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

Matthew J. Millington*

October 25, 2023

Click here for latest version.

Abstract

I build an equilibrium lifecycle model of wages that combines human capital accumulation with on-the-job search in a frictional labor market. In the model, heterogeneous workers endogenously invest in human capital accumulation and search effort, and firms post jobs. I discipline the model using microdata from the SIPP. Using the calibrated model, I show that on-the-job search is the driving force behind lifecycle wage growth, heterogeneous human capital accumulation is the driving force behind lifecycle wage dispersion, and there is significant heterogeneity in how workers attempt to increase their wage. I then use the model as a laboratory for studying tax and transfer progressivity. An increase in progressivity decreases wages, primarily due to reduced on-the-job search effort. Surprisingly, it has little effect on wage dispersion because the human capital and search channels 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 typically frictionless, so a worker's wage equals his marginal product of labor. As a worker accumulates human capital, he becomes more productive, and his wage increases. In order to replicate wage dispersion in the data (in particular the increase lifecycle profile of wage dispersion), human capital theory is supplemented with 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. If a worker moves to a more productive firm, he and the firm bargain over his new 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 the life cycle in the most productive firm, there would be no wage growth. When implemented in a quantitative model, workers are often heterogeneous in fixed skill.

I develop an equilibrium model where both channels interact endogenously. The model combines endogenous human capital accumulation, endogenous on-the-job search, endogenous job posting, and a life cycle. Workers choose how much effort to invest in human capital accumulation and job search effort, and firms choose how many job vacancies to post in equilibrium. Both types of investment increase wages, but workers choose an optimal mix of investments according to their fixed learning ability, their current state, and the state of the labor market. To my knowledge, a quantitative model with all of these elements has not been studied. Going forward, I refer to the human capital part as the"human capital channel," and I refer to the search-and-matching part as the "search channel."

What is the benefit of endogenizing all of these elements? Previous research (cited below) shows that both the human capital and search channels matter empirically. It follows that both channels may respond to changes in policy. So, a model with both channels will have more accurate estimates. In this paper, I experiment with tax progressivity.

Furthermore, my model allows for decomposition exercises with regards to the different channels. First, I analyze the benchmark model and show how the channels interact quantitatively for explaining wage growth and wage dispersion. There are meaningful differences in thw investment mix across the wage distribution, which has important implications for redistributive policy.

When I counterfactually adjust tax policy, I further decompose the effects of a tax change between the channels. A priori, it is not clear how the channels will interact quantitatively after a policy change. These are quantitative questions which require a quantitative model.

The model uses a stochastic lifecycle with overlapping generations. In the model, workers are heterogeneous at the beginning of life in fixed learning ability and initial human capital. Firms are heterogeneous in productivity. Importantly, human capital and firm productivity are complements in production. 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 are risk averse and maximize their lifetime utility of consumption. To do so, workers choose how much effort to invest in human capital accumulation or searching for a new job. Each activity is costly in the sense that the worker endures convex disutility. However, both activities are beneficial 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 and switching to a higher-paying job.

Workers also face negative risk from exogenous unemployment shocks. Low-wage workers are more likely to experience unemployment, and human capital depreciates while a worker is unemployed.

Firms enter the model by posting job vacancies. Meeting probabilities are determined by an aggregate matching function, which depends on the number of open job vacancies. In equilibrium, free entry requires that the number of vacancies makes firms indifferent toward posting the marginal 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 vacancy to a filled job, increasing the benefit of job posting and incentivizing more job posting.

Workers face a clear tradeoff: if the worker investments in human capital accumulation or job search effort today, he may increase his wage tomorrow. His optimal mix of human capital investment and search effort depends on his learning ability, age, and current levels of human capital, firm productivity, and the bargaining term. Therefore, workers in different situations will (a) choose different levels of investment and (b) choose different mixes of investment.

For example, I analytically show that workers will adjust their investment mix according to their fixed learning ability. A worker high learning ability will choose to invest more in human capital accumulation and less in job search. The opposite is true for workers with low learning ability. In this sense, workers substitute between the human capital channel and the search channel. Compared to a pure human capital model, this is good news for low-ability workers; they can invest heavily in job search to increase their wages instead of struggling to accumulate human capital. On its own, this force reduces wage dispersion.

On the other hand, since human capital and firm productivity are complements in production, there is also complementarity between the two channels. All else equal, a worker who is employed at more productive firm will invest more in human capital because he has a greater benefit of increased human capital. The result is that wages will grow more quickly for workers who happen to have better initial conditions. This force puts upward pressure on wage dispersion.

I calibrate the model using micro data from the Survey of Income and Program Participation (SIPP) for 1990-2019. The SIPP is a large panel data set which tracks respondents for several years at a time. It is particularly useful for this paper because wages and jobs are reported monthly. Therefore, I can calibrate the model to match monthly wage growth and job changes.

Since I experiment with tax progressivity, I carefully calibrate a government sector. I estimate an average tax function over wages which replicates the income tax code and means-tested transfer payments for the average household in the US. Some workers pay negative taxes; these workers receive more transfers larger than they pay in taxes. In counterfactual exercises, I increase the progressivity of the system. Greater progressivity implies that low-wage workers receive more transfers while high-wage workers pay more taxes.

The benchmark calibrated model successfully replicates lifecycle profiles for mean wages, wage dispersion, and job-to-job transitions. Additionally, the model matches the job switching rate, the unemployment transition rates, and monthly wage growth for workers who stay at the same job or switch jobs. The latter targets are important for identifying wage growth from human capital accumulation versus job switches.

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 Ozkan et al. (2023) and Bagger et al. (2014).

However, there is significant heterogeneity between workers. Heterogeneity in fixed learning ability generates differential returns to human capital investment, and workers with greater learning ability will invest more in human capital accumulation, especially at the beginning of life. Across learning abilities, I see roughly equal job ladder outcomes, but vast differences in human capital accumulation. The result is that 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 my results as suggesting that wage growth comes from the job ladder while wage dispersion comes from human capital.

In the aggregate, there is slight positive correlation between human capital and firm productivity in the benchmark model, and it becomes more pronounced over the life cycle. The positive correlation is driven by two forces. First, there is the fact that human capital and firm productivity are complements in production, the implications of which I describe above. The second force is differential rates of unemployment: low-wage workers are more likely to lose their job and fall several rungs down the job ladder, which portends less success climbing the job ladder over the life cycle.[1]

I then turn to experiments regarding tax progressivity. Greater tax progressivity implies that, when a worker increases his wage, he will retain a smaller part of the wage increase. Thus, workers are disincentivized from growing their wages, which 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 exert less search effort, firms are discouraged from posting jobs, leading to fewer open vacancies in

[1] See Jarosch (2023) for a deep discussion of this mechanism.

equilibrium.

I perform a simple tax policy experiment where I increase the progressivity of the tax and transfer system from the US benchmark to that of a country like Denmark. 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. The reason is that, for the mean worker in the economy, search effort is the primary method for generating wage growth and the most relevant margin for adjustment. About 15% of the decrease is due to a decrease in vacancy posting, implying that equilibrium effects are also significant.

An increase in tax progressivity also decreases lifecycle wage growth, but only slightly. The decrease in lifecycle wage growth is driven by less human capital accumulation for those at the top of wage distribution.

Surprisingly, because the effects of the human capital and search channels offset one another. With regards to the human capital channel, since an increase in progressivity has more "bite" at the top of wage distribution, workers with high wages are relatively more disincentivized from accumulating human capital. So, wages at the top of the distribution decrease more than wages at the bottom, and the gap between high wages and low wages decreases.

In contrast, in a job ladder model, an increase in tax progressivity increases wage dispersion. In the model, workers randomly meet firms from a fixed productivity distribution. With more tax progressivity, workers exert less search effort and make fewer job-to-job transitions. So, 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. First, I contextualize my paper within the literature in Section 2. I present the model in Section 3. I then analytically investigate the model mechanisms in Section 4. My calibration strategy is described in Section 5, after which I analyze the benchmark model in Section 6. I undertake counterfactual taxation experiments in Section 7. Section 8 concludes.

2 Related Literature

To my knowledge, my model is the first to combine a life cycle, endogenous human capital accumulation, endogenous search effort, and endogenous job posting. But there have been papers which combine subsets of these ingredients. Bowlus and Liu (2013) analyze a model with endogenous human capital accumulation and endogenous search effort. I add firms which post jobs in equilibrium, which also requires that I adopt a wage bargaining scheme. The inclusion of the firm side allows for counterfactual policy experiments. My model is similar to Engbom (2022) which features endogenous human capital accumulation and job posting. The key differences are that my model also includes endogenous search effort and

I use my model to investigate taxes and transfers. Finally, Rauh and Santos (2022) build a search-and-matching model with human capital accumulation, endogenous job posting, and a carefully-calibrated government sector, and use the model to investigate transfer payments. However, human capital accumulation is exogenous and there is no on-the-job search.[2]

I rely on well-established methods for modeling human capital accumulation and job search. This is intentional in order to isolate the effects of their interaction. With regards to human capital accumulation, my model builds on Ben-Porath (1967) and Huggett et al. (2011). For modeling on-the-job search, the model borrows from Burdett and Mortensen (1998)[3]. I model endogenous search effort and job posting as in Pissarides (2000) and Mortensen and Pissarides (1994). Finally, the wage bargaining protocol comes from Cahuc et al. (2006) and Bagger et al. (2014).

Quantitative human capital models, such as those in Huggett et al. (2011) and Badel et al. (2020), rely on idiosyncratic shocks to increase dispersion in wages. Such idiosyncratic shocks represent unemployment spells or job-to-job transitions. In a sense, my model (and models like it) provide a theory for those idiosyncratic shocks.

The current paper also draws from a literature which decomposes wage growth and wage inequality between the human capital and search channels (Ozkan et al., 2023; Bagger et al., 2014; Pavan, 2011; Carrillo-Tudela, 2012; Veramendi, 2012; Omer, 2004; Schönberg, 2007; Dustmann and Meghir, 2005).[4] In particular, my model resembles Ozkan et al. (2023) except that human capital accumulation, job finding rates, and labor market tightness are endogenous.[5] The fact that workers and firms can alter their behavior in response to policy changes makes my model an appropriate setting for counterfactual experiments.

In my counterfactual experiments, I contribute to two strands of literature which study the effects of progressive taxes and transfers on lifecycle wage growth and inequality. The first strand of literature investigates labor taxation in models with endogenous on-the-job search and job posting but fixed human capital (Bagger et al., 2021, 2019; Sleet and Yazıcı, 2017; Kreiner et al., 2015; Gentry and Hubbard, 2004).[6] The second strand investigates labor taxation in models of human capital without labor market frictions (Badel et al., 2020; Guvenen et al., 2014; Blandin, 2018; Kapička, 2015, 2006).[7] I show that these channels interact in meaningful ways in response to a change in tax policy.

[2]On the other hand, the model in Rauh and Santos (2022) includes incomplete asset markets, which is an important element.

[3]See also Mortensen (2003).

[4]Acabbi et al. (2023) further the literature by investigating the how human capital and job search interact over the business cycle.

[5]As with all things, there is a tradeoff. In this case, Ozkan et al. (2023) are able incorporate more sophisticated worker heterogeneity.

[6]A related literature shows that some types of labor market regulation and redistribution may be efficient in a frictional labor market (Cubas and Silos, 2020; Lise et al., 2016).

[7]In these papers, workers make human capital decisions each period. One can also study progressive taxation in a model with an endogenous human capital/education choice which only occurs at the beginning of the life cycle (Heathcote et al., 2020, 2017; Esfahani, 2020; Krueger and Ludwig, 2016).

3 Model

3.1 Life Cycle

For computational efficiency, the model is in continuous time. I model overlapping generations using a stochastic lifecycle.[8] There are $I+1$ stages of life: I working stages and a retirement stage. For $i \in \{1, 2, ..., I\}$, workers transition from stage i to $i+1$with probability ζ. Retired workers 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 working life is approximately a decade.

3.2 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.

3.3 Utility

Workers discount the future at discount rate ρ. They enjoy utility from consumption where consumption equals the worker's after-tax-and-transfer wage. (In other words, workers are hand-to-mouth). I assume that utility of consumption is logarithmic, $\ln(c)$. Given the average tax rate function $T(hpr)$ (explained below), $c = [1 - T(hpr)]hpr$.

Workers can be employed or unemployed. Employed workers choose how much effort to invest in human capital accumulation, l (for "learning"), and/or job search, s. Unemployed workers can search but cannot accumulate in human capital. Both types of investment are costly in that workers experience disutility from effort.

Disutility of effort is a convex function. For the employed, disutility of effort is $\phi(l+s)^{1+\gamma_E}$ with $\phi > 0$ and $\gamma_E > 0$. The disutilty of effort for the unemployed is $\phi s^{1+\gamma_U}$ with $\gamma_U > 0$. So, unemployed workers are subject to the same disutility function with a different curvature parameter and no possibility for human capital investment.[9]

[8]One can think of this structure as a perpetual young structure with multiple stages.

[9]The fact that employed and unemployed workers have a different curvatures of disutility reflects that unemployed workers value their time differently than employed workers. Later, I estimate $\gamma_U > \gamma_E$, which means that unemployed workers can exert relatively more search effort before their disutility of effort gets prohibitively steep. Therefore, unemployed workers will tend to exert more search effort and get more job offers. This modeling choice ultimately allows the model to replicate higher job finding rates for unemployed workers than employed workers.

3.4 Human Capital Accumulation

Workers accumulate human capital as in Ben-Porath (1967). Starting from 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.

3.5 Labor Market Frictions

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. Therefore, when a worker exerts search effort, he increases his probability of meeting an outside firm.[10] θ is defined as the number of vacancies v per unit of aggregate search effort S, $\theta = v/S$.[11] 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 given at this point, wage bargaining is over the piece rate r. They will form a match if the worker is better off with the new firm relative to his outside option.

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 decreasing function of wage, $\Lambda(hpr)$.[12]

3.6 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 as well as the worker's reservation strategies for accepting a new job. Every resulting objects is a function which depends on the characteristics of the worker and firm(s). This section develops the functions. Behind every function is a surplus sharing rule where, if the worker

[10] Mukoyama et al. (2018) show empirically that greater search effort (in the form of time spend searching for a job) is significantly and positively correlated with the probability of finding a job.

[11] S is mathematically defined in Appendix A.1 Equation (24), but it does not need to be solved for in computing the model; the relevant object is θ.

[12] 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 life cycle (Ozkan et al., 2023; Jarosch, 2023).

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 3.8. 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 the firm's outside option is zero. So, if $r = 1$, the firm is indifferent between maintaining the match or not.

There are two scenarios where wage bargaining arises: when an employed worker meets an outside firm, and when an unemployed worker meets a firm. Consider first the scenario of a meeting between an employed worker with value $E_{ai}(h,p,r)$ 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.

Within this scenario, there are three possible outcomes. First, 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 from the match.[13]

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. This is

[13] 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) which uses 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, regardless of how the total surplus is shared. So, Equation (2) states 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, and 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) and impose this wage structure even with curvature in utility and disutility over effort. For an approach which uses total surplus but allows for curvature in utility, see Lise et al. (2016).

the second possible outcome. 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

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

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

The third and final outcome is that the outside firm has a lower productivity than the incumbent firm, $p' < p$, but 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 ignores 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 $p' > p$, the worker is poached, and the worker earns the piece rate $R^P_{ai}(h, p, p')$. 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 $p' < q_{ai}(h, p, r)$, the outside firm cannot compete with the current firm and the worker stays with the same firm at the same piece rate.

Next, 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 $U_{ai}(h, w, \iota)$ be the value of unemployment for a worker with learning ability a, in stage of life i, with human capital h, who previously earned the pre-tax wage w, and is eligible for unemployment benefits with indicator ι. 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 $R^U_{ai}(h, w, \iota, p')$ which solves

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

The piece rate is set such that the worker gets the value of his outside option, $U_{ai}(h, w, \iota)$, 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 $r = 1$. Let $z_{ai}(h, w, \iota)$ be the lowest value of p' which will entice the worker to leave unemployment. $z_{ai}(h, w, \iota)$ solves[14]

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

3.7 Government

I model both income taxes and means-tested transfers in a single average tax equation. The tax equation is of the same functional form as Bénabou (2002). At wage hpr, workers pay the average tax rate

$$T(hpr) = 1 - \tau_0 (hpr)^{-\tau_1} \tag{7}$$

which subsumes both income taxes paid and means-tested transfers received. It is possible that workers with low wages pay negative taxes. If so, it implies that transfer payments exceed taxes paid. Going forward, I describe Equation (7) as a tax function for simplicity, though it should be understood that the function is carefully calibrated to combine progressive taxation and means-tested transfers.

The two parameters of the tax function are easily interpretable: 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, and I let τ_0 adjust so that the government budget constraint holds in equilibrium.

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 earned before becoming unemployed. Unemployment benefits expire with probability χ, after which the worker gets a transfer payment T_0. I denote whether unemployed workers qualify for unemployment benefits with the indicator $\iota \in \{0, 1\}$.

Finally, retired workers earn a flat social security payment, SS.

[14] The algebra is as follows: if $p' = z_{ai}(h, w, \iota)$, then unemployed worker is indifferent between remaining unemployed and working for the firm at the highest possible wage. So,

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

which simplifies to Equation (6).

3.8 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

$$\begin{aligned}\rho E_{ai}(h,p,r) = \max_{l,s}\ & u(c) - d_E(l+s) + (a(lh)^\omega - \delta h)\frac{\partial E_{ai}(h,p,r)}{\partial h} \\ +\Lambda(hpr)&\left[U_{ai}(h,hpr,1) - E_{ai}(h,p,r)\right] + \zeta\left[E_{a,i+1}(h,p,r) - E_{ai}(h,p,r)\right] \\ +sm(\theta)&\Bigg(\int_{q_{ai}(h,p,r)}^{p}\Big[E_{ai}\left(h,p,R^R_{ai}(h,p,p')\right) - E_{ai}(h,p,r)\Big]\,dF(p') \\ &+\int_{p}^{\overline{p}}\Big[E_{ai}\left(h,p',R^P_{ai}(h,p,p')\right) - E_{ai}(h,p,r)\Big]\,dF(p')\Bigg)\end{aligned} \tag{8}$$

subject to

$$c = [1 - T(hpr)]hpr. \tag{9}$$

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 with unemployment benefits. With probability ζ, the worker ages to the next stage of life. And 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

$$\begin{aligned}\rho U_{ai}(h,w,\iota) = \max_{s}\ & u(c) + d_U(s) - \delta h\frac{\partial U_{ai}(h,w,\iota)}{\partial h} \\ +\zeta&\left[U_{a,i+1}(h,w,\iota) - U_{ai}(h,w,\iota)\right] + \chi\left[U_{ai}(h,w,\iota) - U_{ai}(h,w,0)\right] \\ +sm(\theta)&\int_{z_{ai}(h,w)}^{\overline{p}}\Big[E_{ai}\left(h,p',R^U_{ai}(h,w,p')\right) - U_{ai}(h,w,\iota)\Big]\,dF(p')\end{aligned} \tag{10}$$

subject to

$$c = \mathbb{1}[w>0](1-T(bw))bw + \mathbb{1}[w=0]T_0 \tag{11}$$

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

$$\left(\rho + \overline{\zeta}\right)\overline{E} = u(SS). \tag{12}$$

Employed and unemployed workers transition to retirement according to

$$E_{a,I+1}(h,p,r) = \overline{E}. \tag{13}$$

and

$$U_{a,I+1}(h,w) = \overline{E}. \tag{14}$$

3.9 Firms

I model firms as one worker-one job matches. 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 will 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 $s_{Uai}(h,w,\iota)$ 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

$$\begin{aligned} \rho J_{ai}(h,p,r) = (1-\tau_b)hp(1-r) + (a(l_{ai}(h,p,r)h)^{\omega} - \delta h)\,\frac{\partial J_{ai}(h,p,r)}{\partial h} \\ +\zeta\left[J_{a,i+1}(h,p,r) - J_{ai}(h,p,r)\right] \\ +s_{Eai}(h,p,r)m(\theta)\int_{q_{ai}(h,p,r)}^{p}\Big[J_{ai}\left(h,p,R^{R}_{ai}(h,p,p')\right) - J_{ai}(h,p,r)\Big]\,dF(p') \\ +\Big[s_{Eai}(h,p,r)m(\theta)\left(F(\overline{p}) - F(p)\right) + \Lambda(hpr)\Big](-J_{ai}(h,p,r)) \end{aligned} \tag{15}$$

Note that Equation (15) resembles Equation (8).

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

$$J_{a,I+1}(h,p,r) = 0. \tag{16}$$

Let $\Psi_E(h,p,r|a,i)$, $\Psi_U(h,w|a,i,\iota)$, and $\Psi_R(a)$ denote the distributions of employed, unemployed, and retired workers, respectively.[15] The free entry condition is

$$\kappa = m_f(\theta)\Bigg[\sum_{a,i}\int_{\Psi_E} s_{Eai}(h,p,r)\int_p^{\overline{p}} J_{ai}\left(h,p',R_{ai}^P(h,p,p')\right)dF(p')d\Psi_E(h,p,r|a,i)$$
$$+\sum_{a,i,\iota}\int_{\Psi_U} s_{Uai}(h,w)\int_{z_{ai}(h,w)}^{\overline{p}} J_{ai}\left(h,p',R_{ai}^U(h,w,p')\right)dF(p')d\Psi_U(h,w|a,i,\iota)\Bigg]. \tag{17}$$

The left had side of Equation (17), κ, is the cost of posting and maintaining an open vacancy. The right hand side of Equation (17) 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.

3.10 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

$$\sum_{a,i}\int_{\Psi_E} T(hpr)hpr\,d\Psi_E(h,p,r|a,i) + \sum_{a,i}\int_{\Psi_U} T(bw)bw\,d\Psi_U(h,w|a,i,\iota=1)$$
$$+\sum_{a,i}\int_{\Psi_E}\tau_b(1-r)hp\,d\Psi_E(h,p,r|a,i) = \sum_{a,i}\int_{\Psi_U} bw\,d\Psi_U(h,w|a,i,\iota=1) \tag{18}$$
$$+\sum_{a,i}\int_{\Psi_U} T_0\,d\Psi_U(h,w|a,i,\iota=0) + \sum_a \Psi_R(a)SS + \overline{g}.$$

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 (18). The right hand side consists of outlays in the form of unemployment benefits, transfers to workers without unemployment benefits, social security payments, and public consumption, $\overline{g}$. 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.

[15] These distributions are defined such that

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

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

My focus is the stationary equilibrium. I define the recursive stationary equilibrium in Appendix C.

4 Insights from a Two-Period Model

Before proceeding to quantitative exercises, I analytically inspect the economics of the worker's problem. To do so, I use slightly simplified discrete-time two-period model. I describe the model formally and write the worker's problem in Appendix B. Suppose workers live for two periods and discount the second period at discount factor β. They experience utility from consumption both periods, but can only invest in human capital accumulation or job search in the first period. Workers earn a constant piece rate $r = 1$ and there is no human capital depreciation. With probability $sm(\theta)$, workers move from firm productivity p to $p + \Delta_p$.

For a worker with in the first period of life with human capital h, firm productivity p, and learning ability a, the first order conditions state that the optimal choices for human capital effort l and search effort s solve

$$\phi(1+\gamma)(l+s)^{\gamma} = \frac{\beta(1-\tau_1)\omega}{(lh)^{1-\omega}/a + l} = \frac{\beta(1-\tau_1)}{p/(m(\theta)\Delta_p) + s}. \tag{19}$$

The left hand side of this equation is the marginal cost of investing in human capital or search effort,[16] the middle part is the marginal benefit of human capital accumulation, and the right hand side is the marginal benefit of search effort. When the worker is behaving optimally, all three are equal. Note that tax progressivity, τ_1, enters directly in the marginal benefit of human capital accumulation and search effort. A small simplification yields

$$\frac{\phi(1+\gamma)(l+s)^{\gamma}}{\beta(1-\tau_1)} = \frac{\omega}{(lh)^{1-\omega}/a + l} = \frac{1}{p/(m(\theta)\Delta_p) + s}. \tag{20}$$

Consider two otherwise identical workers except that one has greater learning ability than the other, $a_1 > a_0$. The marginal benefit of human capital investment, the middle part of Equation (20), is greater for the worker with high a. In order for Equation (20) to hold, the worker will increase l. However, by increasing l, the worker increases the marginal cost of effort on the left hand side. Thus, for Equation (20) to hold, the worker must also decrease s. In all cases, the worker will not adjust l or s without adjusting the other.

It can be shown that the optimal decision rules for the two agents satisfy $l_1 > l_0$, $s_1 < s_0$, and $l_1 + s_1 > l_0 + s_0$.[17] In words, the agent with more learning ability will invest more in human capital accumulation, less in search, and exert more effort overall. Since the worker has greater learning ability, he will substitute away from search toward human capital, but

[16] Because disutility is simply a function of $l+s$, the marginal cost of investing in either activity will always be equivalent.

[17] The mathematical proof comes from the fact that all other possibilities lead to a contradiction.

the substitution is imperfect; the decrease in l exceeds the decrease in s.[18] Within the period, we can therefore say that human capital investment and job search effort are akin to imperfect substitutes.

It follows that, compared to a pure human capital model, endogenous search effort has an equalizing force in my model. Instead of struggling to accumulate human capital, workers with low learning ability can direct their efforts toward on-the-job search.

With symmetry, the same holds for differences in $m(\theta)$. In an economy with more open vacancies and greater $m(\theta)$, workers will substitute away from human capital accumulation towards job search, but imperfectly. Overall, since the substitution is imperfect, there will be more wage growth in an economy where θ is greater.

In a dynamic sense, there is also some complementarity at work in Equation (20). Consider two otherwise identical workers except that one worker works for more productive firm, $p_1 > p_0$. It can be shown that $l_1 > l_0$, $s_1 < s_0$, and $l_1 + s_1 < l_0 + s_0$. In words, the worker at the more productive firm will invest more in human capital accumulation, less in in search, and exert less effort overall.[19] Recall that h and p are complements in production; at a more productive firm, the benefit of increasing human capital is greater. So, the worker is incentivized to invest in h such that the level of h "catches up" with the high level of p. With symmetry, the same logic holds for differences in h. Therefore, contrary to the substitution story above, there is reason to believe that workers with high learning ability - and, therefore, high human capital - may work at more productive firms. This effect, combined with the fact that l and s are imperfect substitutes, implies that the model may generate more wage inequality compared to a model with a single channel.

Finally, consider otherwise identical economies except that tax progressivity is greater in one economy, $\tau_1 > \tau_0$. It can be shown that both l and s will be lower in the economy with higher tax progressivity. More progressive taxes discourage wage growth in all of its forms because, either way, when the worker's wage increases, the benefit of the increased wage is smaller.

5 Calibration

5.1 Data

I rely on microdata from the Survey of Income and Program Participation (SIPP) to parameterize the model. The SIPP is a panel data set with interviews every four months. Within interviews, respondents report on what occurred in the time between interviews.[20] I use every panel between the years 1990 and 2019, 12 panels in total.[21]

[18] Indeed, across learning abilities, we will see imperfect substitution between l and s in the calibrated model. See Figure 4.

[19] Note that $s_1 < s_0$ even without decreasing returns to search as workers move up the job ladder, which we have in the quantitative model. The reason is that the worker is substituting away from search effort.

[20] 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.

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

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 growth across months. Second, the SIPP tracks worker-job matches, which allows me to observe job switches.[22]

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 and drop the self-employed. Observations with wages which are below the federal minimum wage are dropped. For every statistic, I calculate the weighted mean within the panel using panel weights, then take the weighted mean across panels where each panel is weighted by the number of months is covers. I drop the first two and last two months of each panel. Though the SIPP has weekly labor force indicators, I aggregate to a monthly frequency and use the labor force indicator for the second week of the month to mirror the CPS.[23]

5.2 Functional Forms and Distributions

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

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.[24] 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 $\varepsilon \sim LN(0, \sigma_\varepsilon)$.[25] I describe how I discretize the distribution of a in Appendix D.2.

I make one final adjustment to the model. To help the model fit the data, I allow for "godfather shocks." 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.[26] 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.[27] All equations which are affected by the godfather shock are updated in Appendix

[22] 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).

[23] Unemployment-to-employment and employment-to-unemployment rates in the SIPP are consistently lower than the familiar values in the CPS. See the discussion in Footnote 16 in Menzio et al. (2016).

[24] 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.

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

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

[27] See Dorn (2018) and Tjaden and Wellschmied (2014) for investigations into job-to-job transitions with a wage decrease, and see Moscarini and Postel-Vinay (2018) for an example of godfather shocks in use.

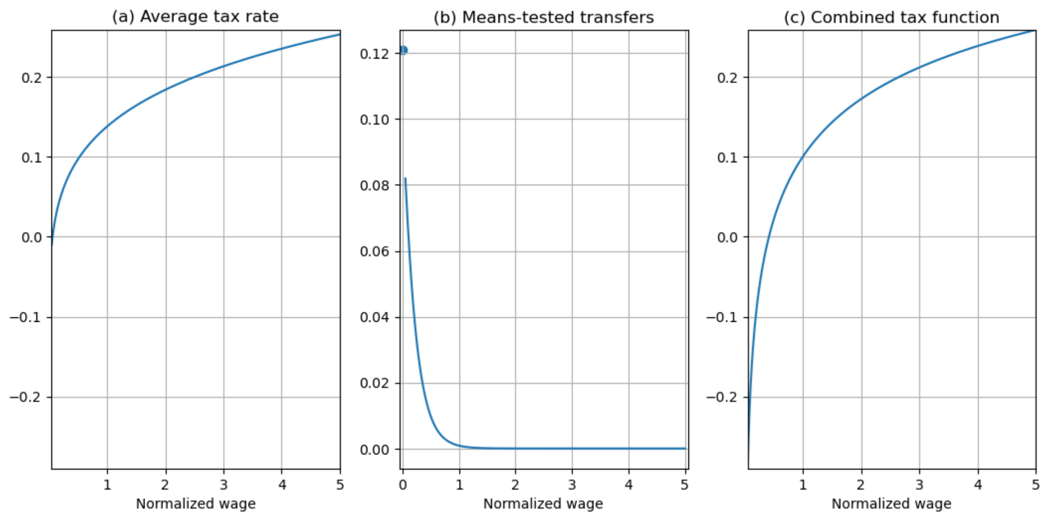

Figure 1: Combined tax function and its ingredients

A.2.

5.3 Taxes and Transfers

My goal in calibrating taxes and transfers is to accurately represent the US system while maintaining a simple structure. I therefore estimate Equation (7) such that it fits both the progressive tax system and the means-tested transfer system in the US. My strategy consists of borrowing an estimated average tax function and an estimated means-tested transfer function from the literature. I then re-estimate Equation (7) on the combination of both functions. The key difference is that including transfers significantly increases progressivity compared to a function without transfers.

For an estimated tax function, I use an estimate of federal income taxes from Guner et al. (2014) for all households (married and unmarried) which takes the earned income tax credit (EITC) into account.[28] To this function, I add a flat state-and-local income tax rate of 5% (Guner et al., 2022). Wages are normalized such that the mean wage in the cross section is one. Using these estimates, for normalized income y, the average tax rate is

$$Tax(y) = -0.294 + 0.382y^{0.164} + 0.05.$$

The average tax function is plotted in Panel (a) of Figure 1.

For an estimated means-tested transfer function, I use an estimated transfer function for all households from (Guner et al., 2023). The means-tested transfer function takes the following programs into account: 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),

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

TANF (Temporary Assistance for Needy Families), and housing. For a discussion of these programs, see Guner et al. (2023). The estimated transfer function is

$$Tr(y) = e^{-2.122}e^{-4.954y}y^{0.044} \tag{21}$$

and is plotted in Panel (b) of Figure 1.

When I combine taxes and transfers and estimate the parameters of Equation (7), I find $\tau_0 = 0.899$ and $\tau_1 = 0.120$.[29] The new tax function is plotted in Panel (c) of Figure 1. Comparing Panel (c) and Panel (a), it is clear that including transfers significantly increases progressivity. Panel (c) states that those who earn about 30% or less of the mean wage benefit from means-tested transfers.

I set the transfer payment for workers without unemployment benefits to $T_0 = 0.121$ as estimated in Guner et al. (2023) . The flat tax rate on firm profit is $\tau_b = 0.243$, the estimated average tax on business income in Cooper et al. (2016).

5.4 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 $\overline{\zeta} = 1/(10 \times 12)$.

Concerning other government parameters, 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 $SS = .5$.[30]

I set $\alpha = 0.5$ (Petrongolo and Pissarides, 2001) and $\eta = 0.4$ (Bagger et al., 2014).[31] 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 $\delta = 0.00116$ to match the decline in wages at the end of life.[32] 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. Recall that the mean wage in the model is normalized to 1. $\Lambda(hpr)$ is presented graphically in Appendix D.1.

[29] These estimates are similar to Holter et al. (2019).

[30] 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.

[31] Engbom (2022) estimates a similar value for this parameter.

[32] Assume that workers do not invest in human capital in the last stage of life, $i = I$. Without job switching or unemployment, wages are multiplied by $1 - \delta$ every 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 $\delta = 0.00116$.

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	$\left(\frac{42\times 12}{I}\right)^{-1}$	Working for 42 years on average (ages 23-65)
$\overline{\zeta}$	Probability of death for the retired	$(10\times 12)^{-1}$	Retired for 10 years on average
Policy			
τ_0	Tax+transfer progressivity	0.899	Guner et al. (2014, 2023)
τ_1	Tax+transfer level	0.120	Guner et al. (2014, 2023)
T_0	Transfer for worker without UI	0.121	Guner et al. (2023)
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

5.5 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.[33]

I target the lifeycle 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 life cycle: 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), and I normalize mean $l + s_E$ to equal 0.1.[34]

5.6 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 clear one-to-one relationships between moments and parameters. γ_U pins down the unemployment-to-employment rate. 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 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.

Next, the level parameter of the learning ability distribution, μ_a, pins down the mean monthly wage increase for job stayers. This leaves ω as the closest parameter to the concave shape of wages over the lifecycle.

Following the logic of Huggett et al. (2011) (and given the parameters and moment accounted for to this point), I associate the dispersion parameters of the learning ability and initial human capital distribution $(\sigma_a, \lambda_a, \eta_1, \sigma_\varepsilon)$ with the variance of wages, skewness of

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

[34] The final normalization is not necessary, but it is convenient because it keeps other parameter levels contained. Since I have level parameters on the disutility of effort (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.

wages, and lifecycle profile of the variance of wages. Given that it determines the skewness of a,λ_a is mostly closely related to the skewness of wages. 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.

That leaves 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. In particular, with ω determining the decreasing returns for human capital accumulation, γ_E determines the decreasing returns to search effort.

One final note. It is important that I match monthly wage growth for job stayers versus job switchers. The difference allows me to differentiate the month-to-month wage growth which comes from the human capital channel versus the search channel. One may object 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 from human capital accumulation or from renegotiation. Thus, it is not the case that the wage growth of job stayers only reflects human capital accumulation. 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 for job switchers is due to the search. Then, for job stayers, the residual wage growth which is unidentified from the search channel comes from human capital growth.

5.7 Calibration Results

The fit between the model and lifecycle moments is illustrated in Figure 2, and the fit between the model and nonlifecycle moments is presented in Table 2. Overall, I am able to match the lifecycle profiles of mean log wages and job changes rates well. The level of wage dispersion is slightly off, but the slope over the lifecycle is good.

The parameters which deliver this fit are listed in Table 3. A few of these estimates deserve a brief discussion. First, my estimate of the curvature parameter in the human capital production function, $\omega = 0.52$, is in line with micro estimates (Browning et al., 1999).[35] Second, the disutility function has significantly more curvature for the unemployed . This means that unemployed workers can invest more in human capital before the marginal cost of investment becomes prohibitively large, which is necessary to match the large difference in unemployment-to-employment rates compared to job switching rates. Third, my calibrated value for the dispersion of learning ability, σ_a, is relatively large compared to the literature.[36] I attribute this to the fact - described above - that endogenous search acts as an equalizer for wages. In order to generate the same level of wage dispersion, compared to a model without endogenous search, I need more dispersion in learning ability.

[35] For discussions of this parameter, see Blandin (2018) and Browning et al. (1999).

[36] Compare with estimates in Badel et al. (2020) and Esfahani (2020).

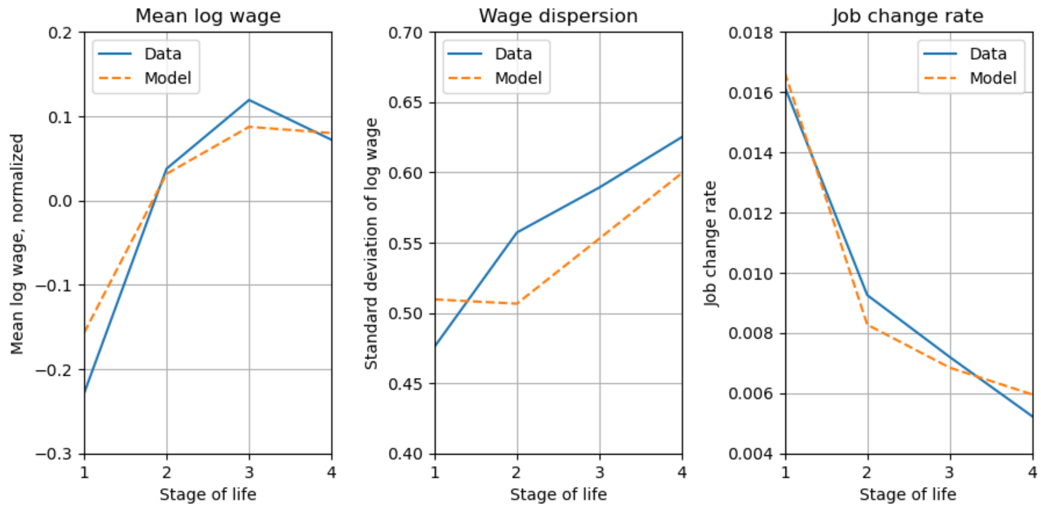

Figure 2: Fit between data and model, lifecycle moments

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

As illustrated in Panel (c) of Figure 2, the job switching rate decreases over the lifecycle. It is important that I replicate this moment because it follows directly from my on-the-job search environment (Burdett and Mortensen, 1998). In a Burdett-Mortensen framework, workers move to more productive firms as they age. But since the firm productivity distribution is fixed, an older the worker is less likely to meet an outside firm which is more productive than his current firm. The implication is that young workers are more likely to switch jobs.

6 Properties of Benchmark Model

Before experimenting with taxes, I describe some properties of the benchmark model.

6.1 Mean Wages

Recall that the worker earns the wage hpr. Given this structure, I can easily decompose log wages between human capital h, firm productivity p, and the bargained piece rate r. The

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

log wage is the sum of the log of the three components,[37]

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

The mean log wage follows,

$$\mathrm{E}[\ln(wage)] = \mathrm{E}[\ln(h)] + \mathrm{E}[\ln(l)] + \mathrm{E}[\ln(r)]. \tag{22}$$

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 3 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.[38]

However, the means mask significant heterogeneity in the ways that workers grow wages. A worker's choice of l and s depend on his state, age, and fixed learning ability. In practice, learning ability is especially important. As described in Section 4, workers with high learning ability will substitute away from search effort toward human capital accumulation, and vice versa.

The result is that human capital growth is significantly heterogeneous, and the heterogeneity is driven by in human capital. In Figure 4, and going forward, I split my distribution

[37]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.

[38]Bayer and Kuhn (2019) find a similar result.

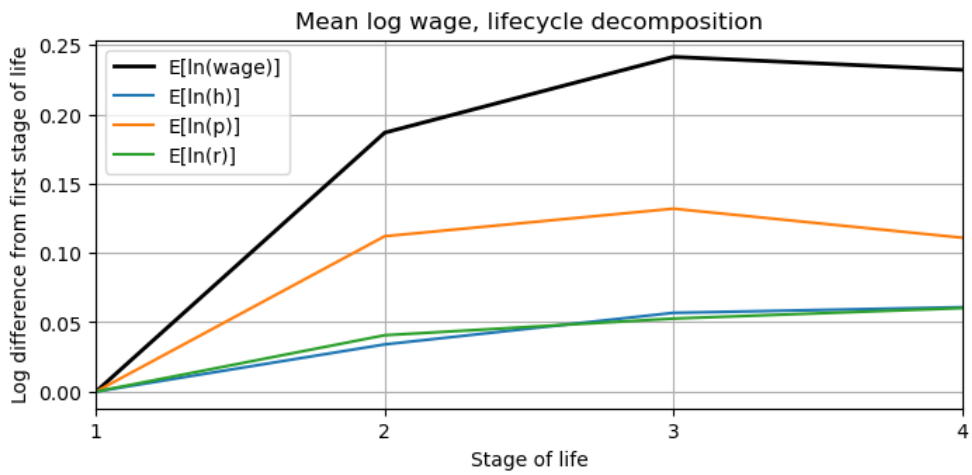

Figure 3: Additive decomposition of log wage growth

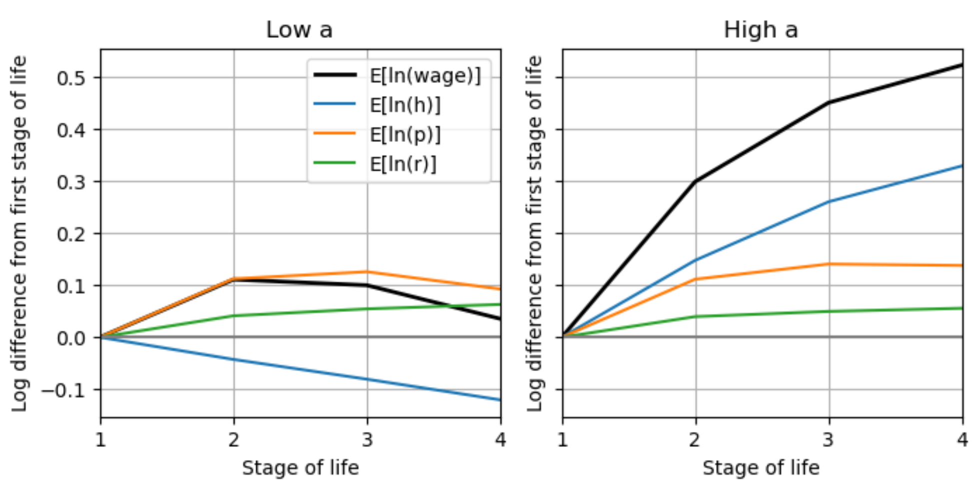

Figure 4: Additive decomposition of log wage growth by learning ability

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

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

of a values into two bins: the lower 60% and the upper 40%.[39] Figure 4 replicates the same decomposition of mean log wages for each of the two learning ability bins. Clearly, human capital growth differentiates the two groups. In fact, the growth of p and r is quite similar across groups. Furthermore, those with high learning ability rely mostly on human capital for wage growth, while those with low learning ability rely entirely on search. Those with low learning ability even experience a loss of human capital over the life cycle.[40] In general, this finding supports the conclusions of Ozkan et al. (2023) and Bagger et al. (2014), which show 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. The difference is completely due to human capital.

6.2 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 log component plus an interaction term for each pair of log components,

$$\begin{aligned}\mathrm{Var}[\ln(wage)] = &\mathrm{Var}[\ln(h)] + \mathrm{Var}[\ln(p)] + \mathrm{Var}[\ln(r)] \\ &+ 2\,\mathrm{Cov}[\ln(h), \ln(p)] + 2\,\mathrm{Cov}[\ln(h), \ln(r)] + 2\,\mathrm{Cov}[\ln(p), \ln(r)]\end{aligned} \tag{23}$$

Using this decomposition, I find that the increase in wage dispersion over the lifecycle is completely driven by human capital. In Figure 5, I decompose wage dispersion over the lifecycle into the six terms in Equation (23). 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

[39] These values translate conveniently to how I discretize the distribution of a. See Appendix D.2.

[40] One way to rationalize the negative human capital growth is that my calibration begins at age 23. It is not entirely unreasonable to presume that low-wage workers have more skills when they are young than when they are old. It may be the case that they forget what they learned in school or their skills depreciate as technology changes.

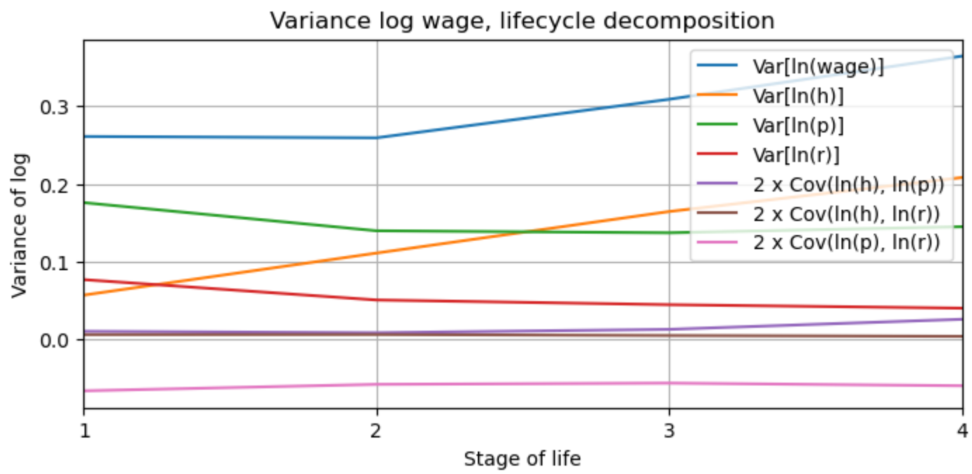

Figure 5: Simple decomposition of the variance of log wages

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. Over the lifecycle, the gap between workers with high learning ability and those with low learning ability increases. 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.

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 close to complements that substitutes at the micro 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 6, 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.

Despite the substitutability of human capital and search, Figure 7 shows that the correlation between h and p is positive and U-shaped over the lifecycle.[41] Recall that (a) h and

[41] The literature following Abowd et al. (1999) has tended to find a positive correlation between estimated

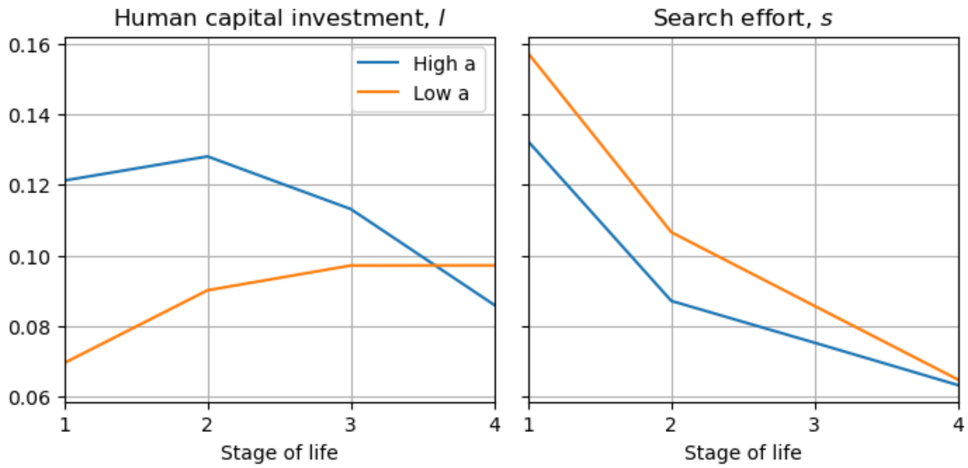

Figure 6: Mean policy functions across learning abilities

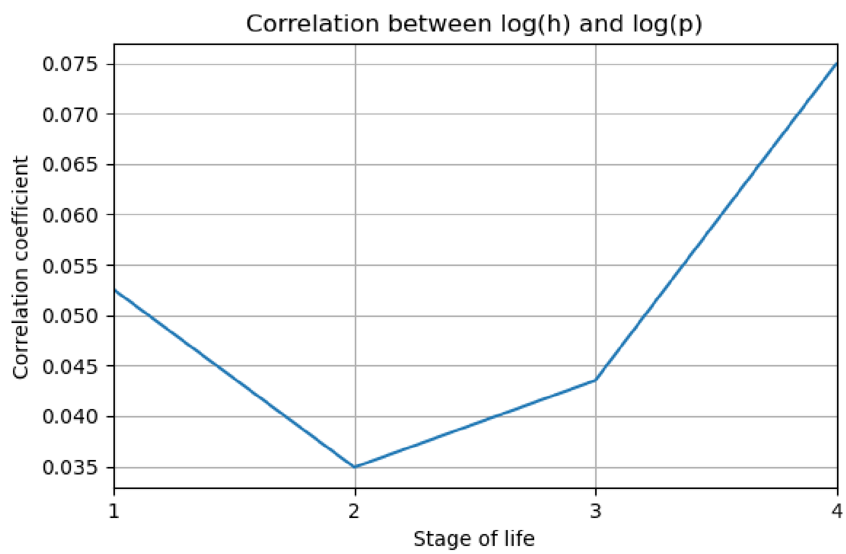

Figure 7: Correlation of h and p

p are complements in production and (b) h and p are not perfect substitutes. Both forces imply that workers with high learning ability and high h are still incentivized to climb the job ladder.

Why is the correlation between h and p U-shaped? 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 6, workers with low learning ability exert more search effort. So, those with low learning ability 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 (Jarosch, 2023). Thus, over the course of the lifecycle, the correlation between h and p increases.

7 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, if a worker increases their wage, then they take home a smaller part of the wage. 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. 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 objects which adjust so that model is in equilibrium under the new τ_1: θ, which adjusts such that the free entry condition (17) holds, and τ_0, which adjusts such that government budget constraint (18) holds.

7.1 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 with double the progressivity, $\tau_1 = 0.24$. Setting $\tau_1 = 0.24$ puts the level of tax-and-transfer progressivity in line with countries such as Denmark, Finland, Germany, the Netherlands, and Sweden.[42]

I plot the average tax rate for each tax system in Panel (a) Figure 8.[43] As τ_1 increases, the average tax curve increases in curvature; those with low wages earn more transfers, and those with high wages pay more taxes. The increase in τ_1 also increases the threshold at which workers are net tax payers from about 1/3 of the mean wage to 2/3 of the mean wage.

worker and firm productivity (though the original study did not); see Card et al. (2018) for a discussion. In a search-and-matching model with the same flexible wage bargaining structure and on-the-job search, Bagger and Lentz (2019) find that a substantial portion of wage dispersion is due to positive sorting.

[42]See the appendix of Holter et al. (2019).

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

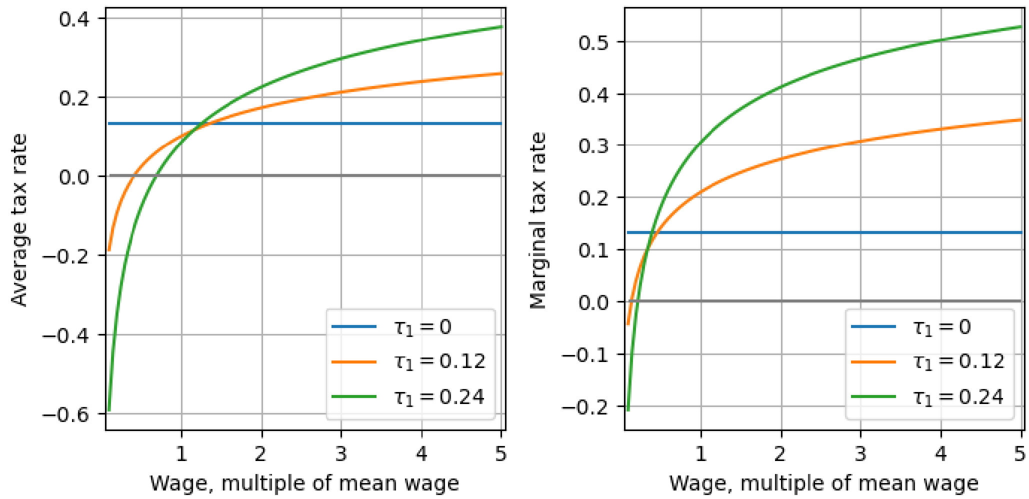

Figure 8: Tax rates by wage

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

Marginal tax rates are plotted in Panel (b) of Figure 8. 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 9 and Table 5. In Figure 9, 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 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

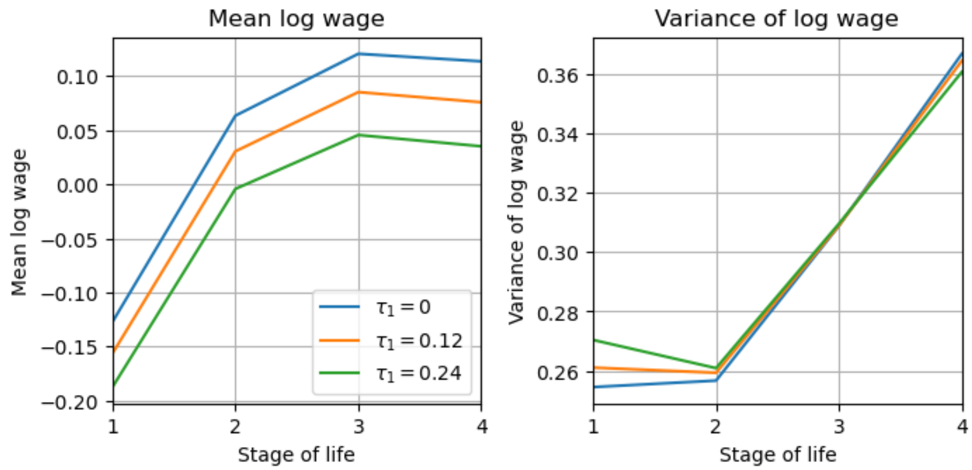

Figure 9: Effect of tax progressivity on lifecycle profiles of mean log wage and variance of log wages

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.

7.2 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 (22). According to the overall differences of $\mathrm{E}[h]$, $\mathrm{E}[p]$, and $\mathrm{E}[r]$ (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 $\mathrm{E}[\ln(p)]$ is twice as large as the decrease in $\mathrm{E}[\ln(h)]$.

Earlier in the paper, I showed that moving to more productive firms is the primary factor behind lifecycle wage growth for the mean worker. 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 policy response 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 $\mathrm{E}[\ln(h)]$ is over 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. This is related to the result in the previous section that only workers with high learning ability rely heavily on accumulating human capital for lifecycle wage growth while all workers are roughly equally successful in climbing the job ladder.

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

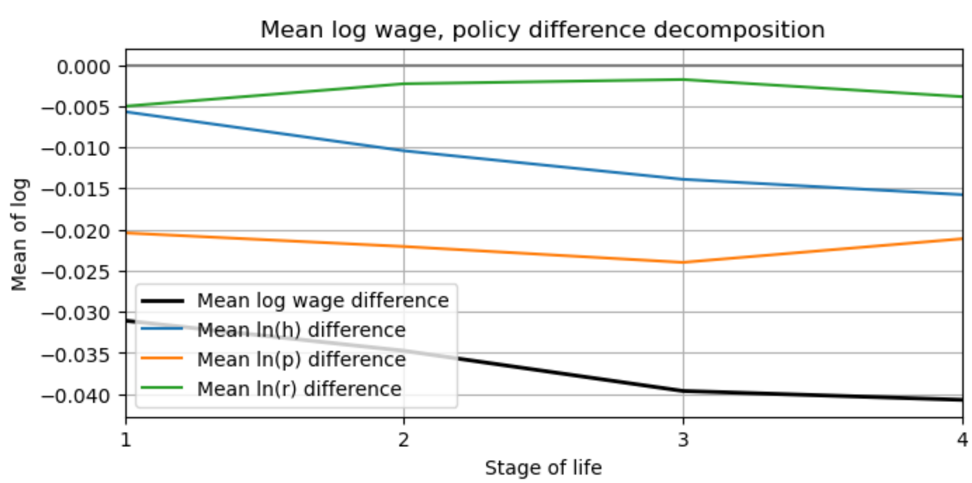

Figure 10: Mean log wage, difference from tax chance, decomposition of change

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, $\text{Var}[\ln(h)]$	0.136	0.129	-0.007
Variance from search channel	0.149	0.156	0.008
$\text{Var}[\ln(p)]$	0.152	0.155	0.003
$\text{Var}[\ln(r)]$	0.054	0.058	0.004
$2\,\text{Cov}[\ln(p), \ln(r)]$	-0.057	-0.057	0.000
Variance of interaction	0.024	0.024	0.000
$2\,\text{Cov}[\ln(h), \ln(p)]$	0.017	0.018	0.001
$2\,\text{Cov}[\ln(h), \ln(r)]$	0.007	0.006	0.000
Variance within groups	0.230	0.234	0.004
Variance between groups	0.078	0.075	-0.003

With regards to lifecycle wage growth, its decrease in response to an increase in tax progressivity is driven by a decrease in human capital accumulation. Figure 10 shows the difference between 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 decrease in the rate at which $\ln(p)$ grows, 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. So, 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 offset one another. In Table 7, I decompose the effects of increased tax progressivity on the variance of log wages using Equation (23). I group the additive components of wage variance on the right hand side of Equation (23) into three parts: (1) the variance from the human capital channel, $\text{Var}[\ln(h)]$; the variance from the search channel, $\text{Var}[\ln(p)] + \text{Var}[\ln(r)] + 2\,\text{Cov}[\ln(p), \ln(r)]$; and the variance from interactions between human capital and search, $2\,\text{Cov}[\ln(h), \ln(p)] + 2\,\text{Cov}[\ln(h), \ln(r)]$.

When I increase tax progressivity, the variance from the human capital channel decreases while the variance from the search channel increases. Table 6 illustrates why the variance of human capital channel decreases. Human capital levels decrease for all workers, the but 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. Since wage dispersion decreases as a result of high wages decreasing, we can say that the human capital channel decreases wage dispersion 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 more at the top of the wage distribution, and the variance of the search component is less than it was initially.

When tax progressivity increases, workers exert less search effort, are less likely to meet outside firms, and do not climb the job ladder as quickly. 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.

7.3 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 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.[44]

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).[45]

In the second out-of-equilibrium model, I fix l^* and θ^*. Using the same logic as above,

[44] Badel et al. (2020) does a similar experiment.

[45] 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

this is a model where the path of human capital is fixed, and workers can only respond to the policy by decreasing 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 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 policy response for the mean worker. 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 that the two channels offset one another.

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. This is a result of the assumption that h and p are complements in production.

Second, endogenous job posting is of some importance for the response of the wage

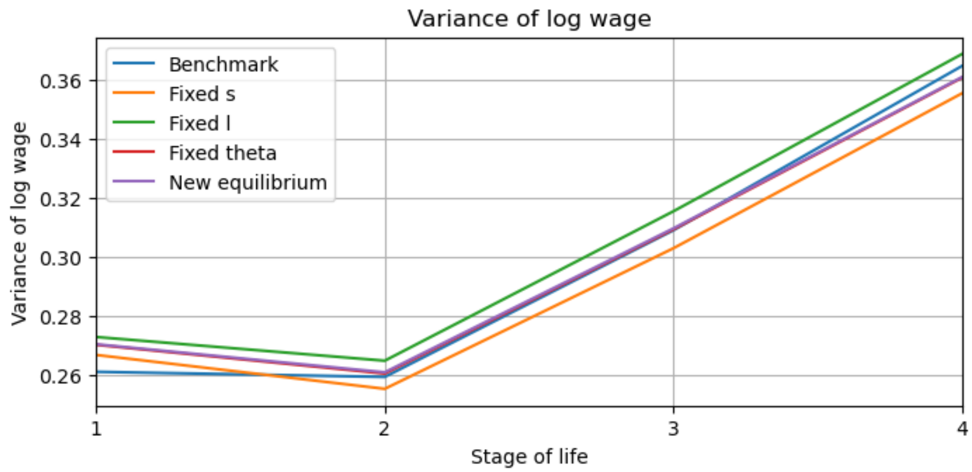

Figure 11: Variance of log wage for out-of-equilibrium models

level. Compare the mean log wage in columns (4) and (5). Column (5) states that, when 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%[46] 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 11 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.

8 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. I endogenize human capital accumulation, on-the-job search, and job posting. I show that these elements interact in ways which are quantitatively important.

In the benchmark calibration, job search is the driving force of lifecycle wage growth for the average worker. However, human capital is the driving force of wage dispersion, particularly as it relates to the top of the wage distribution. Workers take different approaches to growing their wages. In particular, younger workers and workers with low learning ability invest more in search effort, while working with high learning ability invest more in human capital accumulation. Because workers can vary their approach to growing wages, the model requires more dispersion in learning ability in order to generate empirically accurate levels

[46] $(3.7 - 3.3)/3.7 \approx 0.15$.

of wage dispersion.

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 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.

Appendix

A Equations

A.1 Aggregate Search Effort, S

Aggregate search effort is defined as

$$S = \sum_{a,i} \int_{\Psi_E} s_{Eai}(h,p,r) d\Psi_E(h,p,r|a,i) + \sum_{a,i} \int_{\Psi_U} s_{Uai}(h,w) d\Psi_U(h,w|a,i). \tag{24}$$

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 (8), the worker's problem adds a new line like so:

$$\begin{aligned}
\rho E_{ai}(h,p,r) = \max_{l,s}\ & u(c) - d_E(l+s) + (a(lh)^\omega - \delta h) \frac{\partial E_{ai}(h,p,r)}{\partial h} \\
&+ \Lambda(rhp)\left[U_{ai}(h,hpr,\iota) - E_{ai}(h,p,r)\right] + \zeta\left[E_{a,i+1}(h,p,r) - E_{ai}(h,p,r)\right] \\
&+ sm(\theta)\Bigg(\int_{q_{ai}(h,p,r)}^{p} \Big[E_{ai}\left(h,p,R^R_{ai}(h,p,p')\right) - E_{ai}(h,p,r)\Big]\, dF(p') \\
&+ \int_{p}^{\overline{p}} \Big[E_{ai}\left(h,p',R^P_{ai}(h,p,p')\right) - E_{ai}(h,p,r)\Big]\, dF(p') \Bigg) \\
&+ \psi \int_{\underline{p}}^{\overline{p}} \Big[E_{ai}\left(h,p',R^G_{ai}(h,hpr,p')\right) - E_{ai}(h,p,r)\Big]\, dF(p')
\end{aligned} \tag{25}$$

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 (15) like so:

$$\begin{aligned}
\rho J_{ai}(h,p,r) = &\ hp(1-r)(1-\tau_b) + (a(l_{ai}(h,p,r)h)^\omega - \delta h) \frac{\partial J_{ai}(h,p,r)}{\partial h} \\
&+ \zeta\left[J_{a,i+1}(h,p,r) - J_{ai}(h,p,r)\right] \\
&+ s_{Eai}(h,p,r)m(\theta) \int_{q_{ai}(h,p,r)}^{p} \Big[J_{ai}\left(h,p,R^R_{ai}(h,p,p')\right) - J_{ai}(h,p,r)\Big]\, dF(p') \\
&+ \Big[s_{Eai}(h,p,r)m(\theta)\left(F(\overline{p}) - F(p)\right) + \Lambda(hpr) + \psi\Big](-J_{ai}(h,p,r))
\end{aligned} \tag{26}$$

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,

$$\kappa = m_f(\theta)\Bigg[\sum_{a,i}\int_{\Psi_E} s_{Eai}(h,p,r)\int_p^{\overline{p}} J_{ai}\left(h,p',R^P_{ai}(h,p,p')\right)dF(p')d\Psi_E(h,p,r|a,i)$$

$$+\sum_{a,i}\int_{\Psi_U} s_{Uai}(h,w)\int_{z_{ai}(h,w)}^{\overline{p}} J_{ai}\left(h,p',R^U_{ai}(h,w,p')\right)dF(p')d\Psi_U(h,w|a,i)\Bigg] \tag{27}$$

$$+\psi\sum_{a,i}\int_{\Psi_E}\int_{\underline{p}}^{\overline{p}} J_{ai}\left(h,p',R^G_{ai}(h,hpr,p')\right)dF(p')d\Psi_E(h,p,r|a,i).$$

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

$$E_{ai}\left(h,p',R^G_{ai}(h,hpr,p')\right) = U_{ai}(h,0) + \eta\left[E_{ai}(h,p',1) - U_{ai}(h,0)\right]. \tag{28}$$

B Simple Two-Period Model

I write this model in discrete time. Agents live for two periods, $i = \{0,1\}$, and discount the second period at discount factor β. In the first period, agents work, consume, and make investments in human capital accumulation and job search effort; in the second, agents only work and consume.

Assume that workers earn the piece rate $r = 1$ and that there is no human capital depreciation, $\delta = 0$. Otherwise, the human capital accumulation equation is as in the full model. When workers put in search effort s, then, with probability $m(\theta)$, they move from their current p to $p + \Delta_p$.

In the first period, a worker with learning ability a, human capital h, and firm productivity p solves

$$W_{a0}(h,p) = \max_{l,s} u(c) - d_E(l+s) + \beta W_{a1}(h',p') \tag{29}$$

subject to

$$c = [1 - T(hp)]hp$$
$$h' = h + a(lh)^{\omega}$$
$$p' = p + sm\Delta_p.$$

In the second period, the worker simply enjoys the fruits of his labor,

$$W_{a1}(h,p) = u(c)$$

subject to

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

When we solve for the first order conditions from Equation (29), we get Equation (19).

C Equilibrium Definition

The recursive stationary equilibrium consists of a set of value functions $\{E_{ai}(h,p,r), U_{ai}(h,w,\iota), \overline{E}, J_{ai}(h,p,r)\}$, a human capital investment policy function $l_{Eai}(h,p,r)$, a set of search effort policy functions $\{s_{Eai}(h,p,r), s_{Uai}(h,w,\iota)\}$, a set of job search cutoff rules $\{q_{ai}(h,p,r), z_{ai}(h,w,\iota)\}$, a set of wage functions $\{R^P_{ai}(h,p,p'), R^R_{ai}(h,p,p'), R^U_{ai}(h,w,\iota,p'), R^G_{ai}(h,hpr,p')\}$, vacancies v, aggregate search effort S, labor market tightness $\theta = v/S$, a distribution of workers $\Psi = \{\Psi_E(h,p,r|a,i), \Psi_U(h,w|a,i,\iota), \Psi_R(a)\}$, and a set of government policies parameters $\mathcal{G} = \{\tau_0, \tau_1, \tau_b, b, \chi, SS, \overline{g}\}$ 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 $\mathcal{G}$, $E_{ai}(h,p,r)$ solves (25) subject to (9) and (13) with associated decision rules $l_{Eai}(h,p,r)$ and $s_{Eai}(h,p,r)$.

2. Unemployed worker optimization: Given $R^U_{ai}(h,w,p')$, $z_{ai}(h,w,\iota)$, θ, and $\mathcal{G}$, $U_{ai}(h,w,\iota)$ solves (10) subject to (11) and (14) with associated decision rule $s_{Uai}(h,w)$.

3. Retired workers: $\overline{E}$ solves (12).

4. Filled jobs: Given $R^R_{ai}(h,p,p')$, $l_{ai}(h,p,r)$, $s_{Eai}(h,p,r)$, θ, and $\mathcal{G}$, $J_{ai}(h,p,r)$ solves (26) and (16).

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

   (d) $R^G_{ai}(h,hpr,p')$ solves (28).

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

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

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

7. Free entry: Given $J_{ai}(h,p,r)$, $R^P_{ai}(h,p,p')$, $R^U_{ai}(h,w,\iota,p')$, $R^G_{ai}(h,hpr,p')$, $s_{Eai}(h,p,r)$, $s_{Uai}(h,w,\iota)$, S, Ψ, and $\mathcal{G}$, θ solves (27).

8. Government budget constraint: $\mathcal{G}$ satisfies (18).

9. Aggregate search effort: Given $s_{Eai}(h,p,r)$, $s_{Eai}(h,w)$, and Ψ, S satisfies (24).

10. Consistency: Ψ is the stationary distribution.

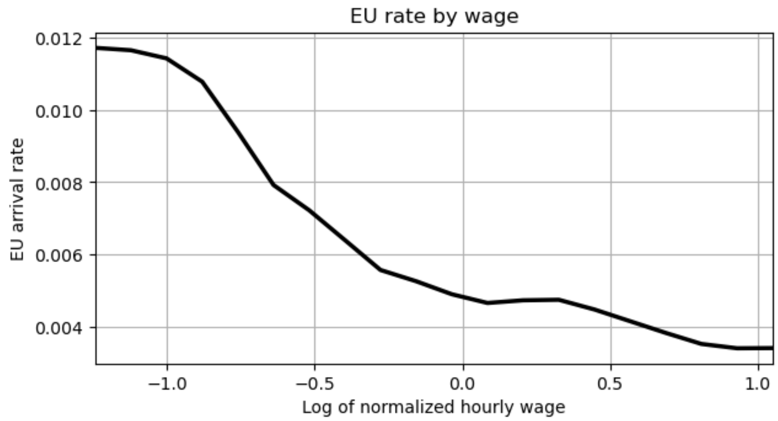

Figure 12: Estimate of exogenous job destruction function, $\Lambda(wage)$

D Calibration Details

D.1 Λ Function

I estimate $\Lambda(hpr)$ directly from SIPP data using interpolation. See Figure 12 for an illustration. As with the tax function, the mean wage is normalized to one.

D.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

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High Wage Workers and High Wage Firms. *Econometrica 67*(2), 251–333.

Acabbi, E. M., A. Alati, and L. Mazzone (2023). A Labor Market Sorting Model of Hysteresis and Scarring.

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.

Bagger, J., R. Holloway, M. Hejlesen, K. Sumiya, and R. Vejlin (2019). Income Taxation and the Equilibrium Allocation of Labor.

Bagger, J. and R. Lentz (2019). An Empirical Model of Wage Dispersion with Sorting. *The Review of Economic Studies 86*(1), 153–190.

Bagger, J., E. R. Moen, and R. Vejlin (2021). Equilibrium Worker-Firm Allocations and the Deadweight Losses of Taxation.

Bayer, C. and M. Kuhn (2019). Which Ladder to Climb? Decomposing Life Cycle Wage Dynamics.

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.

Blandin, A. (2018). Learning by Doing and Ben-Porath: Life-cycle Predictions and Policy Implications. *Journal of Economic Dynamics and Control 90*(C), 220–235.

Bowlus, A. J. and H. Liu (2013). The contributions of search and human capital to earnings growth over the life cycle. *European Economic Review 64*, 305–331.

Browning, M., L. P. Hansen, and J. J. Heckman (1999). Micro data and general equilibrium models. In *Handbook of Macroeconomics*, Volume 1, pp. 543–633.

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.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and Labor Market Inequality: Evidence and Some Theory. *Journal of Labor Economics 36*(S1), S13–S70.

Carrillo-Tudela, C. (2012). Job Search, Human Capital and Wage Inequality.

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.

Cubas, G. and P. Silos (2020). Social Insurance and Occupational Mobility. *International Economic Review 61*(1), 219–240.

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

Dustmann, C. and C. Meghir (2005). Wages, Experience and Seniority. *The Review of Economic Studies 72*(1), 77–108.

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

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

Gentry, W. M. and R. G. Hubbard (2004). The effects of progressive income taxation on job turnover. *Journal of Public Economics 88*(11), 2301–2322.

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., R. Kaygusuz, and G. Ventura (2022). The Looming Fiscal Reckoning: Tax Distortions, Top Earners, and Revenues.

Guner, N., C. Rauh, and G. Ventura (2023). Means-Tested Transfers in the US: Facts and Parametric Estimates.

Guvenen, F., B. Kuruscu, and S. Ozkan (2014). Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis. *The Review of Economic Studies 81*(2 (287)), 818–850.

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

Heathcote, J., K. Storesletten, and G. L. Violante (2017). Optimal Tax Progressivity: An Analytical Framework. *The Quarterly Journal of Economics 132*(4), 1693–1754.

Heathcote, J., K. Storesletten, and G. L. Violante (2020). How Should Tax Progressivity Respond to Rising Income Inequality? *Journal of the European Economic Association 18*(6), 2715–2754.

Holter, H. A., D. Krueger, and S. Stepanchuk (2019). How do tax progressivity and household heterogeneity affect Laffer curves? *Quantitative Economics 10*(4), 1317–1356.

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

Jarosch, G. (2023). Searching for Job Security and the Consequences of Job Loss. *Econometrica 91*(3), 903–942.

Kapička, M. (2006). Optimal income taxation with human capital accumulation and limited record keeping. *Review of Economic Dynamics 9*(4), 612–639.

Kapička, M. (2015). Optimal Mirrleesean Taxation in a Ben-Porath Economy. *American Economic Journal: Macroeconomics 7*(2), 219–248.

Kreiner, C. T., J. R. Munch, and H. J. Whitta-Jacobsen (2015). Taxation and the long run allocation of labor: Theory and Danish evidence. *Journal of Public Economics 127*, 74–86.

Krueger, D. and A. Ludwig (2016). On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium. *Journal of Monetary Economics 77*(C), 72–98.

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

Menzio, G., I. A. Telyukova, and L. Visschers (2016). Directed search over the life cycle. *Review of Economic Dynamics 19*, 38–62.

Mortensen, D. T. (2003). *Wage Dispersion: Why Are Similar Workers Paid Differently?* Cambridge: The MIT Press.

Mortensen, D. T. and C. A. Pissarides (1994). Job Creation and Job Destruction in the Theory of Unemployment. *The Review of Economic Studies 61*(3), 397–415.

Moscarini, G. and F. Postel-Vinay (2018). The Cyclical Job Ladder. *Annual Review of Economics 10*(1), 165–188.

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.

Mukoyama, T., C. Patterson, and A. Şahin (2018). Job Search Behavior over the Business Cycle. *American Economic Journal: Macroeconomics 10*(1), 190–215.

Omer, V. (2004). Wage Growth, Search and Experience: Theory and Evidence.

Ozkan, S., J. Song, and F. Karahan (2023). Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job-Ladder Risk versus Human Capital. *Journal of Political Economy Macroeconomics 1*(3), 506–550.

Pavan, R. (2011). Career Choice and Wage Growth. *Journal of Labor Economics 29*(3), 549–587.

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.

Pissarides, C. A. (2000). *Equilibrium Unemployment Theory* (Second ed.). Cambridge: The MIT Press.

Rauh, C. and M. R. Santos (2022). How Do Transfers and Universal Basic Income Impact the Labor Market and Inequality?

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.

Schönberg, U. (2007). Wage Growth Due to Human Capital Accumulation and Job Search: A Comparison Between the United States and Germany. *ILR Review 60*(4), 562–586.

Sleet, C. and H. Yazıcı (2017). Taxation, Redistribution and Frictional Labor Supply.

Tjaden, V. and F. Wellschmied (2014). Quantifying the Contribution of Search to Wage Inequality. *American Economic Journal: Macroeconomics 6*(1), 134–161.

Veramendi, G. (2012). Labor Market Dynamics: A Model of Search and Human Capital Accumulation.