A Labor Market Sorting Model of Scarring and Hysteresis *

Edoardo Maria Acabbi[1], Andrea Alati[2], and Luca Mazzone[3]

[1]*Universidad Carlos III de Madrid*
[2]*Bank of England*
[3]*International Monetary Fund*

January 2023

Abstract

This paper contributes a new framework to account for the interactions between labor market sorting and business cycle fluctuations. We propose a tractable search equilibrium model of the labor market with aggregate risk, firm and worker heterogeneity, life-cycle dynamics and endogenous human capital accumulation. The model is estimated using Italian administrative matched employer-employee data. We show that sorting of workers to firms is a key factor in explaining the increased persistence of fluctuations, directly relating labor reallocation to economic hysteresis. Changes in the cross sectional distribution of worker and firm characteristics explain why "double dips" can be more painful than isolated recessions. Finally, the model can be used to account for the increased length of recessions in recent decades.

*The authors thank David Andolfatto, Alessandro Dovis, Romain Duval, Simone Ferro, Cecilia Garcia-Peñalosa, Kyle Herkenhoff, Thibaut Lamadon, Paolo Martellini, Antonio Mele, Kurt Mitman, Matteo Paradisi, Facundo Piguillem, Josep Pijoan-Mas, Silvia Vannutelli, Liangjie Wu and the participants to the Visitinps, EIEF, CEMFI, HEC Lausanne, IMF, UC3M, University of Milan, Bank of England seminars and Northwestern University, to the Bonn CRC workshop, to the SMYE 2021 conference and to the AEA 2022 poster session and to the 20^{th} Brucchi-Luchino workshop for useful comments and conversations. Acabbi gratefully acknowledges the financial support from the Comunidad de Madrid (Programa Excelencia para el Profesorado Universitario, convenio con Universidad Carlos III de Madrid, V Plan Regional de Investigación Científica e Innovación Tecnológica), the Spanish Ministry of Science and Innovation, project PID2020-114108GB-I00, and the Fundacion Ramon Areces. This study uses anonymous data from the Italian Social Security Institute (INPS). Data access was provided as part of the VisintINPS Scholars initiative. We are grateful to Alfredo Arpinelli, Massimo Antichi, Edoardo Di Porto, Pietro Garibaldi, Monia Monachini, Paolo Naticchioni, Marcella Nunzi and all the staff of Direzione Centrale Studi e Ricerche at INPS for making the VisitINPS programme possible and to Massimo Ascione and Vadim Bottoni for their help with the data. The findings and conclusions expressed are solely those of the authors and do not represent the views of INPS, the Bank of England or the IMF. All errors are our own.
Acabbi: eacabbi@emp.uc3m.es
Alati: andrea.alati@bankofengland.co.uk
Mazzone: lmazzone@imf.org

1 Introduction

Business cycle fluctuations strongly interact with the functioning of labor markets. These interactions are not only evident in the cyclicality of flows between jobs, and from employment to unemployment, but also in the sorting of workers and firms (Lise and Robin, 2017, Haltiwanger et al., 2021). Research has shown that differences in the allocation of resources can account for differences in income levels across countries (Restuccia and Rogerson, 2017), but little is known regarding how changes in the sorting between workers and firms interact with business cycle dynamics and how they can explain longer run trends in human capital accumulation and job creation.

Empirically, the degree to which more productive workers are assigned to more productive firms is negatively correlated with output at business cycle frequency, but correlates positively with output levels at lower frequencies.[1] Income inequality is instead counter-cyclical, but correlates positively with sorting (Heathcote, Perri and Violante, 2020, Song et al., 2019). Taken together, these results imply that understanding the rich interplay between sorting and aggregate fluctuations can provide a better understanding of both business cycles and labor markets. A deeper understanding of the cyclical properties of sorting is crucial to establish whether recessions are ultimately improve or worsen allocative efficiency in the economy. Such a characterization also provides insights into the causes of persistent losses in output in recessions and in growth thereafter - that is, of output state-dependence, or *hysteresis*.

To this end, we contribute a new theoretical framework that accounts for the interactions between cross sectional labor market dynamics and aggregate fluctuations. Our framework incorporates firm and worker heterogeneity, life-cycle and participation dynamics, search on the job and endogenous human capital accumulation in a directed search model of the labor market with aggregate risk. We estimate the model matching a large set of moments from administrative data. The model yields a realistic characterization of the dynamics of labor market sorting, and allows for match creation to depend dynamically on the evolution of the cross sectional distribution of the productivities of workers and firms as well as the aggregate state, without sacrificing tractability.

The focus on firm-worker sorting for aggregate dynamics in our work is closely related to recent papers by Lise and Robin (2017) and Baley, Figueiredo and Ulbricht (2022). In common with our work, they develop rich models of the labor market highlighting the importance of the interactions between business cycles and skill (mis-)matches. An important difference with these studies and ours is that, by focusing on human capital accumulation and a realistic pattern of labor market flows by worker and firm type, we

[1] See Figure 1.

establish a direct link between labor market sorting and aggregate dynamics. Our focus is thus ideal for understanding aggregate fluctuations at lower frequencies, and the source of longer term changes in labor market structure.

In our model, firms produce according to a technology that yields increasing returns in both worker and firm productivity. Human capital is accumulated at a faster pace in jobs posted by more productive firms, which are harder to get. This induces a job ladder that depends on the accumulation of human capital over the life cycle. Most importantly, we show it induces a monotonically increasing mapping in the search strategies of workers towards firms. Search strategies change over the business cycle, because vacancy creation and thus outside options are responsive to aggregate conditions. As a result, the sorting between workers and firms is not perfect, and depends on structural parameters governing the fluidity of labor markets as well as the impact of job displacement.

In order to capture the role of displacements, we incorporate a two-sided commitment friction that generates inefficient separations. Downward wage rigidity limits the ability of firms to reduce losses when aggregate or idiosyncratic shocks hit them, so they can become insolvent and lay off workers in bad times. In such a framework, whenever aggregate conditions change, firms and workers respond endogenously by adjusting their respective wage and vacancy openings vis-a-vis quit and search policies. The resulting pattern of wage growth, matches and separations is consistent with the data, and reproduces a host of results from the empirical labor literature.

We use our model to decompose the effect of aggregate fluctuations on output through labor market adjustments. We show that displacements constitute the main driver of output dynamics in the short run at the onset of a recession. In the medium run though, the length of recoveries depends on a combination of losses in human capital investment and worsened worker-firm sorting. Recessions have a long run effect for the economy as well; we show that, even after output levels have recovered, cross sectional movements in labor compensation and matches quality following a recession can make the economy more susceptible to additional shocks. This way, we provide a rationale for the severity of *double dip* recessions, successive recessions hitting an economy within a relatively short time span.

The core economic mechanism behind medium and long term effects is the change in sorting. Recent research has shown that advanced economies have experienced a secular trend towards greater sorting in labor markets (see Song et al., 2019). We find that changes in the human capital production technology can explain not only the increased sorting over the past decades, but also contributed to the contemporaneous increase in recession length observed by Fukui, Nakamura and Steinsson (2023). Together with demographic shifts (aging), the increased importance of firm quality for the path of workers' human capital accumulation can explain most of the increased length of recoveries observed in all advanced economies in the last 30 years.

The recessionary dynamics and the longer term effects for the economy determined by distorsions in sorting and human capital accumulation have significant first-order effects on welfare. In our model the welfare cost of business cycle is on average above 2%, orders of magnitude larger than what calculated by Lucas (1987), and broadly in line with the one in Barlevy (2004). These costs are highly heterogeneous across workers, with lower skill and younger workers suffering an almost doubled welfare loss from business cycles. Given this finding, the analysis and introduction of distributional policies becomes quite natural. We analyze two policies: a minimum wage (Dustmann et al., 2021, Berger, Herkenhoff and Mongey, 2022) and a countercyclical transfer to the unemployed, akin to countercyclical unemployment benefits (Nekoei and Weber, 2017). In our model the introduction of a minimum wage reduces welfare because the negative impact on job creation ends up being larger than the gains accruing to workers from the increase in earnings at the bottom of the earnings distribution, and the subsequent upwards adjustment in search strategies. On the other hand, countercyclical unemployment benefits increase long run output and welfare, mostly by reducing the scarring effects of recessions on displaced workers. The expectation of large unemployment transfers improves labor force participation. Large transfers also make workers more "patient", and hence more ambitious, in their search strategy. This improves the overall allocative efficiency and increases the stock of human capital.

As noted above, our quantitative results hinge on estimating the key parameters of the model by matching cross sectional empirical moments. For this purpose, we rely on administrative data covering the universe of worker-firm relationships provided by the Italian Social Security Administration (INPS) for all years between 1996 and 2018. We model the economy with an overlapping-generation structure, where workers are heterogeneous in their human capital level, and in their ability to accumulate it on the job. Workers, both when employed and when unemployed, direct their search towards firms, which are heterogeneous in productivity, trading off the value of each vacancy with the likelihood of matching. The value of matches depends not only on the promised earnings, but also on the implied human capital accumulation path during the worker's tenure at the firm. Similarly to Lise and Postel-Vinay (2020), we assume human capital accumulation is stronger in more productive firms. We validate this mechanism by replicating the results on the importance of past employers on subsequent earnings in Arellano-Bover (2022), Herkenhoff et al. (2018) on both model and data. Because vacancy opening costs are increasing in firm productivity, entrant firms trade off the incentives on attracting the best workers against the probability of filling the match. A free entry condition closes the model. In equilibrium, each firm type will offer a different menu of contract to workers, inducing a search strategy that is monotonically increasing in workers' human capital.

To discipline the dynamics of sorting, we match the age-specific profile of employment to employment transitions, as well as the overall degree of sorting in the economy, and the distribution of firm productivity.

We discipline the earnings' dynamics of the model by matching life cycle income profiles by education. The estimation also targets the age-specific flow rates from employment to unemployment, and the overall unemployment rate. Profit-maximizing firms device a wage protocol with the objective of attracting the most productive workers, and retaining them - see the Balke and Lamadon (2022) treatment of the optimal contract problem of Holmstrom (1983). The resulting state-contingent contract insures workers against match value fluctuations, and backloads compensation to maximize retention. Because workers accumulate human capital on the job, incentives to search on the job vary over job tenure. As a result, the within-match earnings growth is smaller than the between-match growth, and both decrease over the life cycle. In addition, the share of between-match wage growth is even higher when workers are employed by more productive firms.[2]

A sizeable literature identifies substantial resistance to nominal or real wage cuts (Altonji and Devereux, 1999, Agell and Lundborg, 2003, Grigsby, Hurst and Yildirmaz, 2021). Firms cannot adjust wages downward when the continuation value of matches is negative, and lay off workers instead. The model is then able to reproduce empirical separation patterns across age, worker and firm productivity, and matches the correlation of separations with business cycles. Because of cyclical separations and vacancy creations, recessions can have a lasting impact on workers' earnings (Kahn, 2010, Schwandt and von Wachter, 2019, Schmieder, von Wachter and Heining, 2022, Huckfeldt, 2022, Bertheau et al., forthcoming). As a validation of our proposed mechanisms, we verify the model is able to replicate the empirical estimates on *scarring effects* both in magnitude and in duration. Decomposing the scarring effect shows that the main component of scarring effects is the role of displacements, followed by losses in human capital accumulation. Sorting, then, matters in the aggregate but does not impact the earnings of employed workers in the short run.

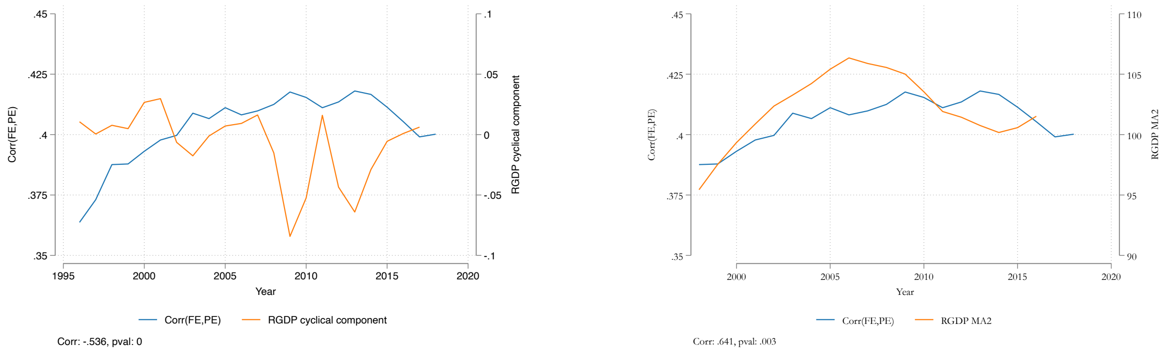
2 Model

In this section we present our model of the labor market. We start by discussing the environment, the timing and the preference structure of the agents populating the economy. We then discuss the features of the frictional labor market with directed

[2]Our research is also importantly related to a strand of literature in labor and finance analyzing the firms' management of liquidity and labor compensation dynamics within the firm (Xiaolan, 2014, Favilukis, Lin and Zhao, 2020, Acabbi, Panetti and Sforza, 2022, Acabbi and Alati, 2022).

Figure 1. Sorting and output correlations

(a) Sorting and the cycle **(b)** Sorting and long run output



Note: The figures depict the correlation between our measure of sorting (the correlation between firm and worker estimated fixed effects) and a measure of cycle and a measure of trend gdp. In panel 1a the cycle component of GDP is obtained by applying a Hamilton filter. In panel 1b GDP trend is a moving average MA(2). Source: Istat (GDP) and INPS-Uniemens (worker-firms panel and AKM fixed effects).

search, and finally we characterize the workers problem, the optimal recursive contract and the equilibrium of the model.

2.1 Environment

Time is discrete, runs forever and is indexed by $t \in \mathbb{N}$. We denote future values in recursive expressions by adding a $'$ to them, or index elements by t in non-recursive ones.

The economy is populated by two kinds of agents: a unit measure of finitely-lived risk-averse workers and a continuum of risk neutral entrepreneurs. The number of operating firms is determined by entrepreneurs' investment decisions. All agents in the economy share the same discount factor $\beta \in (0, 1)$. The economy is populated with $T \geq 2$ overlapping generations of workers, which face both aggregate and idiosyncratic risk. Each household lives for T periods, with age $\tau \in \mathcal{T} \equiv \{1, 2, 3, \ldots, T\}$. Workers are either employed (e) or unemployed (u), are assumed to be hand-to-mouth, and can direct their search towards different submarkets (Shi, 2009, Menzio and Shi, 2010).

Workers maximize lifetime flow-utility from non durable consumption:

$$\mathbb{E}_{t_0}\left(\sum_{\tau=1}^{T} \beta^{\tau} u(c_{\tau,t_0+\tau})\right)$$

where t_0 characterizes the time of entry into the labor market and τ characterizes the age of the agent. $c_{\tau,t_0+\tau}$ thus refers to the consumption of workers of age τ in time $t_0 + \tau$.

Workers are characterized by heterogeneous human capital levels h, with $h \in \mathcal{H} \equiv [\underline{h}, \overline{h}]$ and a fixed type, their education level $\iota \in \mathcal{I} \equiv \{\text{graduate, non-graduate}\}$.

Both types enter the labor market with a baseline level of human capital drawn from type-specific exogenous continuous distributions. Upon entry in the labor market, $\mathbb{E}[h|\text{graduate}] > \mathbb{E}[h|\text{non graduate}]$. To account for the different amount of years schooling, graduate workers entry in the labor market is delayed with respect to non-graduates'. We allow workers to exit from the labor force when their employability deteriorates below a certain utility threshold. This allows us to distinguish long-term unemployment from non-participation.

Workers acquire human capital on-the-job. They are matched with firms characterized by different levels of quality $y \in \mathcal{Y} \equiv [\underline{y}, \overline{y}]$, which is isomorphic to capital levels. Following Lise and Postel-Vinay (2020), we assume human capital accumulation depends on the quality level of the firm y, the level of ability of the worker h and an idiosyncratic human capital shock $\psi \sim \mathcal{N}(0, \sigma_\psi)$. The worker thus accumulates human capital according to a law of motion that is match-specific: $h' = \phi(h, y, \iota, \psi) = g(h, y, \iota) + \psi$, $n : \mathcal{H} \times \mathcal{Y} \times \mathcal{I} \to \mathcal{H}$, where g is the deterministic component of the human capital accumulation dynamics, and ψ constitutes the stochastic component. The function n is concave in h and y. The deterministic component of human capital accumulation is akin to a "catching-up" of the firm's quality, up to a point when the worker will not be able to learn anymore from the match. Consistently with the concept of "mismatch", workers who lose their job and only manage to re-match with a low quality firm see their ability deteriorating with the same g function. The only difference between graduates and non-graduates (indexed by ι) is the speed of the "catching-up". Graduate workers will catch up faster (or lose skills faster when mismatched). Other than this feature, graduate workers face the same problem as non-graduates.

Human capital accumulation is risky: at any period any employed worker is subject to the idiosyncratic human capital shock ψ, which enters additively with respect to the deterministic component.[3] The shock affects workers' ability and can amplify, shrink or even reverse human capital accumulation. We further allow for the possibility that human capital deteriorates while workers are unemployed, according to an arbitrary process g_u.[4]

Firms are modeled as one worker-one job matches, thus abstracting from firm size. Each job match is characterized by a promised utility to the worker $V \in \mathcal{V}$. The determination and properties of the contract space $\mathcal{V}$ are described in the next sections.

We group worker-specific characteristics in a tuple $\chi \in \mathcal{X} \equiv \{\mathcal{H} \times \mathcal{T} \times \mathcal{I}\}$. The aggregate state of the economy Ω is characterized by the productivity level $a \in \mathcal{A} \subseteq \mathbf{R}_0^+$ and by the distribution of agents across states $\mu \in \mathcal{M} : \{e, u\} \times \mathcal{Y} \times \mathcal{X} \to [0, 1]$. Let $\Omega = (a, \mu) \in \mathcal{A} \times \mathcal{M}$ represent the aggregate state of the economy and let $\mathcal{M}$ represent

[3]The additive nature of the shock keeps the properties of monotonicity and uniqueness of workers' search strategies unaltered, which is essential for tractability.

[4]This process might be without loss of generality deterministic or stochastic, and might or might not depend on current human capital h.

the set of distributions μ over the states of the economy. Let $\mu' = \Phi(\Omega, a')$ be the law of motion of the distribution. Aggregate productivity evolves as a stationary monotone increasing Markov process, namely $a' \sim F(a'|a) : \mathcal{A} \to \mathcal{A}$, with the Feller property.

2.2 Labor markets

Search is directed. Each labor market is organized as a continuum of submarkets indexed by the expected lifetime utility offered by firms of type y, $v_y \in \mathcal{V} \equiv [\underline{v}, \overline{v}]$. Workers are indexed by the tuple $\chi = (h, \tau, \iota)$.[5] The process of opening a firm, which amounts to posting a vacancy at a quality-specific cost $\kappa(y)$, will be described in **Section 2.7**.

The search process is characterized by a constant return to scale twice continuously differentiable matching function $M(u, \nu)$ for each submarket. The tightness of each submarket in $\mathcal{X} \times \mathcal{V}$ is defined as $\theta = \nu/u$, with $\theta(\cdot) : \mathcal{X} \times \mathcal{V} \to \mathbf{R}_0^+$. Job finding rates are defined as $p(\theta(\cdot)) = M(u, \nu)/u$, where $p(\cdot) : \mathbf{R}_0^+ \to [0, 1]$ is twice continuously differentiable, strictly increasing and strictly concave function with $p(0) = 0$, $\lim_{\theta \to +\infty} p(\theta) = 1$ and $p'(0) < \infty$. The vacancy-filling probability is in turn defined as $q(\theta(\cdot)) = M(u, \nu)/\nu$, where $q(\cdot) : \mathbf{R}_0^+ \to [0, 1]$ is twice continuously differentiable, strictly decreasing and strictly convex, with $q(0) = 1$, $\lim_{\theta \to +\infty} q(\theta) = 0$ and $q'(0) < 0$. Given these properties $q(\theta) = p(\theta)/\theta$, and $p(q^{-1}(\cdot))$ is concave.

Upon match, workers produce according to the twice-continuous increasing and concave production function $f(h, y; a) + x(a) : \mathcal{A} \times \mathcal{H} \times \mathcal{Y} \to \mathbf{R}_0^+$. The $x(a)$ component of the production function is a fixed cost, which can depend on the aggregate productivity realization.[6] Workers' compensation is determined by means of dynamic contracts through which firms deliver a promised utility, as described in **Section 2.6**.

Workers are always allowed to search while unemployed and search while employed with probability λ_e. Matches are destroyed at an exogenous rate λ_τ each period, with the exogenous separation rate possibly varying by age. Moreover, matches separate endogenously if the worker is poached by another firm, if they voluntarily decide to quit to unemployment (quit) or if the value of the match for the firm becomes negative (firings). Lastly, unemployed workers whose expected value of re-employment falls below a threshold, $\underline{\mathrm{p}}$, are assumed to permanently exit the labor force.

[5] As in Menzio and Shi (2010) the equilibrium will be separating. Given a menu of offers from any firm, each worker will visit only a particular submarket. For this reason submarkets can then be indexed directly by workers' current characteristics (see **Section 2.3**).

[6] This is a reduced form way of incorporating financial frictions in the model, which make fixed costs loom larger over flow-production in downturns.

2.3 Informational and contractual structure

Firms post contracts to workers. A contract prescribes an action for each realization of the history of the worker-firm match, and is thus fully state-contingent. The state of a match at a generic time t is defined by $s_t = (h_t, \tau_t, \iota, a^t, \mu^t) \in \mathcal{S}^t = \mathcal{X} \times \Omega^t = \mathcal{H} \times \mathcal{T} \times I \times \Omega^t$, that is the worker skill, age, education, the history of aggregate productivity shocks and workers' distributions across their employment history. A given history of realizations between t and k periods ahead is thus $s^{t+k} = (s_t, s_{t+1} ...., s_{t+k})$. The contract defines a transfer of utility from the risk neutral firm to the risk averse worker within the match for all future possible histories of shocks. We define τ_{t_0} as the age at which the worker is hired and T is the retirement age. The history of realizations between t_0, the time of hiring of the worker, and $t_0 + (T - \tau_{t_0})$, the time of maximum duration of the match with the worker before retirement, is thus $s^{t_0+(T-\tau_{t_0})}$.

Histories of workers and productivity shocks are common knowledge, and their future realizations are fully contractible upon. While the contract is state-contingent, markets are incomplete: workers' actions are private knowledge in the search stage, so firms are unable to counter outside offers. The contracts offered by firms are then defined as:

$$\mathcal{C}^{\tau_{t_0}} := (\mathbf{w}, \zeta) \text{ with } \mathbf{w} := \{w_t(s^{\tau_t - \tau_{t_0} + t_0})\}_{t=t_0}^{t_0+(T-\tau_{t_0})}, \text{ and } \zeta := \{v_t(s^{\tau_t - \tau_{t_0} + t_0})\}_{t=t_0}^{t_0+(T-\tau_{t_0})} \quad (1)$$

Firms promise a series of state-contingent wages, and employed workers reply by enacting their own state-contingent search strategies, defined by the series of utility values v_t sought at each node of the history.[7] ζ is the action suggested by the contract, which is bound to be incentive compatible for the worker. The contract is otherwise flexible in the degree to which the firm can determine wage levels and adjustment paths over match histories.

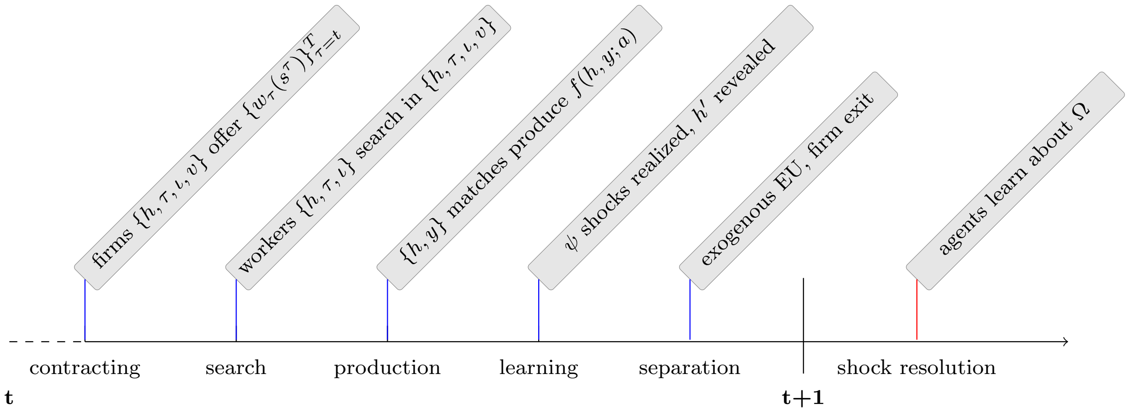
The relationship between workers and firms is characterized by a contract with forward looking constraints. The state space of the worker problem can be expressed in terms of their current lifetime utility, as in Spear and Srivastava (1987), so as to avoid having to keep track of all past histories s^t at each period. The relevant state space is then $\mathcal{X} \times \mathcal{V}$.

2.4 Timing

The timing of each period is represented in Figure 2. At the beginning of each period an aggregate productivity shock is drawn; entrepreneurs open vacancies across submarkets and post their offers; workers search from unemployment or on-the-job, and move to a new job if the search is successful; production takes place; workers accumulate human

[7] Similarly to Menzio and Shi (2010), Balke and Lamadon (2022), in order to guarantee that the problem is well behaved and the firm profit function is concave, the contract will require a randomization, a two-point lottery, which specifies probabilities over the actions prescribed. We omit it here for conciseness.

Figure 2. Timeline of worker-firm match



capital depending on their employment status and idiosyncratic shock realization; an exogenous share of matches breaks down, while some firms endogenously decide to exit.

2.5 Worker problem

Given current lifetime utility V, job seekers with characteristics χ have to decide in which submarket to direct their search. Submarkets are indexed by worker type χ and posted offered utility v. As it will be shown in **Section 2.7**, the choice over v will also indirectly determine which kind of firm y the worker matches with, and thus in expectation the human capital accumulation path. For now, let us assume this (conditional) mapping exists. This amounts to assuming that the function $v(y;\chi,V)$ is an injective function $f_v : \mathcal{Y} \times \mathcal{X} \times \mathcal{V} \to \mathcal{V}$. This means that, upon observing a job offer with utility v, a worker χ with current utility V will be able to infer which firm type y is posting the offer. Notice this also entails that workers implicitly determine their own expected human capital accumulation path, as it depends on the firm quality y they match with.

A worker of type (χ, V) that enters the search stage has lifetime utility $V + \max\{0, R(\chi, V; \Omega\}$, where the second component of the expression embeds the option value of the search, with R being the search value function. R is defined as:

$$R(\chi, V; \Omega) = \sup_{v} \Big[p(\theta(\chi, v; \Omega)) \big[v - V \big] \Big] \tag{2}$$

We denote the solution of the search problem as $v^* = v^*(\chi, V; \Omega)$, and $p^*(\chi, v^*; \Omega) = p(\theta(\chi, v^*; \Omega))$ as the associated optimal job-finding probability.

The lifetime utility of an unemployed worker at the beginning of the production stage can be defined as

$$
\begin{aligned}
U(h,\tau,\iota;\Omega) = \ & u(b(h,\tau)) + \beta\mathbb{E}_{\Omega,\psi}\Big(U(h',\tau+1,\iota;\Omega') \\
& + \max\{0, R(h',\tau+1,\iota,U(h',\tau+1,\iota;\Omega');\Omega')\}\Big)
\end{aligned} \tag{3}
$$

where $b(h,\tau)$ is a (possibly) skill and age dependent unemployment benefit. Given finite workers' lives, $U(h,\tau,\iota;\Omega) = 0 \ \forall(\chi;\Omega) \in \mathcal{X} \times \Omega$ where $\tau > T$.

The corresponding lifetime utility of a worker employed at a firm y with human capital h, age τ, education ι and current promised utility V at the beginning of the production stage can be expressed as:

$$
\begin{aligned}
V(h,\tau,\iota;\Omega) = \ & u(w) + \beta\mathbb{E}_{\Omega,\psi}\Big(\lambda_\tau U(h',\tau+1,\iota;\Omega') + \ (1-\lambda_\tau)\Big[V(h',\tau+1,\iota;\Omega') \\
& + \ \lambda_e \max\{0, R(h',\tau+1,\iota,V(h',\tau+1,\iota;\Omega');\Omega')\}\Big]\Big)
\end{aligned} \tag{4}
$$

where w is the currently promised wage and $V(h',\tau+1,\iota;\Omega')$ is next period's state-contingent promised utility of remaining in the current firm, which becomes the outside option in the search problem.[8] Notice that the state contingent promised utilities are an equilibrium object themselves, as outcomes of the firm dynamic contract optimization.

Workers indirectly impact their current contract with their search strategy. In fact, firms internalize incentives embedded in workers' strategies in their optimization, and post wages and utility offers to maximize profits by optimizing retention. Workers future promised utility incorporates both higher wages and higher option values of search, also through the deterministic component of human capital accumulation dynamics $g(h,y)$.

The policy functions are uniquely defined, and allow to identify target y as long as the injective mapping between the offered utility v and y given χ exists.[9]

The solution of employed workers' on-the-job search problem implicitly defines two policy functions, which incorporate workers' incentive compatibility. In their optimization, firms internalize these optimal responses on the part of workers.

Definition 2.1 (Optimal retention probability and utility return)**.** *The solution of the worker's problem defines a retention function* $\widetilde{p} : \mathcal{X} \times \mathcal{V} \times \Omega \to [(1-\lambda)(1-\lambda_e), 1-\lambda]$

[8]It is here implied that, in case there is an endogenous separation, this future promised value is equivalent to the value of being unemployed.

[9]Proofs of the uniqueness of policy functions and individuals' optimal policy are provided in **Appendix B**.

and a utility return $\widetilde{r}: \mathcal{X} \times \mathcal{V} \times \Omega \rightarrow \mathcal{V}$*:*

$$\widetilde{p}(\chi, V; \Omega) \equiv (1-\lambda_\tau)(1-\lambda_e p^*(\chi, v^*; \Omega)) \tag{5}$$

$$\widetilde{r}(\chi, V; \Omega) \equiv \lambda_\tau U(\chi; \Omega) + (1-\lambda_\tau)\Big[V + \lambda_e \max\{0, R(\chi, V; \Omega)\}\Big] \tag{6}$$

2.6 Firm problem

There is a double sided lack of commitment between parties. Firms commit to the delivery of a utility value to workers, but entrepreneurs can close them when the present value of future profits becomes negative. Worker, instead, cannot credibly commit in any circumstance - they will search for new jobs whenever they have the possibility to do so. Firms cannot observe poaching offers, and cannot thus counteract them. The sequence of past histories s^t is common knowledge, and while the firm cannot observe any of actions of its workers, it has enough information to incorporate their optimal search policy decisions.

We define $J(h, \tau, \iota, W, y; \Omega) \in \mathcal{X} \times \mathcal{V} \times \mathcal{Y} \times \Omega$ as the profit function of a firm, the difference between revenues from production and the promised wage. Given workers' finite lives, $J(h, \tau, \iota, W, y; \Omega) = 0$, $\forall (h, \tau, \iota, W, y; \Omega)$ where $\tau > T$. Incumbent firms make their exit decisions before the realization of aggregate productivity but after the realization idiosyncratic human capital shocks for next period. This implies that at the beginning of a period they already know whether they will exit. Given the current state we can define the following indicator function

Definition 2.2 (Exit policy)**.** *The following indicator takes value one if the firm does not decide to exit in the following period:*

$$\eta(h, \tau, \iota, W, y; \Omega) = \begin{cases} 1 & \textit{if } a \geq \max\{0, a^*\} \\ 0 & \textit{otherwise} \end{cases}$$

with the productivity threshold defined as

$$a^*(h, \tau, \iota, W, y; \Omega, \psi) : \mathbb{E}_\Omega[J(h', \tau+1, \iota, W', y; \Omega')] = 0. \tag{7}$$

Notice that the firm takes its exit decision *before* the realization of a new aggregate shock, but *after* the realization of the worker idiosyncratic human capital shock.[10]

Given $\eta = 1$, the value function of an incumbent firm y in state $(h, \tau, \iota, W_y; \Omega)$ can be

[10] Equivalently, this amounts to having the firm making a state-contingent exit decision in advance of the idiosyncratic shock's realization.

rewritten recursively using the promised utilities as additional state variables as:

$$J(h,\tau,\iota,W,y;\Omega) = \sup_{\pi_i,\{w_i,W_i'\}} \sum_{i=1,2} \pi_i \Bigg(f(y,h;a) - w_i + \beta \mathbb{E}_{\Omega,\psi} \Big[\widetilde{p}(h',\tau+1,\iota,W_i';\Omega')(J(h',\tau+1,\iota,W_i',y;\Omega') \Big] \Bigg) \tag{8}$$

$$s.t.\ W = \sum_{i=1,2} \pi_i \left(u(w_i) + \beta \mathbb{E}_{\Omega,\psi} \left(\widetilde{r}(h',\tau+1,\iota,W_i';\Omega') \right) \right), \tag{9}$$

$$\sum_{i=1,2} \pi_i = 1 \tag{10}$$

where **Equation** (9) is the promise keeping constraint ensuring that the current value of the contract is based on the current wage and future utility promises with $\widetilde{r}_t(\cdot)$. The firm (principal) optimizes over its possible offers taking into account the utility of workers (agent) and their incentive compatible best replies, through the retention probability $\widetilde{p}(\cdot)$ and the expected utility gain $\widetilde{r}(.)$. This yields a subgame perfect Nash equilibrium as the ones identified in leader-follower sequential games (Von Stackelberg (1934)).

2.7 Vacancy opening and free entry

The economy is populated by a continuum of risk-neutral entrepreneurs. Each entrepreneur can invest to reach the desired level of firm quality y. The start-up costs of the firm are priced in terms of the consumption good and they coincide with vacancy posting costs in the frictional labor market.

The cost of each vacancy is positively related to the quality of the firm being created. In order to post a vacancy for the creation of a firm with quality y the entrepreneur must thus pay $c(y)$, a vacancy cost priced in terms of the consumption good. The vacancy cost function $c(y)$ is a strictly convex function of firm quality y.[11]

At a generic time t each entrepreneur chooses in which submarket to post the vacancy selecting a lottery over the offered utility W_y, which maps into the set of firms' qualities $y \in \mathcal{Y}$, and worker characteristics $(\chi, V) \in \mathcal{X} \times \mathcal{V}$.

As entrepreneurs choose the submarkets in which to open a vacancy, they face the

[11] We assume that entrepreneurs can borrow from risk neutral deep pocketed financiers to finance the opening of a vacancy. Herkenhoff (2019) shows that, through this simplifying assumption, the cost of credit for entrepreneurs coincides with the risk-free rate.

following problem:

$$\Pi(h,\tau,\iota,W,y;\Omega) = \sup_{y,h,\tau,\iota,W} \; - c(y) + q(\theta(h,\tau,\iota,W;\Omega))[J(h,\tau,\iota,W,y;\Omega)] \tag{11}$$

Given perfect competition, free entry and the possibility for all entrepreneurs to choose *any* possible firm kind y the expected profits from opening a vacancy are driven down to 0 in submarkets which actually open.[12] This translates into a free entry condition:

$$\Pi(h,\tau,\iota,W,y;\Omega) \leq 0 \;\text{ for } \forall\{h,\tau,\iota,W,y;\Omega\} \in \{\mathcal{X} \times \mathcal{V} \times \mathcal{Y} \times \Omega\} \tag{12}$$

Assuming that $q(\cdot)$ is invertible, the equilibrium tightness in each submarket is:

$$\theta(h,\tau,\iota,W;\Omega) = q^{-1}\left(\frac{c(y)}{J(h,\tau,\iota,W,y;\Omega)}\right). \tag{13}$$

2.8 Equilibrium definition

Recursive Equilibrium. Let $\Theta = \mathcal{A} \times \mathcal{M} \times \mathcal{H} \times \mathcal{T} \times \mathcal{I}$. A recursive equilibrium in this economy consists of a market tightness $\theta : \Theta \times \mathcal{V} \to \mathbb{R}_+$, a search value function $R : \Theta \times \mathcal{V} \to \mathbb{R}$, a search policy function $v^* : \Theta \times \mathcal{V} \to \mathcal{V}$, an unemployment value function $U : \Theta \to \mathbb{R}$, a firm value function, $J : \Theta \times \mathcal{V} \times \mathcal{Y} \to \mathbb{R}$, a series of contract policy functions $\{c_\tau\}_{\tau=1}^{T} : \mathcal{S}^\tau \times \mathcal{Y} \to \mathcal{C}^\tau$, an injective mapping between firm qualities and promised utilities at hiring $f_v : \mathcal{X} \times \mathcal{V} \times \mathcal{Y} \to \mathcal{V}$, an exit threshold for aggregate productivity $a^* : \mathcal{X} \times \mathcal{V} \times \mathcal{Y} \to \mathcal{A}$, a human capital accumulation process $\phi(h,y,\iota,\psi)$, $\mathcal{H} \times \mathcal{Y} \times I \times \Psi \to \mathcal{H}$, a law of motion for the aggregate state of the economy $\Phi_{\Omega,a}$: $\mathcal{A} \times \mathcal{M} \to \mathcal{A} \times \mathcal{M}$ such that:

1. given the mapping f_v, market tightness satisfies **Equation** (13)
2. the unemployment value function solves **Equation** (3)
3. search value functions solve the search problem in **Equation** (2) and v^* is the associated policy function
4. firm value functions and the associated contract policy functions solve **Equation** (8) for each $t \leq T$
5. the exit threshold satisfies **Equation** (7)
6. the law of motion for the aggregate state of the economy respects the search and contract policy functions and the exogenous process of aggregate productivity.

[12] Notice that in this case the expectation does not refer to realizations of the aggregate state Ω or the human capital shock ψ, but to the vacancy-filling probability q.

Definition 2.3 (Block Recursive Equilibrium). *A block recursive equilibrium is a recursive equilibrium such that the value and policy functions depend on the aggregate state only through aggregate productivity, $a \in \mathcal{A}$ and not through the distribution of agents across states $\mu \in \mathcal{M}$.*

We provide a proof for the existence of a BRE equilibrium in **Appendix F**.

3 Discussion

What does the model imply for the distribution of vacancies types and worker-firm matches along the business cycle? This section will discuss some closed-form results that will pave the way for the quantitative analysis. We refer the reader to Appendix A for the proofs to all propositions in this Section.

The objective of our model of dynamic sorting is to understand the properties of firm creation and worker search in a setting with two-sided heterogeneity. The following properties guarantee a high degree of tractability.

Property 3.1 (Unique Injective Mapping). *Upon matching, firm quality y and utility promises in vacancy postings v are related by an injective mapping conditional on the aggregate state of the economy, Ω, and workers characteristics (χ, V).*

The previous proposition establishes that workers' directed search towards promised values is equivalent to directed search towards firms' types. We can then focus on the properties of the search strategy to get a complete view of how sorting works in equilibrium.

Property 3.2 (Search Monotonicity and Uniqueness). *The optimal search strategy when unemployed, conditional on age τ and the aggregate state Ω, is unique and weakly increasing in workers' characteristics (h, ι). The optimal search strategy when employed, conditional on age τ and the aggregate state Ω, is unique and weakly increasing in workers' characteristics (h, ι) and current level of lifetime utility V.*

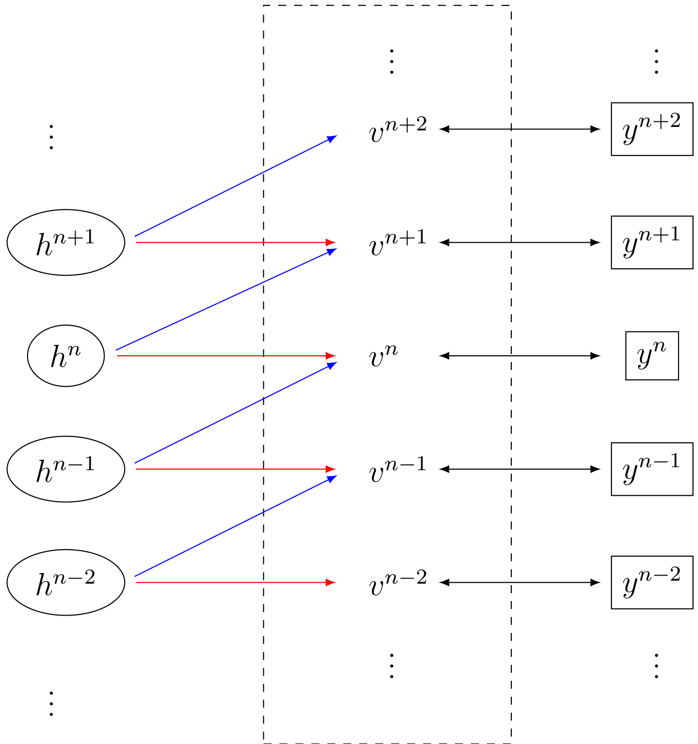
Property 3.2 guarantees that, abstracting from idiosyncratic as well as aggregate shocks, workers sort positively with respect to their education and human capital. **Property 3.1**, in turn, guarantees that workers agree on firms relative ranking. Firms are thus vertically differentiated, and there is a separating equilibrium whereby workers with different characteristics optimally search in distinct firms.

Because we are interested in how aggregate fluctuations in shaping the distribution of matches, we now turn to changes in search strategies across aggregate states.

Property 3.3 (Search in Good and Bad Times)**.** *The optimal search strategy is increasing in the aggregate productivity level, a.*

At this point we are able to illustrate one of the main mechanisms of the model, which is represented in **Figure 3**. The figure highlights one way in which aggregate fluctuations modify sorting in the labor market. The value of vacancies posted by each firm in equilibrium changes with the business cycle, as sub-markets becoming less tight in bad times. Faced with a lower probability of successfully matching with the firm they would aim to in good times, risk averse workers will then adjust their search downwards. In turn, firms will adjust downwards their utility offers given the lower expected values of matches across the board.

Figure 3. Search behavior.



Note: Schematic representation of labor market sorting along the business cycle. Unemployed workers, ordered by human capital levels, search in bad times and good times towards values offered by the (unique) corresponding firm type, presented as an ordered list with respect to order n.

In absence of separations and human capital accumulation, the cross-sectional distribution of matches will directly reflect the history of aggregate shocks. A first element that complicates this relationship is search on the job.

Property 3.4 (Optimal Retention)**.** *Retention probabilities, $\widetilde{p}(h, \tau, \iota, W; \Omega)$ are:*

(i) increasing in the value of promised utilities, W.

(ii) decreasing in aggregate productivity, a

Despite continuation values within the match being pro-cyclical and workers searching more ambitiously in good times, firms are more likely to see workers leave in expansions. This is consistent with the data, as employment-to-employment transitions are strongly pro-cyclical. The first part of **Property 3.4** highlights an important aspect of the incentives that shape the optimal contract designed by firms: retention is increasing in continuation values W.

To close the model, what we need is in fact a rule for surplus sharing between firms and workers. The way in which firms deliver lifetime utility promises depends on the design of the wage protocol.

Property 3.5 (Wage Protocol)**.** *The optimal contract delivers a wage that satisfies:*

$$\frac{\partial \log \widetilde{p}(\chi', W_i'; \Omega')}{\partial W_i'} J(\chi', W_i', y; \Omega') = \frac{1}{u'(w_i')} - \frac{1}{u'(w_i)} \tag{14}$$

with $\chi' \equiv (\phi(h, y, \iota, \psi), \tau + 1, \iota)$ *being the definition of individual characteristics and* w_i' *being the wage paid in the future state, conditional on realizations of idiosyncratic risk* ψ *and aggregate risk* a'*.*

This result extends the wage equation in Balke and Lamadon (2022) to an environment with two-sided heterogeneity. Wage growth is proportional to two elements: the residual continuation value of the match, J and the semi-elasticity of the worker's retention probability to future value promised. Limited liability provides both the foundation of wage rigidity, as it ensures that both elements in equation 14 are weakly positive, and the rationale for inefficient separations.[13]

Property 3.6 (Countercyclical Separations)**.** *Conditional on the existing contract, and on worker and firm types, there exists an aggregate state* a^* *below which firms will not continue the contract. The threshold* a^* *is cæteris paribus decreasing in the value promised to workers, and increasing in worker and firm types.*

A clear implication of **Property 3.6** is that, especially at the onset of recessions, firms will be significantly more likely to lay off workers. In addition, more fragile workers and firms will be more likely to separate in recessions. The counter-cyclicality of separations is a common feature in labor market data, together with the lower job security enjoyed by less productive workers, or provided by less productive firms.

[13] Notice that, in presence of risky human capital accumulation, J will fluctuate together with the human capital levels of the worker even in absence of aggregate fluctuations. However, because the contract provides insurance to workers, changes in their human capital will have asymmetric effects on wage growth.

4 Bringing the Model to the Data

The model features internally and externally calibrated parameters. To estimate the first group of parameters, we target moments from Italian administrative data, provided by the Uniemens dataset by the Italian Social Security Administration (INPS), for all years between 1996 and 2018.[14] To obtain model moments, we simulate a population of overlapping generations working for 45 years (180 quarters, from 18 to 63 years old, the legal retirement age for most years in our period of analysis). We then use a Simulated Method of Moments approach. This section will first present the quantitative setup of the model, then present calibration choices for the parameters that are set externally, and finally discuss our estimation routine.

4.1 Calibration and estimation

Quantitative Setup. All the functional form choices are collected in **Table 1**. We assume a Cobb-Douglass production function, and allow for potentially cyclical maintenance costs, captured by the parameter x. As time is discrete we have to pick a matching function bounded between zero and one. This rules out Cobb-Douglas functions and therefore we follow Schaal (2017) and Menzio and Shi (2010) in picking a CES function in market tightness. Vacancy creation imposes linearly increasing costs, proportional to firm's quality y, governed by the parameter κ. Workers are risk-averse with CRRA utility. The human capital production technology is concave in the (adjusted) firm quality, ξy, and in the existing stock of human capital, h. Future human capital is also subject to additive i.i.d. shocks, $\psi \sim \mathcal{N}(0, \sigma_\psi)$.[15] Home production is increasing in the stock of human capital, according to the parameter ξ_b. Finally, we allow the exogenous separation rate to be age dependent to capture age-specific aspects of worker quality that are unrelated to business cycles but still empirically relevant. The model is characterized by 8 externally calibrated parameters and by 18 jointly estimated parameters. We discuss each in the next two paragraphs.

Calibration. Preference parameters (discount factor β, and agents' risk aversion ν), and the annualized risk-free rate r_f are set in line with the literature. We calibrate the persistence and volatility of the aggregate shock, (ρ_a, σ_a) by estimating an AR(1) on the detrended series of Italian real TFP from the Penn World Tables 10.0. In addition, we need to pick values for the distribution of initial human capital draws. Workers draw their innate ability and human capital upon entry into the market from an initial distribution

[14] Details of data construction and sources are discussed in Appendix

[15] Notice that, by construction of the human capital accumulation process, there will always be a tuple $\{h, y\}$ such that $g(h, y) = h$ - that is to say, there will be a point in time after which the worker will have acquired all the possible human capital gains that position has to offer.

Table 1. Functional Forms

Functions	
Production function	$f(y,h) = Ay^{\alpha}h^{1-\alpha} - x(A-1)$
Job finding probability	$p(\theta) = \theta(1+\theta^{\gamma})^{-\frac{1}{\gamma}}$
Vacancy creation cost	$c(y) = \kappa y$
Utility function	$U(c) = \frac{c^{1-\nu}}{1-\nu}$
Human capital accumulation	$g_{\iota}(h,y) = (\xi y)^{\phi_{\iota}}h^{1-\phi_{\iota}} + \psi$
Home production	$b(h,\tau) = b + \xi_b h$
Exogenous exit rate	$\lambda = \frac{\lambda_b}{\lfloor \tau/4 \rfloor}$

and then exit the market once they reach their retirement age or when the expected value of employment is too low.[16] The initial distribution of human capital for high-school-educated workers is $Beta(\mu_L, \phi_L)$. College-educated workers draw their initial human capital from the same distribution plus a constant spread, ϑ. We estimate the shape and scale of the beta distribution, plus the scaling factor to match the ratio of average initial incomes between the two groups of workers, and the two standard deviations. Finally, in order to properly account for the age distribution of the Italian population, we weighed simulated data according to the age distribution of the working-age population for the year 2010.[17]

[16]Workers either deterministically exit the market (die) at the average Italian retirement age in the data or choose to exit the labor force once the job finding probability is below a constant threshold $\underline{p}$.

[17]Age weights are constructed following the age distribution of the 2010 census from the website of the Italian National Institute of Statistics (ISTAT).

Table 2. Target Moments

Moments	Mean	
	Data	Model
A. Labor Market Flows		
Employment-to-Employment Transition Rate*	1.3%	0.7%
Employment-to-Unemployment Transition Rate*	4.1%	2.9%
Employment-to-Unemployment Correlation w/GDP	0.68	0.38
Employment-to-Unemployment Correlation w/GDP	-0.19	-0.16
B. Earnings		
Earnings Growth (High-School)*	129.7%	134.7%
Earnings Growth (College)*	60.7%	77.1%
Entry Salary Ratio: College to High School	1.46	1.00
C. Other Statistics		
Unemployment Rate	9.7%	9.1%
Labor Market Sorting	0.40	0.55
Inactivity Rate†	21.9%	28.3%
Firm Value Added Distribution**	1.18	1.38

Note: (*): We match the life cycle profiles, with seven age bins for each profile. Reported in the table are average values. (**): We match the ratio between the third and quarter quintiles to the first. (†) The average inactivity rate is obtained from the inactivity rates by age groups from 1996 to 2019 from ISTAT.

Estimation. The routine for jointly estimating the remaining 18 parameters aims to fit a set of standard labor market moments: labor market flows by age, as well as their correlations with aggregate output; the profile of wage growth over workers' careers; the average unemployment rate; the average inactivity rate in the Italian labor market; the average degree of sorting between workers and firms; and the distribution of firms' value added. We define sorting as the average over time of the correlation between the firm and worker fixed effects from an AKM model yearly estimated on the Italian administrative data. In the model, sorting is the correlation between firms' and workers' qualities. The comparison of model and empirical moments are summarized in **Table 2**.

The model fits employment flows by age, capturing labor market dynamism in the data (see Engbom (2020)). We match the cyclical properties of these flows, to account for the jump in job destruction and the drop of job creation in recessions. Characteristics of labor market fluidity and dynamism are further disciplined by matching both the unemployment and the inactivity rate. The model also reproduces the life-cycle wage growth of Italian workers by education levels. Finally, to account for the importance of firm-worker cross-sectional heterogeneity, we target the firm productivity distribution, using quintiles of establishment-level value added per employee in the data and firm

Table 3. Parameter Values

Parameter	Description	Value
	Externally Calibrated	
ν	Risk aversion	2.000
β	Discounting	0.990
r_f	Real interest rate	0.011
$(\mu, \sigma_L, \sigma_H)$	Shape and scale of initial human capital dist.	(2.50,10.00, 0.35)
(ρ_A, σ_A)	Mean and std of TFP process	(0.95, 0.009)
	Jointly Estimated	
α	Production function elasticity to firm quality	0.552
γ	Matching function	1.090
ϕ	Human capital adjustment rate, High School	0.037
ϕ_g	Human capital adjustment rate, College	0.282
b	Unemployment benefit	1.103
λ_b	Exogenous separation prob., initial	0.116
κ	Vacancy cost	2.432
λ_e	On-the-job-search prob.	0.454
ξ	Scaling factor in human capital accumulation	0.640
ξ_b	UB dependence on human capital	0.063
l	Linear loss of humanc capital while unemployed	0.167
τ_{ee}	Human capital retention after EE	0.902
τ_{eu}	Human capital loss after EU	0.771
x	Cyclical component of cost function	-1.647
$\underline{\mathrm{p}}$	Out of labor force threshold	0.037
σ_ψ	Std of idiosyncratic human capital shock	0.684
ϑ	Initial scaling in human capital distribution	0.350
$\underline{\mathrm{y}}$	Lowest bound of firm distribution	2.844

type in the model, and labor market sorting. In order to globally minimize the distance between model-generated and data-generated moments, we adopt two complementary strategies. The first is to minimize the distance between model-generated and data-generated moments using a global solver on the largest feasible domain; for this, we rely on the particle swarm algorithm contained in the library developed by (Blank and Deb, 2020). The second is to solve and simulate the model over a sparse grid of the parameter space.[18] In both cases, we use the combination of parameters that delivers the smallest error as the starting point of a local minimization routine (for this, we rely on the Nelder-Mead algorithm). **Table 3** reports the estimation results.

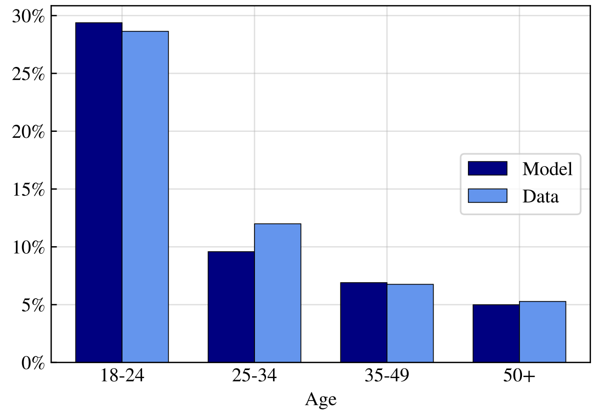
4.2 Fit and untargeted moments

Cross-sectional properties. Despite the model not matching exactly the labor market flows for young workers, we find a very good fit for the age profile for the unemployment rate (untargeted, **Figure 4**). We take these results as evidence of the good performance of the model in matching the cross-sectional characteristics of the labor market.

In order to match the effective retirement age over the life cycle and characterize

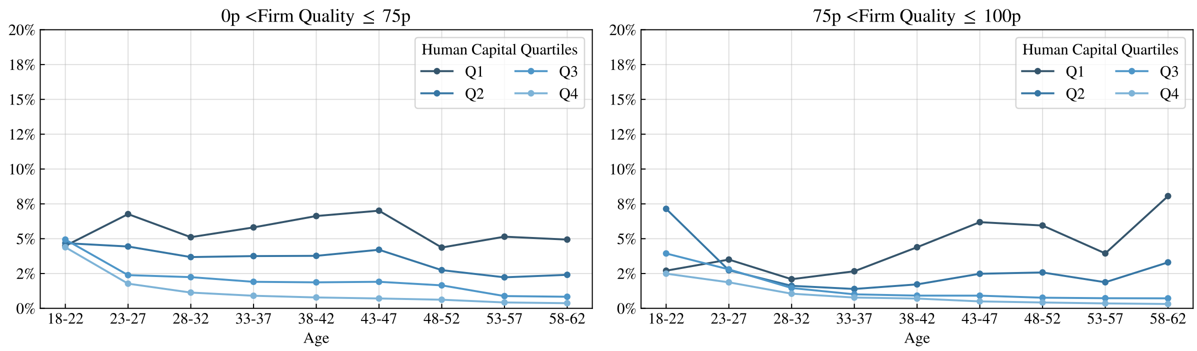
[18]The advantage of selecting the parameters with this procedure is that the grid is built so that the function domain is optimally covered with the least amount of possible points compared to other forms of approximation (e.g. equi-spaced or random grids). The sparse grid is built using the"Tasmanian" libraries (Stoyanov, 2015).

Figure 4. Unemployment rates, model and data



Note: The figure reports the unemployment rate by age groups in model simulations and in the data. *Sources:* unemployment rates are taken from the Italian National Statistical Agency (ISTAT).

Figure 5. Separations rates by age, firm and worker qualities



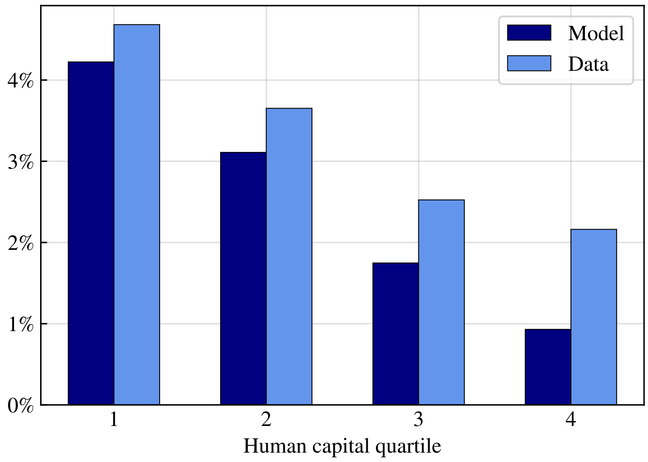
Note: The figure plots the average separation rates by firm quality, human capital quartiles and age in model simulations.

a realistic workers' age distribution, we include the average inactivity rate among our targets. Our model matches this average rate really well. Moreover, the model predicts, as in the data, a U-shaped relationship between participation and age: participation increases as workers complete their education and then starts to decrease steeply after age 45.

We know young workers are more likely to be unemployed, but is it because they have more fragile jobs, or because they are relatively less valuable workers? In **Figure 5** we report the separation rates by firm quality and human capital quartiles. Each panel reports the proportion of matches that are dissolved, on average, by age and human capital in each firm class.[19] The patterns that emerge are informative of the leveraging dynamics at work. In particular, the figure shows how separations are prevalent for old, low-skilled workers in relatively good firms. These matches are

[19]Under this tentative calibration the firm distribution is highly positively skewed so that the median is actually the lower bound of available firm qualities.

Figure 6. Separations rates by worker qualities, model and data



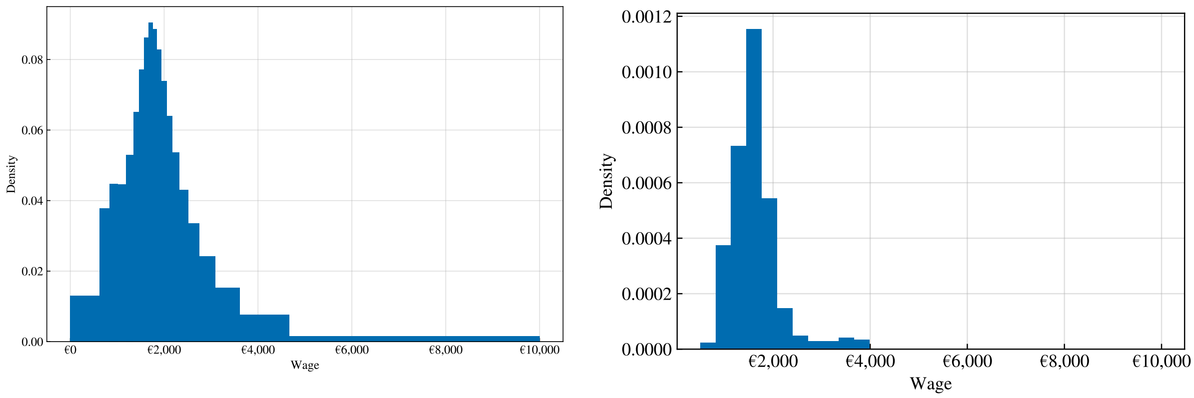
Note: The figure reports the average separation rates by worker quality. In the data worker quality is measured by the worker-specific AKM fixed effect. In the model simulations, worker quality is workers' human capital.

particularly susceptible to recessions as older workers command higher wages thanks to their longer labor market experience, making firms more susceptible to aggregate shocks. This mechanism is stronger for more productive firms as they are also providing a steeper accumulation of human capital and consequently a steeper wage profile. To reconcile these features with the separation rates being substantially higher for younger workers, notice that separation rates are markedly higher for the lowest human capital quartile, and change relatively little across the firm distribution. Hence, we know that the relatively higher fragility of young workers is mostly due to the small endowment of human capital, and changes in the cross-sectional distribution of human capital could increase the overall fragility of the economy to fluctuations, as we will discuss later. **Figure 6** compares the separation rates by worker types in the model and in the data, highlighting how the model is able to capture the higher fragility of low human capital workers very well, despite not targeting directly these moments in the estimation.

An important validation of our model's ability to account for the interactions between cross-sectional worker distribution and business cycle is for the model to be able to reproduce the distribution of earnings at least in a qualitative sense. **Figure 7** displays the cross sectional distribution of earnings in the data and in the model. The model captures the main features of the empirical wage distribution: it is centered at slightly below €2,000, skewed to the left, with most observations below €4,000. What the model fails to generate is the long right tail of wages in the data, which correspond mainly to managerial figures whose earnings command premia that our mechanism struggles to capture.

Figure 7. Wage distributions

(a) Data **(b)** Model



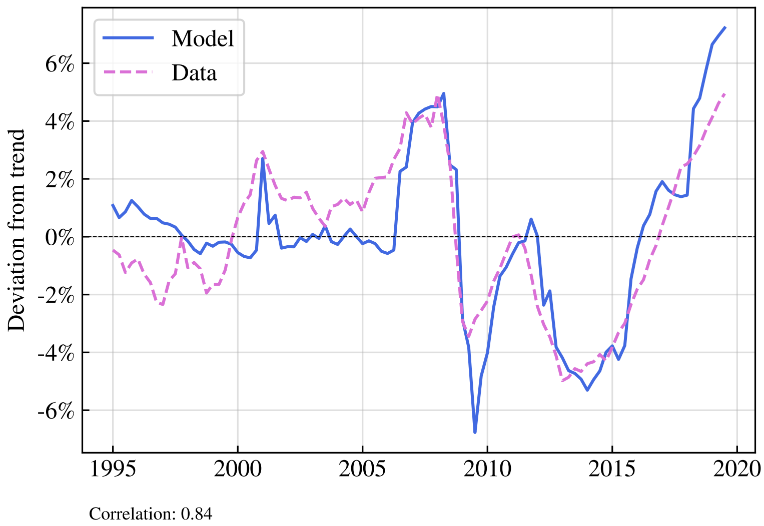
Note: The figure plots the wage distributions in the data and in model simulations.

Time Series Properties. Because structural estimates match micro-level and cross-sectional moments, replicating aggregate time series properties of the data provides a useful validation of the channels in the model. We use the detrended quarterly time series of Italian Total Factor Productivity, and project it on a discrete grid to simulate a series of aggregate shocks in the model. **Figure 8** plots the resulting simulated aggregate production, together with the detrended GDP process from the data. The model replicates not only the amplification channels from productivity shocks to GDP, but also the asymmetry of contractionary shocks being deeper than expansionary ones.

Heathcote, Perri and Violante (2020) show that recessions impact inequality with persistence, affecting the earnings of workers in the left tail of the income distribution the most. **Figure 9** illustrates the dynamics of inequality around business cycles by displaying the pattern of losses across the earnings distribution. Consistently with recent evidence, recessions hit disadvantaged workers the hardest. Another prediction of the model is that the persistence of earnings losses vary across the distribution: while workers with less human capital display more volatility in earnings (mostly due to displacement, see **Figure 5**), the impact on workers with high human capital is dampened but quite persistent.

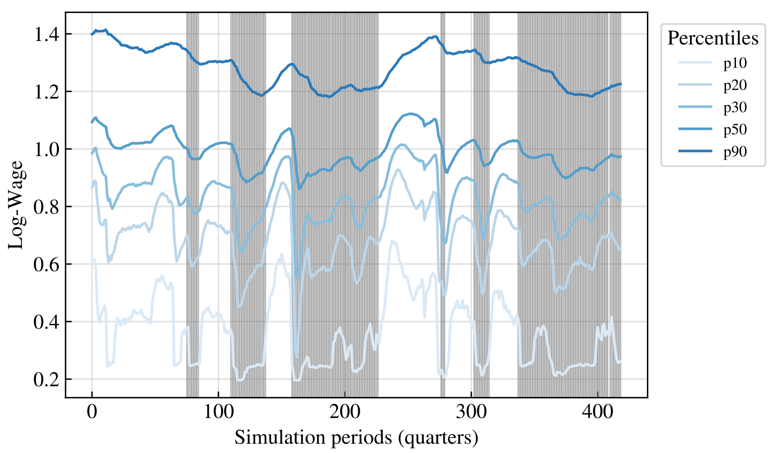
Scarring. The model is also able to replicate the long-run effects of business cycles on workers career outcomes at the micro level. In particular, we adapt the reduced-form models proposed in the literature about the scarring effects of recessions on labor market entrants (e.g. Kahn (2010), Schwandt and von Wachter (2019)) and we run it on both

Figure 8. GDP: Model and Data



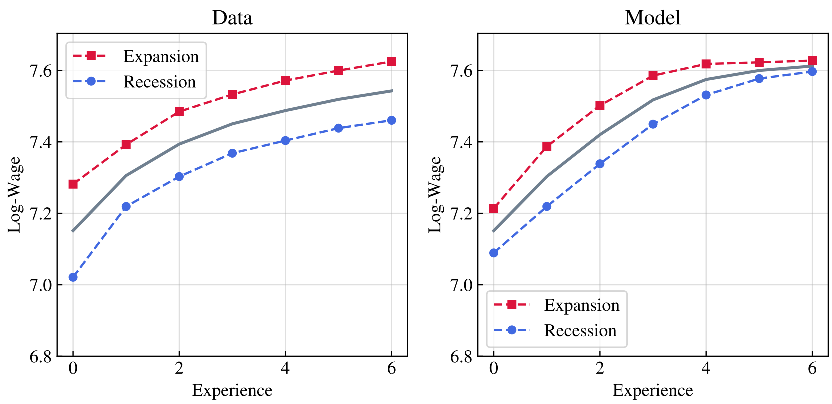
Note: The figure plots the cyclical components of real GDP for Italy and for a model simulation in which the TFP process is matched to the Italian TFP realizations from 1995 to 2019.

Figure 9. Inequality dynamics over the cycle



Note: The figure plots the time series for the different percentiles of the income distribution in model simulations.

Figure 10. Scarring effect of recessions



Note: The figure plots the wage profiles estimated on the data and on model simulations for cohorts of workers entering the labor market. The counterfactual profiles for Expansions (Recessions) are obtained considering a positive (negative) two-standard-deviation realization of cyclical GDP.

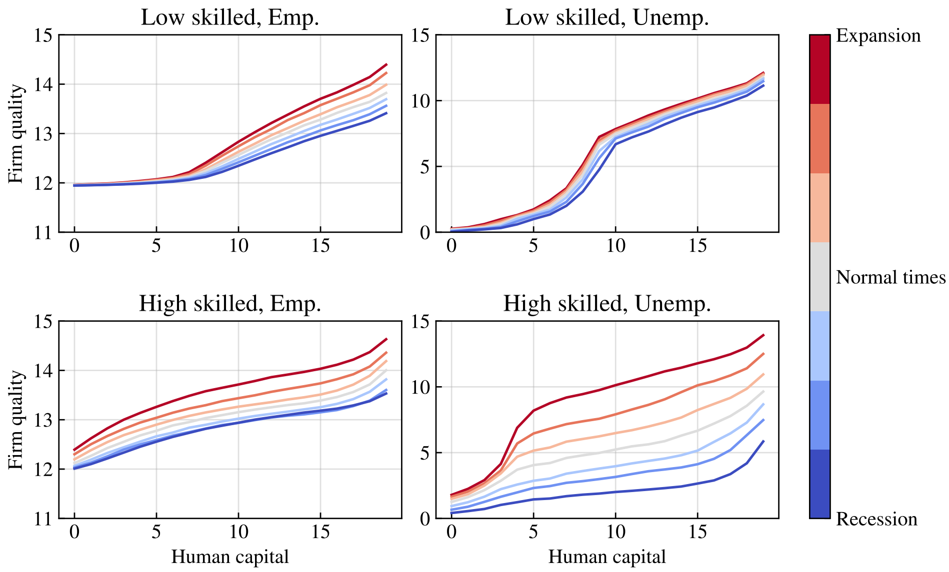
the Italian administrative data and on a model simulated panel.[20] Consistently with the literature, entering the labor market in a downturn is associated with persistent losses in earnings. As shown in **Figure 10**, our model is able to replicate well both the magnitude and the dynamics of these *scarring* effects of business cycles.

Sorting. A central prediction of the model is that workers' search is monotonic in individual characteristics and in the aggregate state (see **Proposition 3.2**). We plot the equilibrium mapping between workers' human capital and their search behavior for different points in the aggregate state. Search is strongly monotonic in both dimensions. The human capital / search mapping is stronger in expansions for very low and very high levels of human capital, where different dimensions are at play: for the first group, expansions open up the possibility of moving towards jobs that imply a growth in human capital - expected earnings don't change much in the next job. On the other hand, improved search opportunities for better workers in expansions are reflected predominantly in higher potential wage gains.

Exploiting again the rich features of the model, in **Figure 12**, we plot a decomposition of the average growth in wages *within* jobs, i.e. within the same contract, and *between* jobs, i.e. after a job-to-job transition. On one hand, the model simulation implies that the bulk of wage growth is due to job-to-job transitions, replicating a well known feature of labor market flows. On the other hand, the average *within* wage growth is declining in age and firm-quality. This is because high quality firms are matched on average with better workers that are paid more on average. In addition, given that the search

[20]These empirical specifications rely on specifications that control for age, period and cohort effects. Following the literature, we address the well-known identification issues in this class of models by proxying cohort fixed effects with the cyclical realization of GDP. We report the empirical estimates in **Table F.5**.

Figure 11. Search behavior



Note: The figure plots the search policy function for each human capital level in each aggregate state, averaged across labor market experience and wage promises.

strategies of employed workers is targeted towards better firms, firms in the right tail of the distribution can offer flatter (i.e. less backloaded) wage contracts as they benefit from a higher retention probability.

The importance of sorting for human capital accumulation and workers' careers is validated also by measuring the correlation of workers *ex-post* wages with their *ex-ante* employer quality after an Employment to Unemployment to Employment transitions.[21] The correlation is increasing in previous firm qualities indicating how workers benefit from employment spells in good firms also once the match is dissolved. **Table F.6** reports the estimated correlations.

5 Aggregate Consequences of Contractions

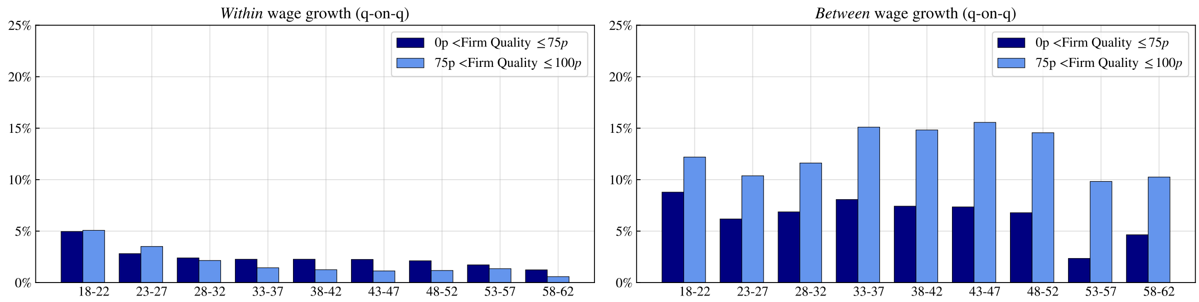
In this section we describe how the mechanisms at the core of our model interact with business cycles, and in particular with recessionary shocks.

5.1 Anatomy of recessions

In order to compute an impulse response function, we compare a simulation of the model without aggregate shocks and a simulation in which we hit the economy with three

[21] This is an adaptation of Herkenhoff et al. (2018) analysis to our model setting. In their paper, they rely on co-workers wages a proxy for firm quality, given the nature of our framework we use AKM fixed effects in the data to measure firm quality and we control for workers pre-transition wage.

Figure 12. Within vs between wage growth by age groups and firm quality in the model



Note: The figure plots the average wage growth, by age and firm quality, *within* employment spell and after EE transitions (*between*). For the *betweeen* component, the firm quality quartiles are computed on the distribution of origin firms.

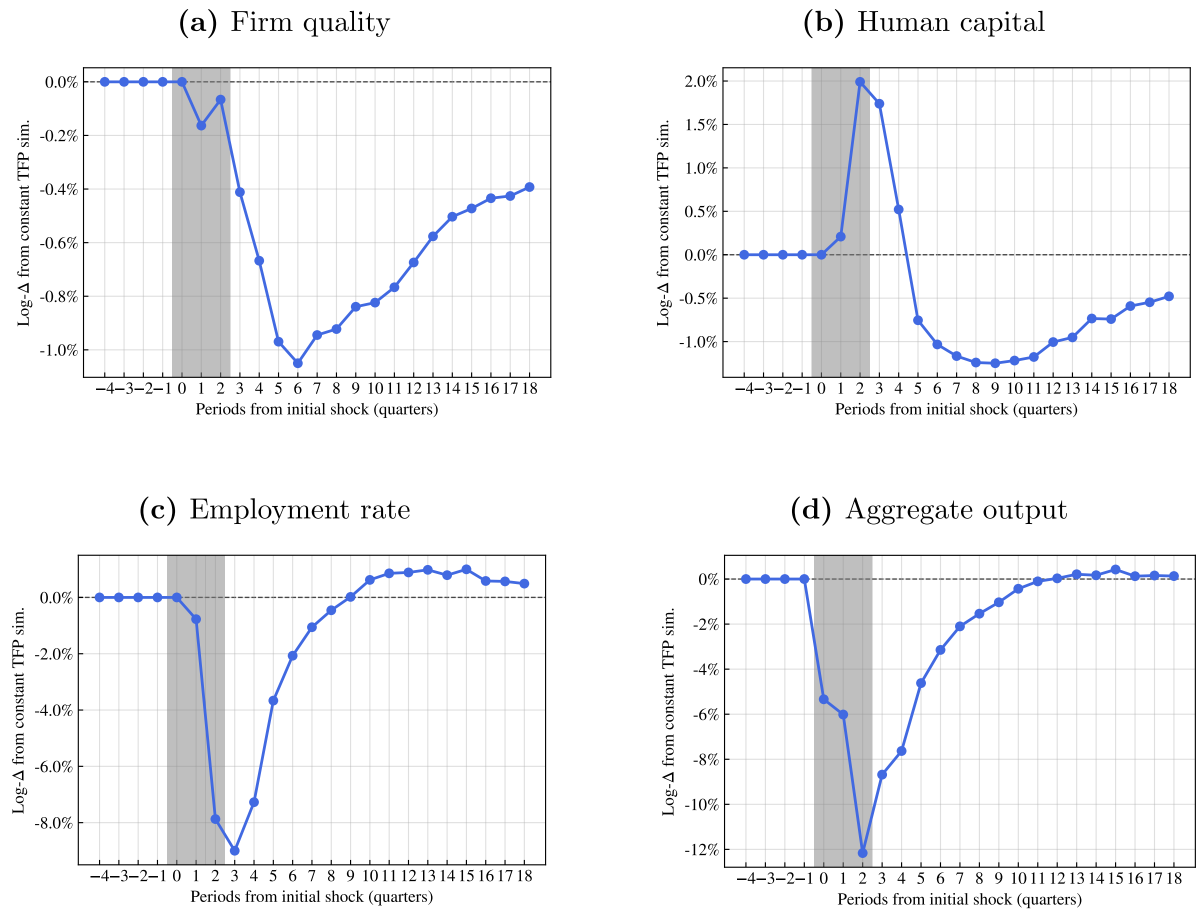
consecutive negative realization of the TFP process.[22] We then look at both labor market outcomes of affected cohorts and the response of aggregate GDP to the recession. The main results are reported in **Figure 13**.

The dynamics of firm quality and average human capital, respectively in **Figure 13a** and **Figure 13b**, offer a clearer picture of the transmission of aggregate shocks in the economy. While the onset of the recession is accompanied by a sort of "Schumpeterian" response, as implied by the initial marginal increase in human capital shown in **Figure 13b**, firm quality is persistently crippled by the recession, with the average quality of firms active in the economy remaining approximately 0.5% below the no-recession economy even two years after the end of the recession. The average human capital in the long run settles to a similar lower level despite the initial increase. This prolonged reduction in the quality of the factors of production increases the persistence of the initial shock on output beyond the original duration of the recession. **Figure 13d**, in fact, shows that even after two years aggregate output its still approximately 1.5% below its counterfactual level.

Decomposing Recessions. What explains the amplification and persistence of recessionary shocks? Different competing channels are at play. The first, which we will call the *human capital channel*, captures the human capital accumulation that does not take place because of the recessionary event. The second one, which we will call the *displacement channel*, captures the job destruction that takes place because of the negative shock and its spillovers. The last one, the *sorting channel*, amounts to the different joint worker-firm distribution that emerges in the periods following the shock, both because of different search strategies and because of additional unemployment spells. We decompose the amplification channels of the model economy by shutting down each channel one at a time, and then comparing the resulting dynamics to the one

[22] The cumulative drop in TFP is approximately 15% (equally split in each quarter) .

Figure 13. Recession experiment

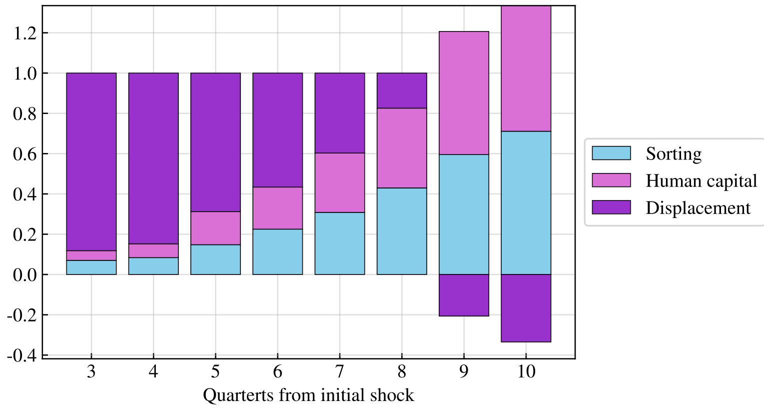


Note: The top panels in the figure plot: (**a**) the behaviour of average firm quality and (**b**) average human capital in an economy with no aggregate shocks, that serves as benchmark, and an economy in which we impose a three-quarter recessions. The bottom panels report: (**c**) the employment rate; (**d**) the aggregate effect of the recession on real GDP.

of a baseline recession. Our decomposition shows that the existing human capital distribution has important implication for the recession dynamics: as human capital increases, recessions become less severe on impact. At the same time, the allocative efficiency worsens, slowing down the speed of recoveries.

Figure 14 shows the evolution of the three channels following a negative shock. Most of the recession intensity depends on the displacement channel. Recovery from displacement, while not immediate because search is frictional, is relatively fast. We find the sorting channel to be the main contributor to the persistence of recessionary events. Interestingly, the two channels play a role in different parts of the workers' cross-sectional distribution; while the displacement channel depends mostly on the size of its left tail, the sorting channel affects predominantly the right tail, for which allocation efficiency matters more than employment *per se*. On either side, the human capital channel is second order for recessionary dynamics. This happens because the main margin of adjustment for the workers with less human capital is mostly on the employment

Figure 14. Decomposing recessions

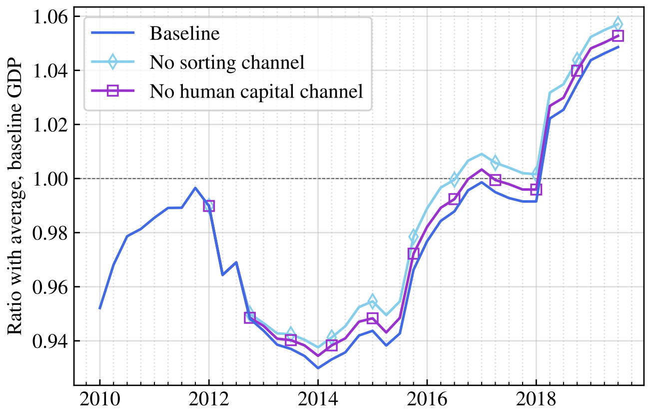


Note: The figure shows the relative importance of each transmission channel compared to the baseline recession for the GDP IRF.

dimension, but their search hardly targets firms that are good enough to provide them with a significant improvement in human capital. For workers with very high human capital, instead, moving to better firms produces more output but does not necessarily increase human capital accumulation, since the human capital level is already high and all firms have concave human capital production function technologies. This has interesting implications for the fragility of the economy: as the human capital distribution shifts to the right, recessions become less severe on impact.

Decomposing the 2011-2013 recessions in Italy. To quantify the importance of each channel in the data, we construct an experiment with the Italian business cycle. We use realizations of our shock process matching the evolution of the Italian TFP, as in **Figure 8**, and we run three counterfactual simulations: i) one in which the 2011-2013 recession does not happen, ii) one in which the recession does happen but the sorting channel during the recession is muted and finally, iii) a simulation in which the human capital accumulation channel during the recession is muted. Therefore, the dynamics of ii) and iii) represent the dynamic of the Italian economy if the recession had no impact on workers' search behavior and human capital accumulation. **Figure 15** shows the baseline evolution of the Italian economy, as well as the three counterfactuals simulations. When the sorting channel is muted, the economy recovers faster, but most importantly enters the following expansion with a higher ability to capitalize on growth: we see that that alone would account for almost a 1.5% difference in the level of GDP as of 2019. Silencing the human capital loss induced by the recession has a similar, but quantitatively smaller effect. Compounding the two, they amount for more than a 2% difference from baseline GDP levels. Even assuming the two channels to have no interactions, policies that address the losses in human capital and in the quality of sorting would pay benefits in the long run.

Figure 15. Decomposing the Italian 2011 - 2013 recession



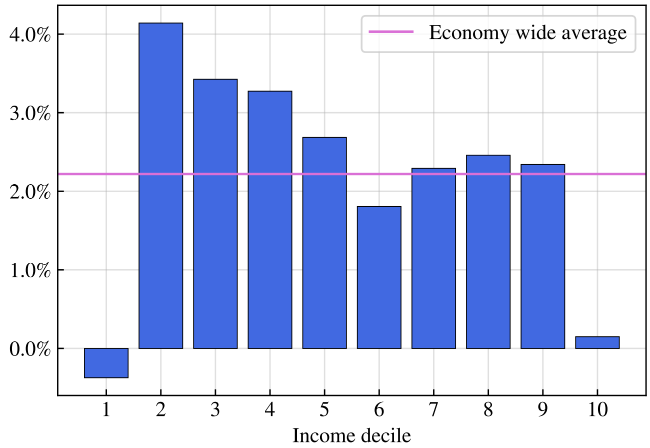
Note: The figure plots the dynamic of the Italian economy with different assumptions on the active channels in the transmission of aggregate shocks to aggregate GDP.

5.2 The importance of inequality for business cycles

Costs of Business Cycles. Influential work by Lucas (1987) argued the welfare gains from reducing business cycle volatility to be negligible, quantifying them in less than 0.1 percentage points of consumption-equivalent units. Barlevy (2004), noting that recessions can have sullying effects, estimates welfare effects to be about two orders of magnitude greater than Lucas' original estimates. We estimate the cost of business cycles closer to Barlevy's on average, at more than two percent of consumption. However, the rich heterogeneity in the model allows us to look at the costs of fluctuations across the income distribution. **Figure 16** plots the cost of business cycles for income deciles. The first decile being comprised almost entirely of unemployed workers, it comes to little surprise that costs of fluctuations are actually rather small for them. Less intuitive is the U-shaped cost across the income distribution for employed workers. Up until the seventh decile, the costs of aggregate fluctuations decrease with income: lower incomes are associated to more fragile jobs, and hence to higher likelihood of unemployment in recessions. For the last three deciles, costs of business cycles are mostly due to long-run impacts on careers, and on the ability to move to higher paying jobs. While fluctuations in utility are smaller during recessions, the impact on the sorting of high-human capital workers ends up being so persistent that its effects amount to a higher percentage of lifetime consumption than for workers on the left tail of the human capital distribution.

Can Double Dips be More Painful? Our model economy exhibits significant state dependence, and can thus speak to the different impact of shocks when the economy is

Figure 16. Heterogeneity and the cost of business cycles

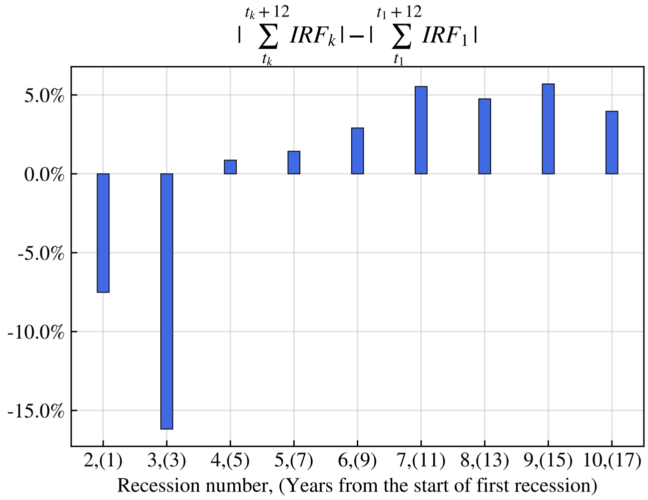


Note: Reduction in consumption-equivalent utility due to aggregate fluctuations

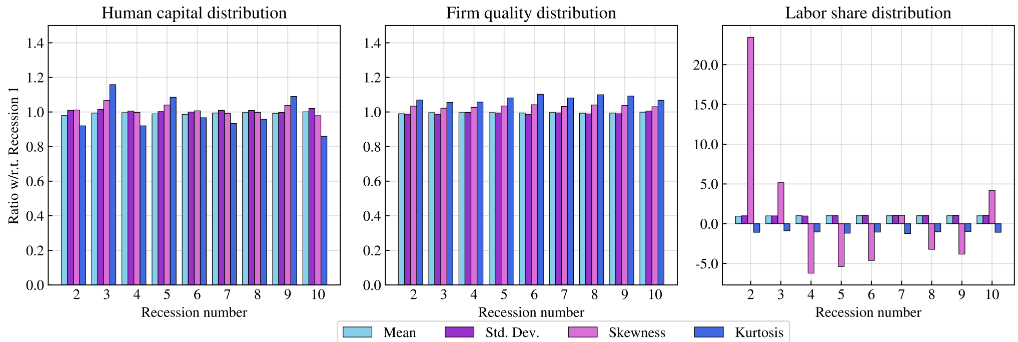
in a better or worse state. Thus, we can contribute to the debate on whether the depth of recessions is increasing in either built-in fragility from preceding long expansions, or from the misallocating effects of a previous recession. To understand this, we perform a "double recession" experiment. The experiment works as follows: at time 0, we hit the economy with a recession identical to the one discussed in **Section 5.1**, then we track the evolution of human capital, firm quality, and labor share by collecting the first four moments of their distributions, absent other aggregate shocks. The three bottom panels in **Figure 17** report their values for nine different points in time following the first recession, starting one year after the return of aggregate shock to trend and then every two years for the following eight. Starting from the cross-sectional distribution computed this way, we hit the economy again with an identical recession, separately at each time horizon specified above. The top panel in **Figure 17** reports the difference of cumulated output reponses between the baseline and the "second" recession. The first two recessions, that happen within five years from the first, have smaller cumulated losses. Fragility builds up over time, and recessions become more severe, peaking at the 11-year interval between the two shocks. This rather surprising result can be interpreted in light of the first decomposition of recessions, that shows the initial severity to depend mostly on the impact of displacements: when the economy is recovering from a previous shocks, even if unemployment has been re-absorbed, the labor share is unusually skewed to the left. This means that more workers have low wages compared to their baseline status, making the leverage problem in their match less severe. Similarly, when double dips hurt the most (for recessions number 7, 8 and 9) two things emerge: the human capital distribution tails are less flat, but the labor share tails are fatter. This imbalance results in more separations and a more severe recession. While more research is definitely needed on this specific issue, we believe this result points to the usefulness of tracing

Figure 17. State dependence and business cycles

(a) Cumulative losses



(b) Match-related distributions



Note: The figure plots a comparison between a series of simulations with two recessionary shocks of the same magnitude, one successive the other but at different horizons. The top panel shows the cumulative difference between the two recessions. The bottom panels report the ratios of the first four moments of selected variables in the quarter before each recession hit the economy. For skewness and kurtosis the denominator is the absolute value of the first recession moment.

the evolution of labor markets micro dynamics to highlight potential fragilities of the economy.

5.3 What can account for the increased length of recessions?

The time economies take in order to recover from recessions has not always been constant. **Figure 18** shows the average number of quarters aggregate GDP has been below trend during recessions for a subset of advanced economies. Notably, from the mid-80s and early 90s there has been a widespread increase in this measure.[23] This is consistent with

[23]Specifically, in the data, we define a recession as occurring after two consecutive quarters of negative GDP growth, and within each recession, we count the number of quarters in which GDP realizations are below trend. We obtain a similar picture if we look at alternative definitions of recession lengths, such as the number of quarters that are needed to reach pre-recession GDP levels and the number of consecutive

evidence presented by Fukui, Nakamura and Steinsson (2023) on the slower recovery of employment after recessions. Among other factors, the increase in job polarization and the rise in the skill premium are phenomena contemporaneous in time with this rise in the time economies need to recover from recessions (Goldin and Katz, 2007, Goos, Manning and Salomons, 2009). Our model provides a useful structure to check whether human capital accumulation and the sorting dynamics in the labor market can explain, at least in part, these aggregate developments.

In order to calculate the duration of recessions and the subsequent recoveries in our baseline economy and a counterfactual one where there are no differences between graduated and non-graduates, we consider two simulations of the model, one *with* and another one *without* skill premium. In particular, for the economy *without* skill premium we set the parameter governing the importance of firm quality in the human capital accumulation function equal across agents. This implies that, in this economy, the average earnings for graduates are only 3% higher than non-graduates. In our baseline economy *with* skill premium, instead, graduates enjoy a faster human capital accumulation which translates to earnings approximately 30% higher for graduates than non-graduates, on average.

The presence of a high skill premium leads to recessions that are approximately 21% longer than those occurring without skill premium. In relative terms, the change in the length of recessions is remarkably close to what is observed for the subset of advanced economies in **Figure 18**. For these countries, in fact, the average duration of below-trend GDP realizations in recessions increases by approximately 29% comparing the period before and after 1990.[24]

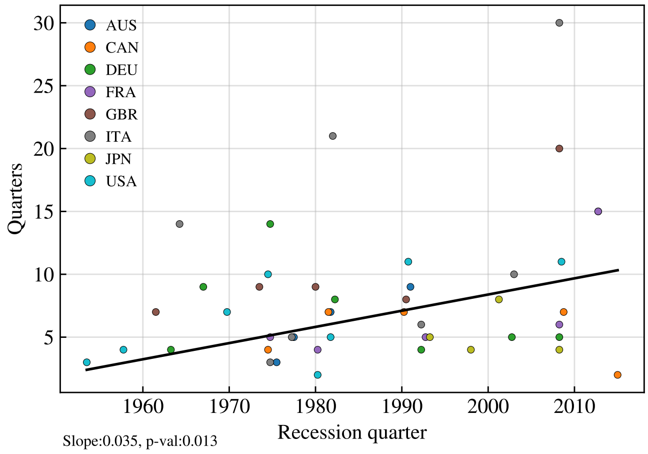
However, increase in the importance of firm quality for the accumulation of human capital is not the only possible channel. Increasing the labor elasticity in the production function by approximately 6%, similar to the change in the share of labor compensation over GDP in Italy after 1990, would not lead to any change in the length of recessions. Another big shift, common to all developed economies, is a substantial aging in population. To see how this affected recession lengths, we simulate the model without skill premium using demographic weights from the 1970 and 2011 Italian censuses, and we calculate the average recession length in these counterfactual scenarios. We find that accounting only for these demographic changes results in a 14% increase in the average recession length, about half of what observed in the data.[25]

quarters with negative GDP growth. However, these definitions cannot be transferred directly to model simulations as they rely on measures of GDP growth, which we do not explicitly include in our model. Therefore, in our simulations, we use the number of consecutive quarters GDP remains below trend as our preferred measure of recession lengths.

[24]For Italy, the increase in this measure has been slightly more pronounced. Recession lengths went from 11 to 15 quarters, approximately a 36% increase.

[25]If we were consider both channels simultaneously it would be sufficient to increase the skill premium

Figure 18. Length of recessions over time



Note: The figure reports the average duration of recessions for a set of OECD countries. Specifically, for each recession, we compute the number of quarters each economy's GDP remains below trend.

6 Policies and the Business Cycle

We showed in the previous sections that business cycles are amplified by contracting frictions, search frictions, and allocative inefficiencies. Since as a result recessions generate persistent welfare losses, there is a natural role for policy. In this section, we discuss the role of two specific policy interventions using our model environment as a laboratory.

Minimum Wage The rationale for minimum wages in our model is simple. Beyond its redistributive nature, minimum wages can induce a flatter job ladder at the beginning, potentially increasing the subsequent accumulation in human capital and making the economy less fragile to fluctuations. A potential cost of minimum wages is the *de facto* disappearance of low paid jobs that effectively act as stepping stones. The equilibrium effect is thus ambiguous, and the overall effects might depend on the level of the minimum wage. We impose a minimum wage equal to the long run median of entry wages (from unemployment), which is equivalent to approximately €880 per month.

Cyclical Transfers A more immediate way to prevent incomes from falling in recessions is to provide direct transfers to unemployed workers. This policy has been used extensively during the Covid recession, in departure from more traditional, means-tested interventions.[26] In this chapter, we link the amount of unemployment transfers to the severity of recessions. Transfers can reduce the welfare losses to

so that earnings for graduates are approximately 24% higher than non-graduates to obtain a relative increase in recession lengths that matches exactly what is observed in the data.

[26]The US government, for instance, spent \$169 million in 2020 and \$402 million in 2021 in one-time stimulus payments and \$119 million in 2020 and \$206 million in 2021 in unemployment benefits (including \$300 per week supplement until September 2021).

displaced workers, and impact their labor market strategies. We calibrate transfers to be increasing in the severity of recessions so that their correlation with aggregate productivity equals -1.5.[27]

Discussion Outcomes of our preferred calibration for each of the policy experiments are presented in **Table 4**. The minimum wage, while effective in affecting the search behavior of more fragile workers, severely impacts job creation. This is reflected in an additional 5.7% of workers leaving the labor force, and in an additional 0.6% of unemployed workers. As a result, employment falls by more than six percent. Output decreases even more. The additional fall in output levels can be attributed to the lower human capital accumulation caused by longer unemployment spells on average. These effects get in the way of minimum wages helping workers building resilience in face of output fluctuations, and the cost of business cycles as a share of consumption grows substantially. The combined impact on unemployment and labor force participation offsets gains in earnings inequality as measured by the Gini index: the estimated zero impact of the policy captures an increase in inequality due to a drop in labor earnings on the extensive margin, together with a compression of the wage distribution, conditional on workers being employed. The discussed effects are amplified for higher levels of minimum wage, mostly due to growing impact on job creation. On the other hand, reducing the level of minimum wages gradually makes it less binding, until no output effects are detectable. In no calibration, however, can output levels grow or overall inequality levels decrease using minimum wages using our estimated model.

Table 4. Impact of Policy Experiments

	Scenario		
	Baseline	Minimum Wage	Cyclical Transfers
Output Level	1.0*	0.9	1.09
Welfare	1.0*	0.89	1.07
Unemployment	9.1%	9.7%	9.5%
Employment	62.6%	56.3%	71.5%
Sorting	0.55	0.62	0.58
Gini Index	0.15	0.15	0.16
Search	1.0*	1.05	1.02
Cost of Business Cycles	2.1%	7.5%	2.2%

Note: Model estimates. For outcomes with an *: baseline values are normalized to 1.

The transfer policy increases welfare by providing broad insurance against the

[27]Because of the nature of our model, we are abstracting from considerations on how the policy is financed. The cost of increasing transfers in bad times is around 4.5% of GDP, which is likely to significantly dampen the reported gains on output and welfare.

fundamental recession-related risk, that is the fall in earnings due to unemployment. In addition, it increases overall employment, because transfers require workers to be active in the labor force. By this channel, unemployment spells increase in duration because search policies operate at a different point in the trade-off between match quality and probability of leaving unemployment. Output then increases both because more workers are active, and because the overall sorting in the economy has improved. Two considerations on the impact of policies on sorting can be made. Notice that the sorting measure is higher than in the baseline model both when minimum wage is introduced or when transfers are implemented. In the first case, the increase in sorting does not reflect an improved allocative efficiency of the economy, but simply the moving into inactivity of part of the workers on the left tail of the human capital distribution. In the second case, instead, sorting is higher because there are more "*good workers in good firms*". Consistently with Song et al. (2019), increases in sorting generate more inequality, in our case captured by a small increase in the Gini index.

7 Conclusions

We have set up a new and tractable model of on-the-job search and human capital accumulation that features heterogeneity both on the worker and on the firm side. Ex-ante heterogeneous workers accumulate on-the-job experience which augments their skills and moves them up in the job ladder. Our contractual framework endogenously accounts for the difference incentives between risk-averse workers and risk-neutral entrepreneurs. We characterize how insurance incentives are of paramount importance in shaping the response of the labor market, the efficiency of workers-firms matches and the overall dynamic of human capital accumulation. Within this framework we show that aggregate fluctuations interact with shifts in the cross-sectional composition of labor markets, in ways consistent with the existence of hysteretic effects of recessions, and strong state dependence in the way the economy responds to aggregate shocks.

Consistent with the data, wage rigidity amplifies negative shocks to firms, and generates inefficient separations. We establish that workers that look for employment in bad economic times direct their search towards less productive firms. Search frictions and aggregate uncertainty prevent an efficient allocation of workers to firms and expose different cohort of workers to different human capital accumulation paths depending on the aggregate state at the time of entry in the labor force. Limits to workers' ability to accumulate human capital imposes a drag on the overall labor productivity of the economy after recessions that persists as long as these cohorts of workers are active in the labor force. Alterations to the sorting induced by recessions are slow to reverse, and contribute not only to slow recoveries, but also to long run changes in the structure of the economy.

We use the model to shed light on two open questions in business cycle literature. First, we show that increases in the importance of firms for workers' human capital accumulation, together with aging, can explain the increased length of recessions in recent decades. Second, we find that distributional channels can explain the different severity of "double dips" - that is, the reason why subsequent recessions are more or less severe when the economy is hit by shocks again within years after a first recession.

The framework lends itself naturally to be used as a laboratory for policy experiments. Our experiments support the idea that increasing the generosity of unemployment transfers in severe recessions can improve the overall allocative efficiency of the economy in the long run. On the contrary, minimum wages impose employment costs that are larger than the benefits generated in terms of reduced wage dispersion.

The flexibility of our proposed framework be exploited to understand how labor markets are shaped by long run trends that we have not included yet, e.g. in interest rates, technology, or firm dynamism. We leave these questions to future research.

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti. 1990. "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." *Econometrica*, 58(5): 1041–1063. 58

Acabbi, Edoardo Maria, and Andrea Alati. 2022. "Defusing leverage: Liquidity management and labor contracts." 4

Acabbi, Edoardo M, Ettore Panetti, and Alessandro Sforza. 2022. "The Financial Channels of Labor Rigidities: Evidence from Portugal." 118. 4

Agell, Jonas, and Per Lundborg. 2003. "Survey evidence on wage rigidity and unemployment: Sweden in the 1990s." *Scandinavian Journal of Economics*, 105(1): 15–30. 4

Altonji, Joseph G, and Paul J Devereux. 1999. "The extent and consequences of downward nominal wage rigidity." 4

Arellano-Bover, Jaime. 2020. "Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size." Working Paper. 43

Arellano-Bover, Jaime. 2022. "The Effect of Labor Market Conditions at Entry on Workers' Long-Term Skills." *The Review of Economics and Statistics*, 104(5): 1028–1045. 3

Arellano-Bover, Jaime, and Fernando Saltiel. 2021. "Differences in On-the-Job Learning Across Firms." 43

Ashenfelter, Orley. 1978. "Estimating the effect of training programs on earnings." *The Review of Economics and Statistics*, 47–57. 44

Baley, Isaac, Ana Figueiredo, and Robert Ulbricht. 2022. "Mismatch Cycles." *Journal of Political Economy*, 720461. 1

Balke, Neele, and Thibaut Lamadon. 2022. "Productivity Shocks, Long-Term Contracts and Earnings Dynamics." *American Economic Review*, 112(7): 2139–2177. 4, 8, 16, 44

Barlevy, Gadi. 2004. "The cost of business cycles under endogenous growth." *American Economic Review*, 94(4): 964–990. 3, 30

Benveniste, L.M., and J.A. Scheinkman. 1979. "On the Differentiability of the Value Function in Dynamic Models of Economics." *Econometrica*, 47(3): 727–732. 57

Berger, David W, Kyle F Herkenhoff, and Simon Mongey. 2022. "Minimum Wages, Efficiency and Welfare." 3

Bertheau, Antoine, Edoardo Maria Acabbi, Cristina Barcelo, Andreas Gulyas, Stefano Lombardi, and Raffaele Saggio. forthcoming. "The Unequal Cost of Job Loss across Countries." *American Economic Review: Insights*. 4

Blank, Julian, and Kalyanmoy Deb. 2020. "Pymoo: Multi-objective optimization in python." *IEEE Access*, 8: 89497–89509. 20

Dustmann, Christian, Attila Lindner, Uta Schönberg, Matthias Umkehrer, and Philipp vom Berge. 2021. "Reallocation Effects of the Minimum Wage." *The Quarterly Journal of Economics*, 137(1): 267–328. 3

Engbom, Niklas. 2020. "Labor market fluidity and human capital accumulation." *New York University mimeo*. 19, 45

Favilukis, Jack, Xiaoji Lin, and Xiaofei Zhao. 2020. "The Elephant in the Room: The Impact of Labor Obligations on Credit Markets." *American Economic Review*, 110(6): 1673–1712. 4

Fukui, Masao, Emi Nakamura, and Jón Steinsson. 2023. "Women, wealth effects, and slow recoveries." *American Economic Journal: Macroeconomics*. 2, 33

Goldin, Claudia, and Lawrence F Katz. 2007. "Long-run changes in the wage structure: narrowing, widening, polarizing." *Brookings Papers on Economic Activity*, , (2): 135–168. 33

Goos, Maarten, Alan Manning, and Anna Salomons. 2009. "Job polarization in Europe." *American economic review*, 99(2): 58–63. 33

Grigsby, John, Erik Hurst, and Ahu Yildirmaz. 2021. "Aggregate nominal wage adjustments: New evidence from administrative payroll data." *American Economic Review*, 111(2): 428–71. 4

Haltiwanger, John C, Henry R Hyatt, Erika McEntarfer, and Matthew Staiger. 2021. "Cyclical Worker Flows: Cleansing vs. Sullying." 48. 1

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante. 2020. "The Rise of US Earnings Inequality: Does the Cycle Drive the Trend?" *Review of Economic Dynamics*, 37: S181–S204. 1, 23

Herkenhoff, Kyle F. 2019. "The impact of consumer credit access on unemployment." *The Review of Economic Studies*, 86(6): 2605–2642. 12

Herkenhoff, Kyle F., Gordon Phillips, and Ethan Cohen-Cole. 2019. "How Credit Constraints Impact Job Finding Rates, Sorting & Aggregate Output." , (22274): 1–78. 63

Herkenhoff, Kyle, Jeremy Lise, Guido Menzio, and Gordon M Phillips. 2018. "Production and learning in teams." 3, 26

Holmstrom, Bengt. 1983. "Equilibrium long-term labor contracts." *The Quarterly Journal of Economics*, 23–54. 4

Huckfeldt, Christopher. 2022. "Understanding the scarring effect of recessions." *American Economic Review*, 112(4): 1273–1310. 4

Kahn, Lisa B. 2010. "The long-term labor market consequences of graduating from college in a bad economy." *Labour Economics*, 17(2): 303–316. 4, 23

Lise, Jeremy, and Fabien Postel-Vinay. 2020. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *American Economic Review*, 110(8): 2328–2376. 3, 6

Lise, Jeremy, and Jean-Marc Robin. 2017. "The macrodynamics of sorting between workers and firms." *American Economic Review.* 1

Lucas, Robert E. 1987. *Models of business cycles.* Vol. 26, Basil Blackwell Oxford. 3, 30

Marcet, Albert, and Ramon Marimon. 2019. "Recursive Contracts." *Econometrica*, 87(5): 1589–1631. 46, 58, 60

Menzio, Guido, and Shouyong Shi. 2010. "Block recursive equilibria for stochastic models of search on the job." *Journal of Economic Theory.* 5, 7, 8, 17, 46, 47, 50

Menzio, Guido, Irina A Telyukova, and Ludo Visschers. 2016. "Directed Search over the Life Cycle ." *Review of Economic Dynamics*, 124(3): 771–825. 63

Nekoei, Arash, and Andrea Weber. 2017. "Does Extending Unemployment Benefits Improve Job Quality?" *American Economic Review*, 107(2): 527–561. 3

Restuccia, Diego, and Richard Rogerson. 2017. "The causes and costs of misallocation." *Journal of Economic Perspectives*, 31(3): 151–74. 1

Rudin, Walter. 1976. *Principles of Mathematical Analysis.* McGraw-Hill Book Company. 54

Schaal, Edouard. 2017. "Uncertainty and Unemployment." *Econometrica*, 85(6): 1675–1721. 17

Schmieder, Johannes F, Till M von Wachter, and Jörg Heining. 2022. "The costs of job displacement over the business cycle and its sources: evidence from Germany." National Bureau of Economic Research. 4

Schwandt, Hannes, and Till von Wachter. 2019. "Unlucky Cohorts: Estimating the Long-Term Effects of Entering the Labor Market in a Recession in Large Cross-Sectional Data Sets." *Journal of Labor Economics*, 37(S1): S161–S198. 4, 23

Shi, Shouyong. 2009. "Directed search for equilibrium wage-tenure contracts." *Econometrica*, 77(2): 561–584. 5, 47

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. 2019.

"Firming up inequality." *The Quarterly journal of economics*, 134(1): 1–50. 1, 2, 36

Spear, Stephen E., and Sanjay Srivastava. 1987. "On Repeated Moral Hazard with Discounting." *The Review of Economic Studies*, 54(4): 599. 8, 58

Stoyanov, M. 2015. "User Manual: TASMANIAN Sparse Grids." Oak Ridge National Laboratory ORNL/TM-2015/596, One Bethel Valley Road, Oak Ridge, TN. 20

Thomas, Jonathan, and Tim Worral. 1988. "Self-enforcing wage contracts ." *The Review of Economic Studies*, 55(4): 541–553. 44

Tsuyuhara, Kunio. 2016. "Dynamic Contracts With Worker Mobility Via Directed on-the-Job Search." *International Economic Review*, 57(4): 1405–1424. 44

Von Stackelberg, Heinrich. 1934. *Marktform und gleichgewicht.* J. springer. 12, 59

Xiaolan, Mindy Z. 2014. "Who Bears Firm-Level Risk? Implications for cash-flow volatility." Working Paper. 4

Appendices

For compactness of notation, we omit the dependence on education level, which is a fixed characteristic, and the idiosyncratic human capital shock, which is additive, from the proof in Appendices. The logic of the proofs follows without loss of generality.

A Discussion

In this section we briefly discuss the properties of the equilibrium of the model economy developed in the previous sections. All propositions and corresponding proofs are reported in **Appendix B** and **C**.

A.1 Workers optimal behavior

In the following proposition we summarize the main results regarding the behavior of the workers and their objective functions.

Proposition A.1. *Given the worker search problem, the following properties hold:*

(i) The returns to search, $p(\theta(h, \tau, \iota, v; \Omega))\big[v - V\big]$, are strictly concave with respect to promised utility, v.

(ii) The optimal search strategy

$$v^*(h, \tau, \iota, V; \Omega) \in \arg\max_v \big\{p(\theta(h, \tau, \iota, v; \Omega))\big[v - V\big]\big\}$$

is unique and weakly increasing in V.

(iii) For all promised utilities, the search gain $R(h, \tau, \iota, V; \Omega)$ is positive, weakly decreasing in V.

(iv) The survival probability of the match, given the optimal choice of the worker, is increasing in the value of current and future promised utilities, so $\widetilde{p}_t(h, \tau, \iota, v; \Omega)$ is increasing in v and V.

Proof. See **Appendix Section B.1**. □

The first statement implies that the marginal returns of searching towards better firms are decreasing. The intuition is that as workers search for work at firms granting better values, their job-finding probability decreases as better employment prospects are also subject to higher competition.

As a consequence of the strict concavity established in the first statement, workers' optimal search strategy is unique. The search strategy is also (weakly) increasing in the value of lifetime utilities granted by the current contract, which is the outside option for the worker.

The third statement follows from the fact that marginal returns to search are decreasing and the set of feasible utility promises is compact. The intuition is that employees at firms with higher utility promises have a relatively fewer chances of improving their position. Given a high outside option, the utility gain from moving is relatively lower, whereas the probability of matching with any firm does not depend on the *current* utility promise per se, but on the future promise offered by the vacancy.

The fourth statement finally follows from considering the implication of the previous ones. Given that the optimal search strategy is increasing in V workers' probability of leaving the firm at any time ends up depending negatively on V. This guarantees a longer expected duration of the match at higher current promised utility V, thus retention probabilities that are increasing in promised utilities v.

As human capital accumulation is tightly linked to the quality of the employer, workers that are able to start their working careers in good times have a greater chance of finding themselves on an higher path of human capital growth. As worker careers are limited and human capital accumulation follows a slow-moving process, business cycle effects on human capital quality fade only slowly and the quality of initial matches, both in terms of lifetime utility and firm quality, bears a long-standing effect on workers' careers.

A.2 Characteristics of the optimal contract

The optimization in the contracting problem balances a trade-off between insurance provision and profit maximization for firms. The contract implicitly takes into account workers' search incentives and their inability to commit to stay. The following proposition characterizes workers' incentives along the business cycle from the firms' standpoint.

Proposition A.2. *The Pareto frontier* $J(h, \tau, \iota, W, y; a, \mu)$ *is increasing in the aggregate productivity shock* a*, while retention probabilities,* $\widetilde{p}(h, \tau, \iota, W; a, \mu)$ *decrease in aggregate productivity.*

Proof. See **Appendix Section C.1**. □

The intuition behind this proposition relies on the observation that higher productivity realization are associated not only with better outcomes on impact but also to better future prospects, given that the productivity process is an increasing Markov chain.

A key property of the model is that it allows to characterize the workers' optimal behaviour along the business cycle. The following proposition summarizes how the search strategy changes depending on the aggregate productivity realization.

Corollary A.1. *The optimal search strategy of the workers is increasing in aggregate productivity.*

Proof. The claim follows directly from **Proposition A.2**. □

Firm value $J(\cdot)$ is increasing in a, while retention is decreasing. As firm make more profits, they can marginally increase offered utilities to workers to maximize retention. This in turn will imply that in better times workers will search for better firms, given their greater outside options. The way in which offered wages w and thus values W are tied to profits J is described in **Proposition A.3**.

Proposition A.2 and Corollary A.1 have an important implication regarding firms' vacancy posting and workers' search decisions. The fact that at the posting stage profits J are increasing in aggregate productivity implies that more entry will take place in good times, and ceteris paribus more entrepreneurs will open up vacancies across the whole firms' distribution.[28] The resulting higher tightness impacts workers' optimal search behaviour as the job finding probability increases in all submarkets. As a consequence, workers respond optimally to the productivity increase searching in submarkets that guarantee higher lifetime utility promises.

Firms utility promises depend on the structure of the optimal contract. The contract provides insurance to workers through wage paths that are downward rigid, and at the same time allows firms to profit as wages only partially adjust to productivity realizations.

The following propositions provide a clear picture of the growth path prescribed by the optimal contract for a continuing firm. First, let us define the productivity threshold that determines whether a worker-firm match does not survive.

Corollary A.2. *There exists a productivity threshold $a^*(h, \tau, \iota, W, y)$ below which firms will not continue the operate.*

The intuition of why this has to be the case is linked to the fact that the Pareto frontier is strictly increasing in a and decreasing in the level of promised utilities to the worker. Hence once the aggregate state realizes a firm is able to perfectly

[28] In our model a better firm is a more productive firm. We do not specifically model the determinant of quality heterogeneity but we take the existence of profound differences in firm quality as a fact (Arellano-Bover, 2020, Arellano-Bover and Saltiel, 2021).

predict whether next period it will exit the market or stay in (given the timing, the decision is based on expected profits, and is thus *not* state-contingent to next period's productivity). The choice is taken *before* new realizations of productivity, so it is possible that a firm makes negative profits for at most one period.

Proposition A.3. *For each state in which the firm is willing to continue the contract, the optimal contract delivers a wage path that follows firms profits according to the wage Euler equation:*

$$\frac{\partial \widetilde{p}(\Theta)}{\partial W_i'} \frac{J'(\Theta)}{\widetilde{p}(\Theta)} = \frac{1}{u'(w_i)} - \frac{1}{u'(w)} \tag{A.1}$$

with $\Theta \equiv (\phi(h, y), \tau + 1, \iota, W_i'; \Omega')$ *being the definition of the relevant state and* w_i *is the wage paid in the future state.*

Proof. See **Appendix Section C.3**. □

The optimal contract links the wage growth to the realization of firms profits. The right hand side of Equation A.1 shows that, in providing insurance to the worker, the firm links wage growth to profits and to the incentive to maximize retention, incorporated in $\frac{\partial \log \widetilde{p}}{\partial W}$, the semi-elasticity of the retention probability to the utility offer. As the production stage takes place *after* exit choices are taken by the incumbent firms, the wage growth related to the continuation value of the contract is bound to be (weakly) positive, hence workers enjoy a non-decreasing wage profile under the optimal contract.[29]

A feature that the optimal contract derived in our model shares with the literature on long-term contracts with lack of commitment on the worker side is thus the backloading of wages.[30] Workers in our model make search decisions that affect the survival probability of the match. They do not however appropriate the full future value of the current match while making these search decisions (unless the firm makes zero profits). This makes it optimal for the firm to front-load profits and back-load wages. The reason is that the firm provides insurance and income smoothing to the worker, but given its risk neutrality it prefers to front-load its profits while providing an increasing compensation path to maximize retention. The contract thus optimally balances the consumption smoothing motives (i.e. the insurance provision of the contract) with the commitment problem of the worker.

Special case with log-utility. The wage Euler equation discussed in **Proposition A.3** can be simplified to a more intuitive interpretation in the

[29]As the exit decision takes place by considering *expected* profits next period, a firm operating at low but positive expected profits might end up, at most for a period, to have a negative continuation value. This would imply that wage growth *can* be negative before a firm's closure, which is actually a common finding in empirical studies (firstly observed in Ashenfelter (1978)).

[30]See for instance, Thomas and Worral (1988), Tsuyuhara (2016) and Balke and Lamadon (2022).

log-utility case. In case of log-utility, in fact, $u'(w_{i,\Omega}) = \frac{1}{w_{i,\Omega}}$. Multiplying and dividing by wage levels and rearranging, we can express the elasticity of retention probability to offered utility as

$$\varepsilon_{\widetilde{p},W_y} = \underbrace{\frac{(w_i - w)}{w}}_{\text{Wage growth}} \quad \underbrace{\frac{w}{J(\Theta)}}_{\substack{\text{Ratio of wage} \\ \text{to match value}}} \quad . \tag{A.2}$$

with $\varepsilon_{\widetilde{p},W} \equiv \frac{\partial \widetilde{p}(\Theta)}{\partial W_i \widetilde{p}(\Theta)}$.

The interpretation of this result is of interest to analyses that relate labor market dynamism to wage dynamics, like Engbom (2020). This is because $\varepsilon_{\widetilde{p},W_y}$, being a function of structural parameters of the matching technology, γ, search frictions λ_e, and measures of labor market tightness θ, provides us with a good proxy of labor market fluidity. The right hand side of (A.2), is composed entirely of observable quantities, as the ratio of wages to match value is a function of factor shares in value added. The quantity can then be used to compare the dynamism of different local, regional or national labor markets.
The next proposition, instead, confirms our initial conjecture that in equilibrium firm qualities and utility promises are related to a one-to-one mapping.

Proposition A.4. *The mapping defined by the function $f_v : \mathcal{Y} \to \mathcal{V}$ is an injective function for each worker characteristic (h, τ).*

Proof. See **Appendix Section C.2**. □

The proof is based on the fact that the Pareto frontier J is concave, the vacancy filling probability q is weakly positive and vacancy costs are both weakly positive and increasing in y. As shown in **Appendix C** these features are enough to guarantee that only one kind of firm y, given workers' characteristics h, τ, can optimally offer a given lifetime utility promise W, and that this mapping is monotonically increasing. We thus obtain a unique monotonic solution in which higher quality firms offer higher lifetime utility promises to workers.

Finally, we provide the alternative recursive formulation for the contracting problem described in the paper. The saddle-point functional equation that can be alternatively used to define the recursive contract in **Equation** (8) is expressed in the following proposition.

Proposition A.5. *The solution to the contracting problem in* ***Equation*** (8) *is the* *same as the solution to the following saddle-point functional equation:*

$$\begin{aligned}\mathcal{P}_t(h_t, \tau_t, y_t, a_t, \gamma_t) = & \underset{\gamma_t}{inf}\, \underset{w_t}{sup}(f(a_t, y_t, h_t) - w_t) + \mu_t^1 W_{y,t} - \gamma_t^1(W_{y,t} - u(w_t)) + \\ & \beta\mathbb{E}_t(\lambda U_{t+1} + (1-\lambda)\lambda_e p_{t+1} v_{t+1}^*) + \beta\mathbb{E}_t \widetilde{p}_{t+1}\mathcal{P}_{t+1}(h_{t+1}, \tau_{t+1}, y_{t+1}, a_{t+1}, \gamma_{t+1})\end{aligned}$$

with $\mu_t = \gamma_{t_1}$ for some starting γ_0.

Proof. See **Appendix D** for the details of the derivation of the SPFE following Marcet and Marimon (2019). □

B Properties of worker optimal behavior

For compactness of notation, we omit the dependence on education level, which is a fixed characteristic, and the idiosyncratic human capital shock, which is additive, from the proofs in Appendices. The logic of the proofs follows without loss of generality.

The following propositions characterize the properties of workers' optimal search strategies that solve the search problem in (2), restated here for convenience:

$$R(h, \tau, V; \Omega) = \sup_v \Big[p(\theta(h, \tau, v; \Omega))\big[v - V\big]\Big]. \tag{B.1}$$

Lemma B.1. *The composite function $p(\theta(h, \tau, v; \Omega))$ is strictly decreasing and strictly concave in v.*

Proof. For this proof we follow closely Menzio and Shi (2010), Lemma 4.1 (ii). From the properties of the matching function we know that $p(\theta)$ is increasing and concave in θ, while $q(\theta)$ is decreasing and convex. Consider that the equilibrium definition of $\theta(\cdot)$ is

$$\theta(h, \tau, v; \Omega) = q^{-1}\left(\frac{c(y)}{J(h, \tau, v, y; \Omega)}\right),$$

and that the first order condition for the wage and the envelope condition on V of the optimal contract problem in (8) implies

$$\frac{\partial J(h, \tau, v, y; \Omega)}{\partial v} = -\frac{1}{u'(w)}.$$

so that as $u'(\cdot) > 0$, $J(\cdot)$ is decreasing in v.

From the equilibrium definition of $\theta(\cdot)$ and noting that $q^{-1}(\cdot)$ is also decreasing due to the properties of the matching function we have that

$$\frac{\partial \theta(h,\tau,v;\Omega)}{\partial v} = \left.\frac{\partial q^{-1}(\xi)}{\partial \xi}\right|_{\xi=\frac{c(y)}{J(h,\tau,v,y;\Omega)}} \cdot \left(-\frac{\partial J(h,\tau,v,y;\Omega)}{\partial v}\right) \cdot \frac{c(y)}{(J(h,\tau,v,y;\Omega))^2} < 0,$$

which, in turn, implies that

$$\frac{\partial p(\theta(h,\tau,v;\Omega))}{\partial v} = \left.\frac{\partial p(\theta)}{\partial \theta}\right|_{\theta=\theta(h,\tau,v;\Omega)} \cdot \frac{\partial \theta(h,\tau,v;\Omega)}{\partial v} < 0.$$

Suppressing dependence on the states (h, τ, y, Ω) for readability, to prove that $p(\theta(v))$ is concave, consider that $J(v)$ is concave[31] and a generic function $\frac{c}{v}$ is strictly convex in v. This implies that with $\alpha \in [0, 1]$ and $v_1, v_2 \in \mathcal{V}$, $v_1 \neq v_2$:

$$\frac{c}{J(\alpha v_1 + (1-\alpha)v_2)} \leq \frac{c}{\alpha J(v_1) + (1-\alpha)J(v_2)} < \alpha\frac{c}{J(v_1)} + (1-\alpha)\frac{c}{J(v_2)}.$$

As $p(q^{-1}(\cdot))$ is strictly decreasing the inequality implies that

$$\begin{aligned} p\left(q^{-1}\left(\tfrac{c}{J(\alpha v_1+(1-\alpha)v_2)}\right)\right) &\geq p\left(q^{-1}\left(\tfrac{c}{\alpha J(v_1)+(1-\alpha)J(v_2)}\right)\right) \\ &> \alpha p\left(q^{-1}\left(\tfrac{c}{J(v_1)}\right)\right) + (1-\alpha)p\left(q^{-1}\left(\tfrac{c}{J(v_2)}\right)\right), \end{aligned}$$

and as $\theta(v) = q^{-1}(\frac{c}{J(v)})$:

$$p(\theta(\alpha v_1 + (1-\alpha)v_2)) > \alpha p(\theta(v_1)) + (1-\alpha)p(\theta(v_2))$$

so that $p(\theta(v))$ is strictly concave in v. □

B.1 Proof of Proposition A.1

Proof. The proofs follow closely Shi (2009), Lemma 3.1 and Menzio and Shi (2010), Corollary 4.4. More formally, for each triplet (h, τ, Ω) given at each search stage, we can re-define the search objective function as $K(v, V) = p(\theta(v))(v - V)$ and $v^*(V) \in \arg\max_v K(v, V)$ as the function that maximises the search returns (i.e. the optimal search strategy of the worker) and prove the following

(i) To show that $K(v, V)$ is strictly concave in v consider two values for v, $v_1, v_2 \in \mathcal{V}$ such that $v_2 > v_1$ and define $v_\alpha = \alpha v_1 + (1-\alpha)v_2$ for $\alpha \in [0, 1]$.

[31] $J(\cdot)$ concave give the two-point lottery in the structure of the contract. See Menzio and Shi (2010) Lemma F.1.

Then by definition:

$$\begin{aligned}K(v_\alpha, V) &= p(\theta(v_\alpha))(v_\alpha - V)\\ &\geq [\alpha p(\theta(v_1)) + (1-\alpha)p(\theta(v_2))][\alpha(v_1 - V) + (1-\alpha)(v_2 - V)]\\ &= \alpha K(v_1, V) + (1-\alpha)K(v_2, V) + \alpha(1-\alpha)[(p(\theta(v_1)) - p(\theta(v_2))](v_2 - v_1)\\ &> \alpha K(v_1, V) + (1-\alpha)K(v_2, V)\end{aligned}$$

where the first inequality follows from the concavity of $p(\theta(\cdot))$ (this is true if $J(\cdot)$ concave with respect to V) and the second inequality stems from the fact that $p(\theta(\cdot))$ is strictly decreasing hence $\alpha(1-\alpha)[(p(\theta(v_1)) - p(\theta(v_2))](v_2 - v_1) > 0$.

(ii) **Weakly Increasing.** Consider a worker employed in a job that gives lifetime utility V. Given that $v \in [\underline{v}, \overline{v}]$, and that submarkets are going to open depending on realizations of the aggregate productivity, a, there is only one region in the set of promised utilities where the search gain is positive. This set is $[V, v(a)]$ with $v(a)$ being the highest possible offer that a firm makes in the submarket for the worker (h, τ). Any submarket that promises higher than $v(a)$ is going to have zero tightness. Therefore, the optimal search strategy for $V \geq v(a)$ is $v^*(V) = V$, as $K(V, v(a)) = K(V, V) = K(\overline{v}, V) = 0$ (the search gain is null given the current lifetime utility V). For $V \in [V, v(a)]$, instead, as $K(v, V)$ is bounded and continuous, the solution $v^*(V)$ has to be interior and therefore respect the following first order condition

$$V = v^*(V) + \frac{p(\theta(v^*(V)))}{p'(\theta(v^*(V)) \cdot \theta'(v^*(V))}. \tag{B.2}$$

Now consider two arbitrary values V_1 and V_2, $V_1 < V_2 < \overline{v}$ and their associated solutions $W_i = v^*(V_i)$ for $i = 1, 2$. Then, V_1 and V_2 have to generate two different values for the right-hand side of (B.2). Hence, $v^*(V_1) \cap v^*(V_2) = \emptyset$ when $V_1 \neq V_2$.

This also implies that the search gain evaluated at the optimal search strategy is higher than the gain at any other arbitrary strategy so that $K(W_i, V_i) > K(W_j, V_i)$ for $i \neq j$. This implies that

$$\begin{aligned}0 >& [K(W_2, V_1) - K(W_1, V_1)] + [K(W_1, V_2) - K(W_2, V_2)]\\ &= (p(\theta(W_2)) - p(\theta(W_1)))(V_2 - V_1),\end{aligned}$$

thus, $p(\theta(W_2)) < p(\theta(W_1))$. As $p(\theta(\cdot))$ is strictly decreasing (see Lemma B.1), then $v^*(V_1) < v^*(V_2)$.

Lipschitz continuous. Consider generic v_2, v_1 such that $v_i = v^*(V_i)$. Given that $p(\cdot)$ is a concave function, one can write $p(v_1) + p'(v_1)(v_2 - v_1) > p(v_2)$.

Using equation(B.2), we obtain $V_2 - V_1$:

$$
\begin{aligned}
V_2 - V_1 &= v_2 - v_1 + \frac{p(v_2)}{p'(v_2)} - \frac{p(v_1)}{p'(v_1)} \\
&> 2(v_2 - v_1) + p(v_2)\frac{p(v_1') - p(v_2')}{p'(v_2)p'(v_1)} > 2(v_2 - v_1)
\end{aligned}
$$

This condition proves that v^* is Lipschitz.

Unique. Uniqueness follows directly from strict concavity shown in (i).

(iii) The Bellman equation for the search problem is:

$$
R(h,\tau,V;\Omega) = \sup_{v} \Big[p(\theta(h,\tau,v;\Omega))\big[v - V\big]\Big]
$$

hence a simple envelope argument shows that

$$
\frac{\partial R(h,\tau,V;\Omega)}{\partial V} = -p(\theta(h,\tau,v;\Omega)) \leq 0,
$$

as the job finding probability is weakly positive for all utility promises.

As $p(\theta(\cdot)) \geq 0$, $v^*(\cdot) \in [\underline{v}, \overline{v}]$ then $R(\cdot) \geq 0$.

(iv) Given the optimal search strategy, $v^*(h,\tau,V;\Omega)$, we can define the survival probability of the match as in (5):

$$
\widetilde{p}(h,\tau,V;\Omega) \equiv (1-\lambda)(1-\lambda_e p(\theta(h,\tau,v^*;\Omega))).
$$

Then, given (h,τ,Ω)

$$
\frac{\partial \widetilde{p}(V)}{\partial V} = -\beta(1-\lambda)\lambda_e \left.\frac{\partial p(\theta)}{\partial \theta}\right|_{\theta=\theta(v^*)} \left.\frac{\partial \theta(v)}{\partial v}\right|_{v=v^*(V)} \frac{\partial v^*(V)}{\partial V} > 0,
$$

because $p(\cdot)$ and $v^*(\cdot)$ are both increasing functions while $\theta(\cdot)$ is a decreasing function in promised utilities.

□

C Properties of the optimal contract

Lemma C.1. *The Pareto frontier $J(h,\tau,W_y,y;\Omega)$ is concave in W_y.*

Proof. This is a direct consequence of using a two-point lottery for $\{w_i, W'_{iy}\}$ as shown by Menzio and Shi (2010), Lemma F.1. $\square$

Lemma C.2. *The Pareto frontier $J(h, \tau, W_y, y; \Omega)$ is increasing in y.*

Proof. The intuition for this proof follows the fact that a higher y firm, once the match exists, can always deliver a certain promise V and have resources left over. Within a dynamic contract, future retention is already optimized as the match is formed. This means that the promise V can be delivered by the greater capacity on the part of producing with respect to a close y firm. In presence of human capital accumulation, the worker is compensated through greater option values in the future, which again means that, even with lower retention, the firm cashes in more profits while decreasing wages (and respecting the V promise).

One can get to the same conclusion by starting from time T, noticing that the function J is trivially increasing in y in the last period, and the stepping back. At $T-1$, given V, any higher y function can make greater profits with the same delivery of value V, given the contract's optimal promise, which is a fortiori true with human capital accumulation (the option value is greater, so the firm can decrease w as a response). A more formal argument goes as follows: start from the optimal policies, as per Eq. (8) of a firm that has $y = \bar{y}$ and assume that one could exogenously increase its installed capital to $\bar{y}+\varepsilon$. We want to know whether, keeping policies constant, this would increase the flow of profits while keeping the worker indifferent. If this is true, then *a fortiori* it will be true that the firm value function J will be increasing in y. With a slight abuse of notation and for conciseness, we refer to the future J in Eq. (8) as J'.

This amounts to calculating:[32]

$$\left.\frac{d\bar{J}}{dy}\right|_{W_y,\{\pi_i,w_i,W'_{iy}\}} = \frac{\partial f(\cdot)}{\partial y} + \beta \mathbb{E}_{\Omega}\left[\left.\frac{\partial \tilde{p}(\cdot)}{\partial y}\right|_{W_y,\{\pi_i,w_i,W'_{iy}\}} \bar{J}'\right]$$

This first order condition presents the trade-off discussed in words above, namely that an increase in y will be instantaneously beneficial to production, but might also potentially have a longer term adverse effect on profits through decreased retention. Finding the sign of the derivative on the LHS hinges on understanding the sign of the derivative of the second element on the RHS, since $\frac{\partial f(\cdot)}{\partial y} > 0$ given the properties of the production function $f(\cdot)$. The change in y would affect search objective v_y through the variation in h due to the human capital accumulation dynamics, even taking the

[32]Without loss of generality, we assume that J' is constant with respect to y. Alternatively, one may start proving the result for contracts offered to workers just one period before retirement (for which $\frac{\partial J'}{\partial y}$ is trivially positive), and generalize the result with backward induction.

current firm policies as given.

Notice that:

$$\frac{\partial \tilde{p}(\cdot)}{\partial y} = -\lambda\lambda_E \frac{\partial p(\cdot)}{\partial \theta}\frac{\partial \theta(\cdot)}{\partial y}$$

Remember that $\frac{\partial p(\cdot)}{\partial \theta} > 0$ due to the properties of the job-finding probability function $p(\cdot)$. Assume first the case in which $\frac{\partial \theta(\cdot)}{\partial y} \leq 0$. In this case $\frac{\partial \tilde{p}(\cdot)}{\partial y} > 0$, which in turn implies, as argued, that J has to be increasing in y.

Now consider the second case, namely $\frac{\partial \theta(\cdot)}{\partial y} > 0$. By the free entry condition, we obtain:

$$\frac{\partial \theta(\cdot)}{\partial y} = \frac{\partial q^{-1}(c(y)/J(y))}{\partial y} = \frac{1}{q'\left(q^{-1}\left(c(y)/J(y)\right)\right)} \frac{c'(y)J(y) - \frac{\partial J(y)}{\partial y}c(y)}{J(y)^2}$$

The first term in the result is negative, given the properties of function $q(\cdot)$. Given our assumption on $\frac{\partial \theta(\cdot)}{\partial y}$ the second term in the result has to be negative as well. This requires: $c'(y)J(y) - \frac{\partial J(y)}{\partial y}c(y) < 0$, or $\frac{\partial J(y)}{\partial y} > \frac{c'(y)J(y)}{c(y)} > 0$. □

Corollary C.1. *The retention probability* $\widetilde{p}(h, \tau, V_{y,\Omega}; \Omega)$ *is decreasing in* h.

Proof. Similarly to what shown above:

$$\frac{\partial \tilde{p}(\cdot)}{\partial h} = -\lambda\lambda_E \frac{\partial p(\cdot)}{\partial \theta}\frac{\partial \theta(\cdot)}{\partial h}$$

The sign of $\frac{\partial \theta(\cdot)}{\partial h}$ can be again obtained by looking at the free entry condition. This time:

$$\frac{\partial \theta(\cdot)}{\partial h} = \frac{1}{q'\left(q^{-1}\left(c(y)/J(\cdot)\right)\right)} \frac{-\frac{\partial J(h)}{\partial h}c(y)}{J(\cdot)^2} > 0$$

Both terms are negative - it is straightforward to show that $\frac{\partial J(h)}{\partial h} > 0$ using a similar logic as in **Lemma C.2**. Intuitively: if $\frac{\partial J(h)}{\partial h} < 0$ then retention increases in h. But so do flow revenues, which would contradict the initial assumption. Bringing these elements together we get $\frac{\partial \theta(\cdot)}{\partial h} < 0$ which proves the overall result. □

C.1 Proof of Proposition A.2

Proof. We proceed by backward induction.[33] Following the logic of the proof of **Lemma C.2**, the proposition is trivially true for workers T periods old. Given that

[33] For compactness of notation, we omit without loss of generality the two-point lottery in the equations in the proof.

the firm increases its production while keeping the worker at least indifferent, J is at least weakly increasing in a. However, the firm can also feasibly increase the worker's wage by ε, with $\varepsilon < \frac{\partial f(\cdot)}{\partial a}$. J is thus strictly increasing in y.

Consider now a worker who is $T-1$ periods old. A firm matched to a worker in submarket $\{h, T-1, y, W_{y,\Omega}\}$ will face the following Pareto frontier

$$J(h, T-1, y, W_y; a, \mu) = \sup_{w, W'_y} \Big(f(h, y; a) - w \\ + \mathbb{E}_\Omega \left[\widetilde{p}(h', T, W'_y; a', \mu')(f(h', y; a') - w') \right] \Big)$$

Analogously to the proof in **Lemma C.2**, assume that aggregate productivity increases from $\bar{a}$ to $\bar{a} + \varepsilon$. Assume that the firm keeps its policies constant once again. We aim at proving that, even in such a case, firm value increases while keeping the worker at least indifferent. If this is the case, it is *a fortiori* true that J increases in a after reoptimizing firms' policies.

We are now interested in the sign of:[34]

$$\left.\frac{\partial \bar{J}}{\partial a}\right|_{W_y, \pi, w, \{W'_y\}} = \frac{\partial f(\cdot)}{\partial a} + \beta \mathbb{E}_\Omega \left[\left.\frac{\partial \tilde{p}(\cdot)}{\partial a}\right|_{W_y, \pi_i, w, \{W'_y\}} \bar{J}' \right]$$

Now notice that, in equilibrium,

$$\frac{\partial \widetilde{p}(\theta)}{\partial a} \propto -\frac{\partial p(\theta)}{\partial a} = \underbrace{\frac{\partial p(\theta)}{\partial \theta}}_{>0} \cdot \underbrace{\frac{\partial \theta}{\partial J(\cdot)}}_{>0} \cdot \frac{\partial J(\cdot)}{\partial a}$$

where the sign of the second derivative on the right hand side comes from the free entry condition and the properties of vacancy filling probability function $q(\cdot)$. Given this, it has to be that $\frac{\partial p(\theta)}{\partial a}$ and $\frac{\partial J(\cdot)}{\partial a}$ have the same sign in equilibrium. Now, if both are strictly positive, both statements of our proposition are immediately true. Let's now assume they are both negative or zero. If this is the case, then $\frac{\partial \tilde{p}(\cdot)}{\partial a} \geq 0$. But this implies $\frac{\partial \bar{J}}{\partial a} > 0$, which is a contradiction. $\square$

Corollary C.2. *There exists a productivity threshold $a^*(h, \tau, W_y, y)$ below which firms will not continue the operate.*

Proof. The proof follows immediately from **Proposition A.2** and the timing of the shock. Given the timing of the shock, exit is fully determined by the current

[34]We assume that $J' = f(h', y; a') - w'$ is constant with respect to a. It is possible to prove, by backward induction, that this assumption is without loos of generality for the sake of the proof.

productivity shock and incumbent firms know in advance whether they are willing to produce in the next period.

Therefore, as the Pareto frontier is strictly increasing in a, firms are willing to continue the contract if $\mathbb{E}_\Omega[J(h', \tau+1, W'_{y,}, y; a', \mu')|h, \tau, W_y, y, a, \mu] \geq 0$, so that the threshold that determines exit is

$$a^*(h, \tau, W_y, y) : \mathbb{E}_\Omega[J(h', \tau+1, W'_y, y; a', \mu')|h, \tau, W_y, y, a, \mu] = 0.$$

□

Corollary C.3. *The productivity threshold $a^*(h, \tau, y, W_y)$ below which firm y in match with worker (h, τ) and promised utility W_y exits the market in the aggregate state Ω is increasing in y.*

Proof. Consider two firms characterized by y_1, y_2 with $y_1 < y_2$. Consider the threshold for firm y_1, $a_1^* = a^*(h, \tau, , W_{y_1}, y_1)$. Firm y_1 makes 0 profits if state a_1^* materializes next period. Consider firm y_2 trying to mimic the current contract offered by y_1 to (h, τ). We know that J is increasing in y from **Lemma C.2**, which implies that the firm is making a profit at a_1^*. This completes the proof. □

Lemma C.3. *The Pareto frontier $J(h, \tau, W_y, y; \Omega)$ is strictly concave in y.*

Proof. **No human capital accumulation.** The proof without human capital accumulation is straightforward. Assuming there is no dependency of h' on y, one can write:

$$\frac{dJ}{dy} = f_y + \widetilde{p}\frac{dJ'}{dy}$$
$$\frac{d^2J}{dy^2} = f_{yy} + \widetilde{p}\frac{d^2J'}{dy^2}$$

where we take J' to represent next period's firm value $J(\cdot)$, and the dependence of all controls on y is ignored by virtue of the envelope condition. One can readily observe by induction that, given the concavity of $f(\cdot)$ in y, $\frac{d^2J}{dy^2} < 0$.

Human capital accumulation. With human capital accumulation the concavity of the function J on y depends on further assumptions on the concavity of the human capital accumulation function $g(\cdot)$. As these assumptions involve equilibrium objects, we state them here and verify ex-post numerically that they are always respected in our setting.

$$\frac{dJ}{dy} = f_y + g_y\frac{\partial \widetilde{p}J'}{\partial h} + \widetilde{p}\frac{dJ'}{dy}$$

$$\frac{d^2J}{dy^2} = f_{yy} + g_{yy}\frac{\partial \widetilde{p}J'}{\partial h'} + g_y^2\frac{\partial^2\widetilde{p}J'}{\partial h'^2} + 2g_y\left(\frac{\partial \widetilde{p}}{\partial h'}f_y + \widetilde{p}f_{hy}\right) + \widetilde{p}\frac{d^2J'}{dy^2}$$

This condition implies that the concavity of $J(\cdot)$ necessarily imposes limits in equilibrium on the properties of $g(\cdot)$, the deterministic component of human capital accumulation. We find this condition to be respected for reasonable parameterization and choices of functional form.

□

Corollary C.4. *As $J(h,\tau,\iota,W,y;\Omega)$ is concave, the tangent line at a generic $y_0 \in \mathcal{Y}$ is above the graph of $J(h,\tau,\iota,W,y;\Omega)$ so that*

$$J(h,\tau,\iota,W,y_0;\Omega) + \left.\frac{\partial J(h,\tau,\iota,W,y;\Omega)}{\partial y}\right|_{y=y_0}(y-y_0) \geq J(h,\tau,\iota,W,y;\Omega).$$

Proof. Dropping dependence on $(h,\tau,\iota,W;\Omega)$, consider two values for firm quality $y_0 < y_1$ both in $\mathcal{Y}$. Then, as $J_t(\cdot)$ is concave in y, taking $\alpha \in [0,1]$ the following inequalities are true:

$$\begin{aligned}
&J(\alpha y_0 + (1-\alpha)y_1) \geq \alpha J(y_0) + (1-\alpha)J(y_1)\\
\Rightarrow &J(\alpha y_0 + (1-\alpha)y_1) - J(y_0) \geq (1-\alpha)(J(y_1) - J(y_0))\\
\Rightarrow &\frac{J(\alpha y_0 + (1-\alpha)y_1) - J(y_0)}{\alpha y_0 + (1-\alpha)y_1 - y_0} \geq \frac{J(y_1) - J(y_0)}{y_1 - y_0}.
\end{aligned}$$

where the third inequality comes from noting that $y_1 > y_0$ and $\alpha y_0 + (1-\alpha)y_1 - y_0 = (1-\alpha)(y_1 - y_0)$.

Taking the limit for $\alpha \to 1$, we have that the left hand side tends to $\left.\frac{\partial J_t(y)}{\partial y}\right|_{y=y_0}$ and hence

$$J(y_0) + \left.\frac{\partial J(y)}{\partial y}\right|_{y=y_0}(y_1 - y_0) \geq J(y_1). \tag{C.1}$$

Note that if $y_0 > y_1$ then $\frac{J(\alpha y_0+(1-\alpha)y_1)-J(y_0)}{\alpha y_0+(1-\alpha)y_1-y_0} \leq \frac{J(y_1)-J(y_0)}{y_1-y_0}$ but multiplying again the left hand side and the right hand side for $(y_1 - y_0) < 0$ still delivers (C.1). □

C.2 Proof of Proposition A.4

Proof. Note: throughout the proof we drop the dependence of the functions to the state (h,τ,Ω) to ease readability.

If the function f_v is an injective function then it defines a one-to-one mapping between $\mathcal{Y}$ and $\mathcal{V}$ so that for $(y_1, y_2) \in \mathcal{Y}$, and $f_v(y_1) = W_1$ and $f_v(y_2) = W_2$, $(W_1, W_2) \in \mathcal{V}$, $f_v(y_1) = f_v(y_2) \Rightarrow y_1 = y_2$.[35] We proceed by contradiction. To begin,

[35] As the contrapositive of Definition 2.2 in Rudin (1976), that defines a one-to-one mapping for

assume that $f_v(y_1) = f_v(y_2)$ and $y_1 \neq y_2$.

As the optimal contract is a concave function in firm quality, we know that the tangents at each point are above the graph of the function. Thus, we can define the tangents at the two points y_1, y_2 as

$$T_1(y) \equiv J(y_1) + \frac{\partial J(y)}{\partial y}\bigg|_{y=y_1} (y - y_1) \quad \text{and} \quad T_2(y) \equiv J(y_2) + \frac{\partial J(y)}{\partial y}\bigg|_{y=y_2} (y - y_2).$$

Without loss of generality, consider the case in which $y_2 > y_1$. Knowing that $T_i(y) \geq J(y)$ for $i = 1, 2$ due to the concavity of $J(\cdot)$, we can define the following inequalities:

$$T_1(y_2) - J(y_2) \geq 0 \quad \text{and} \quad T_2(y_1) - J(y_1) \geq 0.$$

Using the definitions for the tangents at y_1 and y_2 they imply that

$$\frac{J(y_2) - J(y_1)}{y_2 - y_1} \leq \frac{\partial J(y)}{\partial y}\bigg|_{y=y_1} \quad \text{and} \quad \frac{J(y_2) - J(y_1)}{y_2 - y_1} \geq \frac{\partial J(y)}{\partial y}\bigg|_{y=y_2},$$

hence combining the inequalities we get that

$$\frac{\partial J(y)}{\partial y}\bigg|_{y=y_2} \leq \frac{J(y_2) - J(y_1)}{y_2 - y_1} \leq \frac{\partial J(y)}{\partial y}\bigg|_{y=y_1}. \tag{C.2}$$

However, the free-entry condition in vacancy posting implies that in the submarket (h, τ, W) both firms must be respecting $c(y_i) = q(\theta)\beta J(y_i)$ for $i = 1, 2$. As $c(y_i)$ is a linear function of firm quality $\frac{\partial c(y_i)}{\partial y_i} = c$ for $i = 1, 2$ and therefore from the free-entry condition:

$$c = q(\theta)\beta \frac{\partial J(y)}{\partial y}\bigg|_{y=y_i}$$

which is a contradiction of the slopes of the two tangents being decreasing as shown in Equation (C.2). Note that if $c(y)$ is convex and twice differentiable, then the derivatives of $c(y)$ are increasing in y while the derivatives of $J(\cdot)$ are decreasing leading again to a contradiction. The proof for the case in which $y_1 > y_2$ follows the same arguments and leads to a similar contradiction on the implied slopes of the optimal contract and those implied by the free entry condition. □

Lemma C.4. *Given a state (y, h, τ, W) the optimal contract implies that*

$$-\frac{\partial J_t(h, \tau, y, W; \Omega)}{\partial W} = \frac{1}{u'(w)}$$

so that promised utilities and wages move in the same direction.

$(x_1, x_2) \in A$ as $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$.

Proof. The proof follows directly from the envelope theorem and the concavity of the utility function $u(\cdot)$, as discussed in the proof of Proposition A.3. □

Corollary C.5. *The Pareto frontier* $J(h, \tau, y, W; \Omega)$ *is decreasing in promised utilities* W.

Proof. The envelope condition in **Lemma C.4** and note that $u'(\cdot) \geq 0$. □

Proposition C.1. *Then utility promises are unique and increasing in* y, $\frac{\partial W}{\partial y} > 0$.

Proof. Uniqueness follow directly from **Lemma C.3** and **Proposition A.4**, given the existence of an injective mapping $f_v(\cdot)$ and the strict concavity of the objective function in Equation (11). One can prove that W is increasing in y by analyzing the first order condition in y of the problem.

Conditional on χ the entrepreneur has to choose the optimal value y to offer to workers.

For the rest of the proof we consider as given the dependence of the functions on (χ) and consider directly the function $q(\theta(W))$ as $q(W)$. The optimization involves a trade-off which respects the following first order condition:

$$c_y - qJ_y = 0 \tag{C.3}$$

We also know that the second order condition is strictly negative:

$$-c_{yy} + qJ_{yy} < 0 \tag{C.4}$$

By the implicit function theorem, the derivative of quation C.3 with respect to W is:

$$(-c_{yy} + qJ_{yy} < 0)\frac{\partial y}{\partial W} + q_W J_y + qJ_{Wy} = 0 \tag{C.5}$$

The first term in parenthesis is negative, as shown above. The second term is positive, given **Lemma C.2** and the fact that q_W is positive (**Lemma B.1**). The third term is 0 because of the envelope condition for the firm problem being independent of y (**Lemma C.4**). This means that, in order for the equality to be respected, $\frac{\partial y}{\partial W} > 0$, which given $f_v(\cdot)$ also implies $\frac{\partial W}{\partial y} > 0$.

□

C.3 Proof of Proposition A.3

Proof. Consider the firm problem in Equation (8), restated here for convenience

$$J(h,\tau,W,y;\Omega)=\sup_{\{\pi_i,w_i,W_i\}}\sum_{i=1,2}\pi_i\Big(f(y,h;\Omega)-w_i$$
$$+\mathbb{E}_{\Omega}\left[\widetilde{p}(h',\tau+1,W_i;\Omega')J(h',\tau+1,W_i,y;\Omega')\right]\Big)$$

$$s.t.\ [\lambda]\ W=\sum_{i=1,2}\pi_i\left(u(w_i)+\mathbb{E}_{\Omega}\widetilde{r}(h',\tau+1,W_i;\Omega')\right),$$
$$\sum_{i=1,2}\pi_i=1,\quad h'=\phi(h,y).$$

For $i=1,2$, the first order conditions with respect to the wage and the promised utilities are:

$$[w_i]:\ \lambda=\frac{1}{u'(w_i)} \tag{C.6}$$
$$[W_i]:\ \frac{\partial\widetilde{p}(\cdot)}{\partial W_i}J(\cdot)+\widetilde{p}(\cdot)\frac{\partial J(\cdot)}{\partial W_i}+\lambda\frac{\partial\widetilde{r}(\cdot)}{\partial W_i}=0. \tag{C.7}$$

Note that by definition,

$$\widetilde{r}(h,\tau,V;\Omega)\equiv\lambda U(h,\tau;\Omega)+(1-\lambda)\Big[W+\lambda_e\max\{0,R(h,\tau,V;\Omega)\}\Big]$$

therefore we can use the envelope theorem as in Benveniste and Scheinkman (1979), Theorem 1 and the definition in Equation (5) to derive an expression for the derivative of the employment value in $t+1$ as the period ahead of the following:

$$\frac{\partial\widetilde{r}(h,\tau,W;\Omega)}{\partial W}=\widetilde{p}(h,\tau,W;\Omega).$$

Similarly, using the envelope condition on the firm problem and the first order condition for the wage, we can establish that

$$\frac{\partial J(h,\tau,y,W;\Omega))}{\partial W}=-\lambda\ \therefore\ \frac{\partial J(h,\tau,W,y;\Omega))}{\partial W}=-\frac{1}{u'(w)}. \tag{C.8}$$

Moving these two expressions one period ahead, substituting them in (C.7), focusing on $\widetilde{p(\cdot)}>0$ and $\pi_i>0$ and rearranging we have that:

$$\frac{\partial\widetilde{p}(\Theta)}{\partial W_i}\frac{J(\Theta)}{\widetilde{p}(\Theta)}=\frac{1}{u'(w_i')}-\frac{1}{u'(w)},$$

with $\Theta\equiv(\phi(h,y),\tau+1,W;\Omega')$ and where w' is the wage next period in state Ω'. $\square$

D Derivation of recursive contract SPFE

Solving the optimal contract and the overall model given the recursive structure obtained by following the promised utility method of Spear and Srivastava (1987) is computationally infeasible. This is due to the fact that the optimal contract requires to define a valid recursive domain and codomain of promised values that respects all the future forward looking constraints. Known solution methods for these kinds of models (Abreu, Pearce and Stacchetti, 1990), although robust, easily become computationally unmanageable as the number of states of the model increases. We thus follow Marcet and Marimon (2019) in deriving a recursive expression for the optimal contract in which the Lagrange multiplier for the promise keeping constraint **Equation C.8** is added as a co-state of the model, and allows us to circumvent the problem of searching for valid promised values domains altogether.

The reason why the recursive contracts method in Marcet and Marimon (2019) simplifies our problem is simple. As shown in **Equation C.8**, wage growth and levels in any next period and at every node are determined by the state-contingent multiplier on tomorrow's promise keeping constraints. This considerably reduces the complexity of the problem, as by definition Lagrange multipliers are defined over $\mathbb{R}^+$.

We follow Marcet and Marimon (2019) (hereby MM) and their terminology to define how a recursive saddle point functional equation (SPFE) can be obtained from the sequential formulation of the problem. For the present exposition of the constructive method to obtain the SPFE, for simplicity and without loos of generality, we ignore the randomization of the contract over the lotteries and the limited liability constraint. The latter choice, in particular, does not create any problem in terms of thinking about of developing the sequential problem over time: our choice of timing of exit decision is such as that exiting firms know form the start of their period whether the productivity level is below the critical one $a^*_{h,\tau,\iota,y}$ for the match (h, τ, ι, W, y), and thus whether they will exit or not. The lack of uncertainty and optimization over the next periods makes the problem of these firms, at some low states, equivalent to the problem of a firm with a lower maximum length (which is T, the retirement age, in general). At an exiting state t the firm knows *with certainty* that any $J_j = 0$ for $j > t$, match with a worker of age T.

Consider the problem[36]

[36]Without loss of generality, we ignore the level of education ι in the proof, as it is a fixed worker characteristic.

$$J_t(h_t, \tau_t, y_t, W_t, a_t) = \sup_{w_t, \{W_{s^{t+1}}\}} \Big(f(a_t, y_t, h_t) - w_t$$
$$+\mathbb{E}_{s^t}\left[\widetilde{p}(h_{t+1}, \tau_{t+1}, W_{s^{t+1}}, a_{s^{t+1}})(J_{t+1}(h_{t+1}, \tau_{t+1}, y_t+1, W_{s^{t+1}}, a_{s^{t+1}})\right] \Big) \tag{D.1}$$

$$s.t.\ W_t = u(w_t) + \beta \mathbb{E}_{s^t} \Big(\lambda U_t(h_{t+1}, \tau_{t+1}, a_{t+1}) +$$
$$(1-\lambda)(\lambda_e p_{t+1}(h_{t+1}, \tau_{t+1}, W_{s^{t+1}}, a_{s^{t+1}}) v^*(h_{t+1}, \tau_{t+1}, W_{s^{t+1}}, a_{s^{t+1}})$$
$$+ (1 - \lambda_e p_{t+1}(h_{t+1}, \tau_{t+1}, W_{s^{t+1}}, a_{s^{t+1}})) W_{s^{t+1}}) \Big) \tag{D.2}$$

We define as endogenous states $\mathbf{x}_t = [h_t, \tau_t, y_t, W_t]$, controls $\mathbf{c}_t = [w_t, W_{s^{t+1}}]\ \forall t, s^{t+1}$, whereas the only exogenous state is a_t. The endogenous states follow the law of motion

$$\mathbf{x}_{t+1} = \begin{bmatrix} h_{t+1} \\ \tau_{t+1} \\ y_{t+1} \\ W_{s^{t+1}} \end{bmatrix} = l(\mathbf{x}_t, \mathbf{c}_t, a_{s^{t+1}}) = \begin{bmatrix} \phi(h_t, y_t) \\ \tau_t + 1 \\ y_t \\ W_{s^{t+1}} \end{bmatrix} \tag{D.3}$$

In the subsequent notation, where appropriate, we omit listing all states on which elements in the equation, and subsume their dependence under just listing the time t. J can be rewritten, by developing forward the recursion until time T, at which the match surely dissolves, as

$$J_t(\{h_t, \tau_t, y_t, W_t, a_t\}_{t=t_0}^{T-t_0}) = \mathbb{E}_{t_0} \sum_{t=t_0}^{T} \beta^{t-t_0} \prod_{i=0}^{t-t_0} \widetilde{p}_{t_o+i} \Big(f(a_t, y_t, h_t) - w_t \Big) \tag{D.4}$$

where $\widetilde{p}_{t_0} = 1$. Notice that the forward-looking constraint in **Equation D.2** is state contingent and an instance of it applies at *every* node of any possible history $s^t\ \forall t$ given the prevailing W_y promised at that node. The equilibrium is an instance of subgame perfect Nash equilibrium in which an agent chooses its strategies while anticipating the best response of the following agent, as common in dynamic games with a leader-follower component introduced by Von Stackelberg (1934). The structure of the problem and the solution also shares some commonality with Ramsey optimal policy problems in which a policy maker (in this case the firm) optimizes the utility of all agents according to some weights and taking into account their optimal behavior.[37]

[37] In the terminology of MM, we treat constraints coming from **Equation D.2** as a set of one period ahead forward looking constraint, which makes the analysis of our case akin to their case where one have $j = 1$ forward looking constraints, and $N_1 = 0$. The difference with their problems, however, is that our problem features finite time, and thus each one period ahead forward looking constraint

We can redefine the problem:

$$V_{t_0}(\mathbf{x}_t, a_t) = \underset{\{w_{s^t}, W_{s^t}\}}{sup} \mathbb{E}_{t_0} \sum_{t=t_0}^{T} \beta^{t-t_0} \prod_{i=0}^{t-t_0} \widetilde{p}_{t_o+i} \Big(f(a_t, y_t, h_t) - w_t \Big) \tag{D.5}$$

$$s.t. \; [j=0]: \sum_{t=t_0}^{T} \beta^{t-t_0} \prod_{i=0}^{t-t_0} \widetilde{p}_{t_o+i} \Big(f(a_t, y_t, h_t) - w_t \Big) - R \geq 0 \tag{D.6}$$

$$[j=1, s^t]: \; W_{s^t} - u(w_{s^t}) - \beta \mathbb{E}_{s^t} \Big(\lambda U_{s^{t+1}} + (1-\lambda)(\lambda_e p_{s^{t+1}} v^*_{s^{t+1}}) + \widetilde{p}_{s^{t+1}} W_{s^{t+1}}) \Big) \geq 0 \tag{D.7}$$

where the constraint D.13 is a slack participation constraint for a sufficiently small R, so that the principal (the firm) is willing to enter the contract in the first place.

In the terminology of MM we can label

$$h_0^0(\mathbf{x}_t, \mathbf{c}_t, a_t) = f(a_t, y_t, h_t) - w_t \tag{D.8}$$

$$h_1^0(\mathbf{x}_t, \mathbf{c}_t, a_t) = f(a_t, y_t, h_t) - w_t - R \tag{D.9}$$

$$h_0^1(\mathbf{x}_t, \mathbf{c}_t, a_t) = W_t \tag{D.10}$$

$$h_1^1(\mathbf{x}_t, \mathbf{c}_t, a_t) = W_t - u(w_t) + \beta \mathbb{E}_t(\lambda U_{t+1} + (1-\lambda)\lambda_e p_{t+1} v^*_{y,t+1}) \tag{D.11}$$

and define the Pareto problem ($\mathbf{PP}_\mu$)

$$\mathbf{PP}_\mu : V_{\mu,t_0}(\mathbf{x}_t, a_t) = \underset{\{w_{s^t}, W_{s^t}\}}{sup} \mathbb{E}_{t_0} \sum_{t=t_0}^{T} \beta^{t-t_0} \prod_{i=0}^{t-t_0} \widetilde{p}_{t_0+i} \mu^0 \Big(f(a_t, y_t, h_t) - w_t \Big) + \mu^1 W_{t_0} \tag{D.12}$$

$$s.t. \; [j=0;\; \gamma^0]: \sum_{t=t_0}^{T} \beta^{t-t_0} \prod_{i=0}^{t-t_0} \widetilde{p}_{t_o+i} \Big(f(a_t, y_t, h_t) - w_t \Big) - R \geq 0 \tag{D.13}$$

$$[j=1, s^t;\; \gamma^1_{s^t}]: \; W_{s^t} - u(w_{s^t}) - \beta \mathbb{E}_{s^t} \Big(\lambda U_{s^{t+1}} + (1-\lambda)(\lambda_e p_{s^{t+1}} v^*_{s^{t+1}}) + \widetilde{p}_{s^{t+1}} W_{s^{t+1}}) \Big) \geq 0 \tag{D.14}$$

Still following the notation from Marcet and Marimon (2019), we can define the Saddle Point Problem ($\mathbf{SPP}_\mu$) as:

technically applies to a *different* function j_t (indexed by t).

$$
\begin{aligned}
\mathbf{SPP}_{\mu}: SV_{\mu,t_0}(\mathbf{x}_{t_0}, a_{t_0}) = & \underset{\{\gamma\in\mathbb{R}^l_+\}}{inf} \underset{\{w_{st}, W_{st_0}\}}{sup} \mu^0\Big(f(a_{t_0}, y_{t_0}, h_{t_0}) - w_{t_0}\Big) + \mu^1 W_{t_0} + \\
& + \beta\mathbb{E}_t\left(\phi(\mu,\gamma)\sum_{i=0}^{T-t_0}\left[\beta^{t_0+i}\prod_{i=0}^{T-t_0-1}\widetilde{p}_{t_0+1+i}\left(f(a_{t_0+i}, y_{t_0+i}, h_{t_0+i}) - w_{t_0+i}\right) + W_{t_0+i}\right]\right) + \\
& + \gamma^1\left(u(w_{t_0} + \beta\mathbb{E}_{t_0}\left(\lambda U_{t_0+1} + (1-\lambda)\lambda_e p_{t_0+1} v^*_{y,t_0+1}\right)\right) + \\
& + \gamma^0\left(f(a_{t_0}, y_{t_0}, h_{t_0}) - w_{t_0} - R\right) \qquad \text{(D.15)}
\end{aligned}
$$

The problem can be restated as a saddle-point problem over a Lagrangian equation

$$
\begin{aligned}
\underset{\gamma_t}{inf} \underset{\{w_{st}, W_{st}\}}{sup} & \mu^0\Big(f(a_{t_0}, y_{t_0}, h_{t_0}) - w_{t_0}\Big) + \mu^1 W_{t_0} + \\
& \gamma^0\Big(\big(f(a_{t_0}, y_{t_0}, h_{t_0}) - w_{t_0}\Big) - R) + \\
& \gamma^1_{t_0}\Big(- W_{t_0} + u(w_{t_0}) + \beta\mathbb{E}_{t_0}(\lambda U_{t_0+1} + (1-\lambda)(\lambda_e p_{t_0+1} v^*_{t_0+1} + \widetilde{p}_{t_0+1} W_{t_0+1})\Big) + \\
& \beta\mathbb{E}_{t_0}\Bigg[(\mu_0 + \gamma_0)\sum_{t=t_0+1}^{T}\beta^{t-t_0-1}\prod_{i=0}^{T-t_0-1}\widetilde{p}_{t_0+1+i}\Big(f(a_t, y_t, h_t) - w_t\Big) + \\
& \sum_{t=t_0+1}^{T}\mathbb{E}_t\beta^{t-t_0-1}\prod_{i=0}^{t-t_0-1}\widetilde{p}_{t_0+1+i}\gamma^1_t\Big(- W_t + u(w_t) + \\
& \beta(\lambda U_{t+1} + (1-\lambda)(\lambda_e p_{t+1} v^*_{t+1} + \widetilde{p}_{t+1} W_{t+1})\Big)\Bigg] \qquad \text{(D.16)}
\end{aligned}
$$

which, thanks to some algebra and the law of iterated expectations becomes

$$
\begin{aligned}
\underset{\gamma_t}{inf} \underset{\{w_{st}, W_{st}\}}{sup} & - \gamma^0 R + \mathbb{E}_{t_0}\sum_{t=t_0}^{T}\beta^{t-t_0}\prod_{i=0}^{t-t_0}\widetilde{p}_{t_0+i}\Bigg[\Big(\mu^0_t + \gamma^0_t\Big)\Big(f(a_t, y_t, h_t) - w_t\Big) + \mu^1_t W_t - \\
& \gamma^1_t\Big(W_t - u(w_t) - \beta(\lambda U_{t+1} - (1-\lambda)\lambda_e p_{t+1} v^*_{t+1}\Big)\Bigg] \qquad \text{(D.17)}
\end{aligned}
$$

where $\mu^0_t = \mu^0 = 1$, $\gamma^0_t = \gamma_0 = 0$, $\mu^1_t = \gamma^1_{t-1}$ for some starting $\gamma^1_{t_0-1}$.

The problem can now be written in recursive form. Define

$$
\mathcal{P}_t(h_t, \tau_t, y_t, a_t, \gamma_t) = \underset{W_t}{sup}\, J_t(h_t, \tau_t, y_t, W_t, a_t) + \mu^1_t W_t \qquad \text{(D.18)}
$$

Given **Equation D.17** the SPFE of the problem can be written as

$$\mathcal{P}_t(h_t,\tau_t,y_t,a_t,\gamma_t) = \inf_{\gamma_t} \sup_{w_t} (f(a_t,y_t,h_t) - w_t) + \mu_t^1 W_t - \gamma_t(W_t - u(w_t)) + \beta\mathbb{E}_t(\lambda U_{t+1} + (1-\lambda)\lambda_e p_{t+1} v^*_{t+1}) + \beta\mathbb{E}_t \widetilde{p}_{t+1}\mathcal{P}_{t+1}(h_{t+1},\tau_{t+1},y_{t+1},a_{t+1},\gamma_{t+1}) \tag{D.19}$$

One can easily verify that the solution of this equation is the same we found in the maximization of **Equation 8** in the main text. Take the first order conditions and compute the envelope condition:

$$[FOC\ w_t]:\ -1 + \gamma_t u'(w_t) = 0 \tag{D.20}$$

$$[ENV\ W_t]:\ \frac{\partial \mathcal{P}_t}{\partial W_t} = \mu_t^1 - \gamma_t \tag{D.21}$$

$$[FOC\ W_{t+1}]:\ -\widetilde{p}_{t+1} W_{t+1}\gamma_t + \frac{\partial \widetilde{p}_{t+1}}{\partial W_{t+1}}\mathcal{P}_{t+1} + \widetilde{p}_{t+1}\frac{\partial \mathcal{P}_{t+1}}{\partial W_{t+1}} = 0 \tag{D.22}$$

where **EquationD.22** is obtained by adding and subtracting from **Equation D.19** $\beta\gamma_t\widetilde{p}_{t+1}W_{t+1}$. The reader should also keep in mind that the condition in **Equation D.22** is actually state contingent and applied to *all* future states next period, with a different set of co-states $\gamma_{s^{t+1}}$ for each realization of a_{t+1}.

Some rearranging of the **Equation D.22** leads to the following result

$$\frac{\partial \log \widetilde{p}_{t+1}}{\partial W_{t+1}}\bigg(\mathcal{P}_{t+1} - \gamma_t W_{t+1}\bigg) = \gamma_{t+1} - \mu^1_{t+1} \tag{D.23}$$

which, given the law of motion of the co-states and the definition in **Equation D.18** can be re-written as:

$$\frac{\partial \log \widetilde{p}_{t+1}}{\partial W_{t+1}} J_{t+1} = \frac{1}{u'(w_{t+1})} - \frac{1}{u'(w_t)} \tag{D.24}$$

which is exactly **Equation 14**, namely the Euler equation that governs the behavior of wage setting and disciplines the provision of insurance within the contract.

E Existence of a Block Recursive Equilibrium

In order to show that a Block Recursive Equilibrium (BRE) exists in our model we need to show that the equilibrium contracts, the workers' and the entrepreneurs value and policy functions do not depend on the distribution of employed and unemployed workers.

Most of the results are tightly linked to our search protocol, directed versus random search, and our contracting structure whereby workers have finite lives and therefore contracts end in finite time. The intuition for why directed search is paramount for the existence of a BRE is linked to the fact that, with directed search, workers that are matched with a particular job accept that job with certainty as they are actively looking for it in the labor market. This fact, together with the large (potentially unbounded) supply of vacancies opening each period up to the point in which their expected value is nil, makes it possible to pin down the tightness in each submarket regardless of the distribution of workers across states. This implies that the only element of the aggregate state that matters for a firm when making an hiring decision is the state of aggregate productivity but not the distribution of worker types (e.g. employed vs unemployed).

Proposition E.1. *A block recursive equilibrium as defined in Definition 2.3 exists.*

Proof. We follow the approach in Menzio, Telyukova and Visschers (2016), Herkenhoff, Phillips and Cohen-Cole (2019) and prove the existence of a BRE using backward induction.

Consider the lifetime values of an unemployed and an employed worker before the production stage in the last period of households lives with $\tau = T$:

$$U(h, T, \iota; \Omega) = u(b(h, T)) \tag{E.1}$$

$$V(h, T, \iota; \Omega) = u(w(a)), \tag{E.2}$$

their values trivially do not depend on the distribution of types as both valuations are 0 from $T+1$ onward. Hence, $U(h, T, \iota; \Omega) = U(h, T, \iota; a)$ and $V(h, T, \iota; \Omega) = V(h, T, \iota; a)$.

The optimal contract for agents aged $\tau = T$, instead, solves the following problem

$$J(h, T, W, y; \Omega) = \sup_{w} [f(y, h; a) - w] \quad s.t. \; W = u(w),$$

that clearly does not depend on the distribution of worker types due to the directed search protocol and where the aggregate state only affects the promised utility and the optimal wage through realization of the aggregate productivity processes. Therefore, $J(h, T, \iota, W, y; \Omega) = J(h, T, \iota, W, y; a)$.

This also implies that the equilibrium market tightness

$$\theta(h, T, \iota, W; \Omega) = q^{-1} \left(\frac{c(y)}{J(h, T, \iota, W, y; a)} \right)$$

is independent from the distribution of worker types and it is only affected by realization of aggregate productivity, so $\theta(h, T, W; a)$.

This in turn implies that the search problem workers face at the beginning of the last period of their lives depends on the aggregate state only through aggregate productivity a:

$$R(h, T, \iota, V; a) = \sup_{v} \Big[p(\theta(h, T, \iota, v; a)) \big[v - V \big] \Big],$$

does not depend on the distribution of worker types.

Stepping back at $\tau = T - 1$, the value functions for the unemployed and the employed agents are solutions to the following dynamic programs

$$\begin{aligned}
&\sup_{v} u(b(h, T-1)) + \beta \mathbb{E}_{\Omega,\psi} \Big(U(h', T, \iota; a') + p(\theta(h, T, \iota, v; a')) \big[v - U(h', T, \iota, ; a') \big] \Big) \\
&u(w) + \beta \mathbb{E}_{\Omega,\psi} \left(\begin{array}{c} \lambda U(h', T, \iota; a') + \beta(1-\lambda) W + \\ + \beta(1-\lambda)\lambda_e \max(0, R(h', T, \iota, , W); a')]] \Big] \end{array} \right),
\end{aligned}$$

where both do not depend on the distribution of worker types.

The optimal contract at this step is a solution to

$$\begin{aligned}
J_t(h, T-1, \iota, V, y; a) = \sup_{\{\pi_i, w_i, W_i\}} \sum_{i=1,2} \pi_i \Big(& f(y, h; a) - w_i \\
& + \mathbb{E}_{\Omega,\psi} \left[\widetilde{p}(h', T, W_{i,\Omega'}; a')(J(h', T, y, W_i; a') \right] \Big)
\end{aligned}$$

$$\begin{aligned}
s.t.\ V = \sum_{i=1,2} \pi_i \left(u(w_i) + \mathbb{E}_{\Omega,\psi} \widetilde{r}(h', T, W_i; a') \right),\ h' = \phi(h, y, \iota, \psi) \\
\mathbb{E}_{\Omega,\psi} \sum_{i=1,2} \pi_i \left(\mathbb{E}_{\Omega,\psi} J(h', T, \iota, W_i, y; a') \right) \geq 0 \text{ and } t \leq T
\end{aligned}$$

which does not depend on types distribution.

Therefore, also the equilibrium tightness and the search gain at $T-1$ are independent from types' distributions, as

$$\begin{aligned}
\theta(h, T-1, \iota, W; a) &= q^{-1} \left(\frac{c(y)}{J(h, T-1, \iota, W, y; a)} \right) \\
R(h, T-1, \iota, V; a) &= \sup_{v} \Big[p(\theta(h, T-1, \iota, v; a)) \big[v - V \big] \Big].
\end{aligned}$$

Stepping back from $\tau = T-1, ..., 1$ and repeating the arguments above completes the proof. $\square$

F Additional Figures and Tables

In order to estimate the effect of entering the labor market in a recession we use an age-cohort-period model in which we break the collinearity among the three set of fixed effects by proxying the cohort fixed effects with the cyclical component of real GDP (Hamilton filtered). In particular, we estimate the following yearly model:

$$\log(w)_{t,c,e} = \Phi_t + \Phi_e + \beta_e \widetilde{Y}_c \times \Phi_e + u_{t,c,e}, \tag{F.1}$$

where Φ_t, Φ_e, are dummies for calendar years and labor market experience, and $\widetilde{Y}_c$ is the cyclical realization of real GDP for cohort c at time of their labor market entry. The set of coefficients β_e, therefore, estimate the effect of aggregate conditions on real wages at each year of labor market experience.

Table F.5. Effects of initial aggregate conditions along the experience profile and experience growth profile

Dep.Variable: Log-Wage	Experience × Cycle	Experience
Experience Dummy		
0	1.968	
	(0.452)	
1	1.463	0.150
	(0.238)	(0.011)
2	1.473	0.246
	(0.283)	(0.014)
3	1.239	0.304
	(0.343)	(0.014)
4	1.250	0.342
	(0.349)	(0.015)
5	1.239	0.375
	(0.349)	(0.015)
6	1.301	0.399
	(0.287)	(0.015)
Age FE	✓	✓
Year FE	✓	✓
Sex FE	✓	✓
LLM FE	✓	✓
R^2	0.89	0.89
N	254,000,000	254,000,000

Note: The table reports regression coefficients for the empirical estimates in the data used to construct the profiles in Figure 10.

Table F.6. E-U-E transitions

(a) Data

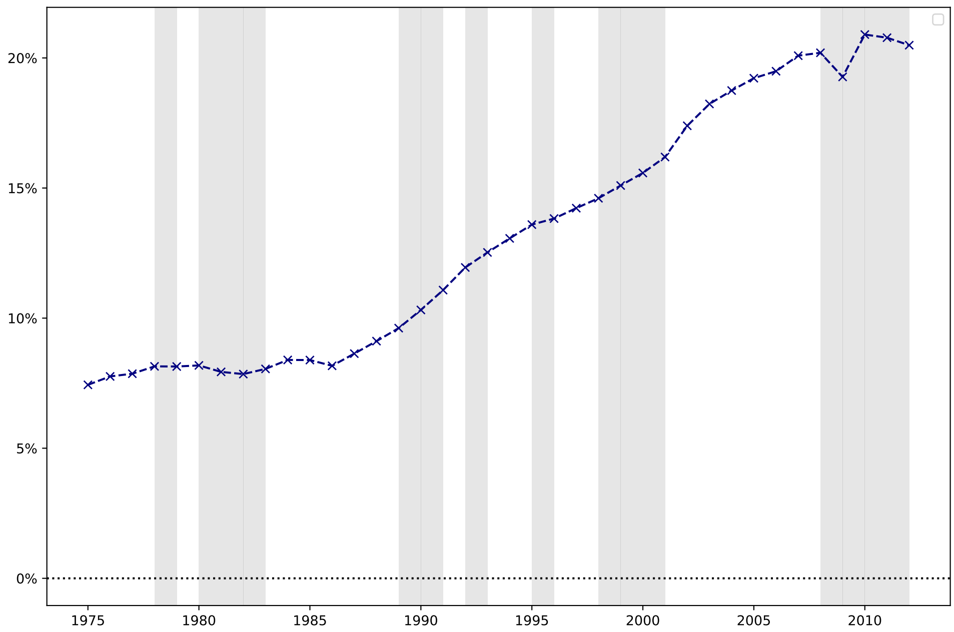
Dep.Variable: Log-wage after E-U-E transition	(1)	(2)
Quality of origin firm (FQ):		
2^{nd} quint.	0.100	0.128
	(0.002)	(0.004)
3^{rd} quint.	0.143	0.210
	(0.002)	(0.004)
4^{th} quint.	0.153	0.182
	(0.002)	(0.004)
5^{th} quint.	0.230	0.260
	(0.002)	(0.004)
Log-wage at origin	0.669	0.609
	(0.001)	(0.002)
Experience controls	✓	✓
Sex FE	✓	✓
Year FE	✓	✓
Contract type FE	✓	✓
Full- & Part-Time FE	✓	✓
Justified dismissals	✓	
R^2	0.48	0.36
N	955,602	338,975

(b) Model

Dep. Variable: Log-wage after E-U-E transition	(1)
Quality of origin firm (FQ):	
$p0 \leq FQ < 75p$	0.743
	(0.007)
$p75 \leq FQ < 100p$	0.783
	(0.008)
Log-wage at origin	0.117
	(0.009)
Controls	✓
R^2	0.069
N	18,669

Note: Standard errors in parentheses. The tables report a specification on datasets based on workers that experience experience an Employment to Unemployment to Employment transitions (E-U-E). In Panel **(a)**, column (2) excludes separations that are justified in the Italian labor law (*giusta causa*). In the model, *Controls* include the pre-transition human capital and a polynomial in labor market experience. Referenced on page(s) [26] .

Figure F.19. Tertiary School Enrollment and the Business Cycle



Note: The figure plots the ratio of students enrolled in tertiary education to population aged 16-29 in every year. Source: Italian National Institute of Statistics (ISTAT). Shaded areas indicate the OECD based Recession Indicators for Italy.