ and $d_{IC} = d$. Now, in an optimal allocation, w_{IC} must be chosen so that the promise-keeping constraint is binding. The worker's expected utility is

$$\begin{aligned}\mathbf{W}(s, z, \{w, \tau, x, d, W'\}) =& w + \beta\mathbb{E}\Big[\big(d + (1-d)\,\tau\big)\mathbf{U}(s') + (1-d)\,(1-\tau)\lambda p\left(\theta\left(s', x\right)\right) x \\ & + (1-d)\,(1-\tau)\left(1 - \lambda p\left(\theta\left(s', x\right)\right)\right) W'\Big].\end{aligned}$$

Given $\{\tau_{IC}, x_{IC}, d_{IC}, W'_{IC}\}$, there exists a unique wage w_{IC} that matches exactly the promised utility. This does not affect the joint surplus, which is maximized by assumption. From proposition 1, the firm's profit is maximized when the level of promised utility is exactly achieved. We have thus found a contract that implements the optimal allocation.

Step 2. I will now proceed to the second step of the proof and show that the participation constraint is satisfied by $W'(s', z'; x)$. Let us first have a look at the problem faced by the worker choosing whether or not to leave the firm at the time of separation. The participation constraint is satisfied if

$$\max_{x} \lambda p(\theta\left(s', x\right))x + (1 - \lambda p(\theta\left(s', x\right)))W'\left(s', z'; x\right) \geq \mathbf{U}(s).$$

Abusing notation slightly, denote $p(s', x) \equiv p(\theta(s', x))$, we can derive the first-order condition for the worker:

$$\lambda p'(s', x)(x - W') + \lambda p(s', x) = 0.$$

Turning back to the joint surplus maximization, the terms related to x and τ are

$$\begin{aligned}&\mathbf{U}\left(s'\right) \int_0^n \tau \mathrm{d}j + \lambda p\left(s', x\right) x \int_0^n (1-\tau)\mathrm{d}j \\ &+ \mathbf{V}(s', z', \int_0^n (1 - \lambda p\left(s', x\right))(1-\tau)\mathrm{d}j + n_i).\end{aligned}$$

To simplify the notatinon, write $nT = \int_0^n \tau dj$, T being the total fraction of layoffs. We can rewrite the above term as

$$nT\mathbf{U}\left(s'\right) + n(1-T)\lambda p\left(s',x\right)x + \mathbf{V}\left(s',z',n\left(1-T\right)\left(1-\lambda p\left(s',x\right)\right)+n_i\right).$$

The first-order condition with respect to x is

$$n(1-T)\lambda p'(s',x)\Big(x - \mathbf{V}_n\big(s',z',n(1-T)(1-\lambda p(s',x))+n_i\big)\Big) + n(1-T)\lambda p(s',x) = 0.$$

Notice that it is possible to identify W' from the two first-order conditions. The incentive compatible contract must be such that

$$W'\left(s',z'\right) = \mathbf{V}_n\Big(s',z',n(1-T)(1-\lambda p(s',x))+n_i\Big).$$

To verify whether the participation constraint is satisfied, it is informative to look at the first-order condition with respect to T (ignoring the irrelevant case where $T=1$):

$$n\mathbf{U}\left(s'\right) - n\Big(\lambda p(s',x)x + (1-\lambda p(s',x))\mathbf{V}_n\Big) \leq 0,$$

which is exactly equivalent to the participation constraint

$$\lambda p(s',x)x + (1-\lambda p(s',x))W'\left(s',z';x\right) \geq \mathbf{U}\left(s'\right).$$

The incentive-compatible contract therefore satisfies the participation constraint. $\square$