Procyclical Wage Losses from Skill Mismatch*

Jie Duan†

National University of Singapore

October 26, 2024

Click for the Latest Version

Abstract

This paper investigates the cyclicality of wage losses from skill mismatch and the underlying mechanisms. Using the NLSY97, I find that wage losses due to skill mismatch are procyclical, with a one percentage point increase in the unemployment rate leading to a 1.3% reduction in wage losses, equivalent to a 14.6% recovery. I document that this effect is more pronounced in industries and occupations with higher training incidence, and that training provision is procyclical. To explain these findings, I develop a directed search model with skill mismatch and endogenous training that encompasses a productivity and training channel. The model demonstrates that declines in labor productivity and reduced training during recessions interact to reduce wage losses from skill mismatch.

JEL Classification: E32, J24, J31

Keywords: Skill Mismatch; Wage Losses; Business Cycle; Human Capital

*I thank Tasso Adamopoulos, Titan Alon, Julieta Caunedo, Ryan Chahrour, Ying Feng, Ana Figueiredo, Fatih Guvenen, Paul Jackson, Ronald Wolthoff, Shu Lin Wee and Randall Wright for their helpful comments and suggestions.

†Department of Economics, National University of Singapore. Blk AS1, #01-02, 1 Arts Link, Singapore 117570. *Email*: duanjie@u.nus.edu.

1 Introduction

Skill mismatch, interpreted as the discrepancy between the portfolio of abilities possessed by worker and skills required by the job, is posited as an artifact of labor market misallocation.[1] The literature consistently demonstrates that skill mismatch is associated with wage losses (Guvenen et al., 2020). Meanwhile, both fundamental labor productivity and human capital investment choices, like on-the-job training (Kim and Lee, 2007), fluctuate over business cycles, which potentially alters the output profile of mismatched worker-job pairs. Surprisingly, little research has explored how output losses from skill mismatch interact with these fluctuating determinants, and how such interactions determine the cyclicality of wage losses from skill mismatch.

This paper investigates the cyclical nature of wage losses from skill mismatch and explores the underlying mechanisms. Most prominently, I document that wage losses from skill mismatch are procyclical, with the losses easing during recessions. To explain this, I develop a directed search model with endogenous training, and identify two channels: the productivity channel and the training channel. I then characterize analytically and quantitatively how these channels influence the cyclicality of wage losses.

Using the monthly employment histories of white male college graduates from the National Longitudinal Survey of Youth 1997 (NLSY97), I estimate the cyclicality of wage losses from skill mismatch by introducing the interaction between skill mismatch and unemployment rates into a Mincer-style regression. I document three key findings. First, wage losses from skill mismatch decrease during recessions, with a one percentage point increase in the unemployment rate leading to a 1.3% reduction in wage losses, equivalent to a 14.6% recovery in wage losses from skill mismatch. To corroborate this finding, I conduct a set of robustness tests, including: (i) using alternative business cycle indicators; (ii) re-estimating the effect for subgroups insulated from sorting dynamics; and (iii) measuring skill mismatch with multi-year occupational descriptors. Second, the mitigation of wage losses is more pronounced in industries and occupations with a higher likelihood of providing training to employees. Third, training provision is also procyclical, with employers less likely to offer training during economic downturns. These findings collectively suggest that the lower incidence of training in recessions reduces the wage losses due to skill mismatch.

To explain the procyclicality of mismatch costs, I develop a directed search model with endogenous training. Building on the human capital accumulation framework of Rosen (1972) and Guvenen et al. (2020), where discrepancies between a worker's learning ability and a job's skill requirements result in losses in human capital acquisition through on-the-job learning, my model extends their framework by introducing two key components. First, in addition to on-the-job learning, I introduce another form of human capital acquisition: training for employees, which is jointly determined by workers and firms and complements the on-the-job learning. Second, I incorporate fluctuations in aggregate productivity, enabling an exploration of the cyclicality of mismatch costs within this context.

This model encompasses two channels through which wage losses from skill mismatch are alleviated during economic downturns. First, the decline in aggregate labor productivity during recessions lowers the marginal product of human capital, thus mitigating the marginal output loss caused by inefficiencies in human capital accumulation through on-the-job learning. Second, less training is provided during recessions due to reduced marginal returns on human capital. Since training complements on-the-job learning, this

[1] Worker-job skill mismatch can be further decomposed into positive (when a worker's skills exceed job requirements) and negative (when they fall short). In addition to skill gaps, worker-job mismatch can also be assessed by comparing workers with jobs in terms of field of study, required education levels, or peer characteristics (Montt, 2017; Flisi et al., 2017; Fredriksson et al., 2018). This paper focuses on skill mismatch because it is directly related to human capital development, and can be quantitatively measured in the data.

reduction in training leads to a decline in the marginal accumulation of human capital through on-the-job learning, thereby reducing the marginal losses in human capital accumulation due to skill mismatch.

I next solve the model numerically and demonstrate how each channel mitigates wage losses from skill mismatch. Comparative statics on the productivity parameter and training amount reveal that both channels effectively cushion wage losses as the economy shifts from a good to a bad state. Notably, the scenario with an active training channel generates a greater reduction in wage losses compared to when the training channel is inactive. This suggests that reduced training magnifies the impact of lower labor productivity on wage losses, which underscores the crucial role of training in explaining the cyclicality of wage costs associated with skill mismatch.

This paper contributes to the growing literature on skill mismatch between workers and jobs, and its implications for the labor market. The notion that only the skills utilized in a job are rewarded through wages dates back to Tinbergen (1956), and any misalignment between workers' skills and job requirements can lead to earning losses. For example, Guvenen et al. (2020) show that skill mismatch not only suppresses wage growth in the current job but also leaves a lingering wage scar in future jobs. Similarly, Lise and Postel-Vinay (2020) demonstrate that the misalignment affects the skill adjustment through on-the-job learning, which in turn affects workers' lifetime output. Moreover, Lindenlaub (2017) find that the task-biased technological change re-allocates workers with manual know-how into high-cognitive-demand jobs, which fuels wage inequality along the cognitive skill as worker-job complementarities in cognitive skill increase. Building on this literature, this paper examines the cyclicality of skill mismatch costs. In particular, I investigate whether wage or output losses from skill mismatch are amplified or mitigated during recessions characterized by declines in labor productivity and shifts in on-the-job training.

This paper draws on the literature examining the procyclicality of labor productivity. The procyclicality of aggregate labor productivity has been well documented since Hultgren (1960), with the consensus explanation centered on labor hoarding (Okun, 1963). However, recent studies suggest that the procyclicality of labor productivity has been waning. For instance, Stiroh (2009) highlights the reduced volatility of labor productivity, while Fernald and Wang (2016) report a weak correlation (around 0.2) with output since the mid-1980s. They attribute this shift to factors such as increased labor market flexibility, selective layoffs, a reduced share of manufacturing, and reallocation effects during downturns. Likewise, Galí and Van Rens (2021) attribute the decline in procyclicality to a less frictional labor market, aided by improvements in information regarding match quality. In this paper, the procyclicality of labor productivity is treated as exogenous, as its sources lie outside the scope of this study. However, given its diminishing trend, the procyclicality of labor productivity alone may not sufficiently explain the observed reduction in wage losses from skill mismatch. To address this, I introduce the training channel, which works alongside the productivity channel and potentially amplifies its effect.

This paper also contributes to the literature on training provision and its fluctuations over the business cycle. The procyclicality of training was first documented by Lynch (1992). Using NLSY79 data from 1988-1996, Majumdar (2007) finds that both the incidence and duration of firm-provided training decrease as the unemployment rate rises. Expanding on this, Méndez and Sepúlveda (2012) examine the cyclicality of different types of skill acquisition, revealing that firm-sponsored training is strongly procyclical and significantly shaped by financial constraints. Likewise, Lüthi and Wolter (2020) provide consistent evidence using Swiss cantons panel data (1987-2016).[2] This paper contributes by providing empirical evidence supporting

[2]The training cyclicality is, in fact, debated. Training may appear countercyclical when (i) firms, anticipating future recovery, retain and train idle employees to enhance furture productivity; or (ii) workers have fewer outside options, which reduces their hold-up incentives. Conversely, training tend to be procyclical when, in periods of higher unemployment, firms opt to hire skilled labor

the procyclicality of training. Additionally, it establishes a link between training and on-the-job learning, and explore how their interaction drives the cyclicalility of costs from skill misallocation.

Finally, this paper relates to the growing literature on sorting dynamics over the business cycle. The extent of skill mismatch over business cycles is determined by two opposing forces: it may decrease as matches with under-qualified workers, particularly those at the bottom rung of the job ladder, are destroyed (the ''cleansing effect'') (Mortensen and Pissarides, 1994; Lise and Robin, 2017), or it may worsen due to the creation of new matches with over-qualified workers (the ''sullying effect'') (Barlevy, 2002; Barnichon and Zylberberg, 2019). However, recessions have a less pronounced effect on the employment profile of college graduates compared to their less-educated counterparts. Hoynes et al. (2012) find that a one percentage point increase in unemployment leads to nearly a two percentage point rise in unemployment for workers without a high school degree, whereas the impact on college graduates is less than half a percentage point. Similarly, Hershbein and Kahn (2018) and the broader literature on polarization demonstrate that employment losses during recessions are concentrated in routine-manual occupations, disproportionately affecting low-skilled workers. This paper focuses on college graduates, who are largely insulated from the sorting dynamics, and investigates how output losses from skill mismatch are shaped by fundamental labor productivity changes and training adjustments over business cycles.

The paper proceeds as follows. Section 2 introduces the data and skill mismatch measurement. Section 3 documents the procyclicality of both wage losses from skill mismatch and training. Section 4 presents the model particulars and the key channels. Finally, Section 5 concludes.

2 Data and Measurement

This section introduce the primary datasets, sample selection process, skill mismatch measurement and training variables used in this paper.

2.1 Data

This paper primarily use two datasets. The first is the O*NET dataset, which contains detailed descriptors for nearly 1,000 occupations across the entire U.S. economy. As a key source of occupational information, the O*NET project characterizes the mix of knowledge, skills, and abilities (KSA) commonly required to perform tasks in each occupation. For each occupation, O*NET analysts assign importance scores to 277 descriptors. In this paper, I use the analyst's dataset version 17.0, released in July 2012.[3]

The second dataset is the National Longitudinal Survey of Youth (NLSY97), a widely-used dataset that tracks a nationally representative sample of 8,984 American youth, aged 12 to 17 at the time of their first interview. It provides comprehensive information on respondents' work-related activities, allowing me to track their consecutive employment and training histories since they entered the labor market. Additionally, it includes a complete record of demographic information, encompassing both cognitive and non-cognitive test scores. In particular, all respondents took the Armed Services Vocational Aptitude Battery (ASVAB) at the start of the survey, along with a series of personality assessments, such as Goldberg's

externally rather than invest in in-house training, or when declining profits prompt firms to cut training budgets. Which dynamic dominates depends on institutional structure, workforce composition, and other factors.

[3] This version was selected due to its proximity to the midpoint of the sample period, although the choice of version does not materially affect the main conclusions. On the one hand, the skill requirements of occupations remain relatively stable over the sample period. On the other hand, the key findings are robust even when multiple versions of the O*NET data are used, as detailed in the robustness tests in Section 3.1.3.

Big Five Personality Assessment in round 6 and Ten-item Personality Inventory Scale in round 12, among other personality scales.[4] These variables make it possible to quantify workers' innate learning abilities, thereby enabling the measurement of skill mismatch.

The seasonally-adjusted nationwide monthly unemployment rates are sourced from the U.S. Bureau of Labor Statistics, while the estimated quarterly natural rates of unemployment are retrieved from the Congressional Budget Office.

2.2 Sample Selection

In this paper, I construct a sample of 493 white male graduates with at least a bachelor's degree, with consecutive monthly employment histories spanning from 1997 to 2018. This sample possesses several characteristics that allow the mechanisms of interest to stand out. First, it is largely insulated from labor market discrimination with regard to race and gender. Second, the employment volatility of college workers remains relatively stable over business cycles, ensuring that the estimated wage effects are rarely affected by sorting dynamics. Third, the jobs held by college graduates are more likely to involve training, allowing me to examine how training varies over business cycles in the analysis that follows.

To align with the structure of business cycle indicators, I construct a sample of monthly employment histories, as outlined in Table 1. Starting with original weekly employment histories, I identify each individual's primary jobs for each month as the one for which they worked the most hours, following the approach of Neal (1999) and Pavan (2011). This yields an initial sample of 733 white male college graduates with $219,016$ monthly employment observations. I then refine the sample by excluding individuals who worked more than $1,200$ hours in the initial survey year (1996) to focus on those who began working after the survey started. I retain only those who worked at least $1,200$ hours for two consecutive years and exclude respondents who have ever served in the military. Additionally, I limit observations to those after each individual's last school enrollment, focusing solely on post-graduation employment. I also exclude individuals with weak labor market attachment, defined as being out of the labor force for more than 24 months during their first ten potential working years. After applying these criteria, I arrive at a final sample of 493 individuals with complete employment records and valid cognitive and non-cognitive test scores. Summary statistics for this selective sample are presented in Table A1.

2.3 Skill Mismatch Measurement

This section concentrates on the measurement of skill mismatch, following the approach of Guvenen et al. (2020). Simply put, skill mismatch refers to the discrepancy between the portfolio of skills required by an occupation and abilities possessed by a worker for learning those skills. To quantify the skill mismatch for each worker-occupation pair, I first measure both worker's aptitudes and the job's demand along three skill dimensions: (cognitive) verbal skills, (cognitive) math skills, and (non-cognitive) social skills.

Worker's Abilities Measurement The Armed Services Vocational Aptitude Battery (ASVAB) evaluates respondents' knowledge and skills across 12 areas, but I focus on four subtests most relevant to cognitive

[4]In round 12, respondents provided self-ratings on 10 personality traits. On a scale of 1 to 7, they rated how well the following paired traits applied to them: extraverted/enthusiastic; critical/quarrelsome; dependable/self-disciplined; anxious/easily upset; open/complex; reserved/quiet; sympathetic/warm; disorganized/careless; calm/emotionally stable; conventional/uncreative. These question are identified by the prefix "YTEL-TIPIA."

Table 1: Sample Selection in NLSY97, 1997-2018

Criterion	No. Respondents	No. Observations
White male college graduates	733	219,016
Start career after the survey started	721	215,430
Work more than $1,200$ hours for two consecutive years	690	118,351
Drop respondents who ever served in military	659	112,520
Post-graduation employment	659	88,463
Drop respondents who are weakly attached to labor force	654	80,077
Valid occupation and industry code	601	73,118
Valid ASVAB cognitive test scores	529	65,274
Valid non-cognitive test scores	493	61,564

Note: Top 0.1% and bottom 0.1% in the wage distribution have been encoded into missing.

abilities: word knowledge, paragraph comprehension, arithmetic reasoning, and mathematics knowledge.[5]

To measure cognitive ability, I start with individuals who have complete scores for these four subtests. Since test-taking age may systematically affect ASVAB scores, I control for this by normalizing each test score within each age cohort, following the procedure of Altonji et al. (2012). I then apply the *Principal Component Analysis (PCA)* separately to the verbal (word knowledge and paragraph comprehension) and math (arithmetic reasoning and mathematics knowledge) subtests, extracting the first principal component from each to represent verbal and math abilities. Finally, these ability measures are then converted into percentile ranks across the whole sample, as the scales of principal components are somewhat arbitrary.

To measure social ability, I use the Ten-Item Personality Inventory (TIPI) Scale that assesses ten personality traits. I restrict the analysis to respondents with complete scores for both *extraverted* and *reserved*, and control for test age effects as done previously. The *reserved* score is transformed into an *animated* score by calculating (7 *reserved*). Social ability is then derived by applying PCA to these two standardized scores, with the first principal component representing it.[6] As with cognitive abilities, the social ability scores are converted into percentile ranks due to the arbitrary scales of the principal components.

Ultimately, individual i's learning ability is represented by a three-dimension ability vector $(a_{i,v}, a_{i,m}, a_{i,s})$, where each component corresponds to their verbal, math, and social abilities, respectively.

Skill Requirements Measurement Now turn to the measurement of skill levels required by each occupation. To determine the verbal and math skill requirements, the first step is to construct four ASVAB-comparable test categories for each occupation by weighted summing 26 relevant descriptors, chosen by the Defense Manpower Data Center (DMDC) and listed in Panel A of Table A2. After normalizing the standard deviation of each score to one, these four ASVAB-comparable categories are further reduced to two composite variables, verbal and math, using PCA. Specifically, verbal skill is the first principle component of word knowledge and paragraph comprehension, while math skill is the first principle component of arithmetic reasoning and mathematics knowledge. A percentile transformation is also applied here.

[5]The full ASVAB includes subtests in arithmetic reasoning, electronics information, numerical operations, assembling objects, general science, paragraph comprehension, auto information, mathematics knowledge, shop information, coding speed, mechanical comprehension, and word knowledge. These scores reflect the respondent's performance in each area, with higher scores indicating higher ability.

[6]The non-cognitive skill tests I used are different from Guvenen et al. (2020) who employed the Rotter Locus of Control Scale and Rosenberg Self-Esteem Scale for NLSY79 respondents, as these tests were not administered in the NLSY97 dataset.

The same process is used to identify social skill requirements. By applying PCA on six scaled O*NET descriptors, I construct a single index representing social skill, followed by a percentile transformation. The six descriptors used are listed in Panel B of Table A2.

Finally, occupation o is characterized by a three-dimension skill vector $(r_{o,v}, r_{o,m}, r_{o,s})$, where each component corresponds to the required levels of verbal, math, and social skills, respectively.

Mismatch Measurement To quantify the skill mismatch, I measure the distance between a worker's ability and the skill requirements of their current job.

First, I calculate the mismatch for each worker-occupation pair along each skill dimension. Specifically, the absolute $(m_{i,j,o})$, positive $(m^{+}_{i,j,o})$, and negative mismatch $(m^{-}_{i,j,o})$ in skill dimension j for the match between worker i and occupation o are defined as follows:

$$m_{i,j,o} = \left| a_{i,j} - r_{o,j} \right|, \quad m^{+}_{i,j,o} = \max[a_{i,j} - r_{o,j}, 0], \quad m^{-}_{i,j,o} = \min[a_{i,j} - r_{o,j}, 0],$$

where $a_{i,j}$ is the percentile rank of worker i's learning ability in skill j, and $r_{o,j}$ is the percentile rank of the requirement of occupation o in skill j. Using these measures, I then calculate the aggregate absolute $(m_{i,o})$, positive $(m^{+}_{i,o})$ and negative mismatch $(m^{-}_{i,o})$ for the match between worker i and occupation o across all skill dimensions:

$$m_{i,o} = \sum_j \{ w_{j,o} \left| a_{i,j} - r_{o,j} \right| \}, \quad m^{+}_{i,o} = \sum_j \{ w_{j,o} \max[a_{i,j} - r_{o,j}, 0] \}, \quad m^{-}_{i,o} = \sum_j \{ w_{j,o} \min[a_{i,j} - r_{o,j}, 0] \}.$$

The weights $w_{j,o}$, derived from the factor loadings of the normalized first principal component, reflect the relative importance of each skill dimension j for occupation o.

2.4 Training

The NLSY97 dataset provides information on formal training beyond regular schooling, including the start and end dates, type, duration, and financial support (e.g., government funding, self-payment, financial aid/loans, or employer financing) for each training program. It also records whether the training resulted in a certificate, license, or degree.[7]

To link the training data to monthly employment history, I focus on employer-financed training with valid start and end dates.[8] In the restricted sample, 207 out of 493 (42%) college graduates received at least one training sponsored by the their employers. This aligns with Ma et al. (2024), which reports that 20% to 50% of workers are exposed to training based on EU-CVT data for European developed countries, and with Loewenstein and Spletzer (2000), which finds a training incidence range of 17% to 38% using NLSY data.

3 Empirical Evidence

In this section, I present three key finding. First, wage losses from skill mismatch decline during economic recessions marked by higher unemployment rates. Second, this easing is particularly pronounced in in-

[7] Before 2003, training duration was reported in days per week and hours per day; from 2004 onwards, it switched to hours per week. For consistency, I converted pre-2003 training durations to hours per week.

[8] The reasons I focus on employer-financed training are twofold. First, firm-sponsored training is the most productivity-driven and, therefore, closely aligns with the endogenous training decision in the model. Second, firm-provided training is the primary source of adult education beyond formal schooling, as noted in Ma et al. (2024).

dustries and occupations with higher training incidence. Third, firms are less likely to invest in training during economic recessions. These findings collectively suggest that wage losses from skill mismatch are alleviated, potentially through adjustments in training investment over business cycles.

3.1 Fact I: The Procyclical Wage Losses

This section examines how wage losses from skill mismatch change over business cycles. In the baseline estimation, I use the monthly unemployment rate as an indicator of business cycles, where a higher unemployment rate indicates a slacker labor market and serves as a proxy for economic downturns. To estimate that, I use the following specification:

$$\begin{aligned}\ln(wage_{i,e,o,t}) = b_0 + b_1 mm_{i,o} + b_2 urate_t + b_3 mm_{i,o} \times urate_t + b_4 \bar{A}_i + b_5 \bar{A}_i \times T_{i,o,t} + b_6 \bar{r}_o \\ + b_7 \bar{r}_o \times T_{i,o,t} + b_8 OJ_{i,e,t} + \underbrace{f(T_{i,o,t}) + f(E_{i,e,t}) + f(Exp_{i,t})}_{\text{Tenure Variables}} + \underbrace{\gamma_{ind2} + \gamma_{occ2}}_{\text{Fixed Effects}} + \varepsilon_{i,e,o,t},\end{aligned} \tag{1}$$

where $wage_{i,e,o,t}$ is the wage of worker i with employer e in occupation o at time t, $mm_{i,o}$ is the mismatch between worker i's abilities and the skill requirements of occupation o, while $urate_t$ denotes the monthly unemployment rate at time t. The coefficient b_1 captures the semi-elasticity of wages with respect to skill mismatch, while b_3, which is of particular interest, measures how the effect of mismatch on wages changes with the level of the monthly unemployment rate, essentially capturing the cyclicality of wage effects from skill mismatch.

Next, I explain control variables in Equation (1). First, I control for worker's innate ability ($\bar{A}_i$) and its interaction with occupational tenure ($T_{i,o,t}$), as the mismatch could be correlated with the worker's innate ability, and worker's ability is expected to positively affect wages. Similarly, I control for the occupational skill requirement ($\bar{r}_o$) and its interaction with occupational tenure. Second, $OJ_{i,e,t}$ is a dummy variable indicating whether the current job is a continuation with the same employer. Third, $f(\cdot)$ represents third-degree polynomials in occupational tenure ($T_{i,o,t}$), employer tenure ($E_{i,e,t}$), and potential experience ($Exp_{i,t}$). Fourth, I include 2-digit industry and occupation fixed effects to control for time-invariant factors. Finally, $\varepsilon_{i,e,o,t}$ is the error term, which includes any mismatch not accounted for by the constructed measure $mm_{i,o}$, since $mm_{i,o}$ serves merely as a proxy for skill mismatch between workers and their jobs.

One concern is that occupational tenure might be endogenous. The probability of occupational switches could be correlated with the magnitude of mismatch and, given that, occupational tenure might correlate with any uncaptured mismatch present in the error term. To address this, I adopt the approach proposed by Altonji and Shakotko (1987) and used by Garcia-Louzao et al. (2023) and Kambourov and Manovskii (2009) to construct a valid instrumental variable for occupational tenure. The core idea is to remove the match-specific component by deviating the original tenure variable from its mean across all observations for a given match. Specifically, the instrumental variable for occupational tenure is calculated as $T^{IV}_{i,o,t} = T_{i,o,t} - \bar{T}_{i,o}$, where $\bar{T}_{i,o}$ is the mean occupational tenure between individual i and occupation o. This instrument is valid because: (i) by construction, $Cov(T^{IV}_{i,o,t}, T_{i,o,t}) \neq 0$ (relevance condition); (ii) $Cov(T^{IV}_{i,o,t}, \varepsilon_{i,o}) = 0$ since $T^{IV}_{i,o,t}$ sums to zero over periods when individual i is employed in occupation o (exogeneity condition). For interaction terms involving occupational tenure, I construct their instrumental counterparts by replacing the original tenure variables with their instrumented ones. By the same token, I construct instrumental variables for employer tenure and potential experience.

Table 2 presents the results of the baseline estimation. The first (last) four columns report the estimation with (without) instrumental variables for tenure. I primarily discuss the results with instrumental variables, as the main conclusions remain consistent in the estimation without them. The first two columns show the wage losses from absolute, positive, and negative mismatch. The coefficient b_1 for absolute mismatch is statistically significant at 0.0131, indicating that workers with a mismatch one standard deviation above the mean are predicted to have wages approximately 2.6% lower than those one standard deviation below the mean. This estimated wage loss is about 65% of the value reported in Guvenen et al. (2020), which is not surprising given that the college graduates in NLSY97 are relatively younger, and wage losses tend to be smaller for younger worker. Similarly, the estimated wage losses from positive and negative mismatch are 1.1% and 3.2%, respectively.

In columns (3) and (4), I introduce the interaction between mismatch and the monthly unemployment rate to explore how wage losses respond to changes in the unemployment rate. The estimated positive coefficient, b_3, indicates that wage losses from skill mismatch decrease as unemployment rates rise. Specifically, a one percentage point (pp) increase in the unemployment rate is associated with a 1.3% reduction in (marginal) wage losses from absolute mismatch. In other words, an additional pp of unemployment rate can offset nearly 14.6%$(=\frac{0.013}{0.089})$ of the wage losses generated by one standard deviation of absolute skill mismatch above the mean. Moreover, the reduction in wage losses holds for both over- and under-matches. According to column (4), a one pp increase in the monthly unemployment rate is associated with a 1.8% (1.0%) lower wage losses from positive (negative) mismatch, equating to a recovery of 15.0% (10.7%). Given the relatively symmetric recovery in wage losses for both types of mismatch, I will concentrate on the aggregate absolute mismatch throughout this paper, rather than differentiate between the two.

Table 2: Baseline Regression – Wage Losses and Unemployment Rate

	(1) IV	(2) IV	(3) IV	(4) IV	(5) OLS	(6) OLS	(7) OLS	(8) OLS
Mismatch	-0.01307		-0.08876		-0.00621		-0.08084	
	(0.00262)		(0.00911)		(0.00257)		(0.00877)	
Urate	-0.00460	-0.00690	-0.02580	-0.03064	-0.01563	-0.01656	-0.03690	-0.03932
	(0.00169)	(0.00172)	(0.00297)	(0.00314)	(0.00170)	(0.00173)	(0.00303)	(0.00321)
Mismatch Urate			0.01258				0.01244	
			(0.00144)				(0.00146)	
Positive Mismatch		-0.01077		-0.11959		0.00130		-0.11119
		(0.00333)		(0.01116)		(0.00327)		(0.01197)
Negative Mismatch		0.03169		0.08977		0.02501		0.07310
		(0.00335)		(0.01101)		(0.00319)		(0.00943)
Positive Mismatch Urate				0.01791				0.01853
				(0.00175)				(0.00196)
Negative Mismatch Urate				-0.00963				-0.00801
				(0.00172)				(0.00153)
N	52,569	52,569	52,569	52,569	52,569	52,569	52,569	52,569
R^2					0.297	0.272	0.299	0.274

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses. *($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

The key finding of this paper is that wage losses from skill mismatch decrease when the unemployment rate rises, which is referred to as the procyclicality of wage losses. To my knowledge, this is the first study to document this procyclical feature. To ensure its robustness, I conduct a large set of tests, which fall into three main categories: (i) using alternative business cycle indicators instead of the unemployment rate; (ii) restricting the analysis to subgroups insulated from the cyclicality of match quality; and (iii) employing multiple-year O*NET descriptors to measure occupational requirements and the associated skill mismatch.

3.1.1 Robustness Test I: Alternative Indicators of Business Cycle

In this section, I present estimation results of Equation (1) using alternative business cycles indicators in place of the monthly unemployment rate. The key coefficients from the IV estimations are summarized in Table 3, where X represents the alternative indicators, with their specific titles displayed in the top row. I report the estimation results for absolute mismatch in the main text, as the findings for positive and negative mismatch resemble those for absolute mismatch. To conserve space, the full regression results are provided in Appendix A.2.1.

Table 3: Robustness Test I – Alternative Indicators of Business Cycles

	(1) Urate Gap	(2) HP-Filtered Urate	(3) Urate Recession	(4) NBER Recession	(5) HP-Filtered GDP
Mismatch	-0.02746	-0.01020	-0.02798	-0.01625	-0.01071
	(0.00317)	(0.00263)	(0.00333)	(0.00275)	(0.00263)
X	-0.03096	-0.04300	-0.08531	-0.05740	0.01658
	(0.00380)	(0.00662)	(0.01013)	(0.01549)	(0.00498)
Mismatch X	0.01589	0.02576	0.03808	0.03996	-0.01301
	(0.00186)	(0.00337)	(0.00497)	(0.00793)	(0.00255)
N	52,569	52,569	52,569	52,569	52,569

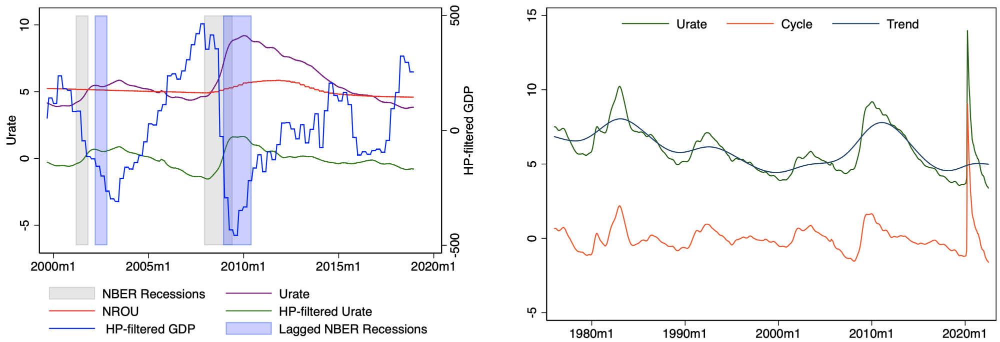
Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses. *($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

Unemployment Gap One alternative measure of labor market states are the difference between the actual unemployment rate and the natural rate of unemployment (NROU).[9] Figure 1(a) shows the actual monthly unemployment over the sample periods (the purple curve), alongside the natural rate of unemployment (the red curve). The unemployment gap, defined as the difference between the actual adjusted unemployment rate and the natural rate, reflects labor market slack, with a larger (positive) gap indicating a slacker economy.

With the unemployment gap, I re-estimate the effect of labor market slack on wage losses from skill mismatch. As shown in column (1), the marginal wage loss from skill mismatch decreases when the economy becomes slacker. Specifically, a one pp increase in the unemployment gap is associated with a 1.6% higher semi-elasticity of wages with respect to skill mismatch.

HP-Filtered Unemployment Rate The rise in the unemployment rate might be driven by long-term trends rather than labor market volatility. To account for this, I apply the Hodrick-Prescott (HP) filter to separate the cyclical component of the monthly unemployment rate from its trend, with a smoothing parameter of 129,600. The cyclical component of unemployment rates thus effectively captures labor market booms and busts. Figure 1(b) illustrates both the trend (the blue curve) and cyclical components (the red curve). The HP-filtered unemployment rate ranges from 1.5 to 1.7 and averages around 0.05.

[9]Natural unemployment is the minimum unemployment rate resulting from real or voluntary economic forces.



(a) Time Series of Recession Indicators

(b) Raw and Filtered Monthly Unemployment Rate

Figure 1: Robustness Check I

As shown in column (2) of Table 3, the finding aligns with the baseline: each additional percentage point of HP-filtered unemployment rate is associated with a 2.6% increase in the wage semi-elasticity with respect to skill mismatch. While the recovery fraction exceeds one, the small magnitude of the HP-filtered unemployment rate suggests that it hardly fully offsets the wage losses from skill mismatch.

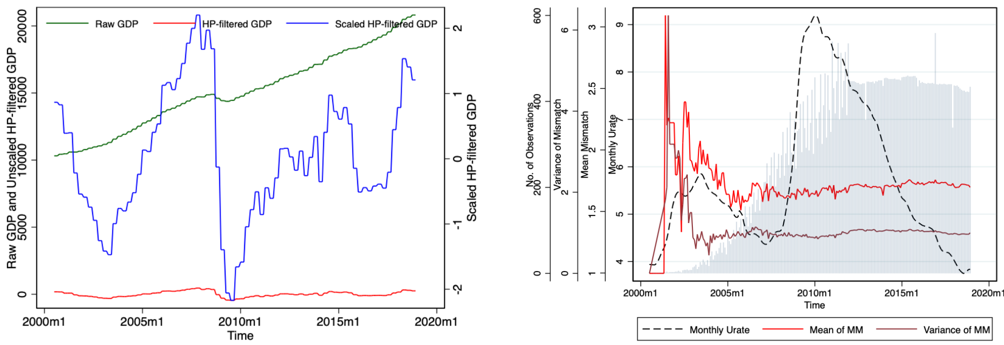
Unemployment-Defined Recession Beyond the continuous measure of unemployment rates, binary indicators of recessions can also capture economic slack. One such measure is to define a recession as any period when the HP-filtered unemployment rate exceeds zero. As shown in column (3) of Table 3, wage losses due to skill mismatch are alleviated, even reversed during recessions.

NBER Recession As shown by the gray shaded areas in Figure 1(a), NBER identifies two recessions during the sample period: March 2001 to November 2001 and December 2007 to June 2009.[10] Comparing these periods with the monthly unemployment rate reveals a lag in the unemployment response, where joblessness peaks after the stated end of the NBER recession. This hysteresis was particularly noticeable in the second recession, where unemployment peaked about a year after December 2007, potentially due to wage rigidity. To account for this delay, I construct a one-year-lagged NBER recession indicator, represented by the blue shaded areas in Figure 1(a).

As shown in column (4) of Table 3, the semi-elasticity of wages with respect to skill mismatch increases by about 4% during recessions, suggesting that wage losses due to mismatch are mitigated and even reversed during severe recessions. This aligns with results based on unemployment-defined recessions.

HP-filtered GDP One standard definition of a recession is consecutive declines in gross domestic product (GDP). Given that, I construct the HP-filtered GDP by applying the HP filter to remove the trend component from the raw quarterly GDP. A lower (higher) HP-filtered GDP reflects a worsening (improving) economic situation. To maintain comparability across different indicators, this measure is standardized and yields values between 2.17 to 2.20. Figure 2(a) demonstrates the raw, unscaled HP-filtered, and standardized (scaled) HP-filtered GDP over the sample period.

[10]The National Bureau of Economic Research (NBER) identifies U.S. business cycles with dates of peaks and troughs, where recessions are determined based on a range of monthly measures of aggregate real economic activity such as real personal income less transfers, non-farm payroll employment, employment as measured by the household survey, real personal consumption expenditures, wholesale-retail sales adjusted for price changes, and industrial production.



(a) Quarterly Gross Domestic Product (GDP)

(b) Inertia of Match Quality

Figure 2: Robustness Check II

The estimation result is presented in column (5) of Table 3. In particular, a one standard deviation decrease in HP-filtered GDP is associated with a 1.3% higher semi-elasticity of wages with respect to skill mismatch.

3.1.2 Robustness Test II: Ruling Out Sorting Dynamics

Table 4: Robustness Test II – Sorting Dynamics and Dynamic O*NET

	(1) Job Stayer	(2) Occupation Stayer	(3) Experienced Worker	(4) Worker-Job FE	(5) Dynamic O*NET
Mismatch	-0.09665	-0.09843	-0.09082	-	-0.07736
	(0.01019)	(0.01006)	(0.01013)	-	(0.00919)
Urate	-0.03122	-0.03205	-0.04175	-0.02126	-0.02260
	(0.00329)	(0.00325)	(0.00337)	(0.00181)	(0.00285)
Mismatch Urate	0.01376	0.01410	0.01206	0.00760	0.01065
	(0.00162)	(0.00160)	(0.00162)	(0.00091)	(0.00145)
N	38,137	39,620	52,569	52,507	51,740

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses. *($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

The effect of business cycles on wages might be driven by the cyclicality of skill mismatch. During recessions, the *sullying effect* can create poor-matched worker-job pairs while the *cleansing effect* can destroy some existing bad matches. This cyclicality of match quality over business cycles is referred to as ''sorting dynamics.'' (Figueiredo, 2020; Baley et al., 2022) However, sorting dynamics are argued to have a limited impact on college graduates or highly skilled workers, whose employment stability tends to be more resilient during economic downturns. Figure 2(b) demonstrates the irresponsiveness of match quality over business cycles. The shaded bar in the background represents the number of employment observations, while the red (purple) curve represents the mean (variance) of mismatch for college graduates. Notably, when the monthly sample size is fairly large, both the mean and variance of match quality show little responsiveness to labor market slackness, as proxied by the monthly unemployment rates. This is further supported by conditional estimates, where an additional percentage point of unemployment rate is associated with only a negligible reduction of 0.02 in skill mismatch, which accounts for just 1% of the mean skill mismatch (1.69) in the sample.[11]

To formally isolate the estimated alleviation of wage losses from sorting dynamics, I focus on subsamples less exposed to frequent job changes and, therefore, less affected by sorting dynamics over business cycles. Building on the findings of Black and Figueiredo (2022) that wage effects of mismatch cyclicality are more pronounced among new hires and job switchers, I limit the estimation in Equation (1) to: (i) job stayers with at least one year of employer tenure; (ii) occupational stayers with at least one year of occupational tenure; and (iii) experienced workers with at least two years of work experience.[12]

For brevity, the key coefficients are presented in Table 4, while the full regression results are delegated into Appendix A.2.2. In particular, an additional percentage point of monthly unemployment rate is associated with a 1.4%/1.4%/1.2% increase in the semi-elasticity of wage with respect to skill mismatch for job stayers, occupation stayers, experienced workers, respectively. In other words, this offsets about 14%/14%/13% of wage losses from skill mismatch.

A stricter approach to rule out sorting dynamics is to impose worker-job fixed effects. As shown in column (4) of Table 4, a one percentage point increase in the unemployment rate is associated with a 0.8% increase in the semi-elasticity of wages with respect to skill mismatch.

3.1.3 Robustness Test III: Dynamic O*NET

Given the nearly 20-year span of the restricted sample (1997-2018), job skill requirements likely evolved significantly, potentially affecting the accuracy of the skill mismatch measure. To account for this, I use the average of O*NET descriptors published across the sample years to measure job requirements.[13]

As shown in column (5) of Table 4, an additional percentage point of unemployment is associated with a 1.1% increase in the semi-elasticity of wage with respect to skill mismatch, offsetting roughly 14% of wage losses from skill mismatch. This suggests that the estimated procyclicality remains robust to the measure of skill mismatch.

3.2 Fact II: Heterogeneity in Procyclicality of Wage Losses

In this section, I explore the heterogeneity in the estimated alleviation of wage losses from skill mismatch across industries and occupations with different tendencies to provide firm-sponsored training to incumbent workers. In summary, I find that wage loss alleviation during slack labor markets is more pronounced in industries and occupations with a higher training incidence.

To quantify the differences in training provision across industries and occupations, I calculate the likelihood of firm-sponsored training being provided within each industry or occupation. Specifically, the training probability is calculated as the ratio of firm-sponsored training instances to the number of employment observations. Among the 13 one-digit industries, the most training-intensive are Educational, Health, and Social Services, as well as Professional, Scientific, Management, Administrative, and Waste

[11] Following Guvenen et al. (2020), the mismatch term has been standardized using $(mm \quad min(mm))/sd(mm)$, so the standardized mm ranges from 0.00 to 5.16.

[12] Experienced workers typically have lower occupational mobility and fewer job switches. Figure A1 shows the average number of job and occupation switches, along with 95% confidence interval. The frequency of switches peaks within the first two working years and declines to below 0.3 thereafter.

[13] Skill data prior to 2003 was excluded due to methodological differences in data collection. Before 2003, occupational analysts supplied the data, whereas from O*NET 5.0 onward, a multi-method approach involving job incumbents, experts, and big data was used. Researchers are encouraged to start with the O*NET 5.0 database for longitudinal studies. Note that some obsolete occupations lack recent skill data, while some newer occupations were not evaluated in earlier surveys.

Management Services. In contrast, the least training-intensive industries are Transportation, Warehousing, and Utilities.[14]

Similarly, among the six broad occupational groups, the most training-involved occupations are Construction, Extraction, Maintenance, and Repair, as well as Management, Professional, and Related Occupations. In contrast, the least training-intensive occupations are Farming, Fishing, and Forestry, and Production, Transportation, and Material Moving. The full lists are available in Appendix Table A13.

To estimate the heterogeneous effect of the procyclical nature of wage losses from skill mismatch, I group industries and occupations into two broad categories based on their calculated training likelihood: seven industries (three occupations) with a high occurrence of firm-sponsored training, labeled as high-training industries (occupations), and six industries (three occupations) with a lower occurrence of firm-sponsored training, labeled as low-training industries (occupations).

First, I re-estimate Equation (1) within each disaggregated subgroup, and results are presented in Table 5. In brief, wage losses from skill mismatches become smaller in industries and occupations with higher occurrence of firm-sponsored training. In particular, an additional percentage point in the unemployment rate can recover nearly 15.0% (3.9%) of wage losses from one standard deviation higher of skill mismatch in high- (low-) training industries. Similarly, an additional percentage point in the unemployment rate can offset approximately 16.7% (6.2%) of wage losses from skill mismatch in high- (low-) training occupations. Notably, the estimated coefficients are statistically significant only for high-training industries and occupations, while not for the low-training groups.

Table 5: Heterogenous Estimation I – Subgroup Estimation

	By Industries				By Occupations			
	High-Training		Low-Training		High-Training		Low-Training	
	IV	OLS	IV	OLS	IV	OLS	IV	OLS
MM	-0.10293	-0.09275	-0.05000	-0.03595	-0.08792	-0.08438	-0.11196	-0.04067
	(0.01179)	(0.01066)	(0.01461)	(0.01613)	(0.00965)	(0.00942)	(0.02996)	(0.02862)
Urate	-0.04301	-0.05362	0.02785	0.02062	-0.03439	-0.04595	0.02648	0.02668
	(0.00366)	(0.00381)	(0.00516)	(0.00495)	(0.00313)	(0.00326)	(0.00965)	(0.00877)
MM Urate	0.01548	0.01489	0.00196	0.00153	0.01469	0.01483	0.00699	0.00032
	(0.00186)	(0.00176)	(0.00227)	(0.00259)	(0.00152)	(0.00158)	(0.00448)	(0.00421)
N	36,907	36,907	15,662	15,662	47,542	47,542	5,027	5,027
R^2		0.252		0.423		0.264		0.481

Note: MM is an abbreviation for mismatch. All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses. *($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

Alternatively, I introduce the interaction between training occurrence measure and the cyclical term into the baseline specification to examine whether there is a stronger recovery of wage losses when the training

[14]This ranking is consistent with the Association for Talent Development (ATD) 2014 State of the Industry Report, which indicates that healthcare and pharmaceutical organizations led in per-employee training expenditures ($1,392), followed by finance, insurance, and real estate ($1,107), consolidated industries ($1,028), and manufacturing ($535). See https://associationsnow.com/2014/11/commitment-growth-companies-continue-spend-employee-training/.

ratio in the industry or occupation worked is higher. In particular, the modified specification becomes:

$$\ln(wage_{i,e,o,t}) = b_0 + b_1 mm_{i,o} + b_2 urate_t + b_3 mm_{i,o} \quad urate_t + g\ mm_{i,o} \quad urate_t \quad TRN + b_4 \bar{A}_i + b_5 \bar{A}_i \quad T_{i,o,t}$$
$$+ b_6 \bar{r}_o + b_7 \bar{r}_o \quad T_{i,o,t} + b_8 OJ_{i,e,t} + f(T_{i,o,t}) + f(E_{i,e,t}) + f(Exp_{i,t}) + Q_{ind2} + Q_{occ2} + e_{i,e,o,t}, \tag{2}$$

where TRN represents the training ratio in the industry i or occupation o where worker i is employed at time t, and g captures the response of procyclical term ($mm_{i,o} \quad urate_t$) to the training ratio.

Table 6 reports the regression results. It is obvious that the recovery of wage losses from skill mismatch strengthens in occupations or industries with a higher likelihood of providing firm-sponsored training. In particular, working in an industry (or occupation) with a 1% higher training occurrence can increase the marginal wage recovery from skill mismatch by 0.09% (0.70%).

Table 6: Heterogenous Estimation II - Interaction-Term Estimation

	Industries		**Occupations**	
	IV	OLS	IV	OLS
Mismatch	-0.07650	-0.06299	-0.16143	-0.08737
	(0.02526)	(0.02803)	(0.07519)	(0.07470)
Mismatch Urate	0.00681	0.00576	-0.02049	-0.02584
	(0.00393)	(0.00460)	(0.01217)	(0.01284)
Mismatch Urate TRN	0.08656	0.10586	0.69860	0.81153
	(0.08412)	(0.09606)	(0.25939)	(0.27660)
Observations	52,569	52,569	52,569	52,569
R^2		0.302		0.302

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses. *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

3.3 Fact III: The Procyclical Training Provision

In this section, I explore the cyclicality of training provision. Firm-sponsored training, or more broadly, on-the-job training, is a crucial component of human capital acquisition, particularly among skilled workers with at least a bachelor's degree. For example, Majumdar (2007) finds that human capital investment through training is considerably more common among well-educated workers: 22% of college graduates in the NLSY79 (1988-1996) received company-sponsored training, in stark contrast to less than 5% of high-school dropouts. Similar findings has been documented in Méndez and Sepúlveda (2012). Moreover, large firms , which typically employ more skilled workers, tend to provide both more frequent and longer training. These facts highlight the pivotal role that firm-sponsored training plays in shaping human capital and productivity for college graduates.

To examine the cyclical nature of training, I restrict the sample to college graduates who have participated in at least one employer-sponsored training program from their labor market entry until 2018. I apply the same specification as Equation (1), but with the dependent variable now being an indicator for receiving training $TRN_{i,e,o,t}$:

$$TRN_{i,e,o,t} = b_0 + b_1 mm_{i,o} + b_2 urate_t + b_3 mm_{i,o} \times urate_t + b_4 \bar{A}_i + b_5 \bar{A}_i \times T_{i,o,t} + b_6 \bar{r}_o$$
$$+ b_7 \bar{r}_o \times T_{i,o,t} + b_8 OJ_{i,e,t} + \underbrace{f(T_{i,o,t}) + f(E_{i,e,t}) + f(Exp_{i,t})}_{\text{Tenure Variables}} + \underbrace{Q_{ind2} + Q_{occ2}}_{\text{Fixed Effects}} + e_{i,e,o,t}, \quad (3)$$

where $TRN_{i,e,o,t}$ is a binary variable that takes the value of 1 if the worker i participates in an employer-financed training program during month t. The coefficient b_2, of particular interest, captures the effect of labor market slack on the likelihood of receiving training. By introducing the interaction between mismatch and the unemployment rate, that effect is allowed to differentiate across different levels of mismatch.

Table 7 presents the findings, with the left (right) four columns showing estimation with (without) instruments for tenure variables. As shown in the odd columns, the coefficient b_2 is consistently statistically significant and negative, indicating that training provision is procyclical. In particular, a one percentage point increase in the unemployment rate is associated with 0.38% – 0.76% lower likelihood of receiving firm-sponsored training, holding other factors constant. The result remain robust even when when accounting for differential effects across match quality, as indicated in the even columns.

Beyond the baseline regression, I conduct robustness checks for the training cyclicality, including (i) considering alternative indicators of the business cycle; and (ii) re-estimate in subgroups that is largely isolated from sorting dynamics over business cycles. Further details can be found in Appendix Section A.4.

Table 7: Baseline Regression - Training over Business Cycles

	IV				OLS			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Urate	-0.00377	0.00003	-0.00451	-0.00036	-0.00764	-0.00443	-0.00451	-0.00484
	(0.00140)	(0.00240)	(0.00125)	(0.00229)	(0.00134)	(0.00235)	(0.00125)	(0.00244)
MM	-0.01319	0.00032			-0.01065	0.00078		
	(0.00209)	(0.00708)			(0.00201)	(0.00698)		
Positive MM			-0.01612	-0.00318			-0.01612	-0.00750
			(0.00248)	(0.00825)			(0.00248)	(0.00779)
Negative MM			0.00911	-0.00680			0.00911	-0.00522
			(0.00256)	(0.00794)			(0.00256)	(0.00893)
Urate × MM		-0.00222				-0.00187		
		(0.00110)				(0.00108)		
Urate × Positive MM				-0.00212				-0.00101
				(0.00128)				(0.00120)
Urate × Negative MM				0.00263				0.00223
				(0.00124)				(0.00137)
N	26,568	26,568	26,568	26,568	26,568	26,568	26,568	26,568
R^2					0.062	0.062	0.014	0.061

Note: MM is an abbreviation for mismatch. All regressions include a constant, 2-digit industry FE, 2-digit occupation FE, as well as its interaction with the business cycle indicator. This sample includes 23 occupational and 20 industrial groups. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses.*($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

3.4 From Empirics to Model

This paper documents the novel fact that wage losses from skill mismatch are procyclical, with this nature being more pronounced in industries and occupations with higher occurrence of training. Additionally, the provision of firm-sponsored training is also procyclical. Together, these findings suggest that wage losses from skill mismatch are alleviated by the decline in training during recessions.

Existing research establishes that discrepancies between occupational skill requirements and workers' learning abilities lead to inefficiencies in human capital accumulation, which, in turn, result in wage losses. This paper extends to understand how these inefficiencies fluctuate over business cycles characterized by different levels of aggregate labor productivity, and how they interact with training adjustments over business cycles. To explain this, I develop a directed search model with skill mismatch and endogenous training, and incorporate two key components. First, alongside on-the-job learning contingent on match quality, training serves as an additional source of human capital acquisition. Second, the model incorporates exogenous labor productivity, with its level corresponding to distinct economy states.

4 Model

This section develops a directed search model with skill mismatch and endogenous training. Section 4.1 introduces the environment. Section 4.2 characterizes the equilibrium and Section 4.3 details the mechanisms that drive the procyclicality of wage losses.

4.1 Environment

Time, Agents, and Preferences Time is discrete and continuous forever. At period $t = 0$, there is a unit measure of workers and a large measure of homogeneous firms. All agents are risk neutral and infinitely lived, and discount the future with factor $\beta \in (0, 1)$.

Each firm corresponds to one job that either filled or vacant. There always exists a large measure of idle firms, and the measure of active firms is endogenous.

Workers are endowed with an indivisible unit of labor, and are ex-ante heterogeneous in their learning ability a, which is determined by nature and drawn from a log-normal distribution where $\log(a) \sim \mathcal{N}(\mu_a, \sigma_a^2)$. The innate learning ability is fixed over time. Upon entering the labor market, workers possess an initial stock of human capital, h_0, drawn from a log-normal distribution where $\log(h_0) \sim \mathcal{N}(a, \sigma_h^2)$, with their innate ability as the distribution mean. Workers are either unemployed or employed in the labor market.

Production Technology Firms operate a technology that maps one unit of labor from a worker with human capital h into y units of output,

$$y = zh,$$

where z is aggregate labor productivity and follows a Markov process with $z \in Z = \{z_1, z_2, ..., z_N\}$, where N is a positive integer. At the beginning of each period, z is drawn from the probability distribution $\Gamma(z|z_{-1})$. All firms and workers can observe z once it is realized.

Human Capital Accumulation Technology The human capital accumulation of a worker with learning ability a, who receives training t while employed by a firm with skill intensity r, denoted by $H(a, t)$, is:

$$H(a,t) = \left[a t^{e} + (1 - a)\left(m \; exp\left(-\frac{(a - r)^2}{2n^2}\right)\right)^{e} \right]^{\frac{1}{e}}, \tag{4}$$

where a is the share parameter and $\frac{1}{1-e}$ is the elasticity of substitution between training and on-the-job learning. The accumulated human capital consists of two components. The first part comes from training, t, which is jointly decided by firms and workers to maximize the match value. The second part is on-the-job learning, which can be understand as a by-product of working in a job, with the amount of skill acquisition depending on the worker's learning ability, a, and the firm's skill intensity, r. The Gaussian form of on-the-job learning is characterized by an amplitude parameter m and a standard deviation parameter n.[15] In particular, this implies that any mismatch between the worker's learning ability and the firm's skill intensity results in losses in human capital accumulation. From equation (4), training is complementary to the on-the-job learning in developing human capital when $e < 1$.

Information I assume the perfect information, whereby everything is observable upon realization.

The Labor Market The labor market is organized in a continuum of submarkets indexed by $w = (a, h, c)$. In submarket w, firms search for workers with learning ability a and human capital h, and offer a contract that delivers fraction c of the match surplus to the workers.

When a firm and a worker meet in submarket $w = (a, h, c)$, a worker whose innate ability $a' \neq a$ or human capital $h' \neq h$ is automatically rejected. Meanwhile, a worker who exactly satisfies the submarket pre-requisite requirement (a, h) and chooses to accept the offer begins production within the same period.

Timing Each period is divided into six stages: entry, separation, search and matching, human capital investment, production, and human capital depreciation. At the beginning of period t, the aggregate labor productivity, z, is realized and observed by both workers and firms.

Stage 1: Entry Idle firm decides whether to create a vacancy and, if so, which submarket to post it in. If a firm decides to post a vacancy, it incurs a vacancy posting cost of k.

Stage 2: Separation At the separation stage, an existing match ends with probability $d \in \{d, 1\}$. In particular, a matched pair is agreeably dissolved when the match surplus turns negative and d represents separations occurring for exogenous reasons. A worker who loses their job in the separation stage must wait one period before they can search for another job.

Stage 3: Search and Matching Workers who begin the period unemployed search with probability one. There is no on-the-job search.

Workers and firms who search in the same submarket are brought together by a constant returns to scale matching technology. Let $u(w)$ denote the measure of vacancies and $u(w)$ the measure of unemployed workers in submarket w. The number of matches in submarket w is given by the matching function

[15] The amplitude parameter, m, governs the maximum amount of learning. The standard deviation parameter, n, governs the magnitude of marginal losses from skill mismatch.

$F(u(w), u(w))$. Define $q(w) = u(w)/u(w)$ as the tightness in submarket w, and job-finding and job-filling probabilities depend on the labor market tightness. The probability of finding a job $p(q(w)) = F/u(w)$, where $p : \mathbb{R}_+ \to [0,1]$ is twice-differentiable, strictly increasing and strictly concave in q with boundary conditions $p(0) = 0$ and $p(\infty) = 1$. A firm fills a job with probability $q(q(w)) = \frac{p(q)}{q}$, where $q : \mathbb{R}_+ \to [0,1]$ is twice continuously differentiable, strictly decreasing, strictly convex in q, $q(0) = 1$ and $q(\infty) = 0$.

Stage 4: Human Capital Investment The employed worker, together with his incumbent firm, jointly decides on the optimal amount of training $t(w,z)$ by maximizing the match value, taking into account of the training cost $c(t)$, which is assumed to be everywhere strictly increasing, differentiable, convex and with the boundary condition $c'(0) = 0$.

For a worker with innate learning ability a, human capital h_t at the start of period t, and matched with a firm with skill intensity r, the upgraded human capital is given by $h_t + H(a, t^*)$ where $t^* = argmax\ M(a,h,z)$, which can be utilized for production instantaneously.

Stage 5: Production In the production stage, unemployed workers produce b units of output. Matches between a worker with upgraded human capital $(h + H)$ and a firm produce $y = z(h + H)$ units of output.

Stage 6: Depreciation Human capital is assumed to be perfectly transferable across firms, and evolves over time. In particular, the law of motion of human capital is given by

$$h_{t+1} = (1 - r)(h_t + H), \tag{5}$$

where r is the depreciation rate of human capital. If unemployed, the human capital at the start of next period is equal to h_t adjusted for depreciation, $h_{t+1} = (1 - r)h_t$.

4.2 Equilibrium

Unemployed Worker Consider an unemployed worker with learning ability a and human capital h at the start of human capital investment stage. In the current period, he produces output b, and search in the next period's search and matching stage. If he searches in submarket $\hat{w} = (a, (1 - r)h, c)$, he finds a job with probability $p(q(\hat{w}, z'))$ and the continuation value is the employment value, $\mathbb{E}_{z'|z} V(\hat{w}, z')$, with the expectation taken with regard to the uncertainty about the aggregate labor productivity tomorrow. If he doesn't find a job, the continuation value is $\mathbb{E}_{z'|z} U(a, (1 - r)h, z')$. The worker's value of being unemployed, $U(a,h,z)$, satisfies:

$$U(a,h,z) = b + b \Big\{ \mathbb{E}_{z'|z} U(a,(1-r)h,z') + p(\hat{w},z') \Big[\mathbb{E}_{z'|z} W(\hat{w},z') - \mathbb{E}_{z'|z} U(a,(1-r)h,z') \Big] \Big\}. \tag{6}$$

Employed Worker Now consider a match between a firm and a worker with learning ability a and human capital h under the realized aggregate labor productivity z at the human capital investment stage. In the current period, the worker earns $w(w,z)$ in the submarket $w = (a,h,c)$. In the following separation stage, the match is destroyed with probability d, and the continuation value is $\mathbb{E}_{z'|z} U(a,h',z')$ with human capital evolves by (5). With probability $1 - d$, the match is not destroyed, and the continuation value is

$\mathbb{E}_{z'|z}W(w',z')$ where $w' = (a, h', \omega)$. Therefore, the employment value for workers, $V(w,z)$, satisfies:

$$W(w,z) = \max_{d \in \{\delta,1\}} w(w,z) + \beta \Big[(1 - d(w',z'))\mathbb{E}_{z'|z}W(w',z') + d(w',z')\mathbb{E}_{z'|z}U(a,h',z') \Big]. \tag{7}$$

Active Firm Consider the same match as the above. In the current period, firm gets output net of training cost and wages. If the match is destroyed in the following separation stage, the firm's continuation value is zero due to free entry. However, if the match remains intact, the continuation value is $\mathbb{E}_{z'|z}J(w',z')$. Thus, the value of a filled job for firms, $J(w,z)$, satisfies:

$$J(w,z) = z \cdot (h + H(a,\tau)) - c(\tau) - w(w,z) + \beta(1 - d(w',z'))\mathbb{E}_{z'|z}J(w',z'). \tag{8}$$

In any submarket visited by a positive number of workers, tightness is consistent with the firm's incentives to create vacancies if and only if

$$\kappa \geq q(\theta(w,z))J(w,z), \tag{9}$$

and $\theta(w,z) \geq 0$ with complementary slackness.

The Match Value The joint match value is the sum of the worker's value and the firm's value, as given by

$$M(a,h,z) = \max_{\tau,d} z \cdot (h + H(a,\tau)) - c(\tau) + \beta \Big[(1 - d(w',z'))\mathbb{E}_{z'|z}M(a,h',z') + d(w',z')\mathbb{E}_{z'|z}U(a,h',z') \Big], \tag{10}$$

where the optimal amount of training, τ^*, is jointly determine by firms and workers to maximize the match value, that satisfies:

$$\tau^* = argmax \; M(a,h,z). \tag{11}$$

Wages By definition, the worker gets a fraction ω of the total match surplus:

$$W(a,h,\omega,z) = U(a,h,z) + \omega[M(a,h,z) - U(a,h,z)], \tag{12}$$

where ω is determined by firm's maximization problem and solved by the market utility approach:

$$W = \max_{\omega,\theta} p(\theta)\,[U + \omega(M - U)] + (1 - p(\theta)) \cdot U,$$
$$s.t. \qquad \frac{p(\theta)}{\theta}(1 - \omega)(M - U) = J.$$

Given θ, the constraint pins down the share uniquely: $\omega = 1 - \theta p'(\theta)/p(\theta)$, where $\theta p'(\theta)/p(\theta)$ is the elasticity of matching function with respect to market tightness.[16] The surplus-splitting equation implies that the firm and the worker both want a positive match surplus, otherwise there would be a mutual agreement on endogenous separation:

$$d(a,h,z) = \begin{cases} 1 & \text{if } M(a,h,z) - U(a,h,z) < 0, \\ \delta & \text{if } M(a,h,z) - U(a,h,z) \geq 0. \end{cases} \tag{13}$$

[16]The derivation can be found in Appendix Section B.2.

Finally, it is straightforward to show that the wage equation is given by:[17]

$$
\begin{aligned}
w(a,h,\chi,z) = \chi\Big(z\,(h+H) - c(\tau) + \beta\kappa\theta(a,(1-\rho)h,z')\Big) \\
+ (1-\chi)\Big[b - \beta\Big(\mathbb{E}_{z'|z}U(a,h',z') - \mathbb{E}_{z'|z}U(a,(1-\rho)h,z')\Big)\Big],
\end{aligned}
\tag{14}
$$

where $\chi = 1 - \theta p'(\theta)/p(\theta)$. The wage equation imply that the worker and the firm share the training cost in accordance with their surplus splitting shares.

Definition A recursive equilibrium consists of a tightness function $\theta(\omega)$, value function for unemployed workers $U(a,h,z)$, joint value and policy functions, $M(\omega,z)$ and $d(\omega,z)$, endogenous training decision $\tau^*(a,h,z)$, and a distribution of workers that satisfies the following conditions. First, $\theta(\omega)$ satisfies (9) for all ω. Second, the value function of unemployed workers satisfies equation (6). Third, the joint value and associated policy functions for a match satisfy equations (10) and (13). Fourth, the optimal training decision satisfies (11). Finally, the distribution of workers satisfies the laws of motion specified in Appendix B.1.

As established by Menzio and Shi (2011) for directed search models with free entry and bilateral efficiency, a recursive equilibrium exists and is block-recursive (BRE). As workers self-select into submarkets based on their observable characteristics, firms know they will only meet one type of worker in each submarket. Hence, tightness in submarket, value and policy functions, depend on the aggregate state of the economy only through the realization of the aggregate labor productivity, while independent of the distribution of workers across learning ability and human capital stock.

4.3 Procyclical Wage Losses from Skill Mismatch

This section details two distinct sources of the procyclicality of wage costs due to skill mismatch. The first source is changes in fundamental aggregate labor productivity, z, over business cycles. The second is adjustments of training, τ^*, over business cycles.

Proposition 1. *If $H(a,\tau)$ is increasing and concave in τ, then the training provided is positively associated with aggregate labor productivity z.*

This proposition states that less training is provided when the economy state is bad, or equivalently, when aggregate labor productivity is low. To understand this, consider the optimal training decision from equation (11):

$$
\frac{\partial c(\tau)}{\partial \tau} = z\frac{\partial H(a,\tau)}{\partial \tau} + \beta\left[-\frac{\partial d(\theta(\omega',z'))}{\partial \tau}\Big(\mathbb{E}M(a,h',z') - \mathbb{E}U(a,h',z')\Big) + \frac{\partial \mathbb{E}M(a,h',z')}{\partial \tau} - d(\theta(\omega',z'))\left(\frac{\partial \mathbb{E}M(a,h',z')}{\partial \tau} - \frac{\partial \mathbb{E}U(a,h',z')}{\partial \tau}\right)\right].
\tag{15}
$$

This equation relates the marginal cost to the marginal benefit of providing one additional unit of training. Intuitively, when the marginal output of human capital (z) declines, the incentive to invest in training decreases because the marginal benefit becomes smaller. For the concavity condition to hold, the elasticity should necessarily be less than one, i.e., $\varepsilon < 1$; in other words, training and on-the-job learning are com-

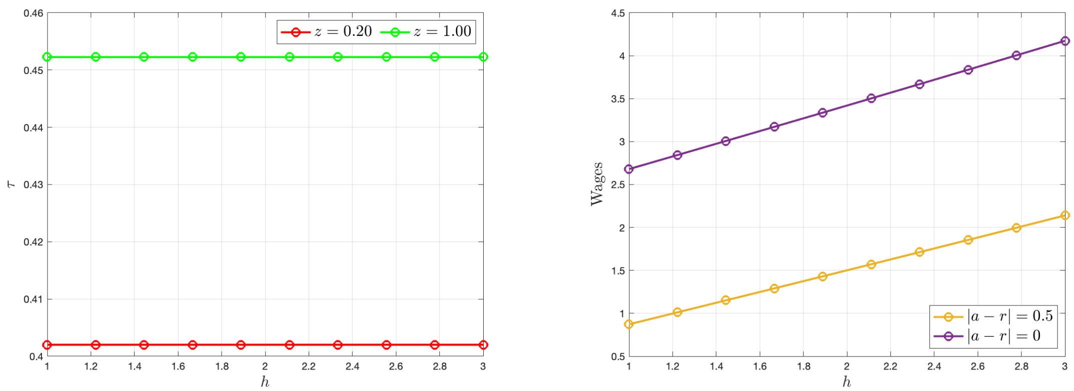
[17] Details can be found in Appendix Section B.3. Wages can also be expressed as:

$$
w = W(\omega,z) - \beta\Big[(1-d(\omega',z'))\mathbb{E}_{z'|z}W(\omega',z') + d(\omega',z')\mathbb{E}_{z'|z}U(a,h',z')\Big].
$$

In the model solution, these two forms of wage expressions result in the same wage values.

plementary to each other in developing human capital.[18] Figure 3(a) graphically illustrates the amount of training over human capital stock when economy is either in a good state ($z = 1.00$) or a bad state ($z = 0.20$). In particular, the amount of training provided is 0.452 (0.402) when the economy is good (bad).

To demonstrate the mechanisms behind mitigating wage losses, it is essential to first understand the sources of wage losses from skill mismatch. Turning to equation (4), any divergence between a worker's learning ability and the firm's skill intensity is associated with some losses in human capital accumulation through on-the-job learning. Since the output technology takes the form of $y = zh$, these losses in human capital accumulation will translate into losses in wages, governed by the factor z. This framework has been well established in Guvenen et al. (2020). Figure 3(b) visualizes the wage gap between a perfectly matched worker-job pair with no mismatch, represented by the purple curve, and a mismatched pair, represented by the yellow curve, when the economy is in the good state. Clearly, the wage for the perfectly matched pair is higher than that of the mismatched pair, regardless of their current human capital stock.



(a) Training and z

(b) Wage Losses

Figure 3: Numerical Example

Notes: The numerical example assumes $r = 0.10$ for the depreciation rate, $a = 0.5$, $e = 10$, $m = 1$ and $n = 0.25$ for the human capital technology parameters, $z = 0.05$ and $h = 0.5$ for the matching function parameters, $k = 0.5$ for the vacancy posting cost, and $b = 0$ for home production. Training cost is assumed to be $c(t) = gt^2$ with cost parameter $g = 0.05$. The firm's required skill level is fixed at 0.5. For simplicity, the economy states are either good ($z = 1.00$) or bad ($z = 0.20$) with transition matrix $\mathbf{P} = [0.9, 0.1; 0.1, 0.9]$. All parameters remain fixed at the stated values unless specified otherwise.

Next I proceed to demonstrate the mechanisms that work to alleviate wage losses from skill mismatch when the economy transitions from a good state to a bad state. The first mechanism is the direct effect of declining aggregate labor productivity (z), referred to as the productivity channel. As labor productivity declines from a high to a low level, the marginal output of human capital becomes smaller, and so does the marginal output loss from skill accumulation inefficiencies. To illustrate this, I simplify the model by deactivating the training channel by setting $a = 0$, leaving only the on-the-job learning component in the human capital accumulation technology. Accordingly, the wage equation now becomes

$$w(a,h,c,z) = c \quad z \quad (h + \tilde{H}) + bkq(a,(1 \quad r)h,z') \quad + (1 \quad c)b, \tag{16}$$

[18] The complementarities between training and on-the-job learning can be rationalized as follows. First, the content acquired from each differs. Training provides general or foundational knowledge, while on-the-job learning, as a byproduct of working, develops specialized and practical skills. Thus, the foundational knowledge from training complements the hands-on skills gained through on-the-job learning. Second, training offers conceptual or theoretical guidance, which becomes productive only when combined with practical on-the-job learning. Last, from a behavioral perspective, firm-sponsored training signals the company's commitment to the worker's development, which would enhance employee's engagement, and in turn, boosts the efficiency of on-the-job learning.

where $\tilde{H}(a) = m \exp(-\frac{(a-r)^2}{2n^2})$. From equation (16), the marginal output of human capital is governed by the aggregate labor productivity factor z. It is evident that a lower z, or equivalently, a lower marginal product of human capital, results in reduced output costs for the lost accumulated human capital in $\tilde{H}$. Figure 4(a) shows this mechanism in a numerical example by plotting wage losses as a function of human capital stock under different economy states. With only the productivity channel active, when the economy transitions from a good state (the green curve) to a bad state (the red curve), wage losses due to skill mismatch decrease from [-2.2,-2] to [-0.95,-0.9].

(a) Productivity Channel

(b) Training Channel

Figure 4: Mechanisms Illustration

Notes: The numerical example on the left-hand side represents the model without training, while the right-hand side illustrates the model with xed $z = 1.00$ but different amounts of training.

The second mechanism operates through the changes in the amount of training during business cycles. Due to the complementarities between training and on-the-job learning, the marginal human capital accumulation through on-the-job learning depends on the level of training. Consequently, the marginal losses of accumulated human capital embedded in the on-the-job learning are also affected by the training level. In particular, the marginal losses in human capital accumulation from skill mismatch are expressed as:

$$\frac{\partial H(\cdot)}{\partial (a-r)^2} = \frac{(1-a)\, m \exp\left(-\frac{(a-r)^2}{2n^2}\right)^{e}}{2n^2} \left(a\, t^{e} + (1-a)\, m \exp\left(-\frac{(a-r)^2}{2n^2}\right)^{e} \right)^{\frac{1-e}{e}},$$

where higher levels of training (t ") are associated with greater inef ciency in human capital accumulation from skill mismatch when $e < 1$.[19] As training declines during transitions from good to bad economic states (from Proposition 1), the marginal losses in human capital accumulation induced by skill mismatch decrease, thereby reducing output losses in the end. Figure 4(b) presents the comparative statics of wage losses with respect to training. It clearly demonstrates that, with aggregate labor productivity held constant, a reduction in training shifts the wage loss curve upward from the yellow curve with $t = 3.00$, to the green curve with $t = 2.00$, and then to the blue curve with $t = 1.00$, indicating a mitigation of wage losses.

Up to this point, I have demonstrated two distinct mechanisms: the productivity channel and the training channel, and how each functions to alleviate wage losses from skill mismatch. A further analysis is to examine the combined effect of both channels. Figure 5(a) compares wage losses between good and bad

[19] A numerical illustration can be found in Appendix Figure B2.

economic states in the full model, where both channels are activated. As the economy transitions from a good to a bad state, wage losses from skill mismatch shrink from [-2.05, -1.80] to [-0.70, -0.62].

(a) Full Model

(b) Fraction of Wage Losses Recovery

A natural question that follows is: How does the pro le of wage losses differ when both channels are active versus when the training channel is turned off? Figure 5(b) compares the performance of the full model with that of the model with only the productivity channel in mitigating wage losses. It shows the full model reduces wage losses by about 65%, while the model without the training channel yields at most a 57.5% reduction when transitioning from good to bad economic states. This underscores the critical role of reduced training during recessions in driving the overall mitigation of wage losses from skill mismatch.

5 Conclusion

This paper documents the procyclicality of wage losses from skill mismatch and introduces the training channel as a novel mechanism to explain that. Using the NLSY97, I nd that the wage losses due to skill mismatch decrease by 14.6% with a one percentage point increase in the unemployment rate, with this effect being more pronounced in industries and occupations where training is more prevalent. Beside that, training provision also follows a procyclical pattern, with employers less likely to offer training during recessions.

Next, I develop a directed search model augmented with endogenous training to formalize the channels through which wage losses are mitigated. The model identi es two distinct channels: the productivity channel and the training channel. In particular, a decline in aggregate labor productivity, combined with a reduction in training during recessions, collectively diminishes the output loss induced by inef ciencies in human capital accumulation.

This paper contributes to the literature by innovatively identifying the cyclical nature of wage losses due to skill mismatch and proposing a new channel to explain this feature. The ndings underscores the importance of human capital investment over the business cycle in shaping the misallocation costs. However, fully quantifying its impact and determining the optimal human capital investment across business cycles open avenues for future research.

References

ALTONJI, J. G., P. BHARADWAJ, AND F. LANGE (2012): "Changes in the characteristics of American youth: Implications for adult outcomes," Journal of Labor Economics, 30, 783–828.

ALTONJI, J. G. AND R. A. SHAKOTKO (1987): "Do wages rise with job seniority?" The Review of Economic Studies, 54, 437–459.

BALEY, I., A. FIGUEIREDO, AND R. ULBRICHT (2022): "Mismatch cycles," Journal of Political Economy, 130, 2943–2984.

BARLEVY, G. (2002): "The sullying effect of recessions," The Review of Economic Studies, 69, 65–96.

BARNICHON, R. AND Y. ZYLBERBERG (2019): "Underemployment and the trickle-down of unemployment," American Economic Journal: Macroeconomics, 11, 40–78.

BLACK, I. AND A. FIGUEIREDO (2022): "The false illusion of wage cyclicality," Working paper.

FERNALD, J. G. AND J. C. WANG (2016): "Why has the cyclicality of productivity changed? What does it mean?" Annual Review of Economics, 8, 465–496.

FIGUEIREDO, A. (2020): "Wage cyclicality and labor market sorting," American Economic Review: Insights.

FLISI, S., V. GOGLIO, E. C. MERONI, M. RODRIGUES, AND E. VERA-TOSCANO (2017): "Measuring occupational mismatch: Overeducation and overskill in Europe — Evidence from PIAAC," Social Indicators Research, 131, 1211–1249.

FREDRIKSSON, P., L. HENSVIK, AND O. N. SKANS (2018): "Mismatch of talent: Evidence on match quality, entry wages, and job mobility," American Economic Review, 108, 3303–38.

GALÍ, J. AND T. VAN RENS (2021): "The vanishing procyclicality of labour productivity," The Economic Journal, 131, 302–326.

GARCIA-LOUZAO, J., L. HOSPIDO, AND A. RUGGIERI (2023): "Dual returns to experience," Labour Economics, 80, 102290.

GUVENEN, F., B. KURUSCU, S. TANAKA, AND D. WICZER (2020): "Multidimensional skill mismatch," American Economic Journal: Macroeconomics, 12, 210–44.

HERSHBEIN, B. AND L. B. KAHN (2018): "Do recessions accelerate routine-biased technological change? Evidence from vacancy postings," American Economic Review, 108, 1737–1772.

HOYNES, H., D. L. MILLER, AND J. SCHALLER (2012): "Who suffers during recessions?" Journal of Economic Perspectives, 26, 27–48.

HULTGREN, T. (1960): "Changes in labor costs during cycles in production and business," NBER Paper.

KAMBOUROV, G. AND I. MANOVSKII (2009): "Occupational speci city of human capital," International Economic Review, 50, 63–115.

KIM, D. AND C.-I. LEE (2007): "On-the-job human capital accumulation in a real business cycle model: Implications for intertemporal substitution elasticity and labor hoarding," Review of Economic Dynamics, 10, 494–518.

LINDENLAUB, I. (2017): "Sorting multidimensional types: Theory and application," The Review of Economic Studies, 84, 718–789.

LISE, J. AND F. POSTEL-VINAY (2020): "Multidimensional skills, sorting, and human capital accumulation," American Economic Review, 110, 2328–76.

LISE, J. AND J.-M. ROBIN (2017): "The Macrodynamics of sorting between workers and firms," American Economic Review, 107, 1104–35.

LOEWENSTEIN, M. A. AND J. R. SPLETZER (2000): "Formal and informal training: Evidence from the NLSY," in Research in labor economics, Emerald Group Publishing Limited, 403–438.

LÜTHI, S. AND S. C. WOLTER (2020): "Are apprenticeships business cycle proof?" Swiss Journal of Economics and Statistics, 156, 1–11.

LYNCH, L. M. (1992): "Private-sector training and the earnings of young workers," The American Economic Review, 82, 299–312.

MA, X., A. NAKAB, AND D. VIDART (2024): "Human capital investment and development: The role of on-the-job training," Journal of Political Economy: Macroeconomics, 2, 107–148.

MAJUMDAR, S. (2007): "Market conditions and worker training: How does it affect and whom?" Labour Economics, 14, 1–23.

MÉNDEZ, F. AND F. SEPÚLVEDA (2012): "The cyclicality of skill acquisition: Evidence from panel data," American Economic Journal: Macroeconomics, 4, 128–152.

MENZIO, G. AND S. SHI (2011): "Efficient search on the job and the business cycle," Journal of Political Economy, 119, 468–510.

MONTT, G. (2017): "Field-of-study mismatch and overqualification: Labour market correlates and their wage penalty," IZA Journal of Labor Economics, 6, 1–20.

MORTENSEN, D. T. AND C. A. PISSARIDES (1994): "Job creation and job destruction in the theory of unemployment," The Review of Economic Studies, 61, 397–415.

NEAL, D. (1999): "The complexity of job mobility among young men," Journal of labor Economics, 17, 237–261.

OKUN, A. M. (1963): Potential GNP: Its measurement and significance, Cowles Foundation for Research in Economics at Yale University.

PAVAN, R. (2011): "Career choice and wage growth," Journal of Labor Economics, 29, 549–587.

ROSEN, S. (1972): "Learning and experience in the labor market," Journal of Human Resources, 326–342.

STIROH, K. J. (2009): "Volatility accounting: A production perspective on increased economic stability," Journal of the European Economic Association, 7, 671–696.

TINBERGEN, J. (1956): "On the theory of income distribution," Weltwirtschaftliches archiv, 77, 155–175.

A Empirical Appendix

A.1 Descriptive Tables

Table A1: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	No. of Respondents
Birth Year	1982.14	1.38	1980	1984	493
ASVAB Score Percentile	77.64	19.06	12.62	100	493
Highest Education	4.53	0.84	4	7	493
– Bachelor's degree (BA, BS)	4	-	-	-	312
– Master's degree (MA, MS)	5	-	-	-	133
– PhD	6	-	-	-	16
– Professional degree (DDS, JD, MD)	7	-	-	-	32
Age	30.45	3.77	19	38	61564
Monthly working hrs	189.17	74.37	0.02	840	61,564
Hourly pay rate	21.31	24.38	0.01	977.22	58,898
Potential working years	6.21	3.64	1	18	61,564

Notes: Some monthly working hours exceed 672 hours because certain months covers ve continuous weeks.

Table A2: List of Descriptors in O*NET 17.0

Panel A: Cognitive Skills		
Oral Comprehension	Written Comprehension	Deductive Reasoning
Inductive Reasoning	Information Ordering	Mathematical Reasoning
Number Facility	Reading Comprehension	Mathematics Skill
Science	Technology Design	Equipment Selection
Installation	Operation and Control	Equipment Maintenance
Troubleshooting	Repairing	Computers and Electronics
Engineering and Technology	Building and Construction	Mechanical
Mathematics Knowledge	Physics	Chemistry
Biology	English Language	
Panel B: Social Skills		
Social Perceptiveness	Coordination	Persuasion
Negotiation	Instructing	Service Orientation

A.2 Fact I: Robustness Tests

A.2.1 Alternative Indicators

Table A3: Wage Losses and Business Cycles (Unemployment Gap)

	(1) IV	(2) IV	(3) OLS	(4) OLS
Mismatch	-0.02746		-0.02012	
	(0.00317)		(0.00290)	
Unemp. Gap	-0.03096	-0.03691	-0.04401	-0.04712
	(0.00380)	(0.00403)	(0.00390)	(0.00412)
Mismatch Unemp. Gap	0.01589		0.01564	
	(0.00186)		(0.00190)	
Positive Mismatch		-0.03236		-0.02102
		(0.00397)		(0.00387)
Negative Mismatch		0.04239		0.03347
		(0.00399)		(0.00353)
Positive Mismatch Unemp. Gap		0.02269		0.02351
		(0.00227)		(0.00257)
Negative Mismatch Unemp. Gap		-0.01178		-0.00965
		(0.00223)		(0.00197)
Observations	52,569	52,569	52,569	52,569
R^2			0.298	0.274

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses. *($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

Table A4: Wage Losses and Business Cycles (HP- ltered Unemployment Rate)

	(1) IV	(2) IV	(3) OLS	(4) OLS
Mismatch	-0.01020		-0.00288	
	(0.00263)		(0.00261)	
HP- ltered Urate	-0.04300	-0.04993	-0.05456	-0.05908
	(0.00662)	(0.00704)	(0.00693)	(0.00731)
Mismatch HP- ltered Urate	0.02576		0.02505	
	(0.00337)		(0.00359)	
Positive Mismatch		-0.00731		0.00495
		(0.00335)		(0.00332)
Negative Mismatch		0.02950		0.02251
		(0.00335)		(0.00323)
Positive Mismatch HP- ltered Urate		0.03328		0.03427
		(0.00403)		(0.00495)
Negative Mismatch HP- ltered Urate		-0.02158		-0.01867
		(0.00412)		(0.00352)
Observations	52,569	52,569	52,569	52,569
R^2			0.297	0.272

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses. *($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

Table A5: Wage Losses and Business Cycles (Unemployment-Defined Recession)

	(1) IV	(2) IV	(3) OLS	(4) OLS
Mismatch	-0.02798		-0.02048	
	(0.00333)		(0.00294)	
Recession	-0.08531	-0.09854	-0.11599	-0.12281
	(0.01013)	(0.01075)	(0.01037)	(0.01104)
Mismatch Recession	0.03808		0.03759	
	(0.00497)		(0.00500)	
Positive Mismatch		-0.03370		-0.02198
		(0.00413)		(0.00392)
Negative Mismatch		0.04259		0.03368
		(0.00420)		(0.00356)
Positive Mismatch Recession		0.05641		0.05759
		(0.00599)		(0.00653)
Negative Mismatch Recession		-0.02760		-0.02304
		(0.00607)		(0.00534)
Observations	52,569	52,569	52,569	52,569
R^2			0.299	0.274

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses. *($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

Table A6: Wage Losses and Business Cycles (One-year Lagged NBER Recession)

	(1) IV	(2) IV	(3) OLS	(4) OLS
Mismatch	-0.01625		-0.00843	
	(0.00275)		(0.00270)	
Lagged NBER Recession	-0.05740	-0.06516	-0.07858	-0.08217
	(0.01549)	(0.01650)	(0.01658)	(0.01754)
Mismatch Lagged NBER Recession	0.03996		0.03634	
	(0.00793)		(0.00812)	
Positive Mismatch		-0.01520		-0.00260
		(0.00348)		(0.00345)
Negative Mismatch		0.03489		0.02724
		(0.00350)		(0.00332)
Positive Mismatch Lagged NBER Recession		0.04747		0.04414
		(0.00967)		(0.01135)
Negative Mismatch Lagged NBER Recession		-0.03687		-0.03100
		(0.00948)		(0.00801)
Observations	52,569	52,569	52,569	52,569
R^2			0.297	0.271

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses.*($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

Table A7: Wage Losses and Business Cycles (Scaled HP- ltered GDP)

	(1) IV	(2) IV	(3) OLS	(4) OLS
Mismatch	-0.01071		-0.00328	
	(0.00263)		(0.00260)	
HP- ltered GDP	0.01658	0.02039	0.02161	0.02461
	(0.00498)	(0.00531)	(0.00493)	(0.00522)
Mismatch GDP	-0.01301		-0.01302	
	(0.00255)		(0.00261)	
Positive Mismatch		-0.00791		0.00441
		(0.00335)		(0.00330)
Negative Mismatch		0.02999		0.02292
		(0.00336)		(0.00322)
Positive Mismatch GDP		-0.01779		-0.01890
		(0.00303)		(0.00356)
Negative Mismatch GDP		0.00992		0.00919
		(0.00313)		(0.00265)
Observations	52,569	52,569	52,569	52,569
R^2			0.297	0.272

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses. *($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

A.2.2 Sorting Dynamics

Table A8: Wage Losses and Business Cycles (Job Stayers)

	(1) IV	(2) IV	(3) OLS	(4) OLS
Mismatch	-0.09665		-0.08922	
	(0.01019)		(0.00986)	
Urate	-0.03122	-0.03648	-0.04283	-0.04572
	(0.00329)	(0.00350)	(0.00318)	(0.00340)
Mismatch Urate	0.01376		0.01335	
	(0.00162)		(0.00164)	
Positive Mismatch		-0.13179		-0.12267
		(0.01277)		(0.01382)
Negative Mismatch		0.10879		0.09008
		(0.01217)		(0.01068)
Positive Mismatch Urate		0.02097		0.02114
		(0.00202)		(0.00229)
Negative Mismatch Urate		-0.01044		-0.00833
		(0.00193)		(0.00170)
Observations	38,137	38,137	38,137	38,137
R^2			0.313	0.281

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses. *($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

Table A9: Wage Losses and Business Cycles (Occupational Stayers)

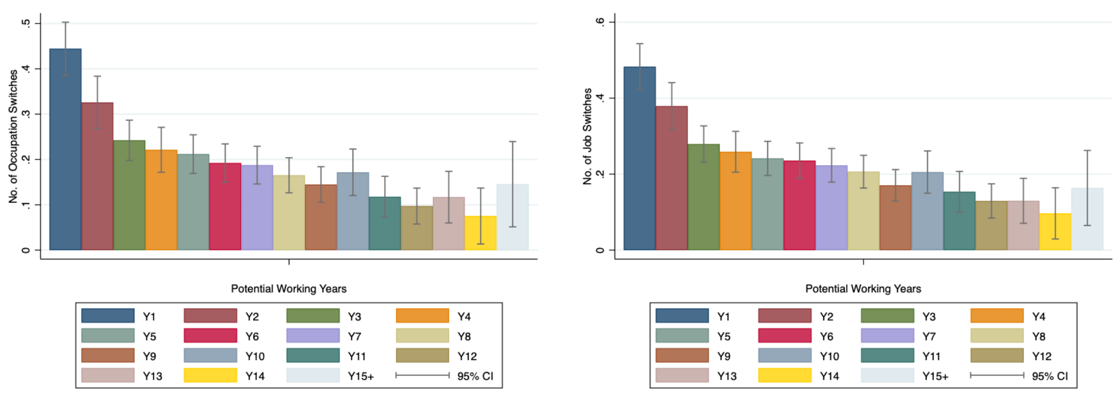
	(1) IV	(2) IV	(3) OLS	(4) OLS
Mismatch	-0.09843		-0.09161	
	(0.01006)		(0.00991)	
Urate	-0.03205	-0.03661	-0.04317	-0.04538
	(0.00325)	(0.00346)	(0.00333)	(0.00355)
Mismatch Urate	0.01410		0.01372	
	(0.00160)		(0.00166)	
Positive Mismatch		-0.13664		-0.12834
		(0.01259)		(0.01373)
Negative Mismatch		0.10304		0.08700
		(0.01202)		(0.01063)
Positive Mismatch Urate		0.02176		0.02195
		(0.00200)		(0.00229)
Negative Mismatch Urate		-0.00973		-0.00792
		(0.00190)		(0.00171)
Observations	39,620	39,620	39,620	39,620
R^2			0.305	0.276

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses. *($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

Table A10: Wage Losses and Business Cycles (Experienced Workers)

	(1) IV	(2) IV	(3) OLS	(4) OLS
Mismatch	-0.09082		-0.08319	
	(0.01013)		(0.00993)	
Urate	-0.04175	-0.04373	-0.05038	-0.05033
	(0.00337)	(0.00358)	(0.00359)	(0.00381)
Mismatch Urate	0.01206		0.01178	
	(0.00162)		(0.00169)	
Positive Mismatch		-0.11366		-0.10516
		(0.01246)		(0.01312)
Negative Mismatch		0.09487		0.07856
		(0.01227)		(0.01098)
Positive Mismatch Urate		0.01573		0.01606
		(0.00196)		(0.00219)
Negative Mismatch Urate		-0.00952		-0.00771
		(0.00198)		(0.00184)
Observations	42,440	42,440	42,440	42,440
R^2			0.291	0.261

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance. *($p < 0.10$),**($p < 0.05$),***($p < 0.01$).



(c) Occupation Switches

(d) Job Switches

Figure A1: The Average Number of Switches over Potential Experience Years

Table A11: Wage Losses and Business Cycles (Worker-job FE)

	(1) IV	(2) IV	(3) OLS	(4) OLS
Unemp. Rate	-0.02126	-0.02416	-0.02154	-0.02444
	(0.00181)	(0.00189)	(0.00246)	(0.00258)
Mismatch Unemp. Rate	0.00760		0.00780	
	(0.00091)		(0.00112)	
Positive Mismatch Unemp. Rate		0.01225		0.01251
		(0.00109)		(0.00160)
Negative Mismatch Unemp. Rate		-0.00582		-0.00598
		(0.00109)		(0.00114)
Observations	52,507	52,507	52,507	52,507
R^2			0.866	0.865

Note: All regressions include a constant, 2-digit industry FE, 2-digit occupation FE and job-worker FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses.*($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

A.2.3 Dynamic O*NET

Table A12: Wage Losses and Business Cycles (Dynamic O*NET)

	(1) IV	(2) IV	(3) OLS	(4) OLS
Mismatch	-0.07736		-0.06794	
	(0.00919)		(0.00935)	
Unemp. Rate	-0.02260	-0.02568	-0.03352	-0.03525
	(0.00285)	(0.00299)	(0.00293)	(0.00308)
Mismatch Unemp. Rate	0.01065		0.01042	
	(0.00145)		(0.00154)	
Positive Mismatch		-0.09448		-0.08906
		(0.01116)		(0.01278)
Negative Mismatch		0.07291		0.06325
		(0.01095)		(0.00973)
Positive Mismatch Unemp. Rate		0.01523		0.01638
		(0.00176)		(0.00212)
Negative Mismatch Unemp. Rate		-0.00536		-0.00448
		(0.00172)		(0.00155)
Observations	51,740	51,740	51,740	51,740
R^2			0.298	0.278

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses.*($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

A.3 Fact II: Additional Tables

Table A13: Firm-Sponsored Training Ratio by Industries and Occupations

	No. of Obs	**No. of Training**	**Training Ratio**
Panel A: By Industries			
Less Training Industries:			
Transportation and Warehousing; Utilities	1,050	9	0.01
Arts, Entertainment, Recreation, Accommodations, and Food Services	3,764	42	0.01
Agriculture, Forestry, Fishing and Hunting; Mining	891	17	0.02
Manufacturing	4,626	130	0.03
Retail Trade	4,651	138	0.03
Information and Communications	1,879	79	0.04
More Training Industries:			
Educational, Health and Social Services	10,990	538	0.05
Professional, Scientific, Mgmt, Admin., and Waste Mgmt Services	11,962	611	0.05
Construction	2,925	150	0.05
Wholesale Trade	1,827	98	0.05
Public Administration	3,137	183	0.06
Finance, Insurance, Real Estate, and Rental and Leasing	8,003	485	0.06
Other Services (Except Public Administration)	1,445	142	0.10
Panel B: By Occupations			
Less Training Occupations:			
Farming, Fishing, and Forestry	398	3	0.01
Production, Transportation, and Material Moving	1,925	50	0.03
Service	3,143	136	0.04
More Training Occupations:			
Sales and Office	12,255	576	0.05
Management, Professional, and Related	40,408	1936	0.05
Construction, Extraction, Maintenance, and Repair	2,088	123	0.06

A.4 Fact III: Robustness Tests

Table A14: Robustness Check I – Training over Business Cycles

	IV				OLS			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Panel A: Unemployment Gap								
Gap	-0.00391	0.00087	-0.00484	0.00047	-0.00839	-0.00427	-0.00847	-0.00477
	(0.00157)	(0.00280)	(0.00156)	(0.00291)	(0.00170)	(0.00301)	(0.00170)	(0.00314)
MM	-0.01319	-0.01049			-0.01066	-0.00832		
	(0.00202)	(0.00241)			(0.00202)	(0.00248)		
Pos. MM			-0.01615	-0.01319			-0.01369	-0.01213
			(0.00248)	(0.00297)			(0.00228)	(0.00281)
Neg. MM			0.00905	0.00614			0.00822	0.00574
			(0.00256)	(0.00297)			(0.00284)	(0.00339)
Gap MM		-0.00279				-0.00242		
		(0.00136)				(0.00142)		
Gap Pos. MM				-0.00302				-0.00165
				(0.00163)				(0.00157)
Gap Neg. MM				0.00308				0.00262
				(0.00159)				(0.00180)
Observations	26,568	26,568	26,568	26,568	26,568	26,568	26,568	26,568
R^2					0.062	0.062	0.060	0.060
Panel B: Unemployment-Defined Recession								
Recess	-0.01569	-0.00576	-0.01773	-0.00640	-0.02610	-0.01720	-0.02620	-0.01816
	(0.00411)	(0.00738)	(0.00409)	(0.00768)	(0.00422)	(0.00756)	(0.00422)	(0.00789)
MM	-0.01308	-0.01059			-0.01054	-0.00831		
	(0.00202)	(0.00255)			(0.00201)	(0.00253)		
Pos. MM			-0.01596	-0.01378			-0.01354	-0.01260
			(0.00248)	(0.00312)			(0.00228)	(0.00291)
Neg. MM			0.00899	0.00567			0.00810	0.00517
			(0.00255)	(0.00313)			(0.00283)	(0.00343)
Recess MM		-0.00587				-0.00527		
		(0.00363)				(0.00352)		
Recess Pos. MM				-0.00522				-0.00243
				(0.00431)				(0.00397)
Recess Neg. MM				0.00804				0.00709
				(0.00435)				(0.00449)
Observations	26,568	26,568	26,568	26,568	26,568	26,568	26,568	26,568
R^2					0.062	0.062	0.061	0.061

Note: All regressions include a constant, 2-digit industry FE, 2-digit occupation FE, as well as its interaction with the economic situation indicator. This sample includes 23 occupational and 20 industrial groups. All regressions use the robust or sandwich estimator of variance, and standard errors are in parentheses.*($p < 0.10$),**($p < 0.05$),***($p < 0.01$).

Table A15: Robustness Check II – Training over Business Cycles

	Job Stayers		Occupational Stayers		Experienced Workers	
	IV	IV	IV	IV	IV	IV
Urate	-0.00401	-0.00443	-0.00372	-0.00410	-0.00619	-0.00611
	(0.00153)	(0.00152)	(0.00151)	(0.00150)	(0.00141)	(0.00141)
Mismatch	-0.01524		-0.01480		-0.01931	
	(0.00254)		(0.00251)		(0.00222)	
Positive Mismatch		-0.01810		-0.01913		-0.02077
		(0.00316)		(0.00312)		(0.00271)
Negative Mismatch		0.01406		0.01206		0.01793
		(0.00312)		(0.00308)		(0.00279)
Observations	19,191	19,191	19,814	19,814	21,917	21,917

Note: All regressions include a constant, 2-digit industry FE, and 2-digit occupation FE, as well as their interactions with the economic situation indicator. *($p < 0.10$), **($p < 0.05$), ***($p < 0.01$).

B Model Appendix

B.1 Laws of Motion

Let $u(a,h,z)$ denote the measure of unemployed workers with learning ability a and human capital stock h when the aggregate labor productivity is z at the beginning of the human capital accumulation stage. Further, $n(a,h,z)$ denotes the measure of employed workers with learning ability a and human capital stock h when aggregate labor productivity is z.

The law of motion for unemployed workers is

$$\begin{aligned} u(a,h,z) = & \sum_{z_{-1}\in\mathcal{Z}} u\left(a,\frac{h}{1-\rho}\right)\phi(z|z_{-1})\left[1-f(a,h,z)\right] + \\ & \sum_{h_{-1}}\sum_{z_{-1}\in\mathcal{Z}} n(a,h_{-1},z_{-1})\mathbb{1}\{[h_{-1}+H(a,h_{-1},z_{-1})](1-\rho)=h\}\phi(z|z_{-1})d(a,h,z)dF(h_{-1}), \end{aligned} \tag{B.1}$$

where $f(\cdot)$ and $d(\cdot)$ represent the job finding probability and separation decision. In Equation (B.1), the first term represents unemployed workers who do not find a job, and the aggregate labor productivity updates from z_{-1} to z. The second term is employed workers whose human capital stock updates from h_{-1} to h do not lose their job, and aggregate labor productivity is updated from z_{-1} to z.

The law of motion for employed worker is

$$\begin{aligned} n(a,h,z) = & \sum_{z_{-1}\in\mathcal{Z}} u\left(a,\frac{h}{1-\rho}\right)\phi(z|z_{-1})f(a,h,z) + \\ & \sum_{h_{-1}}\sum_{z_{-1}\in\mathcal{Z}} n(a,h_{-1},z_{-1})\mathbb{1}\{[h_{-1}+H(a,h_{-1},z_{-1})](1-\rho)=h\}\phi(z|z_{-1})(1-d(a,h,z))dF(h_{-1}). \end{aligned} \tag{B.2}$$

Equation (B.2) has a similar interpretation as (B.1), except that the measure of employed workers consists of unemployed workers who find a job and employed workers who do not lose their job.

B.2 Surplus Splitting Share Derivation

Firm decides on the splitting share to maximize expected profits by posting (c, q), whereby the terms posted by the firm must ensure that workers receive their market utility W_c. J_c is an equilibrium object that is taken as given by workers. We can frame the problem as workers posting the terms of trade subject to the firm's participation constraint:

$$W_c = \max_{c,q} p(q)\,[U + c(M - U)] + (1 - p(q))\,U,$$
$$s.t. \qquad \frac{p(q)}{q}(1 - c)(M - U) = J.$$

Substitute the constraint to eliminate c from the objective function:

$$W_c = \max_{q} U + p(q)(M - U) - qJ$$

The first-order condition gives:

$$p'(q)(M - U) = J.$$

Given q, the constraint pins down the splitting share c uniquely:

$$1 - c = \frac{q p'(q)}{p(q)}.$$

B.3 Wage Derivation

By combining (7), (8) and (12), we obtain:

$$U(a,h,z) = \frac{1}{1-c} w - \frac{1}{1-c}[z(h+H) - c(t)] + b\,[1 - d(q(w',z'))]\,\mathrm{E}_{z'|z} W(w',z') + b d(q(w',z'))\mathrm{E}_{z'|z} U(a,h',z')$$
$$- b\frac{c}{1-c}\mathrm{E}_{z'|z} J(w',z').$$

By combining (6) and (12), we obtain:

$$U(a,h,z) = b + b\mathrm{E}_{z'|z} U(a,(1-r)h,z') + b p(q(a,(1-r)h,c,z'))\frac{c}{1-c}\mathrm{E}_{z'|z} J(a,(1-r)h,c,z').$$

Combining these two equations and substituting $J(w,z)$ with (9) yields:

$$w(a,h,c,z) = c\left[z(h+H) - c(t) + bkq(a,(1-r)h,z')\right] + (1-c)\left[b - b\left(\mathrm{E}_{z'|z} U(a,h',z') - \mathrm{E}_{z'|z} U(a,(1-r)h,z')\right)\right].$$

B.4 Proof of Proposition 1

The correlation between t and z can be proved by contradiction. Suppose that the optimal training t is negatively related to labor productivity z. If z increases, by assumption, t would decline. As a result, the left-hand side of (15) decreases, since the training cost function is increasing and convex. On the right-hand side of (15), $\frac{\partial H}{\partial t}$ increases due to the concavity of $H(a,t)$. Together with the rise in z, the right-hand side becomes larger. This contradicts the decrease on the left-hand side.

To ensure that $H(a,t)$ is increasing and concave in t, we can first derive the first- and second-order derivatives:

$$\frac{\P H(a,t)}{\P t} = a t^{e \quad 1} \quad a t^{e} + (1 \quad a)(m \quad exp(\quad \frac{(a \quad r)^2}{2n^2}))^{e} {}^{\frac{1 \quad e}{e}}, \tag{B.3}$$

$$\frac{\P^2 H(a,t)}{\P t^2} = a(e \quad 1) t^{e \quad 2} \quad a t^{e} + (1 \quad a)(m \quad exp(\quad \frac{(a \quad r)^2}{2n^2}))^{e} {}^{\frac{1}{e} \quad 2} \quad (1 \quad a)(m \quad exp(\quad \frac{(a \quad r)^2}{2n^2}))^{e} \quad . \tag{B.4}$$

From equation B.3, $H(a,t)$ is increasing in t. To guarantee the concavity of $H(a,t)$, we require that $e < 1$, or equivalently, that the elasticity $\frac{1}{1 \quad e} > 0$. This condition implies that training and on-the-job learning are complementary to each other.

B.5 Mechanism Demonstration

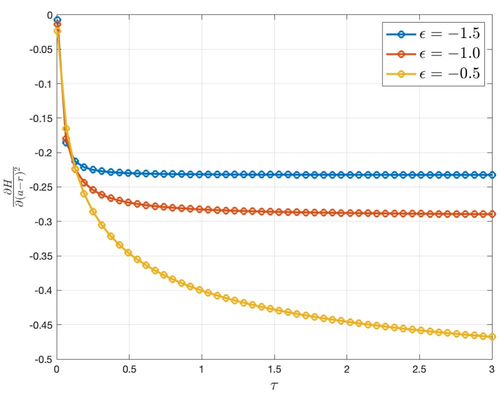


Figure B2: Responsiveness of $\frac{\P H}{\P(a \quad r)^2}$ to t