On the Job Search, Human Capital Formation, and Lifecycle Wages

Matthew J. Millington*

October 12, 2023

Click here for latest version.

Abstract

I build an equilibrium lifecycle model of wages that combines on-the-job search in a frictional labor market with human capital accumulation. In the model, heterogeneous workers invest in human capital accumulation and search effort. Wages and their dispersion rise over the life cycle due to two forces: heterogeneous workers climb the job ladder of firm productivity while investing in human capital. I discipline the model using microdata from the SIPP. Using the calibrated model, I find that on-the-job search is the driving force of lifecycle wage growth. In contrast, wage dispersion is mainly a function of heterogeneous human capital formation. I engineer an increase in tax progressivity and find that it decreases wages, primarily due to reduced to on-the-job search effort. Surprisingly, it has little effect on wage dispersion as human capital accumulation and search tend to offset one another.

*I would like to thank Domenico Ferraro, Gustavo Ventura, and Hector Chade for their guidance.

1 Introduction

Macroeconomists tend to focus on two general theories for wage growth: human capital theory and job ladder theory. In human capital theory (Ben-Porath, 1967), workers increase their wages by accumulating skills. The labor market is frictionless, so a worker's wage equals his marginal product of labor. As a worker accumulates human capital, he becomes more productive, which increases his wage. However, in a quantitative model, in order to match wage dispersion in the data (in particular the increase lifecycle profile of wage dispersion), the human capital theory relies on idiosyncratic shocks, heterogeneity in learning ability, and heterogeneity in initial human capital (Huggett et al., 2011).

In job ladder theory (Burdett and Mortensen, 1998), workers increase their wages by making job-to-job transitions to more productive firms. When a worker moves to a more productive firm, he and the firm bargains over his wage. Because the firm is more productive, the firm can pay the worker a higher wage. The theory relies on search frictions; if workers could start in the most productive firm, there would be no wage growth. I refer to this as the search channel.

In this paper, I develop a unifying theory with both paradigms. I write a macroeconomic model of wages with combines endogenous human capital accumulation, endogenous search effort, endogenous job posting, and a life cycle. To my knowledge, a quantitative model with all of these elements has not been studied. Going forward, I refer to the endogenous human capital part as the"human capital channel," and I refer to the search-and-matching part with endogenous search and a job ladder as the "search channel." It is a rich model with which I can perform a number of policy experiments. In this paper, I experiment with tax progressivity.

What do we gain by making both human capital accumulation and search effort endogenous? First, both channels matter empirically (Karahan et al., 2022). It follows that both channels may respond to changes in policy. Furthermore, the economics are realistic. It is reasonable to assume that, to some degree, workers are aware of both channels of wage growth. The interplay between the channels affects worker decisions at the micro level with macroeconomic consequences.

Second, my model allows us to decompose how the channels contribute to wage levels and dispersion as well as how the channels interconnectedly respond to policy changes. A priori, it is not clear how the channels will interact quantitatively. It may be that one channel dominates. Or, it may be that the interaction between both channels increases or decreases the response to tax progressivity. These are quantitative questions which require a quantitative model.

In the model, workers are heterogeneous at the beginning of life in fixed learning ability and initial human capital. Firms are heterogeneous in productivity. When employed by a firm, a worker earns a wage which is the product of three components: the worker's human capital, the firm's productivity, and an endogenous bargaining component (similar to a rental rate of human capital). The bargaining component arises from a surplus sharing rule as in Cahuc et al. (2006).

Workers maximize their lifetime utility of consumption. To do so, workers choose how much effort to invest in human capital accumulation of searching for a new job. Both activities are costly in the sense that the worker experiences disutility. However, both activities benefit the worker because they lead to wage growth and increased future consumption. In the case of human capital investment, a worker accumulates human capital according to a law of motion as in Ben-Porath (1967). When a worker invests in search effort, he increases the probability of meeting an outside firm. Upon meeting an outside firm, the new firm is more productive than his current firm, he will leave his current firm to work for the new firm.

Workers face negative risk from exogenous unemployment shocks, which is necessary for generating the sharp negative income shocks in the data. Low-wage workers are more likely to experience unemployment. Human capital depreciates while a worker is unemployed.

Firms choose how many jobs vacancies to post. In equilibrium, a free entry condition requires that firms are indifferent toward posting another job. From the firm's perspective, the optimal number of job postings depends the distribution of human capital and search effort in the labor market. If workers have more human capital, then matches with workers will be more profitable, incentivizing more job posting. Similarly, if workers invest more effort into search, then firms have a greater probability of converting a vacant job to a filled job, increasing the benefit of job posting and incentivizing more job posting.

I calibrate the model using microdata from the Survey of Income and Program Participation (SIPP) for 1990-2019. The model successfully replicates lifecycle profiles for mean wages, wage dispersion, and job-to-job transitions. Additionally, I target and successfully replicate monthly wage growth for workers who stay at the same job or switch jobs.

For the benchmark model, I estimate an average tax function which replicates the income tax system as well as means-tested transfers in the US. The average tax rate function is the same form as Bénabou (2002); it consists of two parameters, one which determines tax progressivity and one which determines the tax level. When I do counterfactual experiments, I adjust the progressivity parameter and solve for a new equilibrium where the free entry condition and the government budget constraint hold.

Before I conduct policy experiments, I analyze the calibrated benchmark model. I find that, on average, the search channel is the most important driver of lifecycle wage growth; about 70% of mean lifecycle growth is attributable to workers climbing the job ladder, with the remaining coming from an increase in human capital. This result is consistent with Karahan et al. (2022).

However, this masks significant heterogeneity between workers. Workers with high learning ability have a much greater marginal benefit of investing in human capital, so they will invest more in human capital accumulation. Across learning abilities, I see roughly equal job ladder outcomes, but vast differences in human capital. Furthermore, workers with high learning ability accumulate wages faster than those with low learning ability, and the difference is driven entirely by human capital. Thus, over the lifecycle, the increase in wage dispersion is driven by an increase in the variance of human capital. Taken together, we can generalize the model as suggesting that wage growth comes from the job ladder while wage dispersion comes from human capital.

Are human capital accumulation and job search complements or substitutes? There are features in the model going in either direction. In support of complementarity, human capital and firm productivity are complements in production. Therefore, all else equal, the marginal benefit of increasing human capital is greater for a worker at a more productive firm. On the other hand, the disutility that workers experience from investing in human capital and job search is convex. So, a worker who has already invested heavily in human capital will have a greater marginal disutility of investing in job search. In this sense, workers will choose "between" human capital and job search. On a micro level, the latter effect dominates; workers choose "between" human capital and search. However, in the aggregate, there is slight positive correlation between human capital and firm productivity in the benchmark model. It is driven by unemployment; because low-wage workers are more likely to lose their job, they do not climb the job ladder as effectively.

If tax progressivity is increased, it is clear that both the human capital channel and the job search channel will move in the same direction. Greater tax progressivity implies that if a worker increases his wage, he will retain smaller part of the wage increase. Thus, workers are disincentivized from growing their wages. This puts downward pressure on both human capital investment and search effort. There will also be an equilibrium effect; because the distribution of workers has less human capital and workers expend less search effort, firms are discouraged from posting jobs, leading to fewer open vacancies in equilibrium. The extent to which wages decrease, the change in wage dispersion, and the driving forces behind these effects are quantitative questions which the model can answer.

I perform a simple tax policy experiment where I double the progressivity of income taxes, which puts the tax function in line with a country like Sweden. In total, when I increase tax progressivity, the wage level in the economy decreases by about 4%. Approximately 2/3 of the decrease is due to the job search channel, inclusive of the job posting effect, and the remaining 1/3 decrease is due to the human capital channel. Since the job search channel is the main approach for accumulating wages in the economy, it is unsurprising that it is the most responsive channel to policy. About 15% of the decrease is due to a decrease in vacancy posting, implying that equilibrium effects are of some significance.

The increase in tax progressivity also decreases lifecycle wage growth, but only slightly. The decrease is driven by less human capital accumulation for those at the top of wage distribution. Compared to a pure human capital model, this is a small effect.

In a clear example of how the interaction of the two channels affects policy, an increase in tax progressivity does not affect the variance of wages. The reason is that the human capital channel and the job search channel push wage variance in opposite directions, and the effects offset one another. In a pure human capital model, an increase in tax progressivity decreases wage dispersion. The increase in progressivity has more "bite" at the top of wage distribution. So, wages at the top of the distribution decrease more than wages at the bottom of the distribution, and the gap between high wages and low wages decreases. In contrast, in a pure job ladder model, an increase in tax progressivity decreases wage dispersion. Because workers exert less search effort, there is less bunching at high-productivity firms, and the the variance of wages is more spread. Simply put, since there are fewer job-to-job transitions, workers are more likely to remain an unproductive firms. Taken together, the result casts some doubt on the notion that increasing tax progressivity can decrease wage inequality.

The remainder of the paper proceeds as follows. I describe the model in Section 2 and discuss the calibration in Section 3. I analyze the benchmark model in Section 4 before moving to taxation experiments in Section 5. Section 6 concludes.

2 Model

2.1 Life Cycle

For computational efficiency, the model is in continuous time. I therefore use a stochastic lifecycle. There are $I+1$ stages of life, I working stages and a retirement stage. Workers transition from stage i to $i+1$, $i \in \{1,...,I\}$, with probability ζ. In the retirement stage, workers earn a flat social security payment, SS, and die with probability $\overline{\zeta}$, after which they are replaced by newborns in the first working stage. To fix ideas, when I calibrate the model, I set $I=4$ and calibrate the model to ages 23 to 65. So, each stage of life is approximately a decade.

2.2 Wages

Workers are heterogeneous in human capital h; firms are heterogeneous in productivity p. A match between a worker with human capital h and a firm with productivity p produces hp of the numeraire consumption good. Note that human capital and firm productivity are complements in production. Of the total production of the match, workers earn a piece-rate $r \in [0,1]$. Therefore, before taxes and transfers, the worker earns the wage hpr and the firm earns profit $hp(1-r)$. r is endogenously negotiated between workers and firms as described below.

2.3 Utility

Workers are either employed or unemployed and discount the future at discount rate ρ. They enjoy utility from consumption, $u(c)$, where consumption equals the worker's after-tax-and-transfer wage. Given the average tax rate function $T(hpr)$ (explained below), $c = [1-T(hpr)]hpr$.

Employed workers choose how much effort to invest in human capital accumulation, l (for "learning"), and job search, s. Unemployed workers can search but cannot accumulate in human capital.[1] Effort of either type is costly in that workers experience disutility from effort. For the employed, disutility of effort is $d_E(l+s)$; for the unemployed, disutility of effort is $d_U(s)$.

[1] I assume that unemployed workers cannot accumulate human capital in order to better match the data.

Altogether, employed workers experience utility $u(c) - d_E(l+s)$, and unemployed workers experience utility $u(c) - d_U(s)$. Workers face a clear tradeoff: l and s increase wages and consumption tomorrow at the expense of disutility today.

2.4 Human Capital Accumulation

Workers accumulate human capital as in Ben-Porath (1967). At the time of birth, workers are heterogeneous in fixed learning ability, $a > 0$. Human capital evolves according to the law of motion

$$\frac{dh}{dt} = a(lh)^{\omega} - \delta h \tag{1}$$

where $\omega \in (0, 1)$ governs the level of decreasing returns to human capital accumulation and $\delta \in [0, 1]$ is the human capital depreciation rate. All else equal, workers with a higher a have greater returns to human capital investment.

2.5 Labor Market

All workers, employed or unemployed, meet open vacancies at rate $sm(\theta)$ where s is search effort as described above, $m(\theta)$ is a meeting function, and θ is the labor market tightness ratio. θ is defined as the number of vacancies v per unit of aggregate search effort S, $\theta = v/S$.[2] Therefore, when a worker exerts search effort, he increases his probability of meeting an outside firm. Search is random in the sense that workers and firms randomly meet in a single market and neither party can direct their search toward certain types of firms or workers.

Upon meeting an open vacancy, the worker observes the associated firm's productivity, p', drawn from a distribution $F(p')$ over $[\underline{p}, \overline{p}]$. The worker and new firm then bargain over the wage. Since h and p are determined, bargaining is over the piece rate r. They will form a match if the worker is better off with the new firm.

Each firm has a maximum of one worker. So, if a match is destroyed, the firm's production drops to zero. Firms post and maintain jobs at cost κ. From the firm's perspective, an open job vacancy meets a worker with probability $m_f(\theta)$. Matches are subject to exogenous job destruction shocks which are a function of wage, $\Lambda(hpr)$.[3]

Unemployed workers earn unemployment benefits bw where b is the replacement rate (typically 0.5 in the US) and w is the most recent wage the worker was earning before becoming unemployed. Unemployment benefits expire with probability χ, after which the worker gets a transfer payment T_0. I denote workers without unemployment benefits with $w = 0$.

[2] S is mathematically defined in Appendix A.1 Equation (21).

[3] I assume that the job destruction rate depends on the wage for two reasons: (i) it is a clear feature in the data, and (ii) the increased risk of unemployment for low-wage workers is key for accounting for increasing wage dispersion over the lifecycle (Karahan et al., 2022).

2.6 Wage Bargaining

I use the wage bargaining protocol of Cahuc et al. (2006) as applied in Bagger et al. (2014). The protocol determines the values of r when a worker and a firm form a match and the reservation p's such that a worker will accept a job. All the resulting objects are are functions which depend on characteristics of the worker and firm(s). This section develops the functions. Behind every result is a surplus sharing rule where, if the worker has the opportunity to go to a new firm, he extracts the total surplus his outside option plus a fraction $\eta \in [0, 1]$ of the additional worker surplus at the winning firm.

Let $E_{ai}(h, p, r)$ be the value of employment for a worker with learning ability a, in stage of life i, with human capital h, and currently working for a firm with productivity p at the piece rate r. I define $E_{ai}(h, p, r)$ mathematically in Section 2.7. Currently, the worker earns the wage hpr. The greatest wage that the worker can earn while working at the current firm is hp, the wage when $r = 1$. If $r = 1$, then the firm makes zero profit. In equilibrium, the value of every firm's outside option is 0. So, if $r = 1$, the firm is indifferent between maintaining the match or not.

First, consider the scenario of a meeting between an employed worker and an outside firm with productivity p'. The incumbent firm and the outside firm will commence Bertrand competition for the worker, making alternating bids on the worker's wage. As the firms offer progressively higher wages, they reach a point where the firm with lower p cannot pay the worker a higher wage without earning negative profit; for the firm with lower productivity, that wage is given by $r = 1$. At that point, the firm with higher productivity can offer a marginally larger wage and win the worker.

Suppose that the outside firm has greater productivity that the incumbent firm, $p' > p$. In this case, the worker will be poached and will make a job-to-job transition to the outside firm. The poaching firm will pay the worker a piece rate $R^P_{ai}(h, p, p')$ which solves

$$E_{ai}\left(h, p', R^P_{ai}(h, p, p')\right) = E_{ai}(h, p, 1) + \eta \left[E_{ai}(h, p', 1) - E_{ai}(h, p, 1)\right]. \tag{2}$$

After getting poached by the firm with productivity p', the worker earns a piece rate $r = R^P_{ai}(h, p, p')$ such that the value of employment equals the maximum surplus possible from the incumbent firm, $E_{ai}(h, p, 1)$, plus a fraction η of the additional worker surplus created from the match.[4]

[4] A slight clarification is in order. The worker earns a fraction η of the additional *potential worker* surplus from the match, not the additional *total* surplus from the match. This is slightly different from the original scheme in Cahuc et al. (2006). Cahuc et al. (2006) models the bargaining process "using a version of the Rubinstein (1982) infinite-horizon alternating-offers bargaining game." In Cahuc et al. (2006), workers have linear utility over the wage, as do firms; thus, $E_{ai}(h, p, 1)$ is equivalent to the total surplus of the match. So, we can say that the worker earns a fraction η of the additional total surplus of the match. The same is true in Engbom (2022). However, as Bagger et al. (2014) points out, when workers have curvature in utility and firms have linear utility, the total amount of surplus from the match is not independent of r and therefore not fixed. So, the present scheme may not be a Nash equilibrium. My case is further complicated by the fact that workers also have disutility over effort. I elect to follow Bagger et al. (2014) (as well as Karahan et al. (2022)) and impose this wage structure even with curvature in utility. For an approach which uses total surplus but allows for curvature in utility, see Lise et al. (2016).

Next, suppose that the outside firm has lower productivity than the incumbent firm, $p' < p$. In this case, the incumbent firm will retain the worker because it can pay the worker a greater wage while remaining profitable. However, if the outside firm has a high enough p', it is possible that the outside firm could pay the worker a greater wage than it earns now. In this case, the outside firm triggers a renegotiation between the worker and the incumbent firm. The worker will stay at the current firm but get wage increase; the worker will earn a piece rate $R^R_{ai}(h, p, p')$ which solves

$$E_{ai}\left(h, p, R^R_{ai}(h, p, p')\right) = E_{ai}(h, p,' 1) + \eta \left[E_{ai}(h, p, 1) - E_{ai}(h, p', 1)\right]. \tag{3}$$

The worker earns a piece rate $r = R^R_{ai}(h, p, p')$ such that value of employment is the maximum of the what the outside firm could offer, $E_{ai}(h, p', 1)$, plus a fraction of the additional worker surplus. Note that the worker's outside option is the maximum surplus from the unsuccessful outside firm.

The third and final possibility is that the outside firm has a lower productivity than the incumbent firm, $p' < p$, and the outside firm cannot pay the worker a greater wage even if it offers the worker the maximum $r = 1$. In this case, the worker discards the outside firm and stays at the current firm for the same r.

$q_{ai}(h, p, r)$ defines the minimum p' such that, if the worker meets an outside firm with $p' \geq q_{ai}(h, p, r)$, the meeting will trigger a renegotiation with the current firm. $q_{ai}(h, p, r)$ solves

$$E_{ai}(h, p, r) = E_{ai}\left(h, q_{ai}(h, p, r), 1\right) + \eta \left[E_{ai}(h, p, 1) - E_{ai}\left(h, q_{ai}(h, p, r), 1\right)\right]. \tag{4}$$

To summarize the scenario where an employed worker meets an outside firm: If $p' < q_{ai}(h, p, r)$, the outside firm cannot compete with the current firm and the worker stays with the same firm at the same piece rate. If $q_{ai}(h, p, r) < p' < p$, the worker stays at their current firm but leverages the outside offer into a greater wage with piece rate $R^R_{ai}(h, p, p')$. And if $p' > p$, the worker is poached, and the worker earns the piece rate $R^P_{ai}(h, p, p')$.

Second, consider the scenario of a meeting between an unemployed worker and a firm with productivity p'. In this case, the firm is not competing against another firm, but is competing against the worker's outside option of remaining unemployed. Let the $U_{ai}(h, w)$ be the value of unemployment for a worker with learning ability a, in stage of life i, with human capital h, and who previously earned the pre-tax wage w. Suppose that the outside firm has high enough p' such that they can make an offer which will entice the worker to leave unemployment and form a match. The worker earn a piece rate $R^U_{ai}(h, w, p')$ which solves

$$E_{ai}\left(h, p', R^U_{ai}(h, w, p')\right) = U_{ai}(h, w) + \eta \left[E_{ai}(h, p', 1) - U_{ai}(h, w)\right]. \tag{5}$$

The piece rate is set such that the worker gets the value of his outside option, $U_{ai}(h, w)$, plus a fraction η of the additional worker surplus from the match.

It is possible that a firm cannot make the unemployed worker better off, even if the firm pays the worker $r = 1$. $z_{ai}(h, w)$ is the lowest value of p' which will entice the worker to leave unemployment. $z_{ai}(h,w)$ solves[5]

$$U_{ai}(h,w) = E_{ai}\left(h, z_{ai}(h,w), 1\right). \tag{6}$$

2.7 Hamilton-Jacobi-Bellman Equations for Workers

The value of employment for a worker with learning ability a, in stage of life i, with human capital h, and currently working for a firm with productivity p at the piece rate r solves

$$\begin{aligned}
\rho E_{ai}(h,p,r) = \max_{l,s}\; & u(c) - d_E(l+s) + \left(a(lh)^{\omega} - \delta h\right)\frac{\partial E_{ai}(h,p,r)}{\partial h} \\
&+ \Lambda(hpr)\left[U_{ai}(h,hpr) - E_{ai}(h,p,r)\right] + \zeta\left[E_{a,i+1}(h,p,r) - E_{ai}(h,p,r)\right] \\
&+ sm(\theta)\Bigg(\int_{q_{ai}(h,p,r)}^{p}\Big[E_{ai}\left(h,p,R^R_{ai}(h,p,p')\right) - E_{ai}(h,p,r)\Big]\,dF(p') \\
&+ \int_{p}^{\overline{p}}\Big[E_{ai}\left(h,p',R^P_{ai}(h,p,p')\right) - E_{ai}(h,p,r)\Big]\,dF(p')\Bigg)
\end{aligned} \tag{7}$$

subject to

$$c = [1 - T(hpr)]hpr \tag{8}$$

An employed worker chooses how much effort to invest in skill accumulation and search. The worker consumes the value of his after-tax wage and experiences the disutility associated with human capital accumulation and search effort. With probability $\Lambda(hpr)$, the match is destroyed and the worker becomes unemployed, and with probability ζ, the worker ages to the next stage of life. With probability $sm(\theta)$, the worker meets an outside firm with productivity p' drawn from the distribution $F(p')$. If $p' \in [q_{ai}(h,p,r), p]$, the worker leverages the outside offer and renegotiates a higher wage with the incumbent firm. If $p' \in (p, \overline{p}]$, the worker is poached.

The value of unemployment for a worker with learning ability a, in stage of life i, with human capital h, and who previously earned the pre-tax wage w solves

$$\begin{aligned}
\rho U_{ai}(h,w) = \max_{s}\; & u(c) + d_U(s) - \delta h\frac{\partial U_{ai}(h,w)}{\partial h} \\
&+ \zeta\left[U_{a,i+1}(h,w) - U_{ai}(h,w)\right] + \chi\left[U_{ai}(h,w) - U_{ai}(h,0)\right] \\
&+ sm(\theta)\int_{z_{ai}(h,w)}^{\overline{p}}\Big[E_{ai}\left(h,p',R^U_{ai}(h,w,p')\right) - U_{ai}(h,w)\Big]\,dF(p')
\end{aligned} \tag{9}$$

[5]The algebra is as follows: when faced with a firm of productivity $z_{ai}(h,w)$, the unemployed worker is indifferent between remaining unemployed and working for the firm at the highest possible wage, so

$$U_{ai}(h,w,I_b) = U_{ai}(h,w,I_b) + \eta\left[E_{ai}\left(h, z_{ai}(h,w,I_b), 1\right) - U_{ai}(h,w,I_b)\right],$$

which simplifies to (6).

subject to

$$c = \mathbb{1}[w > 0](1 - T(bw))bw + \mathbb{1}[w = 0]T_0 \tag{10}$$

where $\mathbb{1}$ is an indicator function. An unemployment worker chooses how much effort to invest in search, enjoys utility from consumption, and endures disutility from search effort. The worker's consumption level depends of if he still qualifies for unemployment benefits. With probability ζ, the worker ages to the next stage of life, and with probability χ, his unemployment benefits expire. With probability $sm(\theta)$, the worker meets an outside firm with productivity p', and if $p \in [z_{ai}(h, p, r), \overline{p}]$, the worker will accept a job offer.

Finally, the value of retirement, $\overline{E}$, solves

$$\left(\rho + \overline{\zeta}\right) \overline{E} = u(SS). \tag{11}$$

Employed and unemployed workers transition to retirement according to

$$E_{a,I+1}(h, p, r) = \overline{E}. \tag{12}$$

and

$$U_{a,I+1}(h, w) = \overline{E}. \tag{13}$$

2.8 Firms

Firms have linear utility over after-tax profit. There is a flat tax on profits τ_b, so after-tax profits are $(1 - \tau_b)hp(1 - r)$.

I assume free entry in the labor market. In equilibrium, firms post a number of vacancies v such that firms are indifferent with respect to the marginal job posting. The optimal level of job vacancies will depend on the distribution of workers in the economy. All else equal, if all workers in the economy increase their human capital, then because h and p are complements, filled jobs become more profitable, the benefit of job posting increases, and firms post more jobs. Similarly, if workers increase search effort, then firms have a greater probability of converting an open vacancy to a filled job, the benefit of job posting increases, and firms will post more jobs.

Let $l_{ai}(h, p, r)$, $s_{Eai}(h, p, r)$, and $s_{Uai}(h, w)$ denote the policy functions for employed workers' human capital investment, employed workers' search effort, and unemployed workers' search effort, respectively. For a firm with productivity p, the value of a filled job with a worker of learning ability a, in stage of life i, with human capital h, and earning the piece rate r solves

$$\begin{aligned} \rho J_{ai}(h, p, r) = (1 - \tau_b)hp(1 - r) + (a(l_{ai}(h, p, r)h)^{\omega} - \delta h) \frac{\partial J_{ai}(h, p, r)}{\partial h} \\ + \zeta \left[J_{a,i+1}(h, p, r) - J_{ai}(h, p, r)\right] \\ + s_{Eai}(h, p, r)m(\theta) \int_{q_{ai}(h,p,r)}^{p} \Big[J_{ai}\left(h, p, R^R_{ai}(h, p, p')\right) - J_{ai}(h, p, r)\Big] \, dF(p') \\ + \Big[s_{Eai}(h, p, r)m(\theta) \left(F(\overline{p}) - F(p)\right) + \Lambda(hpr)\Big](-J_{ai}(h, p, r)) \end{aligned} \tag{14}$$

If the worker retires, the firm gets 0 profit,

$$J_{a,I+1}(h,p,r) = 0 \tag{15}$$

Note that Equation (14) resembles Equation (7).

If the match is destroyed, the firm is left with value 0. There are three ways the match can be destroyed: (1) the worker is poached by a firm with higher productivity, which occurs with probability $(s_{Eai}(h,p,r)m(\theta)\left(F(\overline{p}) - F(p)\right)$; (2) the jobs is exogenously destroyed at rate $\Lambda(hpr)$; or (3) the worker retires.

Let $\Psi_E(h,p,r|ai)$, $\Psi_U(h,w|a,i)$, and $\Psi_R(a)$ denote the distributions of employed, unemployed, and retired workers, respectively.[6] The free entry condition is

$$\begin{aligned}\kappa = m_f(\theta)\Bigg[&\sum_{a,i}\int_{\Psi_E} s_{Eai}(h,p,r)\int_p^{\overline{p}} J_{ai}\left(h,p',R^P_{ai}(h,p,p')\right)dF(p')d\Psi_E(h,p,r|a,i)\\ &+\sum_{a,i}\int_{\Psi_U} s_{Uai}(h,w)\int_{z_{ai}(h,w)}^{\overline{p}} J_{ai}\left(h,p',R^U_{ai}(h,w,p')\right)dF(p')d\Psi_U(h,w|a,i)\Bigg].\end{aligned} \tag{16}$$

The left had side of Equation (16), κ, is the cost of posting and maintaining an open vacancy. The right hand side of Equation (16) is the expected benefit of posting an open vacancy. It consists of two terms, both multiplied by the $m_f(\theta)$, the probability that an open job vacancy meets a worker. The first term is the probability and expected value of poaching an employed worker; the second is the probability and expected value of hiring an unemployed worker.

2.9 Government Budget Constraint

When I perform counterfactual experiments with tax policy, I discipline the model such that a government budget constraint must hold in equilibrium. Mathematically, the government budget constraint is

$$\begin{aligned}&\sum_{a,i}\int_{\Psi_E} T(hpr)hpr\,d\Psi_E(h,p,r|a,i) + \sum_{a,i}\int_{\Psi_U} T(bw)bw\,d\Psi_U(h,w|w>0,a,i)\\ &\quad+\sum_{a,i}\int_{\Psi_E}\tau_b(1-r)hp\,d\Psi_E(h,p,r|a,i) = \sum_{a,i}\int_{\Psi_U} bw\,d\Psi_U(h,w|w>0,a,i)\\ &\qquad\qquad+\sum_{a,i}\int_{\Psi_U} T_0\,d\Psi_U(h,w|w=0,a,i) + \sum_a \Psi_R(a)SS + \overline{g}.\end{aligned} \tag{17}$$

[6]These distributions are defined such that

$$1 = \sum_{a,i}\int_{\Psi_E} d\Psi_E(h,p,r|a,i) + \sum_{a,i}\int_{\Psi_U} d\Psi_U(h,w|a,i) + \sum_a \Psi_R(a).$$

Since the distribution of unemployed workers is discontinuous at $w=0$, integrating over w is a slight abuse of notation.

The left hand side consists of tax revenue from employed workers, unemployed workers, and firms. Strictly speaking, since those at the low end of the wage distribution pay negative taxes (receive means-tested transfers), there are also government outlays in the left hand side of (17). The right hand side consists of outlays in the form of unemployment benefits, transfers to workers without unemployment benefits, social security payments, and government spending, $\overline{g}$. I calculate $\overline{g}$ in my benchmark economy such that the government budget constraint holds, and I assume that the government must spend $\overline{g}$ in counterfactual experiments.

3 Calibration

3.1 Data

To parameterize the model, I rely on microdata from the Survey of Income and Program Participation (SIPP). The SIPP is a panel data set with interviews every four months where respondents report on what occurred in the time between interviews.[7] I use every panel between the years 1990 and 2019, 12 panels in total.[8]

Two features of the SIPP make it convenient for my setting. First, respondents report their earnings and hours every month, which allows me to observe hourly wages across months. Second, the SIPP tracks worker-job matches by assigning job IDs to workers, which allows me to observe job switches.[9]

I restrict the data to males between the ages of 23 and 65 who are never out of the labor force. I convert earnings data to hourly wages. For every statistic, I calculate the mean within the panel, then take the mean across panels where each panel is weighted by the number of months it covers. Though the SIPP has weekly labor for indicators, I aggregate to a monthly frequency and use the labor force indicator for the second week of the month to mirror the CPS.

3.2 Functional Forms and Distributions

The utility of consumption is logarithmic, $u(c) = \ln(c)$. The disutility of effort is convex with different curvature parameters for the employed and unemployed. So, $d_E(l+s) = \phi(l+s)^{1+\gamma_E}$ and $d_U(s) = \phi s^{1+\gamma_U}$ with $\phi > 0$ and $\gamma_j > 0$ for $j \in \{E, U\}$.

I use a Cobb-Douglas matching function. So, the total number of meetings in a period is $\xi S^{\alpha} v^{1-\alpha}$ with $\xi > 0$ and $\alpha \in (0, 1)$. From the perspective of a worker, the probability of meeting an open vacancy per unit of search effort is $m(\theta) = \xi\theta^{1-\alpha}$. For a firm, the probability of meeting a worker is $m_f(\theta) = \xi\theta^{-\alpha}$.

[7] In 2018, the SIPP transitioned from interviewing respondents every four months to interviewing respondents every year and using an event history design. I use three panels with this design.

[8] I use the 1990, 1991, 1992, 1993, 1996, 2001, 2004, 2008, 2014, 2018, 2019, and 2020 panels.

[9] I convert monthly transfer rates (unemployment to employment, employment to unemployment, and job switches) to continuous-time arrival rates and correct for time aggregation bias using the methods developed in Mukoyama (2014).

I assume that the distribution of firm productivity, $F(p)$, is Pareto with level parameter μ_p and tail parameter $1/\lambda_p$. Following Badel et al. (2020), learning ability a is drawn from a Pareto lognormal distribution, $a \sim PLN(\mu_a, \sigma_a, 1/\lambda_a)$, where μ_a is the level parameter, σ_a is the dispersion parameter, and $1/\lambda_a$ is the tail parameter.[10] Then, the distribution of initial human capital, h_0, is a linear function of a,

$$\ln(h_0) = \beta_0 + \beta_1 \ln(a) + \ln(\varepsilon)$$

with $\varepsilon \sim LN(0, \sigma_\varepsilon)$.[11] I describe how I discretize the distribution of a in Appendix C.2.

I make one final adjustment to the model. To help the model fit the data, I allow for "godfather shocks" (Dorn, 2018). With probability ψ, an employed worker experiences a godfather shock, which means that they meet an outside firm with productivity $p' \sim F(p')$ and must accept a job from the new firm.[12] Though ad-hoc, godfather shocks are useful for generating one feature in the data: of all workers who switch jobs, 30% earn a lower wage in the new job. In my model, it is impossible for a worker to make a job-to-job transition without a wage increase. All equations which are affected by the godfather shock are updated in Appendix A.2.

3.3 Taxes and Transfers

I assume a progressive average tax function as in Bénabou (2002) which subsumes both income taxes and means-tested transfers in the US. At wage hpr, workers pay the average tax rate

$$T(hpr) = 1 - \tau_0(hpr)^{-\tau_1}. \tag{18}$$

Workers with low wages pay negative taxes, which means that their means-tested transfer payments exceed taxes paid. τ_1 determines the progressivity of the tax and transfer system while τ_0 determines its level. Later in the paper, I will experiment with adjusting τ_1. The mean wage is normalized to 1.

To estimate τ_0 and τ_1, I combine an estimate of the average tax function (Guner et al., 2014) with an estimate of a means-tested transfer function (Guner et al., 2020).[13] I apply both functions to an income grid, then estimate τ_0 and τ_1. As is typical, the Bénabou (2002) tax function fits the data well, even when we take transfers into account. The key difference is that including transfers significantly increases progressivity compared to a function without transfers. I estimate $\tau_0 = 0.899$ and $\tau_1 = 0.120$.

[10] There is more than one type of Pareto lognormal distribution. I use the type which consists of a lognormal distribution with a Pareto right tail.

[11] In the appendix, Badel et al. (2020) show that this implies that h_0 is also distributed Pareto lognormally.

[12] The worker is made "an offer he can't refuse."

[13] I adopt this strategy because I could not find an estimated off-the-shelf tax-and-transfer function which satisfies my conditions. Many tax functions do not include transfers, and those that include transfers may include unemployment benefits or social security payments. I am primarily interest in mean-tested transfers. I choose to combine both taxes and transfers intro a single function so that there a single progressivity parameter to adjust for counterfactuals.

For my estimated tax function from Guner et al. (2014), and I use the estimate for all households (married and unmarried) which takes the earned income tax credit into account.[14] For my estimated means-transfer function from Guner et al. (2020), I use an estimate for all households which takes into the following programs: WIC (the Special Supplemental Nutrition Program for Woman, Infants, and Childen), SSI (Supplemental Security Income, for those with disabilities), SNAP (Supplemental Nutrition Asisstance Program, formerly knows as Food Stamps), TANF (Temporary Assistance for Needy Families), and housing. It does not include Medicaid.

The transfer payment made to workers without unemployment benefits is estimated from Guner et al. (2020) as $T_0 = 0.121$. The flat business tax rate is $\tau_b = 0.243$ as estimated in Cooper et al. (2016).

3.4 External Parameters

External parameter choices are summarized in Table 1. The model is period is monthly. I set the number of working stages of life to $I = 4$ for ages 23 to 65. So, each stage is $(65 - 23)/4 = 10.5$ years, or 126 months, and the rate of transitioning from one stage of life to the next is $\zeta = 1/126$. I set retirement to be 10 years, so $\overline{\zeta} = 1/(10 \times 12)$.

The replacement rate of unemployment benefits is set to $b = 0.5$ in accordance with typical unemployment benefits the US. The expiration rate of unemployment benefits is $\chi = 1/6$ in keeping with the typical rule that unemployment benefits can be collected for a maximum of six months. I normalize $SS = .5$.[15] $\alpha = 0.5$ (Petrongolo and Pissarides, 2001) and $\eta = 0.4$ (Bagger et al., 2014). κ is chosen such that θ is normalized to one in the benchmark equilibrium. I set $\rho = 0.0033$ in accordance with a 4% risk-free annual real interest rate, and I set $\delta = 0.00285$ to match the decline in wages at the end of life.[16] β_0 is normalized such that the lowest possible h_0 is the bottom point on the human capital grid.

Finally, I estimate the job loss function $\Lambda(hpr)$ directly the SIPP. $\Lambda(hpr)$ is presented graphically in Appendix C.1.

3.5 Targeted Moments

I am left with 13 parameters to calibrate internally. Internal parameters are estimated jointly such that the simulated model hits 16 moments from the SIPP. Parameters are estimated using simulated method of moments.[17]

[14] I use the power function specification in Table A5 in the appendix.

[15] The choice of SS is immaterial to my results. If social security payments are equal across workers and workers have no choice but to transition to retirement eventually, retirement does not affect choices during working life.

[16] Assume that workers do not invest in human capital in the last stage of life, $i = I$. Without job switching or unemployment, wage growth is $1 - \delta$ per month. If there are x months in the last stage of life, then $\frac{w_I}{w_{I-1}} = (1 - \delta)^x$. Since each working stage is 126 months on average, and given that $w_I / w_{I-1} = 0.864$ in the data, we have $\delta = 0.00116$.

[17] I solve the SMM minimization problem using MIDACO, a general-purpose ant colony optimization algorithm (Schlüter et al., 2009).

Table 1: Externally calibrated parameters

Parameter	Meaning	Value	Explanation/source
Lifecycle			
I	Stages of life	4	By choice
ζ	Transition pobability from one stage to the next	$\left(\frac{42\times12}{I}\right)^{-1}$	Working for 42 years on average (ages 23-65)
$\overline{\zeta}$	Probability of death for the retired	$(10\times12)^{-1}$	Retired for 10 years on average
Policy			
τ_0	Tax+transfer progressivity	0.899	Guner et al. (2014, 2020)
τ_1	Tax+transfer level	0.120	Guner et al. (2014, 2020)
T_0	Transfer for worker without UI	0.121	Guner et al. (2020)
τ_b	Business tax rate	0.243	Cooper et al. (2016)
b	Unemployment benefit replacement rate	0.5	Standard in US
χ	Unemployment benefit expiration rate	1/6	Standard US maximum of 6 months
SS	Social security payment	0.5	Normalization
$\overline{g}$	Government spending	0.081	Equalizes government budget constraint in benchmark equilibrium
Search			
α	Meeting function elasticity	0.5	Petrongolo and Pissarides (2001)
η	Worker's bargaining power	0.4	Bagger et al. (2014)
κ	Job posting cost	0.124	Normalizes benchmark equilibrium θ to 1
Other			
ρ	Discount rate	0.00330	4 percent annual interest rate
δ	Human capital depreciation rate	0.00116	Matches decline of wages in last stage of life
β_0	Initial human capital intercept	0.324	Normalized such the lowest h_0 is the bottom point on the h grid

I target lifecycle profiles of the mean log wage, variance of log wages, and the job switching rate. For each profile, I target the starting point, the ending point, and the midpoint. I also target five moments which are not associated with the lifecycle: the average unemployment to employment rate, monthly wage growth for those who stay in the same job, monthly wage growth for those who switch jobs, monthly wage growth for those who switch jobs and increase their wage, and cross-sectional log wage skewness. Finally, I target two normalizations: I normalize the mean wage to 1 (in order to be consistent with the job destruction function and tax function, which are functions of wage levels), and I normalize the mean $l + s_E$ equal to 0.1.[18]

3.6 Identification

Before proceeding to my calibration results, I provide an informal discussion about identification. Since I use SMM, all parameters are jointly determined, and most parameters affect more than one moment. Nevertheless, I will describe how each parameter relates to a moment in the data, with the goal of showing that my internal parameters are properly identified. I follow a logic-based approach. The 13 internal parameters are listed and described in Table 3.

First, there are some one-to-one relationships between moments and parameters. There is only one moment relating to unemployment, the unemployment-to-employment rate, and one parameter relating to the unemployed, γ_U. The level parameter for the firm productivity distribution, μ_p is immaterial except for establishing the mean wage in the economy once all other parameters haven been determined. So, μ_p is used to normalize mean log wages to 1.

The two moments associated wage growth for job switchers (wage growth for all job switchers and wage growth for job switchers where the wage increased) are determined by two parameters, λ_p and ψ. λ_p determines the size of wage jumps that workers experience while changing jobs. Conditional on λ_p, ψ identifies the difference between wage growth for those switching jobs with a wage increase and those switching jobs with a wage decrease.

Given the parameters and moments accounted for to this point, there are four parameters - σ_a, λ_a, β_1, and σ_ε - which determine the overall variance level of wages, the lifecycle profile of the variance of wages, and the skewness of wages. Given that λ_a determines the skewness of a, we can associate the skewness of wages with λ_a. With λ_p accounted for, σ_a is identified by the overall level of variance of wages. The lifecycle profile of variance is thus identified by β_1 and σ_ε. Conditional on the distribution of learning ability, these parameters tell us the initial distribution of human capital. In particular, β_1 is closely related to the rate of increase of variance of wages over the lifecycle, and conditional on β_1, σ_ε determines the variance of wages in the first stage of life.

Given the shrinking list of parameters and moments that have not been accounted for, μ_a is most directly tied to the wage growth rate of job stayers, which then leaves ω as the

[18] The final normalization is not necessary, but it is convenient because it keeps other parameter levels contained. Since I have level parameters on disutility (ϕ), the return to human capital investment (a), and the return to job search effort (ξ), l and s can be any level and parameters will adjust to get an equally tight fit. In other words, one of these three parameters can be set to anything, and the other two will adjust.

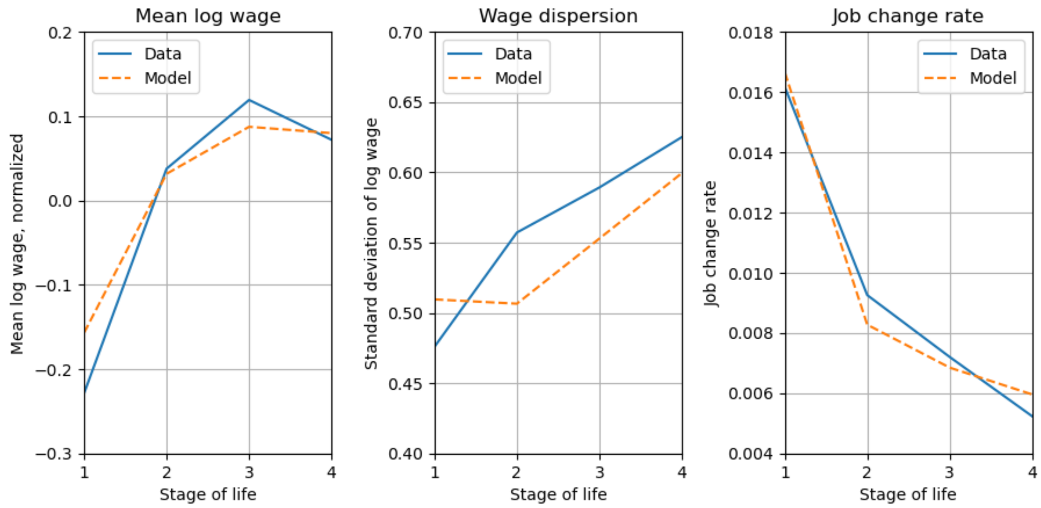

Figure 1: Fit between data and model, lifecycle moments

most directly associated parameter with the shape wages over the lifecycle.

There are three parameters left for three moments: the parameters are ϕ, γ_E, and ξ, and the moments are the level of the job switching rate, the lifecycle profile of the job switching rate, and mean human capital investment plus mean search effort for the employed. ϕ and γ_E are closely related with the normalization of $s_E + l$. All three parameters are closely related to job switching; conditional on the a and ω, the marginal benefit of search is determined by ξ, and the costs are determined by ϕ and γ_E.

A final, general note. It is important that I match monthly wage growth both for job stayers and job switchers, because this allows me to differentiate the month-to-month wage growth which comes from the human capital channel versus the search channel. It is true that in the wage bargaining protocol of Cahuc et al. (2006) and Bagger et al. (2014), workers who stay at the same firm can experience wage growth both from human capital accumulation and from renegotiation. However, it is not necessary that the wage growth of job stayers be purely due to human capital as long as we know that the wage growth from job switchers is due to the search channel. For stayers, the residual wage growth which is unidentified from the search channel comes from human capital growth.

3.7 Calibration Results

Overall, the fit is good. The fit between the model and lifecycle moments is illustrated in Figure 1, and the fit between the model and nonlifecycle moments is presented in Table 2. The parameters which deliver this fit are listed in Table 3.

Table 2: Fit between data and model, nonlifecycle moments

Moment	Target	Model
U to E rate	0.283	0.299
Wage growth, stayers	0.0079	0.0065
Wage growth, switchers	0.106	0.107
Wage growth, switchers with increase	0.350	0.400
Log wage skewness	0.412	0.408

Table 3: Internally calibrated parameters

Parameter	Meaning	Value
ϕ	Disutility level	7.051
γ_E	Disutility curvature, employed	1.542
γ_U	Disutility curvature, unemployed	4.647
ω	Human capital investment curvature	0.521
ξ	Meeting efficiency	1.648
μ_p	Firm productivity distribution level	0.230
λ_p	Firm productivity distribution tail	0.364
μ_a	Learning ability distribution level	0.0048
σ_a	Learning ability distribution dispersion	0.634
λ_a	Learning ability distribution tail	0.018
β_1	Correlation between learning ability a and initial h	0
σ_ε	Initial h dispersion conditional on a	0.108
ψ	Godfather shock rate	0.00036

4 Properties of Benchmark Model

In this section, I describe some properties of the benchmark model which are important for understanding how the model works and the way it will respond to changes in tax policy.

4.1 Mean Wages

Given the model's structure, I can easily decompose wages between human capital h, firm productivity p, and the bargained piece rate r. Recall that the worker earns the wage hpr. The log wage is the sum of the log of the three components,

$$\ln(wage) = \ln(h) + \ln(p) + \ln(r).$$

The mean log wage follows,

$$\mathrm{E}[\ln(wage)] = \mathrm{E}[\ln(h)] + \mathrm{E}[\ln(l)] + \mathrm{E}[\ln(r)]. \tag{19}$$

Therefore, we can decompose the mean log wage into three additive components.[19] Taken together, p and r comprise the contribution of the search channel, and h comprises the contribution of the human capital channel.

In the benchmark model, the growth of wages over the lifecycle is mostly driven by workers moving up the job ladder to more productive firms. Figure 2 plots the growth of mean log wages over the lifecycle along with its three additive components. All three components are significant in accounting for lifecycle wage growth. Approximately half of lifecycle wage growth comes from moving to more productive firms, a quarter comes from accumulating human capital, and a quarter comes from increased bargaining power.

However, there is significant heterogeneity in the way that workers grow wages. In the model, a worker's choice of l and s depend on his state (h, p, r, i) and fixed learning ability a. In practice, learning ability is especially important.

Dispersion in a generates a heterogeneous tradeoff between investing in human capital or search. Recall that the disutility of effort is a function of $l + s$. Therefore, the marginal disutility of an additional unit of search or learning is always equal, so workers compare the relative benefits of l and s. Relative to investing in search effort, workers with high a have a greater benefit from investing in human capital.

The result is that human capital growth is significantly heterogeneous, and heterogeneity in wage growth is driven by heterogeneity in human capital. In Figure 3, Table 4, and going forward, I break my distribution of a values into two bins: the lower 60% and the upper 40%. Figure 3 replicates the same decomposition of mean log wages for each of the two learning ability bins. The growth of p and r is similar across both groups while human capital growth differentiates the two groups. Those with high learning ability rely mostly on human capital

[19]The decomposition is somewhat compromised by the fact that h, p, and r all interact. However, I show later that the implications of this decomposition are the same as when I take a more careful approach, and the additive decomposition is more intuitive.

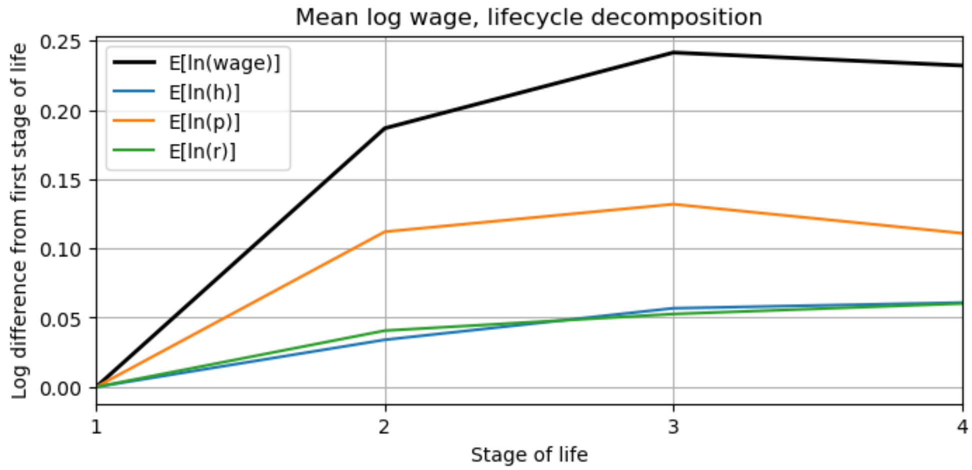

Figure 2: Additive decomposition of log wage growth

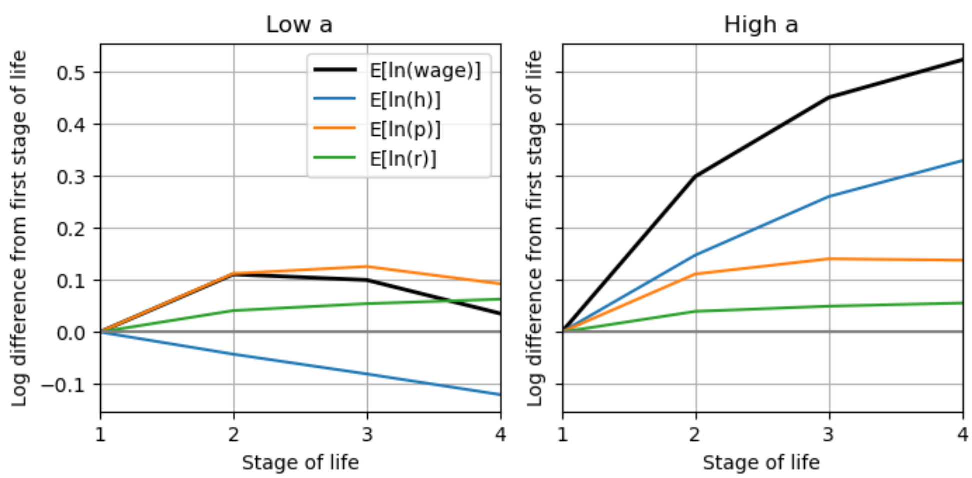

Figure 3: Additive decomposition of log wage growth by learning ability

Table 4: Mean levels of wage and wage components by learning ability

	Low a	High a
$\mathrm{E}[wage]$	0.93	1.64
$\mathrm{E}[h]$	1.29	2.26
$\mathrm{E}[p]$	0.87	0.86
$\mathrm{E}[r]$	0.84	0.84

for wage growth, while those with low learning ability have negative human capital growth.[20] In general, this finding supports the conclusion of Karahan et al. (2022), which finds that job ladder factors are the most important determinant of wage growth for the lower part of the wage distribution, while human capital growth is the most important determinant of wage growth for the upper part of the wage distribution.

Table 4 lists the mean wage and mean h, p, and r for workers with high and low learning abilities. It is clear that differences in human capital drive the different wage levels between groups.

4.2 Wage Dispersion

As with mean log wages, the structure of the model implies a straightforward decomposition for the variance of log wages. The variance of the log wage is the sum of the variance of each component plus an interaction term for each pair of components,

$$\begin{aligned} \mathrm{Var}[\ln(wage)] = &\mathrm{Var}[\ln(h)] + \mathrm{Var}[\ln(p)] + \mathrm{Var}[\ln(r)] \\ &+ 2\,\mathrm{Cov}[\ln(h), \ln(p)] + 2\,\mathrm{Cov}[\ln(h), \ln(r)] + 2\,\mathrm{Cov}[\ln(p), \ln(r)] \end{aligned} \tag{20}$$

Using this decomposition, I find that the increase in wage dispersion over the lifecycle is completely driven by human capital. In Figure 4, I decompose wage dispersion over the lifecycle into the six terms in Equation (20). In the cross section, variance in human capital accounts for about one half of the variance of log wages. But over the lifecycle, the increase in variance of human capital is solely responsible for the increase in wage dispersion.

As Huggett et al. (2011) shows, two key ingredients are required to generate increasing wage dispersion in a human capital model: heterogeneity in learning ability, and positive correlation between learning ability and initial levels of human capital. With heterogeneity in learning ability, those with high learning ability accumulate human capital faster than those with low learning ability. If initial human capital and learning ability are positively correlated, then the gap between workers with high and low learning ability is guaranteed in increase over the lifecycle. Thus, my findings support Huggett et al. (2011); even with an endogenous search component, wage dispersion increases over the lifecycle because workers have different learning abilities. Over the lifecycle, the gap between workers with high learning ability and those with low learning ability increases.

[20] One way to rationalize the negative human capital growth is that my calibration begins at age 23.

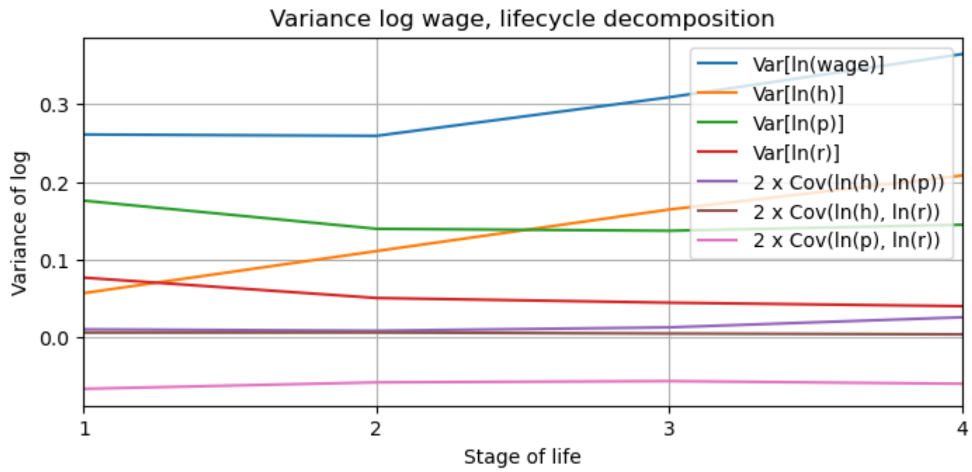

Figure 4: Simple decomposition of the variance of log wages

The benchmark model can be summarized as follows: wage growth is generally a function of the job ladder, while inequality is generally a function of human capital.

Finally, the benchmark model suggests that human capital accumulation and search effort are simultaneously substitutes at the micro level and complements at the macro level. At the micro level, recall that human capital and firm productivity are complements in production. All else equal, the marginal benefit of increasing human capital is greater for a worker at a more productive firm, which suggests complementarity. However, the disutility that workers experience from investing in human capital and job search is convex. So, a worker who has already invested heavily in human capital will have a greater marginal disutility of investing in job search. The result is that workers may choose between human capital and search.

In Figure 5, I plot mean l and mean s for workers in the two learning ability bins from above. At all stages of life, workers with low a invest more in job search effort than workers with high a. Furthermore, workers with low a substitute between l and s over the lifecycle; at the beginning of life, they invest heavily in s and little in l. Over time, as they move up the job ladder and the benefit of search decreases, the lines cross, and they invest more in l than in s.

On the other hand, in Figure 6, I show that the correlation between h and p is positive and U-shaped over the lifecycle. At the beginning of life, workers with high h are more picky, so they are more likely to work at more productive firms. The correlation then drops because, as shown in Figure 5, workers with low a exert more search effort, which implies that they should see more job offers and climb the job ladder faster. But over time, the correlation increases again. The increase is attributable to heterogeneous unemployment risk; because low-wage workers are more likely to lose their job, they do not climb the job ladder as effectively. Thus, over the course of the lifecycle, the correlation between h and p increases.

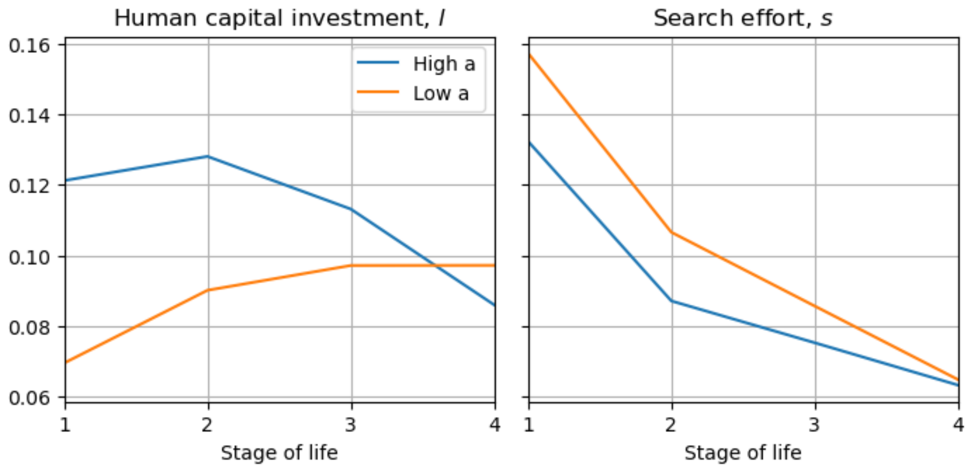

Figure 5: Mean policy functions across learning abilities

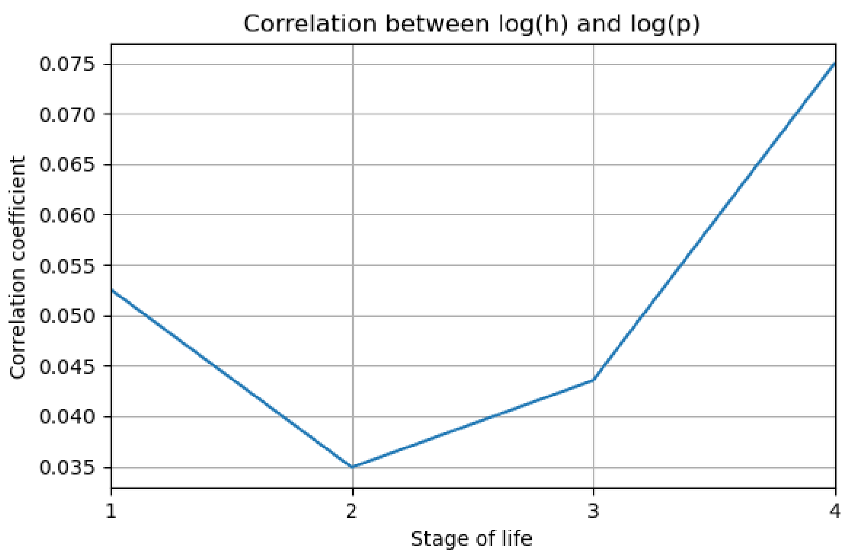

Figure 6: Correlation of h and p

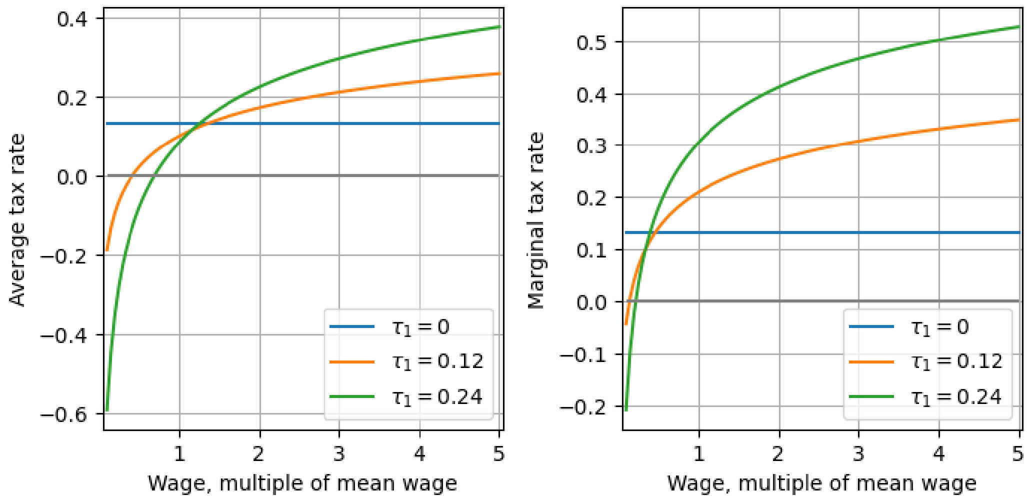

Figure 7: Tax rates by wage

5 Tax Progressivity Experiments

To demonstrate the usefulness of the model, I perform counterfactual experiments on tax progressivity. From the model equations alone, we know that increasing tax progressivity will decrease wages. With higher tax progressivity, when a worker increases their wage, they take home a smaller part of the wage. Human capital accumulation and search effort have the same function: they increase wages. So, since the benefit of effort decreases and the cost is unchanged, l and s will decrease, and wages follow.

What is unclear is the magnitude of the effect and how the channels interact. The model also tells us the extent to which model without both endogenous human capital accumulation and endogenous job search will misstate the effect ot tax progressivity. This analysis requires several decomposition experiments, which I undertake after establishing some basic results.

In counterfactual experiments, I feed the model new values of τ_1. There are two endogenous equilibrium objects which adjust such that model is in equilibrium under the new τ_1: θ, which adjusts such that the free entry condition (16) holds, and τ_0, which adjusts such that government budget constraint (17) holds. I define the recursive stationary equilibrium in Appendix B.

5.1 Basic Results

In the benchmark model, the level of tax progressivity, $\tau_1 = 0.12$, reflects the combined progressivity of the tax system and means-tested transfers in the US. I compare this level of progressivity with flat taxes, $\tau_1 = 0$, and more progressive taxes, $\tau_1 = 0.24$, which roughly corresponds to a Swedish level of progressivity.[21]

[21]See Esfahani (2020) for cross-country comparisons.

Figure 8: Effect of tax progressivity on lifecycle profiles of mean log wage and variance of log wages

Table 5: Summary statistics from increasing tax progressivity

	$\tau_1 = .12$	$\tau_1 = .24$	Difference, $\tau_1 = .24$ model minus $\tau_1 = .12$ model
τ_0, tax level parameter	0.899	0.916	0.016
θ, labor market tightness	1	0.976	-0.024
Mean log wage	0.009	-0.028	-0.037
Variance of log wage	0.308	0.309	0.001
Log wage skewness	0.413	0.207	-0.206
Lifecycle growth of mean log wage	0.232	0.222	-0.010
Lifecycle growth of variance of log wage	0.103	0.090	-0.013
Unemployment rate	0.018	0.018	-0.001
Job switch rate	0.009	0.010	0.000

I plot the average and marginal tax rate for each tax system in Figure 7.[22] If $\tau_1 = 0.12$, then those than earn less than one third of the mean wage earn positive means-tested transfers. As τ_1 increases to 0.24, those with low wages earn more transfers, and those with high wages pay more taxes. When $\tau_1 = 0.12$, the marginal tax rate at the mean wage in the benchmark mode is about 21%; when $\tau_1 = 0.24$, it is about 30%.

A first pass of the experiment yields the results in Figure 8 and Table 5. In Figure 8 I plot the mean log wage and variance of log wage over the lifecycle for all three values of τ_1. In Table 5 and going forward, I only compare the benchmark model ($\tau_1 = 0.12$) with the more progressive model ($\tau_1 = 0.24$).

In total, increasing tax progressivity leads a decrease of mean log wages of 3.7%. It also

[22]For $\tau_1 \neq 0.12$, I use the equilibrium values of τ_0.

Table 6: Effect of tax on mean log wage

	$\tau_1 = .12$	$\tau_1 = .24$	Difference, $\tau_1 = .24$ model minus $\tau_1 = .12$ model
E[ln($wage$)]	0.009	-0.028	-0.037
E[ln($wage$)], low a	-0.178	-0.210	-0.033
E[ln($wage$)], high a	0.286	0.244	-0.042
E[ln(h)]	0.426	0.414	-0.011
E[ln(h)], low a	0.238	0.232	-0.006
E[ln(h)], high a	0.705	0.686	-0.020
E[ln(p)]	-0.218	-0.240	-0.022
E[ln(p)], low a	-0.213	-0.237	-0.024
E[ln(p)], high a	-0.225	-0.244	-0.019
E[ln(r)]	-0.199	-0.202	-0.003
E[ln(r)], low a	-0.202	-0.205	-0.003
E[ln(r)], low a	-0.194	-0.198	-0.004

leads to a slight decrease in the lifecycle growth of mean log wages of 1%. With regards to inequality, increasing tax progressivity has essentially no effect on the variance of log wages, but significantly decreases the skewness of log wages. The lifecycle increase in wage dispersion also slightly decreases. Finally, θ adjusts as expected in equilibrium; with lower levels of human capital and search effort in the labor pool, the benefit of posting a job decreases, firms post fewer jobs, and labor market tightness decreases by 2.4%. There is no change in the unemployment rate or the job switching rate.

5.2 Decomposition Using Wage Components

Wages decrease mostly because workers are employed at less productive firms. In Table 6, I compare log wages using the mean log wage decomposition in Equation (19). According to the overall differences of E[h], E[p], and E[r] (in the top line of each block), the decrease in log wages as a result of the policy change is mostly driven by a decrease in firm productivity, followed by human capital, and then the piece rate. In the fact, the decrease in E[ln(p)] is twice as large as the decrease in E[ln(h)].

Earlier in the paper, I showed that moving to more productive firms is the primary factor behind lifecycle wage growth. One interpretation is that, for most workers, investing in job search is the best way to increase their wages. Since job search is the most relevant margin for wage growth for most workers, the fact that job search is the driving force behind the decrease in wages is unsurprising.

Another reason for the dominance of the search channel is that it is more universal. Recall that workers with high learning ability invest heavily in human capital accumulation while workers with low learning ability do not. In Table 6, the decrease in E[ln(h)] is over

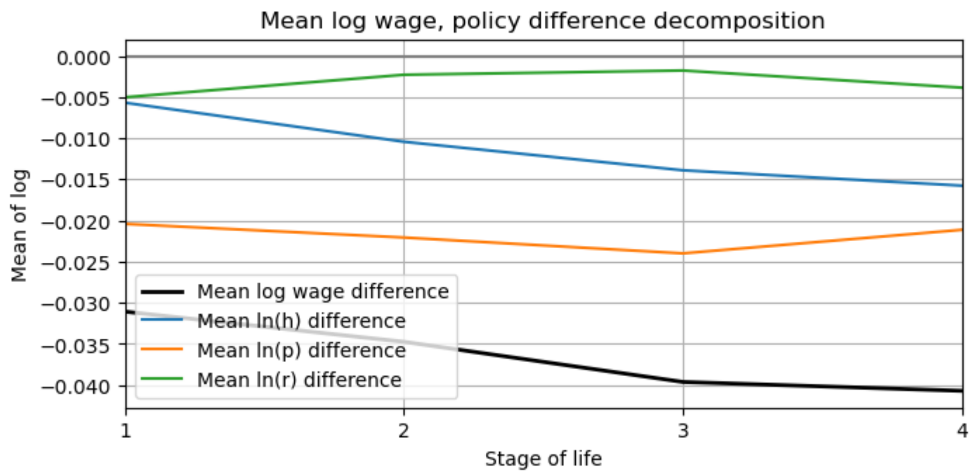

Figure 9: Mean log wage, difference from tax chance, decomposition of change

three times larger for the group with high a than the group with low a. In contrast, the decrease in $\mathrm{E}[\ln(p)]$ is relatively similar across groups.

With regards to lifecycle growth, the decrease in lifecycle growth from the increase in tax progressivity is driven by a decrease in human capital accumulation. Figure 9 shows the difference between of the mean log wage and each of its components between the benchmark and progressive model at each stage of life. As before, the three thin, colorful lines add together to the thick black line. Though we see a difference in the rate at which $\ln(p)$ increases, it is ultimately about one third of the amount that $\ln(h)$ decreases. Between the beginning and end of the life, the difference in $\ln(p)$ is the same. The decrease in the lifecycle growth rate of log wage is entirely a function of $\ln(h)$ increasing at a slower rate.

The implication is that an increase in tax progressivity stymies the growth of human capital. Since workers with high learning ability drive the lifecycle increase in human capital, we can say that the slower rate of wage growth over the lifecycle is driven by workers at the top of the wage distribution acquiring less human capital. Such is the reason that increasing progressivity significantly decreases the skewness of wages in Table 5.

An increase of tax progressivity has no effect on the variance of log wages because the human capital channel and the search channel have opposite implications for wage dispersion and taken together, these effects offset one another. In Table 7, I decompose the effects of increased tax progressivity on the variance of log wages using Equation (20). I group the additive components of wage variance on the right hand side of Equation (20) into three parts: (1) the variance from the human capital channel, $\mathrm{Var}[\ln(h)]$; the variance from the search channel, $\mathrm{Var}[\ln(p)] + \mathrm{Var}[\ln(r)] + 2\,\mathrm{Cov}[\ln(p), \ln(r)]$; and the variance from interactions between human capital and search, $2\,\mathrm{Cov}[\ln(h), \ln(p)] + 2\,\mathrm{Cov}[\ln(h), \ln(r)]$.

When I increase tax progressivity, the variance from the human capital channel decreases, and the variance from the search channel increases. Table 6 explains why the variance of human capital channel decreases. Human capital levels decrease for all workers, the but

Table 7: Effect of tax on variance of log wages

	$\tau_1 = .12$	$\tau_1 = .24$	Difference, $\tau_1 = .24$ model minus $\tau_1 = .12$ model
Var[ln(wage)]	0.308	0.309	0.001
Variance from human capital channel, Var[ln(h)]	0.136	0.129	-0.007
Variance from search channel	0.149	0.156	0.008
Var[ln(p)]	0.152	0.155	0.003
Var[ln(r)]	0.054	0.058	0.004
$2\,\mathrm{Cov}[\ln(p), \ln(r)]$	-0.057	-0.057	0.000
Variance of interaction	0.024	0.024	0.000
$2\,\mathrm{Cov}[\ln(h), \ln(p)]$	0.017	0.018	0.001
$2\,\mathrm{Cov}[\ln(h), \ln(r)]$	0.007	0.006	0.000
Variance within groups	0.230	0.234	0.004
Variance between groups	0.078	0.075	-0.003

magnitude of the decrease is larger for workers with high a. Therefore, the gap in human capital between workers with high a and low a decreases, which decreases the variance of human capital. Wage dispersion decreases from above.

Why do workers with more learning ability experience a larger decrease in human capital? There are two reasons. First, increased tax progressivity has more "bite" at the top part of the wage distribution, so high-wage workers are more disincentivized to grow their wages. Second, those at the higher end of the distribution are the same workers who invest more in human capital, so their relevant margin for adjustment is human capital.

The decrease in variance from the human capital is offset by the search channel. To understand why, recall how on-the-job search works in the model (and in many models of on-the-job search). At the beginning of life (or while unemployed), workers randomly meet a firm with productivity p from the distribution $F(p)$. Ignoring different reservation wages for the moment, the initial distribution of p will roughly mirror $F(p)$. As workers search while on the job, they meet more firms with productivity drawn from $F(p)$. If they meet a firm with higher productivity than their current firm, they switch firms. As workers move up the job ladder, workers tend to bunch at the top of the wage distribution, and the variance of the search component is less than it was initially.

If workers exert search effort, they will be more likely to meet outside firms, and will move up the job ladder more quickly, leading to bunching at the top of the p distribution. When tax progressivity increases, workers exert less search effort. The result is less climbing up the job ladder and less bunching at higher p levels.

Intuitively, with more tax progressivity, workers exert less search effort, so workers are more likely to remain at unproductive firms which would otherwise not be able to retain workers. This increases the dispersion of wages from below. In total, with both channels, the gap between high and low wages is the same, but the level of wages has decreased.

5.3 Decomposition Using Out-of-Equilibrium Models

In this section, I use a different approach to decompose the effect of tax progressivity on log wages. I analyze the model in steps between the benchmark equilibrium and the more progressive equilibrium. To do so, I sequentially hold three endogenous objects constant and re-solve the model. The three objects are equilibrium labor market tightness θ^*, workers' policy functions for human capital accumulation l^*, and workers' policy functions for job search effort for both the employed and the unemployed s^*.

I consider three out-of-equilibrium models between the benchmark equilibrium and the more progressive equilibrium. For each, I impose the policy change of $\tau_1 = 0.24$ and the new value of τ_0 which solves the government budget constraint in the new equilibrium.

In the first out-of-equilibrium model, I fix s^* and θ^* to their values in the benchmark model, thereby only allowing l^*, human capital accumulation, to respond to the policy change. In this setup, search effort, and therefore the the probability of switching jobs, cannot respond to policy; the search channel is exogenous. This roughly corresponds to a human capital model with exogenous shocks (Badel et al., 2020; Huggett et al., 2011).[23]

In the second out-of-equilibrium model, I fix l^* and θ^*. Using the same logic as above, this is a model where the path of human capital is fixed, and workers can only respond to the policy by changing search effort.

Finally, in the third out-of-equilibrium model, I only fix θ^*. In this version, workers can adjust human capital and search effort in response to the change in tax policy, but firms cannot adjust the number of jobs posted. The difference in outcomes between this model and the more progressive equilibrium model show the relative importance of the endogenous job posting channel.

I summarize the decomposition in Table 8. In the interest of readability, I do not report the levels of each variable for each model; I only report the benchmark values and the difference between the value from the model in the question and the benchmark value.

In general, the results in Table 8 support the results in the previous section. With regards to the decrease in the mean log wage, the search channel is more responsive when l is exogenous compared to the responsiveness of the human capital channel when s is exogenous. This confirms that search effort is the more relevant margin of adjustment for workers, and it drives the response of mean wages. The decrease in lifecycle wage growth is larger when s is exogeous, reflecting how h drives the response to lifecycle wage growth. Finally, the effect on variance is goes in opposite directions depending on what is exogenous, confirming the offsetting effects from earlier.

Table 8 also generates some finer points. First, decisions on human capital and search effort are not independent from one another. If that were the case, then, for the mean log wage, we could add columns (2) and (3) together and get the value in column (4). Instead, there is nonlinearity.

Second, endogenous job posting is of some importance for the response of the wage level. Compare the mean log wage in columns (4) and (5). Column (5) states that, when

[23] I also hold $z_{ai}(h, w)$ constant so that unemployment-to-employment transition rates are the same.

Table 8: Out-of-equilibrium decomposition

	(1) Benchmark	(2) Fixed (s^*, θ^*) − benchmark	(3) Fixed (l^*, θ^*) − benchmark	(4) Fixed θ^* − benchmark	(5) More progressive equilibrium − benchmark
θ	1.000	0.000	0.000	0.000	-0.024
Mean log wage	0.009	-0.016	-0.027	-0.033	-0.037
E[ln(h)]	0.426	-0.014	-0.005	-0.011	-0.011
E[ln(p)]	-0.218	-0.001	-0.018	-0.018	-0.022
E[ln(r)]	-0.199	-0.001	-0.003	-0.003	-0.003
Variance of log wage	0.308	-0.004	0.007	0.001	0.001
Lifecycle growth of mean log wage	0.232	-0.013	-0.002	-0.010	-0.010
Lifecycle growth of variance of log wage	0.103	-0.015	-0.008	-0.013	-0.013
Unemployment rate	0.018	0.0002	-0.0008	-0.0008	-0.0006
Job switch rate	0.009	0.0000	-0.0001	-0.0001	0.0002

tax progressivity is increased, the mean wage drops by 3.7%. Without the response of job posting, the decrease is 3.3%. We can therefore conclude that 15%[24] of the decrease in the mean log wage is due to endogenous job posting. Another interpretation is that when the number of job vacancies decrease by 2.5%, mean wages decrease by 0.4%.

Figure 10 provides a clean illustration of how the human capital channels and job search channels offset with regards to the change in wage dispersion. When s^* is fixed and only human capital investment can respond to policy, we can see that the variance of wages decreases. If l^* is fixed and only search effort responds to policy, the variance of wages increases. With both endogenous, the model is back where it started.

6 Conclusion

This paper develops and studies a rich model of wages. Existing literature has mostly consisted of models where either the human capital channel or the search channel is endogenous. In my model, we have both. Combining both channels in a quantitative model is important for economists because different channels have different implications for policy, not only in magnitude but also in direction.

In this paper, I demonstrate the usefulness of the model by investigating income tax progressivity. I find that increasing tax progressivity from the its current level in the US to a level roughly in line with some European countries would decrease the mean wage by 4%, only slightly decrease lifecycle wage growth, and have no effect on wage dispersion. The effect on mean wages is larger than it would be in a model where only one channel is

[24] $0.4/3.7 \approx 0.15$.

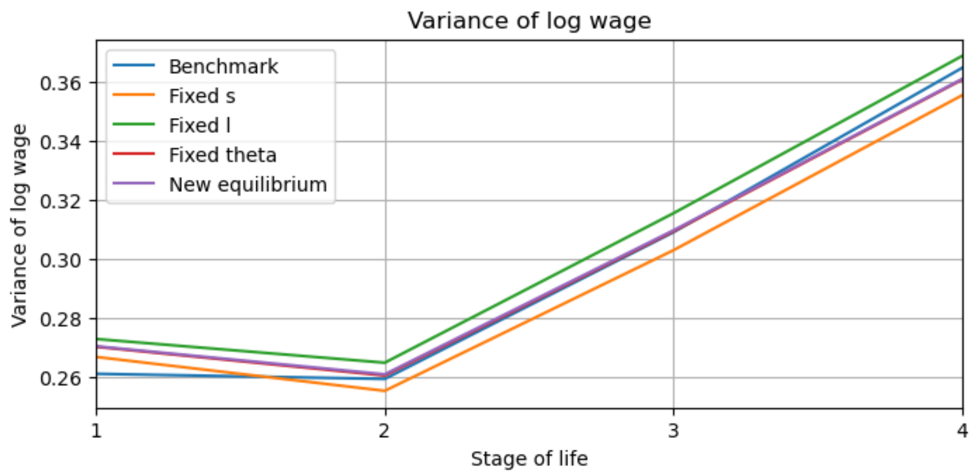

Figure 10: Variance of log wage for out-of-equilibrium models

endogenous. For lifecycle wage growth, the effect is smaller than in a pure human capital model. The lack of a response in the variance of wages is a unique feature arising from the combination of both channels.

The model is ripe for more experiments. Future work could include investigating business taxation, minimum wages, and other labor market regulations.

Appendix

A Equations

A.1 Aggregate Search Effort, S

Aggregate search effort is defined as

$$S = \sum_{a,i} \int_{\Psi_E} s^*_{Eai}(h,p,r) d\Psi_E(h,p,r|a,i) + \sum_{a,i} \int_{\Psi_U} s^*_{Uai}(h,w) d\Psi_U(h,w|a,i). \tag{21}$$

A.2 Equations Updated with Godfather Shocks

When a worker experiences a godfather shock, he randomly meets a firm, and is forced to accept an offer with that firm. When we include godfather shocks in the employed worker's Hamilton-Jacobi-Bellman equation (7), the worker's problem adds a new line like so:

$$\begin{aligned} \rho E_{ai}(h,p,r) = \max_{l,s}\ & u(c) - d_E(l+s) + (a(lh)^\omega - \delta h) \frac{\partial E_{ai}(h,p,r)}{\partial h} \\ &+ \Lambda(rhp)\left[U_{ai}(h,hpr) - E_{ai}(h,p,r)\right] + \zeta\left[E_{a,i+1}(h,p,r) - E_{ai}(h,p,r)\right] \\ &+ sm(\theta)\Bigg(\int_{q_{ai}(h,p,r)}^{p} \Big[E_{ai}\left(h,p,R^R_{ai}(h,p,p')\right) - E_{ai}(h,p,r)\Big]\, dF(p') \\ &+ \int_{p}^{\overline{p}} \Big[E_{ai}\left(h,p',R^P_{ai}(h,p,p')\right) - E_{ai}(h,p,r)\Big]\, dF(p') \Bigg) \\ &+ \psi \int_{\underline{p}}^{\overline{p}} \Big[E_{ai}\left(h,p',R^G_{ai}(h,hpr,p')\right) - E_{ai}(h,p,r)\Big]\, dF(p') \end{aligned} \tag{22}$$

subject to

$$c = [1 - T(hpr)]hpr.$$

Firms have an additional risk of losing their worker to a godfather shock, which adds ψ to the bottom line of (14) like so:

$$\begin{aligned} \rho J_{ai}(h,p,r) = {} & hp(1-r)(1-\tau_b) + (a(l_{ai}(h,p,r)h)^\omega - \delta h) \frac{\partial J_{ai}(h,p,r)}{\partial h} \\ &+ \zeta\left[J_{a,i+1}(h,p,r) - J_{ai}(h,p,r)\right] \\ &+ s_{Eai}(h,p,r)m(\theta) \int_{q_{ai}(h,p,r)}^{p} \Big[J_{ai}\left(h,p,R^R_{ai}(h,p,p')\right) - J_{ai}(h,p,r)\Big]\, dF(p') \\ &+ \Big[s_{Eai}(h,p,r)m(\theta)\left(F(\overline{p}) - F(p)\right) + \Lambda(hpr) + \psi\Big](-J_{ai}(h,p,r)) \end{aligned} \tag{23}$$

On the other hand, firms have additional channel for filling job vacancies; an employed worker might experience a godfather shock and land with the firm. Thus, the free entry condition gets a new term,

$$\kappa = m_f(\theta)\Bigg[\sum_{a,i}\int_{\Psi_E} s_{Eai}(h,p,r)\int_{p}^{\overline{p}} J_{ai}\left(h,p',R^P_{ai}(h,p,p')\right)dF(p')d\Psi_E(h,p,r|a,i)$$
$$+\sum_{a,i}\int_{\Psi_U} s_{Uai}(h,w)\int_{z_{ai}(h,w)}^{\overline{p}} J_{ai}\left(h,p',R^U_{ai}(h,w,p')\right)dF(p')d\Psi_U(h,w|a,i)\Bigg] \quad (24)$$
$$+\psi\sum_{a,i}\int_{\Psi_E}\int_{\underline{p}}^{\overline{p}} J_{ai}\left(h,p',R^G_{ai}(h,hpr,p')\right)dF(p')d\Psi_E(h,p,r|a,i).$$

Finally, I assume that when bargaining over wages, the outside option for workers who have been subject to a godfather shock is unemployment without unemployment benefits. It is impossible that the worker goes to this state, but this establishes the bargaining over r. So, $R^G_{ai}(h,hpr,p')$ solves

$$E_{ai}\left(h,p',R^G_{ai}(h,hpr,p')\right) = U_{ai}(h,0) + \eta\left[E_{ai}(h,p',1) - U_{ai}(h,0)\right]. \quad (25)$$

B Equilibrium Definition

The recursive stationary equilibrium consists of a set of value functions $\{E_{ai}(h,p,r), U_{ai}(h,w), \overline{E}, J_{ai}(h,p,r)\}$, a human capital investment policy function $l_{Eai}(h,p,r)$, a set of search effort policy functions $\{s_{Eai}(h,p,r), s_{Uai}(h,w)\}$, a set of job search cutoff rules $\{q_{ai}(h,p,r), z_{ai}(h,w)\}$, a set of wage functions $\{R^P_{ai}(h,p,p'), R^R_{ai}(h,p,p'), R^U_{ai}(h,w,I_b,p'), R^G_{ai}(h,hpr,p')\}$, vacancies v, aggregate search effort S, labor market tightness $\theta = v/S$, a distribution of workers $\Psi = \{\Psi_E(h,p,r|a,i), \Psi_U(h,w|a,i), \Psi_R(a)\}$, and a set of government policies parameters $G = \{\tau_0, \tau_1, \tau_b, b, \chi, SS, \overline{g}\}$ which satisfy:

1. Employed worker optimization: Given $R^P_{ai}(h,p,p')$, $R^R_{ai}(h,p,p')$, $R^G_{ai}(h,hpr,p')$, $q_{ai}(h,p,r)$, θ, and G, $E_{ai}(h,p,r)$ solves (22) subject to (8) and (12) with associated decision rules $l_{Eai}(h,p,r)$ and $s_{Eai}(h,p,r)$.

2. Unemployed worker optimization: Given $R^U_{ai}(h,w,p')$, $z_{ai}(h,w)$, θ, and G, $U_{ai}(h,w)$ solves (9) subject to (10) and (13) with associated decision rule $s_{Uai}(h,w)$.

3. Retired workers: $\overline{E}$ solves (11).

4. Filled jobs: Given $R^R_{ai}(h,p,p')$, $l_{ai}(h,p,r)$, $s_{Eai}(h,p,r)$, θ, and G, $J_{ai}(h,p,r)$ solves (23) and (15).

5. Wage equations: Given $E_{ai}(h,p,r)$ and $U_{ai}(h,w,I_b)$,

   (a) $R^P_{ai}(h,p,p')$ solves (2),

   (b) $R^R_{ai}(h,p,p')$ solves (3),

   (c) $R^U_{ai}(h,w,p')$ solves (5), and

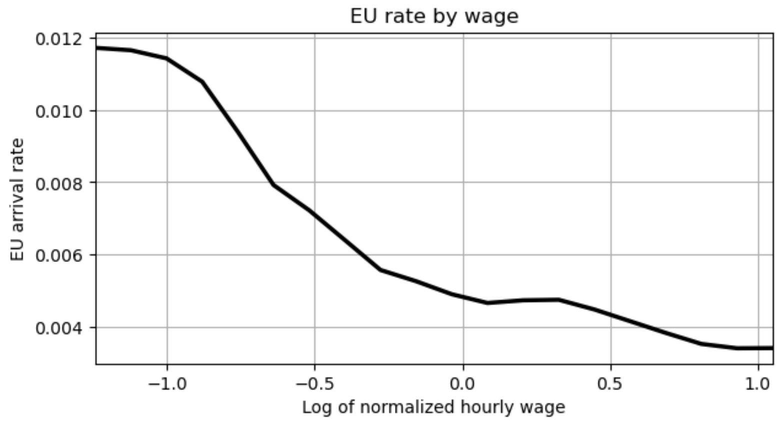

Figure 11: Estimate of exogenous job destruction function, $\Lambda(wage)$

(d) $R^G_{ai}(h, hpr, p')$ solves (25).

6. Job search cutoff rules: Given $E_{ai}(h, p, r)$ and $U_{ai}(h, w)$,

   (a) $q_{ai}(h, p, r)$ solves (4) and

   (b) $z_{ai}(h, w)$ solves (6).

7. Free entry: Given $J_{ai}(h, p, r)$, $R^P_{ai}(h, p, p')$, $R^U_{ai}(h, w, p')$, $R^G_{ai}(h, hpr, p')$, $s_{Eai}(h, p, r)$, $s_{Uai}(h, w)$, S, and Ψ, θ solves (24).

8. Government budget constraint: Given Ψ, G satisfies (17).

9. Aggregate search effort: Given $s_{Eai}(h, p, r)$, $s_{Eai}(h, w)$, and Ψ, S satisfies (21).

10. Consistency: Ψ is the stationary distribution. Ψ is defined such that

$$1 = \sum_{a,i} \int_{\Psi_E} d\Psi_E(h, p, r|a, i) + \sum_{a,i} \int_{\Psi_U} d\Psi_U(h, w|a, i) + \sum_a \Psi_R(a). \tag{26}$$

C Calibration Details

C.1 Λ Function

I estimate $\Lambda(hpr)$ directly from SIPP data using interpolation. See Figure 11 for an illustration. As with the tax function, the mean wage is normalized to one.

C.2 Learning Ability Distribution

I discretize the learning ability distribution as follows. The grid of a points is $\{a_1, ..., a_J\}$ where $a_j = \mathrm{E}[a|a_{P_j} \leq a < a_{P_{i+1}}]$ and a_{P_j} os the P_j-th percentile of the ability distribution. I take the expectation using $PLN(\mu_a, \sigma_a, 1/\lambda_a)$. To calculate the PLN distribution, I use analytical expressions in Hajargasht and Griffiths (2013). Percentiles are chosen to be $(0.0, 0.3, 0.6, 0.9, 0.99, 1.0)$.

References

Badel, A., M. Huggett, and W. Luo (2020). Taxing Top Earners: a Human Capital Perspective. *The Economic Journal 130*(629), 1200–1225.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin (2014). Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics. *The American Economic Review 104*(6), 1551–1596.

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *Journal of Political Economy 75*(4, Part 1), 352–365.

Burdett, K. and D. T. Mortensen (1998). Wage Differentials, Employer Size, and Unemployment. *International Economic Review 39*(2), 257–273.

Bénabou, R. (2002). Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency? *Econometrica 70*(2), 481–517.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006). Wage Bargaining with On-the-Job Search: Theory and Evidence. *Econometrica 74*(2), 323–364.

Cooper, M., J. McClelland, J. Pearce, R. Prisinzano, J. Sullivan, D. Yagan, O. Zidar, and E. Zwick (2016). Business in the United States: Who Owns It, and How Much Tax Do They Pay? *Tax Policy and the Economy 30*(1), 91–128.

Dorn, A. (2018). How Dangerous is Godfather? Job-to-Job Transitions and Wage Cuts.

Engbom, N. (2022). Labor Market Fluidity and Human Capital Accumulation.

Esfahani, M. (2020, October). Inequality Over the Life-cycle: U.S. vs Europe.

Guner, N., R. Kaygusuz, and G. Ventura (2014). Income Taxation of U.S. Households: Facts and Parametric Estimates. *Review of Economic Dynamics 17*(4), 559–581.

Guner, N., C. Rauh, and G. Ventura (2020). Means-Tested Transfers in the US: Facts and Parametric Estimates.

Hajargasht, G. and W. E. Griffiths (2013). Pareto–lognormal distributions: Inequality, poverty, and estimation from grouped income data. *Economic Modelling 33*, 593–604.

Huggett, M., G. Ventura, and A. Yaron (2011). Sources of Lifetime Inequality. *American Economic Review 101*(7), 2923–2954.

Karahan, F., S. Ozkan, and J. Song (2022). Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job Ladder Risk vs. Human Capital. Working Paper 2022-002, Federal Reserve Bank of St. Louis.

Lise, J., C. Meghir, and J.-M. Robin (2016). Matching, sorting and wages. *Review of Economic Dynamics 19*, 63–87.

Mukoyama, T. (2014). The cyclicality of job-to-job transitions and its implications for aggregate productivity. *Journal of Economic Dynamics and Control 39*, 1–17.

Petrongolo, B. and C. A. Pissarides (2001). Looking into the Black Box: A Survey of the Matching Function. *Journal of Economic Literature 39*(2), 390–431.

Rubinstein, A. (1982). Perfect Equilibrium in a Bargaining Model. *Econometrica 50*(1), 97–109.

Schlüter, M., J. A. Egea, and J. R. Banga (2009). Extended ant colony optimization for non-convex mixed integer nonlinear programming. *Computers & Operations Research 36*(7), 2217–2229.