Risk Preference, Occupation Choice, and Earnings Dynamics*

Woosuk Choi†

May 2, 2024

Abstract

Risk preference is a fundamental element in labor market decisions, which inherently entail risks. This paper investigates the role of risk preference in occupation choices and earnings inequality, utilizing data from the NLSY97 and O*NET database. The descriptive findings suggest that workers with greater risk aversion tend to earn lower wages, and this wage gap becomes especially pronounced as their careers progress. Moreover, the evidence indicates that risk-averse individuals tend to choose routine occupations, which provide less opportunity for earnings growth. To investigate mechanisms driving the dynamic impact of risk preferences on earnings, I develop a structural model of occupation choice with human capital accumulation, allowing for heterogeneity in risk preferences. The model is estimated using the Conditional Choice Probability estimator accounting for unobserved heterogeneity. The estimated model underscores that the effect of differential human capital accumulation accounts for 40 percent of earnings inequality between risk attitude groups, highlighting the importance of considering the propagating effect of self-selection on risk attitudes. Using the estimated model, I show that routine-biased technological change contributes to larger disparities in occupation choices across workers with varied risk preferences, enlarging earnings disparities.

*I am deeply grateful to Ronni Pavan, Lisa Kahn, and John Singleton for their invaluable guidance and support. I also thank ..., and all participants in the University of Rochester Applied Reading Group and Student Seminar. All errors are my own.

†Department of Economics, University of Rochseter. Email: woosuk.choi@rochester.edu

1 Introduction

Risk preference is a fundamental element in most economic decisions and this holds especially true for labor market decisions. The majority of choices within labor markets inherently entail risk as their outcomes remain uncertain from the perspectives of workers before their realization. These risks are not diversifiable across alternatives and are not entirely insurable from the financial market. For example, an individual can usually have one full-time occupation over others, yet the precise trajectory of their future earnings remains obscured until gainful employment is secured within their chosen profession. Consequently, the role of risk preferences in decision-making becomes crucial.

Given the pivotal role of risk preference in labor market decisions, the heterogeneity in risk preferences, documented by prior literature[1], likely yields substantial disparities in both labor market decisions and resultant earnings. There has been an extensive body of literature in labor economics devoted to investigating the factors contributing to earnings disparities among workers. This includes the distribution of worker characteristics, such as cognitive ability or social skills, which are closely intertwined with earnings determination (e.g. Deming, 2017). Moreover, since Roy (1951), the concept of workers' self-selection based on those characteristics has been acknowledged as equally, if not more, influential than the variation in worker characteristics. This implies that worker's attitudes toward risks in labor market decisions can also influence their outcomes. Although previous studies have separately shown occupation choices to be influential in wage determination (Sullivan, 2010) and to be correlated with risk preferences (Fouarge et al., 2014), there is a scarcity of empirical studies examining the precise relationship between risk preferences, occupation choices, and individual earnings.

In this paper, I explore the impact of diversity in risk preferences on individuals' occupational choices and the resulting distribution of earnings. Initially, I provide suggestive evidence indicating

[1] Dohmen and Falk (2011) provides evidence of substantial heterogeneity in self-reported risk attitudes. Patnaik et al. (2020) measure the coefficient of risk aversion for college students and show that the distribution of risk aversion is non-degenerate.

a negative correlation between the level of risk aversion and not only the levels of earnings but also their trajectory over workers' lifetimes. This dynamic relationship between risk preference and earnings can be explained by observing the differential occupation choices made by individuals with varying risk preferences: those who are more risk-averse tend to opt for more routine careers characterized by lower earnings and limited earnings growth. Then, I develop an occupation choice model with risk preference heterogeneity, thereby consolidating reduced form components into a coherent framework and rationalizing the dynamic effect of risk preference on earnings inequality. Within this model, more risk-averse workers tend to self-select into occupations with lower variability of unexpected shocks and accumulate different sets of skills over time. I quantify the relative importance of unobserved heterogeneity, occupation choice, and human capital accumulation in driving the effect of risk preference on earnings disparities. Furthermore, I assess how variation in risk aversion can alter the effect of routine-biased technological changes on workers' earnings trajectories.

Using the National Longitudinal Survey of Youth 1997 (NLSY97) and the Occupational Information Network database (O*NET), I first present suggestive evidence that underscores the close relationship between risk preferences and earnings profiles. In the NLSY97, I observe individual responses to hypothetical job lottery choices. This gamble choice enables the classification of individuals into four distinct risk attitude groups. Utilizing this categorization, I show evidence of a significant premium for risk tolerance. Notably, the findings reveal that the most risk-tolerant workers have an average earnings advantage of 0.1 log points relative to their most risk-averse counterparts over their lifetime. Moreover, this wage premium exhibits a modest magnitude at the beginning of one's career, subsequently expanding to as much as 0.15 log points. It is particularly intriguing to note that risk preferences not only correlate with earnings levels but also with earnings growth. This observation suggests that factors beyond mere instantaneous compensation for risk in labor market decisions contribute to differences in earnings trajectories across risk attitude groups. In other words, risk preference appears to be associated with dynamic components of labor market decisions in addition to contemporaneous risk premiums.

I then analyze how occupation choices intervene in the dynamic relationship between risk preferences and earnings. First, I demonstrate that individuals with higher levels of risk aversion tend to self-select into more routine less cognitive careers. This trend is potentially driven by the observation that routine occupations typically exhibit less variability of earnings, as measured by the variance of residual wages. Second, the pursuit of more routine, less cognitive careers relates to limited growth of earnings. To examine this relationship, I regress the log of current earnings on last-period task combinations conditional on current occupations. The results suggest that even when two workers work in the same occupation, one with the more cognitive occupation last period tends to earn more than their counterpart. Recent literature contextualizes these observed patterns as indicative of task-specific skill accumulation; those in more cognitive occupations accumulate cognitive skills more rapidly. Given that cognitive skills are more rewarded in the labor market, they consequently lead to faster growth of earnings over time.

Motivated by the reduced-form evidence on the relationship between risk preference, occupation choice, and earnings dynamics, I build a dynamic occupation choice model that incorporates heterogeneity in risk preferences and human capital accumulation. The usefulness of the model is twofold. First, it facilitates the aggregation of three distinct components potentially correlated with risk preferences and earnings, enabling the quantification of their relative importance in driving risk preference effects. First, risk preference may correlate with both observed and unobserved characteristics that affect earnings. Second, self-selection on risk preferences involves an occupational risk premium for more risk-tolerant workers. Third, human capital accumulation differing across occupations reinforces the earnings gap across workers with varying risk preferences. Integrating these three components into one framework, I can investigate the relative importance of each in contributing to earnings disparities resulting from risk preference heterogeneity.

Another advantage of the model lies in its capacity for conducting counterfactual analysis regarding the role of risk aversion in workers' responses to economic shocks. In particular, I focus on the disproportionate effect of routine-biased technological change (RBTC) on risk-averse workers. The RBTC entails lower wage premiums for routine occupations while providing higher premiums for non-routine cognitive occupations (Cortes, 2016). As a consequence, workers are incentivized to choose cognitive careers as rewards for such occupations become more lucrative. However, transitioning toward more cognitive occupations indicates having a larger variance of shocks within the model, impeding the ability of risk-averse workers to change their career trajectories. In other words, risk aversion operates as a barrier for risk-averse workers to adapt their behaviors in response to the RBTC. With these opposing effects in the occupation choice problem, the heterogeneity in risk preference can result in distributional consequences of the RBTC on workers.

I build the model upon Yamaguchi (2012) and introduce workers' heterogeneity in risk preferences and occupation-specific risks. Workers start their careers with differentiated risk preferences and skill endowments that are allowed to be correlated with each other. Every period, they choose their occupations fully characterized by task combinations. At a given period, they are engaged in gambling choices based on their earnings at the moment and risk preferences. I assume earnings risks are occupation-specific, wherein task combinations dictate the variance of shocks, whether it is transitory or persistent. As a result, risk-averse workers prefer occupations with a lower variance of shocks, all else being equal. Human capital accumulation is also occupation-specific due to learning by doing; if one engages in more cognitive tasks, one accumulates more cognitive skills. This represents an additional pathway through which risk preferences affect occupation choices; risk-averse workers tend to be in more routine careers, in which they accrue more routine skills and consequently can obtain higher returns to their skills. Thus, risk preferences exert a substantial influence on occupation choices not only directly through risk aversion but also indirectly via the differential accumulation of human capital across occupations.

I model workers' preference over earnings to be non-homothetic so that relative risk aversion potentially decreases with earnings. Especially, the utility function of earnings solves the second-order differential equation that defines relative risk aversion as a function of parameters and earnings (Meeuwis, 2020). This assumption is to capture the possibility of a reverse relationship between risk attitudes and earnings in descriptive analysis: those who have higher earnings at the moment of the survey are less risk-averse in gambling choices.

The structural model is estimated using the conditional choice probability (CCP) estimators with unobserved heterogeneity, developed by Arcidiacono and Miller (2011). This approach enables me to address three challenges that make the model intractable. First, the utility function with risk aversion introduces the non-linearity of flow utility, resulting in the absence of an analytical solution for the model. Second, the model involves a large choice set and extensive state spaces, including continuous skills. The CCP estimator mitigates the need for fully solving the recursive model by leveraging the relationship between choice probability and value functions, reducing a substantial computational burden. Finally, the presence of unobserved heterogeneity in risk preferences and skill endowments makes the straightforward application of CCP estimation infeasible. To address this challenge, the expectation-maximization (EM) algorithm is employed to iteratively estimate the structural parameters while accounting for unobserved heterogeneity.

In addition, the model is estimated in the high school graduate sample and the college graduate sample separately. This is to capture that the availability of occupations is substantially different across education levels. For instance, occupations like scientists and engineers may have qualification restrictions that require at least a college degree in the field to work in. Although the education level can be endogenously determined with future occupational qualifications considered, modeling educational decisions is outside of the scope of this paper. The distinction of sample based on a college degree is to abstract away from education decisions and to focus on the occupation choice channel of the risk preference effect on earnings.

The estimates under the CRRA utility assumption indicate that there exists a significant difference in earnings risks across occupations. The estimates of transitory risks present a consistent pattern with the descriptive analysis: earnings variability decreases in both types of tasks with routine tasks having a stronger correlation. The estimated return to skills and tasks also confirms that pursuing cognitive careers is more lucrative relative to routine careers. Finally, there exists a significant association between unobserved initial skills and risk preferences. This implies that the reduced form effect of risk preferences on earnings is likely to absorb the effect of unobserved heterogeneity.

I use the estimated model to decompose the earnings gap across risk attitude groups into three components: correlation with other worker characteristics, contemporaneous return from occupation choices, and human capital accumulation effects. In particular, fixing occupation choices as in the baseline model, I simulate workers' earnings profiles first without any heterogeneity in both observed and unobserved attributes and subsequently with human capital accumulation the same across individuals. I find that around 75% of the gap between the risk attitude group can be attributed to differential occupation choices. Moreover, diverging human capital accounts for 40% of the occupation choice effects. The results highlight the importance of dynamic components in evaluating the effect of risk preferences on labor market outcomes on top of the instantaneous return.

I next conduct a counterfactual analysis that evaluates the role of risk aversion in workers' responses to routine-biased technological change (RBTC). With the RBTC, labor inputs in cognitive tasks become more productive while the return to routine task labor becomes smaller. I assess how workers change their occupation choices with a higher return to cognitive tasks and a lower return to routine tasks. I find that the RBTC reinforces the disparities in earnings across workers with varied risk preferences. There are two channels for a larger earnings gap. First, more risk-averse workers were already in more routine careers in the baseline, so the lower return to routine occupations directly reduces the earnings of risk-averse workers. There are also behavioral consequences. In the counterfactual, the less risk-averse workers have far more cognitive occupations relative to the others, acquiring larger gains from the RBTC.

Related Literature

This paper contributes to the literature on the association between risk aversion and occupation choices. Previous studies document the robust correlation between risk attitudes and occupation choices regarding physical risks (DeLeire and Levy, 2004; Grazier and Sloane, 2008), self-employment (Ahn, 2010), and public sector (Buurman et al., 2012). A small body of work presents evidence of more risk-averse workers working in occupations with a lower variance of earnings (King, 1974; Bonin et al., 2007; Fouarge et al., 2014). My contribution is to provide the first empirical framework that incorporates unobserved heterogeneity interacting with risk preferences and to examine the consequence of self-selection based on risk preferences in earnings profiles. This is in line with the spirit of Shaw (1996) who shows the short-term correlation between risk aversion and individual income growth. In this paper, I provide evidence proposing occupation choices as a direct mechanism driving the growth effects of risk aversion over the life cycle.

I also add to the literature on the effect of occupation choices on earnings dynamics and distributions. The literature on career choice has provided evidence of the contribution of career choices and self-selection on unobserved heterogeneity to wage disparities across different workers (Keane and Wolpin, 1997; Gould, 2002; Kambourov and Manovskii, 2009; Sullivan, 2010; Yamaguchi, 2010; Pavan, 2011). Central to the setting of this paper are task-specific multidimensional skills accumulating over time, which is recently developed and supported by the literature (Lazear, 2009; Gathmann and Schönberg, 2010; Yamaguchi, 2012; Lindenlaub, 2017; Guvenen et al., 2020; Lise and Postel-Vinay, 2020). Especially, Yamaguchi (2012) provides a new approach that accommodates the heterogeneity in human capital across individuals using occupational task combinations as sufficient statistics to describe occupations. I build upon their frameworks to incorporate the heterogeneity in risk preferences and their interaction with skills into the structural model of occupation choices and to characterize occupation-specific risks as a function of tasks.

Finally, my paper also contributes to the literature on risk preference measurement. Recognizing the fundamental role of risk preference in economic decision-making, a large body of literature has made progress in developing an empirical measure of risk preferences with various instruments including self-reported risk attitudes, field experiments, and observed financial decisions (see e.g. Dohmen and Falk, 2011; Barseghyan et al., 2018). I contribute to this literature by estimating relative risk aversion potentially decreasing with earnings using the observed occupation choices with occupation-specific risks.

2 Data

I use the National Longitudinal Survey of Youth 1997 (NLSY97) and the Occupational Information Network (O*NET) database for the analysis. The NLSY97 includes information about demographics and employment histories for a sample of individuals. The O*NET database includes occupation-specific characteristics surveyed through job experts and incumbent workers.

2.1 National Longitudinal Survey of Youth 1997 (NLSY97)

The NLSY97 is an ongoing survey that initially included 8,984 individuals born between 1980 and 1984. The dataset is well-suited for this analysis due to its detailed employment information, including occupation codes, wages, and labor supply, which are available for earlier stages of individuals' careers. More importantly, the NLSY97 provides the answers to a lottery choice experiment that is related to hypothetical job offers. This helps to recover respondents' risk attitudes in the labor market. This rich dataset allows for a thorough examination of the relationship between risk preferences, occupation choices, and earnings outcomes.

I focus on the 1997 cohort and thus expect less labor market detachment of females than 1979 cohort. Among 8,984 individuals, 3,084 individuals whose job lottery choices, the AFQT score, and the education information are missing are removed from the sample. Next, I drop 108 individuals who ever participated in military services. 577 individuals are dropped from the sample who never made long-term transitions into the labor market. A long-term transition means working full-time (more than 30 hours per week) for at least three consecutive years. Then, I removed all observations before the long-term transition. Observations after labor market detachment, or being unemployed for more than two consecutive years, are dropped. Lastly, those who have less than 3 observations or made a long-term transition before age 16 are dropped from the sample. The final sample includes 5,189 individuals with 55,246 yearly observations. If hourly rates of compensation are less than one dollar or larger than 100 dollars, they are considered missing since they are likely to be misreported.

2.1.1 Risk Attitude from Hypothetical Job Lottery Choice

I use individuals' responses to hypothetical job lottery choices to infer their risk preferences. To be specific about hypothetical lotteries, the survey asks respondents at the 14th or 15th wave, when individuals were between 26 and 30 years old:

"Suppose you are the only income earner in the family, but that your current job is ending. You have to choose between two new jobs. The first job would guarantee your current family income for life. The second job is also guaranteed for life and possibly better paying, but the income is less certain. There is a 50-50 chance that the second job will double your current family income for life and a 50-50 chance that it will cut your current family income by a third for life."

Individuals who chose the risky job are then asked:

"Suppose the chances were 50-50 that the second job would double your current family income and 50-50 that it would cut it in half Would you take the first job or the second job?"

On the contrary, those who would not take the risk in the first question are asked:

"Suppose the chances were 50-50 that the second job would double your current family income and 50-50 that it would only cut it by 20 percent. Would you take the first job or the second job?"

Depending on the responses to these hypothetical gambles, samples can be divided into four groups. The first group is most risk-tolerant in the sense that they would take the risk even if it is possible to experience an income cut more than or equal to half. The second group would take the risk if the possible loss is less than half but larger than a third. The third group is the one who would accept the risky job if the possible loss is less than a third but larger than 20 percent. Most risk-averse are the last group who never accept a risky job with more than 20 percent possible loss with half probability.

2.2 Occupational Information Network (O*NET) database

The O*NET database is a comprehensive job description data for 927 occupation categories, based on the Standard Occupation Classification (SOC). It is established to help people to match with better training and occupations based on occupation description and to support employers in finding skilled workers. The data is collected by the National Center for O*NET Development and supported by the US Department of Labor (DOL)/Employment and Training Administration (ETA). A random sample of businesses and incumbent workers in targeted occupations are selected to complete a standardized questionnaire and occupational analysts rate several fields using updated information. It contains 277 descriptors and ratings on knowledge, skills, and abilities required on the occupation, and tasks, activities, and context around the occupations.

Using this data set, I characterize each occupation as a bundle of tasks. Each occupation requires different tasks with different levels of intensity. This approach proves to be useful in capturing the heterogeneity and the similarity between occupations at the narrowly defined level of occupations. For example, in the 2-digit SOC classification, lawyers and legal assistants are in the same occupation group even though their required tasks are considerably different: lawyers' works require more professional knowledge in the field while legal assistants are more engaged in documentation and searching information. On the other hand, administrative assistants are classified in a broad group which is considered as less professional career compared to legal workers despite their structure of work similar to legal assistants. If a legal assistant is to move to either a lawyer or an administrative assistant, one can expect the skills they have been using to be useful in the latter. Characterizing occupations based on tasks enables one to measure the similarity, and thus the transferability of skills, between occupations even across broadly defined occupation groups.

I construct two-dimensional task intensities to characterize occupations: *Cognitive* and *Routine task*. Each task intensity is generated by the principal component analysis as in Yamaguchi (2012) with selected O*NET items.[2] I standardize the first principal component at the occupation level to obtain an index of corresponding task intensity. Finally, I convert all indices into percentile scores so that both tasks share the same range.

The first four columns of Table A2 report the mean and the standard deviation of each task intensity within 2-digit occupations. Occupations related to manufacturing, transportation, and administrative work show the highest values of routine task intensity, and occupations such as education and social services involve the least intensity of routine jobs. Management, social services, and engineering report the highest cognitive task intensity, while construction, repair, and transportation occupations are the most intensive in motor tasks. As the literature has shown, the table indicates that the unidimensional characterization of occupations is not sufficient to illustrate heterogeneity across occupations. For instance, management and social service occupations are similar in cognitive task intensity (0.93 v. 0.933), but routine task intensities of those jobs are different from each other (0.205 v. 0.051). Moreover, the high value of standard deviation implies that there exists considerable heterogeneity within broad categories of occupations. Therefore, the results hint that multidimensional characterization of occupation is essential in analyzing occupational heterogeneity.

In Table 1, I present the summary statistics on demographic and labor market characteristics of the sample. The most risk-averse group accounts for around 52% of the sample. On average, more risk-averse workers consist of more Black, women, less educated, and those with lower AFQT scores. They have lower hourly rates of earnings and work less per year. Finally, they perform less-cognitive tasks while having more-routine tasks on average.

3 Descriptive Evidence on Risk Preferences and Labor Market Outcomes

In this section, I present descriptive evidence of the relationship between risk preference, earnings profiles, and occupation choices. I show that workers with higher levels of risk aversion exhibit

[2]For cognitive task intensity, many items such as *Reading comprehension* and *Mathematics* are included. Routine task intensity is constructed utilizing items including *Spend time making repetitive motions*, *Importance of being exact or accurate*, *Importance of repeating same tasks*, *Pace determined by the speed of equipment*, *Controlling machines and processes*, and the negative of *Structures versus unstructured*.

not just lower lifetime average earnings and volatility but also slower growth of earnings. It is also presented that more risk-averse workers tend to choose more routine and less cognitive occupations which relate to reduced earnings risks. Finally, the evidence reveals that opting for more routine occupations can lead to limited earnings growth, aligning with the hypothesis of task-specific skill development.

3.1 Risk Tolerance Premium in Earnings and Growth

I document the significant correlation between risk preferences and earnings profiles in this subsection. First, I test whether people with varied risk preferences have different levels of earnings and volatility over their lives. If less risk-averse workers pursue riskier careers with premiums in terms of earnings, they would have higher earnings and volatility. To test differences in earnings patterns across risk attitudes, I regress the lifetime average and variability of individual earnings on risk attitude group indicators. Especially, I estimate

$$y_i = \alpha + \sum_{g=1}^{3} \beta_g \mathbf{1}(G_i = g) + X_i'\beta + \epsilon_i \tag{1}$$

where y_i is the dependent variable (lifetime average and standard deviation of earnings), G_i represents the risk attitude group variable taking values from 1, the most tolerant, to 4 the most averse. The omitted group is the 4th group who are classified as the most risk-averse group. The parameter of interest is β_g which measures the average difference in the dependent variable relative to the most averse group. X_i is a vector of observed worker characteristics. I control for various worker characteristics as risk attitude elicited from job lottery choices is correlated with observed worker characteristics, suggested by Table 1. They include race, gender, age-adjusted AFQT scores, and education level indicators as a baseline model. In the other specification, I additionally control for non-cognitive skill measures and parental income variables. In this case, I compare the results with and without the additional controls using the restricted sample that has all the information, to evaluate whether the change of coefficient is driven by additional control or sample restriction.

The results reveal that risk aversion has a significant negative relationship with lifetime earnings as well as volatility. Table 2 reports the coefficient estimates γ_g for each specification. The first three columns show the results on lifetime average earnings. Column (1) indicates that the most risk-tolerant workers have earnings approximately 0.11 log points higher compared to the most averse workers and it gets smaller but still significantly positive as the group gets closer to the most averse group. Columns (2) and (3) show that the results are robust to controlling for additional controls: the most tolerant individuals have around 17 percent higher average earnings. Consistent with the hypothesis above, columns (4)-(6) present that earnings premium for the more risk-tolerant individuals comes with higher earnings volatility. In the baseline specification, the most tolerant individuals experience 0.15 log points higher standard deviation of log earnings.

I next investigate the relationship between risk preference and earnings growth. I estimate experience-earnings profiles by risk attitude groups using the following regression model:

$$lnEarnings_{it} = \alpha + \sum_{g=1}^{3} \beta_g \mathbf{1}(G_i = g) + \sum_{g=1}^{3} \gamma_{gt} \mathbf{1}(G_i = g) + X_i' \beta + \phi_t + \psi_{\tau(i,t)} + \epsilon_{it} \quad (2)$$

where $Earnings_{it}$ is the hourly earnings of individual i with experience t. X_i includes race, gender, AFQT scores, and education level indicators. ϕ_t and $\psi_{\tau(i,t)}$ refer to experience level and year fixed effects, respectively. The parameters of interest are γ_{gt}, the excess earnings growth of group g from experience level 1 to t relative to the most averse workers.

Figure 2 presents that the more risk-tolerant workers have steeper earnings profiles compared to the most averse. In particular, the most risk-tolerant workers experience around 20 percentage points higher growth of earnings over 14 years relative to the most averse workers. This implies that there exist some dynamic components in labor market decisions contributing to the earnings gap on top of contemporaneous premiums for taking risks.

This paper posits that (part of) the observed pattern between risk preference and earnings can be attributed to occupational segregation resulting from heterogeneous risk preferences. In Table A1, I conduct a mediation analysis to examine the extent to which occupation profiles can explain the impact of risk preference on earnings. I first define the major occupation for each individual based on their longest-tenured occupation, while the second major occupation is defined as the one with the second longest tenure. Columns (1) and (4) estimate the same specification of the corresponding columns in Table 2 using the sample with at least two occupations. In the subsequent columns, I sequentially include fixed effects for the major and the second major occupations. The findings show that approximately 20-40 percent of the risk preference effect is mediated through heterogeneity in major occupations, and it diminishes further conditional on the second major occupation. This indicates that not only does occupation choice matter in explaining the relationship between risk preference and earnings but also the trajectory of one's career path could matter. In other words, the dynamic nature of occupation choices may contribute to the observed relationship between risk preferences and earnings.

3.2 Risk Preference and Occupation Choice

I next examine how risk preference is related to occupation choice. If different occupations exhibit different levels of earnings risk, more risk-averse workers theoretically sort into occupations with a safer earnings stream. I first measure occupational earnings risks to characterize safer occupations. Then, based on the relationship between earnings risk measure and task intensity, I provide suggestive evidence of self-selection on the basis of risk aversion into safer occupations.

To illustrate that earnings variability differs across occupations, I utilize a measure of earnings risks derived from residuals in an earnings regression. The regression model includes individual-occupation fixed effects, year-fixed effects, tenures, and experience profiles represented by cubic functions. By calculating the standard deviation of these residuals within the 3-digit SOC level, I construct a proxy for occupation-specific earnings risk. The measures reveal a wide range of values across different occupations, ranging from 0.089 to 0.486. This substantial heterogeneity emphasizes the variability of earnings risks within the occupational landscape. Furthermore, in Table A2, I present the mean and the standard deviation of the measure within the 2-digit occupation level. These findings demonstrate that even within occupational groups, there exist significant variations in earnings risk. Thus, adopting a more nuanced definition of occupations is crucial for comprehending the influence of risk aversion on occupation choices.

Using the constructed risk measure, I explore the relationship between task combinations and earnings risks in Table 3 to characterize safer occupations that would be chosen by risk-averse workers. The risk measure is regressed on task intensity indices, revealing a significant correlation between the two indices. Especially, routine task intensity exhibits a strong negative relationship with residual wage variances: a one standard deviation increase in routine task intensity corresponds to a 0.431 decrease in the standard deviation of residual wage variances by itself. Conditional on cognitive task intensity, the negative relationship gets even stronger. On the other hand, cognitive task intensity presents a significant negative correlation only when conditioned on routine task intensity. When considering the interplay between task intensity and earnings risks, it is suggestive that more risk-averse workers opt for more routine occupations to mitigate exposure to riskier career paths.

I test the hypothesis that more risk-averse workers are more likely to be in more routine careers. I regress individuals' chosen routine (cognitive) task intensity on risk attitude group indicators:

$$y_{it} = \alpha + \sum_{g=1}^{3} \beta_g \mathbf{1}(G_i = g) + X_{it}'\beta + \psi_t + \epsilon_{it} \tag{3}$$

where y_{it} is the dependent variable (task intensity, risk measure, average wage) of the occupation that individual i holds at period t, G_i is the risk attitude group variable, X_{it} includes race, gender, age-adjusted AFQT scores, education level indicators, and experience level indicators, ψ_t is year fixed effects. β_g is the parameter of interest which indicates the excess amount of the dependent variable for group g compared to the most averse workers.

The findings in the first column of Table 4 support the idea of aversion to earnings risks, as it reveals that more risk-averse workers tend to select more routine occupations. For example, the most tolerant workers on average choose occupations with about 0.08 standard deviation lower routine task intensity. Considering the negative correlation between cognitive and routine tasks, choosing a higher routine task intensity may be achieved at the expense of cognitive task intensity. The result on cognitive task intensity, presented in column (2), is suggestive of this trade-off: the more averse workers hold occupations with lower cognitive task intensity. Column (3) reports that more risk-averse workers have occupations with significantly lower earnings risk measures, confirming that selection into more routine careers is closely related to aversion to earnings risks. Finally, the result in column (4) implies a premium for earnings risks: occupations chosen by more risk-averse workers on average pay less.

3.3 Occupational Choice and Earnings Growth

Last, I explore the implication of occupational selection based on risk preferences for earnings growth. Recent literature on human capital emphasizes the existence of task-specific skills acquired through learning-by-doing, transferable across occupations to varying extents (Gathmann and Schönberg, 2010). These multi-dimensional task-specific skills have been shown to effectively explain wage patterns, with each dimension associating different returns (Yamaguchi, 2012). This idea is particularly pertinent in this paper's context. If risk-averse workers have to choose more routine and less cognitive occupations to mitigate earnings risks, they may accumulate a different set of skills whose returns are potentially different. Essentially, they might experience slower accumulation of human capital if routine skill is rewarded less in the market.

To provide the suggestive evidence that selecting more routine occupations may lead to slower growth of earnings, I regress log earnings on last period's task intensities and their interaction with current ones as follows:

$$\begin{aligned}\ln Earnings_{it} =& \beta_c Cog_{it-1} + \beta_r Rou_{it-1} + \beta_{cc} Cog_{it} Cog_{it-1} + \beta_{rr} Rou_{it} Rou_{it-1} \\ &+ \alpha_i + \psi_{j(i,t)} + \lambda_t + \epsilon_{it} \end{aligned} \tag{4}$$

where $Earnings_{it}$ is hourly earnings of individual i at period t. α_i and λ_t represent individual and year fixed effects, respectively. I include $\psi_{j(i,t)}$, current occupation fixed effects. β_c and β_r estimate the return to having worked in cognitive and routine jobs, conditional on current occupations. In other words, even if two workers work in the same occupation today, their earnings can differ due to their working histories. A positive β_c implies that the return to experience in cognitive task is positive. Furthermore, a positive coefficient on the interaction term β_{cc} suggests that cognitive task experience is particularly rewarded in cognitive occupations.

Table 5 indicates that the return to cognitive task experience is positive and significant and it is higher in more cognitive occupations today. Conversely, having worked in routine tasks shows negligible return, but it has a positive premium if in more routine occupations. Consequently, if a worker starts their careers in routine jobs, they are inclined to continue working in similar occupations to get better returns albeit having slower growth of earnings.

In summary, the descriptive analysis suggests the existence of a significant relationship between risk preferences, earnings dynamics, and occupation choices. While suggestive, the interpretation of evidence requires additional caution as hypothetical gamble responses, collected in the middle of individual careers, may have been influenced by their wealth and labor market experiences. For example, under the assumption of decreasing relative risk aversion, those who happen to earn more at the moment of the survey become less risk-averse, which implies the reverse relationship between risk attitudes and earnings. Although previous studies suggest the attitudes elicited through the same job lottery choices remain stable against the change in earnings and employment status,[3] the conclusion about the stability of individual risk preference is still mixed in the literature. To address the possibility of risk preference varying over earnings, the model assumes non-homothetic utility over earnings so that relative risk aversion potentially decreases in earnings.

The analysis is also limited in studying the risk preference effect on earnings through occupation choices since multiple factors may be involved. First, as suggested by summary statistics and previous literature, risk preference may be correlated with worker characteristics whether they

[3] Sahm (2012) uses the panel structure of hypothetical job lottery choices in the Health and Retirement Study (HRS) to examine whether individual risk preferences are stable over their life cycles. They find that while risk preference changes over time, the changes in income and wealth are not the driver of the change in their risk attitudes.

are observed or not. To the extent that correlated characteristics affect occupation choices or productivity, the risk preference effect may absorb those effects. Second, even differential occupation choices can be relevant with various components. Different occupations have different instantaneous returns which include risk premiums. This would lead to the difference in earnings levels across workers with varied risk preferences. Moreover, as discussed above, differences in occupation choices may have a persistent influence on earnings, indicating diverging human capital accumulation. To incorporate various components related to risk preferences into one comprehensive framework, in the next section, I develop a dynamic occupation choice model with risk preference heterogeneity and individual- and occupation-specific human capital accumulation.

4 Occupation Choice Model with Risk Preference Heterogeneity

In this section, I present a dynamic model of occupation choices. The model is built upon the task-based approach suggested by Yamaguchi (2012) in the sense that each occupation is fully characterized by a vector of task intensities and there exist multidimensional skills corresponding to each task. The main novelty of the model is the heterogeneity in risk preferences, occupation-specific distribution of shocks, and non-homothetic preference over earnings. Departing from Yamaguchi (2012), I further assume that individuals choose occupations from a discrete choice set.[4]

At the beginning of their careers, workers are characterized by risk preference type, γ_i, initial skill endowment, s_{i1}, and demographic variables, x_i. For expositional simplicity, the individual subscript i is suppressed hereafter throughout the section. They choose an occupation, j, from a set of occupations, $\{1, \cdots, J\}$. Each occupation of $1, \cdots, J$ can be characterized by K dimensional task intensities, $y_j = (y_{j1}, \cdots, y_{jK})$. The labor market is assumed to be competitive and information on workers' employment history and skill endowments is publicly observed. Workers in occupation j also face corresponding transitory and persistent risks, all of which are formally

[4] Yamaguchi (2012) assumes that individuals choose task bundles directly. This setting is feasible because the linear utility over log earnings allows a closed-form solution to the task choice problem. In this paper, the non-linearity of the utility over earnings is essential in incorporating the concept of risk aversion heterogeneity.

defined in the following subsection.

The timeline of the problem is as follows: each period, a worker chooses an occupation. Earnings are realized with transitory shocks. Skills are accumulated depending on the task intensities of their occupations and persistent shocks. Then, the period ends.

4.1 Utility from Earnings with Risks

The assumption of a competitive labor market implies that workers are paid according to their marginal value of products. Wages are *ex-ante* stochastic with transitory shocks whose distributions vary across occupations. Suppressing the individual subscript, a worker with skill endowment $s_t \in \mathbb{R}^K$ in occupation j obtains

$$\ln w(j; s_t, e_{jt}) = \pi(y_j) + q(y_j, s_t) + e(y_j, e_{jt}) \tag{5}$$

where $e(y_j, e_{jt}) \sim N(0, \sigma^2(y_j))$ represents independent, occupation-specific transitory shocks. $e(y_j)$ can be written as $\sigma(y_j)e_{jt}$ where $\sigma(y) = a_0 + a_1' y$ and e_{jt} follows standard normal distribution. For estimation, $\pi(y_j)$ and $q(y_j, s_t)$ are parameterized as

$$\begin{aligned} \pi(y_j) &= \pi_0 + \pi_1' y_j \\ q(y_j, s_t) &= q_0' s_t + y_j' Q_1 s_t \end{aligned} \tag{6}$$

where π_0 is a scalar, π_1 is a K dimensional vector, q_0 is a K dimensional vector, and Q_1 is a $K \times K$ diagonal matrix. Because occupational risk is a linear function of tasks, $\pi(y_j)$ absorbs the value of occupation-specific output prices and compensating differentials for occupational risks. $q(\cdot)$ represents the matching quality between task intensities and skill endowments. Finally, earnings equation (1) can be written as

$$\ln w(j; s_t, e_{jt}) = \pi_0 + \pi_1' y_j + q_0' s_t + y_j' Q_1 s_t + \sigma(y_j) e_{jt} \tag{7}$$

Wage parameters π_0, π_1, q_0, Q_1, and $\sigma(y_j)$ are known to workers, but the realization of shocks is unknown when they make an occupation decision.

Following the selection of occupations, the duration of work (Wks_{jt}) is stochastically determined through a random distribution. This stochasticity mirrors the reality that workers may experience periods of unemployment within a year. In other words, even if individuals are observed to have occupations at the moment of interview, the possibility of job loss or failure to secure new employment remains present. Similar to wage risks, the risk of employment is specific to each occupation, where the average duration of work weeks is contingent upon task intensity. Moreover, the unemployment duration also depends on individual skills. We assume that wks_{jt} adheres to a Tobit model, truncated at both lower and upper bounds, as follows.

$$\mathrm{wks}^*_{jt} \sim N(\mu_j(s_t), \sigma^2_{\mathrm{wks}}) \quad \text{with } \mu_j(s) = \delta_0 + \delta_1' s_t + \delta_2' y_j$$

$$\mathrm{wks}_{jt} = \begin{cases} \mathrm{wks}^*_{jt} & \text{if} \quad 0 < \mathrm{wks}^*_{jt} < 52 \\ 0 & \text{if} \quad \mathrm{wks}^*_{jt} \leq 0 \\ 52 & \text{if} \quad \mathrm{wks}^*_{jt} \geq 52 \end{cases} \tag{8}$$

During the unemployment spell, workers receive unemployment benefits, proportional to their skills and task intensity.

$$\ln b(j; s_t) = b_0 + b_1' s_t + b_2' y_j \tag{9}$$

The hourly earnings are determined by annual average between wages and unemployment benefits. In other words, workers in occupation j working wks_{jt} obtain the following earnings.

$$E(j; s_t, e_{jt}, \mathrm{wks}_{jt}) = \frac{\mathrm{wks}_{jt}}{52} w(s_t, j, e_{jt}) + \left(1 - \frac{\mathrm{wks}_{jt}}{52}\right) b(s_t, j) \tag{10}$$

Finally, workers are assumed to obtain utility from earnings that allows for non-homothetic preferences defined by the following ordinary differential equation (ODE)[5]:

$$-\frac{u''(E_t;\gamma,\gamma_2)E_t}{u'(E_t;\gamma,\gamma_2)} = \gamma E_t^{-\gamma_2} \tag{11}$$

where γ governs the overall relative risk aversion while γ_2 determines how sensitive the coefficient of relative risk aversion is to earnings. If $\gamma_2 = 0$, the utility takes the form of constant relative risk aversion (CRRA). If $\gamma_2 > 0$, then the relative risk aversion decreases with earnings. Individuals are heterogeneous in risk preferences through varying γ, while γ_2 is common across individuals.

4.2 Human Capital Accumulation

Every period workers accumulate skills based on their current level of skills and current occupations. Define the accumulation technology as follows

$$s_{t+1}(j;x,s_t,\eta_{jt}) = d_0 + D_1 s_t + D_2 y_j + D_3 x + \eta(y_j,\eta_{jt}) \tag{12}$$

where $\eta_t(y) \sim N(0, S(y))$ is a vector of persistent shocks whose distribution varies across occupations. Let $\zeta(y)$ denote the vector of standard errors for $\eta_t(y)$. $\zeta(y)$ is parameterized as a linear function of y as $c_0 + C_1'y$. Shocks of each skill dimension are independent of the other skill dimensions. d_0 is a K-dimensional vector, representing the general accumulation of skills. D_1 is a $K \times K$ diagonal matrix. If all elements D_1 are less than one, it represents that skills could be depreciating over time without working. D_2 is a $K \times K$ diagonal matrix implying learning by doing. In other words, skills are accumulated in proportion to the corresponding task intensities of their current occupations. x is a L dimensional vector of demographic information, and D_3 is $K \times L$ matrix which refers to learning heterogeneity across workers.

[5] There is no closed form solution of equation (11), so it will be numerically solved in the estimation process. This approach is also adopted in Meeuwis (2020).

Initial skill endowments are defined as a function of demographic information.

$$s_1 = h_0 + H_1(x, \gamma_i)' + \xi \tag{13}$$

where h_0 is a K dimensional vector, H_1 is a $K \times (L+1)$ matrix. $\xi \sim N(0, S_0)$ is a K dimensional vector representing unobserved individual heterogeneity. Initial skills are standardized for identification so that the unconditional means of initial skills are zeros and the unconditional variances are set to be one. In other words, $h_0 = -H_1 E(d)$ and the diagonal elements of $Var(\xi)$ are equal to diagonal elements of $I - H_1 Var(d) H_1'$. The covariances of unobserved initial skills are assumed to be zero.

4.3 Non-pecuniary Utility over Tasks

Workers obtain non-pecuniary value from working in an occupation. This captures the possibility that workers with the same skill levels can systematically choose different occupations in reality. Ignoring demographic differences in occupation preferences might lead to biased gaps in skill endowments across worker groups.

Non-pecuniary preference consists of two components: (dis)utility from working in intensive jobs and mobility costs.

$$C(j; s_t, x, j_{t-1}) = (f_0 + F_1 x + F_2 s_t)' y_j - (y_{j_{t-1}} - y_j)' F_3 (y_{j_{t-1}} - y_j) \tag{14}$$

f_0 is a K dimensional vector. F_1 is a $K \times J$ matrix implying systematic differences in job preferences. F_2 is a $K \times K$ diagonal matrix, representing the (dis)utility of intensive tasks dependent on their skills: if an individual is highly skilled in cognitive tasks, they might have lower utility costs of working in cognitive-intensive occupations. F_3 is a $K \times K$ diagonal matrix that represents mobility costs. Moving cost is proportional to how different the new occupation is from the last period's occupation. The difference between the two occupations is determined by the distance in their task intensity vectors.

Before fully transitioning to the labor market, workers may have formed their pre-market careers such as vocational training, part-time work, or internship. This pre-market experience enables workers to get used to certain types of tasks and affects their initial occupation choices through moving costs. Therefore, I assume workers start their careers with initial task intensity as a function of demographic variables as follows.

$$y_0 = \overline{y}_0 + Y_1 x \tag{15}$$

4.4 Occupation Choice Problem

At the beginning of each period, a worker chooses an occupation to maximize their lifetime value. The flow payoff includes the expected utility from earnings, The following equation represents the recursive form of the value function.

$$\begin{aligned} &V_t(\epsilon_t, s_t, \gamma, x, j_{t-1}) = \\ &\max_{j\in\{1,\cdots,J\}} \left\{ E_{e,\text{wks}}\left[u(E(j; s_t, e, \text{wks}); \gamma, \gamma_2)\right] + C(j; s_t, x, j_{t-1}) + \epsilon_{jt} + \beta E_{\epsilon,s}[V_{t+1}(\epsilon, x, s, j)|\gamma, x, s_t, j] \right\} \end{aligned} \tag{16}$$

subject to

$$\begin{aligned} &E(j; s_t, e, \text{wks}) = \frac{\text{wks}}{52} w(j; s_t, e) + \left(1 - \frac{\text{wks}}{52}\right) b(j; s_t) \\ &\ln w(j; s_t, e) = \pi_0 + \pi_1' y_j + q_0' s_t + y_j' Q_1 s_t + e(y_j, e) \\ &\ln b(j; s_t) = b_0 + b_1' s_t + b_2' y_j \\ &s_{t+1}(j; x, s_t, \eta) = d_0 + D_1 s_t + D_2 y_j + D_3 x + \eta(y_j, \eta) \\ &s_1 = h_0 + H_1 x + \xi \\ &C(j; s_t, x, j_{t-1}) = (f_0 + F_1 x + F_2 s_t)' y_j - (y_{j_{t-1}} - y_j)' F_3 (y_{j_{t-1}} - y_j) \\ &y_0 = \overline{y}_0 + Y_1 x \end{aligned} \tag{17}$$

where $\epsilon_t = (\epsilon_{1t}, \cdots, \epsilon_{Jt})$ refers to idiosyncratic preference shocks with Type I Extreme Value distribution.

It is noteworthy to discuss how risk aversion affects workers' lifetime utility in the model. The direct channel through which risk aversion relates to lifetime value is occupation-specific risks and occupation choices. As pointed out above, different occupations have different levels of variance of shocks $(e(y, e), \eta(y, \eta))$, depending on their task intensity, whether they are transitory or persistent. Because workers choose their occupations before they realize the shocks, risk-averse workers would prefer those with lower variance conditional on expected earnings.

This selection behavior on risk aversion can have an indirect influence on future choices due to skill accumulation. In the model, each type of skill is accumulated based on the task intensity to which workers are exposed, which was defined as learning by doing. In other words, if D_2 is positive definite, it means workers who performed more cognitive (routine) tasks should accumulate more of the corresponding skills. By selecting occupations with varied task intensity, risk-averse workers begin to accumulate different sets of skills. Considering that each skill is rewarded more in occupations with higher intensity of corresponding tasks, they are likely to be locked in a career specialized in a certain type of tasks. If different skills and tasks are valued differently in the market, then this specialization leads to diverging patterns of earnings dynamics.

5 Estimation

To quantify the contribution of different channels to the earnings gap, I estimate the model parameters using the NLSY97 and the ONET database. I first propose arguments about the identifiability of model parameters in the following section. Then, I describe an estimation algorithm that helps overcome several complexities residing in the dynamic programming problem with unobserved heterogeneity. I also illustrate additional restrictions on the estimation sample at the end of the section.

5.1 Identification Argument

The identification of parameters is not straightforward because of unobserved heterogeneity in risk preferences and skills. I exploit the panel structure of earnings and dynamic discrete choices as well as one-shot lottery choices to identify the model parameters.

First of all, since skills do not have natural scales and levels, I standardize the initial skill endowments so that unconditional mean and variance are 0 and 1 respectively. Then, the identification of wage and skill parameters comes from the earnings dynamics across individuals. If a certain type of individuals have higher levels of initial earnings conditional on the other characteristics, they would have higher initial skill endowments. For instance, men may have higher earnings in more routine occupations, implying more initial routine skills. If some have faster growth of earnings given employment histories, it speaks to faster growth of skills. If men's earnings growth is faster in routine occupations, it hints at the faster growth of routine skills for men. The same argument can be applied to learning by doing: conditional on employment histories up to two periods ago if a worker having performed more routine tasks last period achieves larger growth of earnings in routine occupations today, it implies routine skills are accumulated faster in routine occupations. Finally, return to skills is identified from the conditional covariance of earnings between the first period and the others. Given that initial skills are normalized, the higher return to skills implies the higher covariance of initial and the other earnings. All the other parameters including risk parameters are also identified from conditional mean and covariance of earnings.

Second, risk preference parameters and the distribution of risk preference types are identified from hypothetical gamble choices and the discrete choice of occupations.[6] Given the observed earnings, individuals' utility over lotteries is solely determined by risk preference parameters. Conditional on earnings, the share of lottery choice groups within observed demographic variables identifies overall risk aversion and baseline distribution of types. The association between lottery group shares and earnings helps the identification of the sensitivity of risk aversion to earnings, and the variation of shares across demographic groups can be used to identify the heterogeneous distribution of risk preference types. Given the identification of wage and skill parameters from observed earnings profiles, the structure of occupation-specific risks is fixed conditional on state variables. Then, the share of occupation choices helps the identification of risk preference parameters and distributions in the same way as lottery choices.

Third, any systematic variation of occupation choices conditional on earnings and occupation histories provides information about non-pecuniary preference over task intensities. For example, if men on average choose more routine occupations relative to women with all characteristics but gender identical, it speaks to the fact that they prefer routine tasks compared to their counterparts. Similarly, initial occupation choices help the identification of initial task inclination.

5.2 Expectation-maximization (EM) Algorithm

Given that the model parameters can be identified, I exploit the EM algorithm to estimate the model, following the method in Arcidiacono and Miller (2011). The main goal of this approach is to overcome challenges in estimating dynamic programming problems with state-dependent unobserved heterogeneity. First, fully solving the problem is computationally costly for a couple of reasons. The utility function over earnings solves the second-order differential equation 11, which does not have a closed-form solution. This indicates that there is no analytical solution to the worker's occupation choice problem. Moreover, there are multiple state variables including last occupations and skills. Since there are many available occupations and skills are continuous, the state space becomes extensive, making the estimation with a fixed point algorithm intractable. To circumvent the necessity to fully solve the model, Hotz and Miller (1993) propose CCP estimators that utilize the relationship between value functions and the probabilities of choosing alternatives.

Although CCP estimators can mitigate the computational costs, their direct application is infeasible because of the second challenge, unobserved heterogeneity. One requirement to implement CCP estimators is the estimated choice probabilities for all possible states, which are not

[6] Although I parametrize the unobserved heterogeneity of risk preference types, the non-parametric identifiability of finite mixture models with dynamic discrete choices has been proven by Kasahara and Shimotsu (2009). Especially, they show that with moderate panel periods ($T \geq 6$), the mixture model of the first-order Markov property is identified.

available in the presence of unobserved state variables. Arcidiacono and Miller (2011) suggest the EM algorithm that iteratively updates the CCP's and estimates the parameter with the Maximum Likelihood Estimation given the updated CCP's. The key insight of their arguments holds in this article although the model departs from their setting in the sense that unobserved heterogeneity transition also depends on choices and other state variables. In the expectation stage, the CCP's are updated. The maximization stage estimates the model parameters given the estimated CCP's from the expectation stage. The following subsections provide greater details of the $(m+1)$th iteration given the mth estimates of CCP's and parameters.

5.2.1 Expectation: Updating Conditional Choice Probability

Given the mth CCP's and parameter estimates as $(\hat{P}^{(m)}, \hat{\theta}^{(m)})$, the update of CCP's directly follows the matrix inversion method of Hotz and Miller (1993). Define the flow payoff of individual i in occupation j as $U_j(s_{it}, x_i, \gamma_i, j_{i,t-1})$. The conditional value function of choosing occupation j can be written as

$$v_j(s_{it}, \gamma_i, x_i, j_{i,t-1}) = U_j(s_{it}, \gamma_i, x_i, j_{i,t-1}) + \beta \int \overline{V}(s', \gamma_i, x_i, j) f(s'|x_i, s_{it}, j) ds' \tag{18}$$

where $\overline{V}(\cdot) = E_\epsilon[V(\epsilon_t, \cdot)]$. With Type 1 Extreme Value distribution of preference shock, the conditional choice probability of occupation j becomes

$$P_j(s_{it}, \gamma_i, x_i, j_{i,t-1}) = \frac{\exp(v_j(s_{it}, \gamma_i, x_i, j_{i,t-1}))}{\sum_{j'} \exp(v_{j'}(s_{it}, \gamma_i, x_i, j_{i,t-1}))} \tag{19}$$

In other words, if one can calculate $\overline{V}$, they can derive the choice probability for each state.

Letting $\overline{V}$ a vector of expected values for each pair of state variables, P_j and U_j vectors of conditional choice probability and the flow payoff for occupation j, and F_j a matrix of skill transition in occupation j, Hotz and Miller (1993) propose that $\overline{V}$ can be reduced into the following form

$$\overline{V} = \left(I - \beta \sum_j (P_j \vec{\mathbf{1}}') \odot F_j \right)^{-1} \left(\sum_j P_j \odot (U_j + \epsilon_j^*) \right) \tag{20}$$

where $\vec{\mathbf{1}}$ is a vector of ones and $\odot$ represents element-wise multiplication. ϵ_j^* indicates the conditional expectation of error terms whose elements have a closed form as $C_{euler} - \ln P_j(\cdot)$.

Given $\hat{P}^{(m)}$, the expected value function can be calculated from equation (20) and thus the optimal policy of occupation choices can be constructed using equation (18) and (19): for all j,

$$\hat{P}_j^{(m+1)}(s_{it}, \gamma_i, x_i, j_{i,t-1}; \hat{P}^{(m)}, \hat{\theta}^m) = \frac{\exp(v_j(s_{it}, \gamma_i, x_i, j_{i,t-1}; \hat{P}^{(m)}, \hat{\theta}^m))}{\sum_{j'} \exp(v_{j'}(s_{it}, \gamma_i, x_i, j_{i,t-1}; \hat{P}^m, \hat{\theta}^m))} \tag{21}$$

5.2.2 Maximization: Maximum Likelihood Estimation

Given the estimates of CCP's, $\hat{P}^{(m+1)}$, the model parameter can be estimated using the maximum likelihood estimation. Before providing a maximization problem, I first accommodate the unobserved risk preference type into the model and define the likelihood function.

Risk Preference Type Risk preferences are heterogeneous across individuals through varying γ_1 in equation 11. However, as a researcher, I do not directly observe γ_1. For simplicity, I assume two types of risk preference: high risk-averse (γ_H) and low risk-averse (γ_L). I also parametrize the distribution of the risk aversion coefficient as a function of demographics, i.e. $P(\gamma_H|x_i) = \rho_0 + x_i'\rho_1$.

Gamble Choice I first define the likelihood function of the hypothetical gamble choices. Individuals answered the gamble experiment over two rounds given their risk preference types and earnings at the moment. As described in section 2, they compared the safe offer, which gives the current level of earnings forever, with the offer with moderate risk, which gives twice or two-thirds of earnings with half probability. In the second round, they weighed the safe offer against another offer whose risks depended on the first round's answer. If they chose the safe one, the second risky offer would provide twice or four-fifths of earnings with half probability; if they chose the risky one, the new offer would provide twice or half of earnings with half probability. Formally, given their current family income W and the utility function on earnings $u(\cdot;\gamma_1,\gamma_2)$, the utility from each offer becomes as follows:

$$\begin{aligned}
U_{safe}(W;\gamma_1,\gamma_2) &= u(W;\gamma_1,\gamma_2)\\
U_{risky}(W;\gamma_1,\gamma_2) &= \Big(u(2W;\gamma_1,\gamma_2) + 2W/3;\gamma_1,\gamma_2)\Big)/2\\
U_{less}(W;\gamma_1,\gamma_2) &= \Big(u(2W;\gamma_1,\gamma_2) + 4W/5;\gamma_1,\gamma_2)\Big)/2\\
U_{risky}(W;\gamma_1,\gamma_2) &= \Big(u(2W;\gamma_1,\gamma_2) + W/2;\gamma_1,\gamma_2)\Big)/2
\end{aligned} \tag{22}$$

Also, the comparison of utility for the first and the second round can be defined as:

$$\begin{aligned}
U_1^*(W;\gamma_1,\gamma_2) &= U_{safe}(W;\gamma_1,\gamma_2) - U_{risky}(W;\gamma_1,\gamma_2)\\
U_2^*(W;\gamma_1,\gamma_2|g_1) &= \begin{cases} U_{safe}(W;\gamma_1,\gamma_2) - U_{more}(W;\gamma_1,\gamma_2) & \text{if } g_1 = 0\\ U_{safe}(W;\gamma_1,\gamma_2) - U_{less}(W;\gamma_1,\gamma_2) & \text{if } g_1 = 1\end{cases}
\end{aligned} \tag{23}$$

The gamble was asked through the survey without any real incentive, so individuals may behave differently from their actual choices. Therefore, I assume the lottery choice involves idiosyncratic noises in utility comparison. Since the utility does not have a natural scale and level, I normalize the noise distribution to follow a normal distribution with mean 0 and standard error 0.1. With the noise at round l given as $\epsilon_l^g \sim N(0, 0.1)$, the lottery choice problem at round l becomes determining whether $U_l^* + \epsilon_l^g > 0$. Define the choice variable as $g_l = 1$ if the inequality holds and vice versa.

The likelihood of gamble choice can be written as

$$P(g_{i1}, g_{i2}|W_i, \gamma_i, \gamma_2) = P(g_{i1}|W_i, \gamma_i, \gamma_2)P(g_{i2}|W_i, \gamma_i, \gamma_2, g_{i1})$$

$$
\begin{aligned}
=&\left(\Phi(-U_1^*(W;\gamma_i,\gamma_2)/0.1)^{1-g_{i1}}\Big(1-\Phi(-U_1^*(W;\gamma_i,\gamma_2)/0.1)\Big)^{g_{i1}}\right)\\
&\left(\Phi(-U_2^*(W;\gamma_i,\gamma_2|g_{i1})/0.1)^{1-g_{i2}}\Big(1-\Phi(-U_1^*(W;\gamma_i,\gamma_2|g_{i1})/0.1)\Big)^{g_{i2}}\right) \qquad (24)
\end{aligned}
$$

where Φ represents the cumulative density function (CDF) of standard normal distribution.

Labor Market Outcomes Let Z_i^t denote the observed history of occupation choices, wages, and working weeks for individual i up to period t. Moreover, $z_{it} = (j_{it}, w_{it}, wks_{it})$ represents the pair of occupation, log wage, and working weeks for individual i at period t.

To calculate the likelihood of observed labor market histories, I use the non-Gaussian State Space approach by Kitagawa (1987). The likelihood of an individual worker's observed profile can be written as

$$
L(Z_i|\gamma_i, x_i;\theta,\hat{P}) = L(z_{i1}|\gamma_i, x_i;\theta,\hat{P})\prod_{t=2}^{T_i} L(z_{it}|Z_i^{t-1},\gamma_i,x_i;\theta,\hat{P}) \qquad (25)
$$

In the first period, the probability of observing a particular job and wage conditional on predetermined characteristics can be derived by summing the product of the probability of initial skills and the likelihood of observed pair of occupation, wage, and working weeks, over all possible unobserved skill levels s. I can disentangle the likelihood of (j_{i1}, w_{i1}) as the product of the probability of j_{i1} and the conditional probability of w_{i1} given j_{i1}. For the expositional purpose, I subsume θ and $\hat{P}$ from the likelihood function.

$$
\begin{aligned}
l(z_{i1}|\gamma_i,x_i) &= \sum_s \pi_1(s|\gamma_i,x_i)P(j_{i1},w_{i1},wks_{it}|s,\gamma_i,x_i)\\
&= \sum_s \pi_1(s|\gamma_i,x_i)P(j_{i1}|s,\gamma_i,x_i)P(w_{i1}|s,j_{i1})P(wks_{i1}|s,j_{i1}) \qquad (26)
\end{aligned}
$$

The likelihood of observing j_{i1} is given from the optimal policy, which can be calculated with equation (21) and $\hat{P}$. Conditional on skills and occupations, the wage follows the normal distribution with conditional mean and variance as follows:

$$\begin{aligned} E[w_{i1}|s,j_{i1}] &= \pi_0 + \pi_1' y_{j_{i1}} + (q_0 + Q_1 y_{j_{i1}})' s \\ Var[w_{i1}|s,j_{i1}] &= \sigma^2(y_{j_{i1}}) \end{aligned} \tag{27}$$

Finally, the number of weeks worked follows the Tobit distribution as in equation (8).

I update the skill distribution $\pi_1(s|x_i, j_{i1})$ utilizing Bayes' rule, incorporating observed occupation choices and wages:

$$\tilde{\pi}_1(s|\gamma_i, x_i, Z_i^1) = \frac{\pi_1(s|\gamma_i, x_i) p(j_{i1}|s, \gamma_i, x_i) \phi\left(\frac{w_{i1} - E[w_{i1}|s, j_{i1}]}{\sigma(y_{j_{i1}})}\middle| s, j_{i1}\right) P(wks_{i1}|s, j_{i1})}{l(y_{i1}|\gamma_i, x_i)} \tag{28}$$

Finally, skills transition with the transition rule, f. Note that s' has the discrete distribution approximating the normal distribution conditional on (s, x_i, j_{i1}).

$$\pi_2(s'|\gamma_i, x_i, Z_i^1) = \sum_s f(s'|s, x_i, j_{i1}) \tilde{\pi}_1(s|\gamma_i, x_i, Z_i^1) \tag{29}$$

This equation describes the evolution of the skill distribution from period 1 to period 2, incorporating transition functions and the updated skill distribution in the current period.

Expanding this framework to subsequent periods involves a similar sequential process. The probability of observing z_{it} given past information Z_i^{t-1} is determined as:

$$\begin{aligned} &L(z_{it}|\gamma_i, x_i, Z_i^{t-1}) \\ &= \sum_s \pi_t(s|\gamma_i, x_i, Z_i^{t-1}) p(j_{it}|s, \gamma_i, x_i, j_{it-1}) \phi\left(\frac{w_{it} - E[w_{it}|s, j_{it}]}{\sigma(y_{j_{it}})}\middle| s, j_{it}\right) P(wks_{it}|s, j_{it}) \end{aligned} \tag{30}$$

The skill distribution is continually updated using Bayes' rule, adapting to the observed job types and past information:

$$\tilde{\pi}_t(s|\gamma_i, x_i, Z_i^t) = \frac{\pi(s|\gamma_i, x_i, Z_i^{t-1}) p(j_{it}|s, \gamma_i, x_i, j_{it-1}) \phi\left(\frac{w_{it} - E[w_{it}|s, j_{it}]}{\sigma(y_{j_{it}})}\middle| s, j_{it}\right) P(wks_{it}|s, j_{it})}{L(j_{it}, w_{it}|\gamma_i, x_i, Z_i^{t-1})} \tag{31}$$

Finally, the iterative process of skill accumulation and update across subsequent periods continues as:

$$\pi_{t+1}(s'|\gamma_i, x_i, Z_i^t) = \sum_s f(s'|s, x_i, j_{it}) \tilde{\pi}_t(s|\gamma_i, x_i, Z_i^t) \tag{32}$$

Maximization Problem Denoting individual profiles of wages, occupations, and weeks worked as $Z_i = \{z_{it}\}_t$, individual contribution to likelihood can be written as

$$L(g_i, Z_i|x_i; \theta, \hat{P}) = \sum_{\gamma \in \{\gamma_H, \gamma_L\}} P(\gamma|x_i; \theta) L(Z_i|\gamma, x_i; \theta, \hat{P}) P(g = g_i|\gamma, x_i, Z_i; \theta) \tag{33}$$

Finally, the maximization likelihood estimation can be formulated as follows:

$$\hat{\theta}^{(m+1)} = \arg\max_\theta \sum_i \log L(g_i, Z_i|x_i; \theta, \hat{P}^{(m+1)} \tag{34}$$

The algorithm proceeds until $\hat{\theta}^m$ converges.

Unobserved skills are continuous variables and must be discretized to estimate the model. Although the more discrete points help better approximation of estimating the model with continuous variables, it also becomes infeasible to estimate the model as each realization of likelihood during estimation still requires calculating the value function with matrix inversion. Larger state space indicates larger matrices to be inversed, making the calculation heavier. I address the problem of the large state space by using value function approximation as suggested by Keane and Wolpin (1994). In particular, I calculate the value function with 5 discrete points for each skill dimension (i.e. 25 points in two dimensions) and approximate the value function with 10 points using regression on the second-order polynomials of skills for each iteration of the likelihood function.

The estimation still calls for a couple of restrictions on the estimation sample to proceed. First, the sample is restricted to those who had positive earnings at the moment of job lottery choices. This is to ensure that job lottery choices are made based primarily on the individual's risk preferences and economic conditions, not on the components outside the model such as parental income. Second, I merge some occupations into the closest ones, determined by broader occupation classifications and task intensities. This reduces the number of 3-digit SOC occupations from 81 to 55. The mapping across occupations is provided in Table A3. Third, I rescale hourly earnings to have 10 dollars as the unit. The choice of 10 dollars is based on the standard deviation of the hourly rate varying between 9.3 for the high school sample and 13 for the college sample.

Lastly, I estimate the model with high school graduates and college graduates separately because the choice set for occupations is likely to differ across education levels. For example, professional occupations such as scientists or physicians require some qualifications that are directly connected to college degrees. Workers may determine educational attainment based on both their risk preferences and the availability of occupations in the future, but studying educational decisions is beyond the scope of this paper. Therefore, I suppose that education level is predetermined, and I distinguish the sample into two groups when estimating the model.

6 Estimation Results

In this section, I present the preliminary estimation results, based on several assumptions to simplify the process. First, I adopt the constant relative risk aversion (CRRA) model for utility over earnings. Second, I determine individual risk aversion coefficients using a framework proposed by Kimball et al. (2008), employing hypothetical lottery choices. These choices are formulated under the CRRA assumption, allowing for the derivation of inequality conditions for risk aversion coefficients. Utilizing maximum likelihood estimation, I estimate the conditional expected values of relative risk aversion coefficients for each risk attitude group, assuming log normality of individual coefficients. These coefficients are then assigned to individuals based on their risk attitude groups. Finally, I divide the parameters into two groups: wage-skill parameters and others. The former encompasses parameters related to wages, initial skills, skill accumulation, and shock distribution, while the latter includes parameters for initial tasks and non-pecuniary preferences. The estimation of wage-skill parameters utilizes observed earnings profiles assuming occupation choices as given, followed by estimating the remaining parameters using the EM algorithm as detailed in the previous section.

The parameter estimates for the college graduate sample are presented in Tables 6 and 7. Notably, there is significant variation in earnings risks across occupations. Occupations with more intensive tasks, irrespective of task types, exhibit lower variance in transitory shocks. Routine tasks show a slightly stronger correlation with transitory risks. Cognitive skill shock variance decreases with task intensities, particularly cognitive tasks, while routine skill shock variance increases with them, especially with routine tasks. Although routine tasks may not effectively reduce persistent risks, their influence on occupation choices might be lesser relative to transitory risks because they interact with the skill prices and discount factor. Therefore, the stronger correlation of routine tasks with transitory risks can aid risk-averse workers in choosing routine careers to mitigate risks.

Furthermore, risk preferences exhibit a significant direct relationship with initial skill heterogeneity. Compared to the most risk-tolerant group, the most risk-averse workers tend to have approximately 0.1 standard deviations lower cognitive skills and around 0.21 standard deviations lower routine skills. This difference between the extreme risk attitude groups exceeds that across genders, with men showing 0.03 standard deviations higher cognitive skills and 0.11 standard deviations higher routine skills at the onset of their careers. Moreover, the negative correlation between routine skills and risk aversion suggests that the sorting of risk-averse workers into more routine occupations is unlikely driven by unobserved heterogeneity in skills.

In Figure 3, the average profiles of log earnings and task intensities from the data are compared with the predicted profiles from the estimated model. While the model's prediction aligns with the wage profiles up to the first 9 years, it tends to overpredict earnings thereafter. This discrepancy arises from the model's assumption of an infinite horizon and its lack of aging components to fully capture the concavity of earnings. Additionally, the model predicts higher cognitive task intensities and lower routine task intensities compared to the data, although it successfully replicates the increasing trend of cognitive task intensities and the decreasing trend of routine task intensities over individuals' lives.

7 Decomposition of Risk Preference Effect on Earnings

The observed earnings gap across different risk attitude groups in our data may be influenced by several factors, including worker attributes correlated with risk preferences, immediate returns to occupation choice, and human capital accumulation. In this section, I employ the estimated model to disentangle the earnings gap between workers with varying risk preferences.

The decomposition process involves three steps. First, I replicate the earnings gap across risk attitude groups using the estimated model by simulating earnings and occupation choice profiles for a sample size five times larger than the actual dataset. To dissect the impact attributed to correlations with worker characteristics and occupation choices, I hold occupation choices constant at the baseline and simulate earnings trajectories for the same individuals, assuming no systematic heterogeneity in initial conditions, including demographic variables and initial skills. Specifically, I assume uniform demographic variables across workers, and set H_1 to zero to eliminate systematic differences in initial skills. Finally, I nullify differential human capital accumulation to isolate the effect stemming from the dynamic interaction between risk preferences and occupation choices from instantaneous returns. This involves setting each of d_0, D_1, D_2, and D_3 to their average values across skill dimensions, thus removing differences in accumulation between cognitive and routine skills. Additionally, I set D_2 to interact with the average level of task intensities rather than actual task intensities, eradicating differential learning-by-doing across occupations and skills. Using these simulated datasets, I compare the regression estimates of log lifetime average earnings on risk attitude group indicators conditional on race, gender, and a dummy variable indicating whether the AFQT score is above the mean.

The results are presented in Table 8. In the first column, I reproduce the earnings gap across risk attitude groups using the actual sample employed in model estimation. Simplifying the analysis by focusing on the two extreme groups, the most risk-tolerant workers, on average, exhibit 16 percent higher lifetime earnings compared to the most risk-averse workers. In the baseline model, this gap is reduced to 13 percent. Moving to the third column, I estimate the same specification using simulated data without any initial condition heterogeneity. Here, the remaining 9.6 percent, or approximately 74 percent of the earnings gap, reflects the effect of risk preferences through occupation choices, encompassing both concurrent differences in returns to occupation and the divergence in skill development through on-the-job learning.

The last column presents estimates without differential human capital accumulation. Here, the earnings gap signifies the differential returns stemming solely from selecting different occupations. Specifically, the most risk-tolerant workers earn 5.5 percent higher earnings due to their occupation choices paying more, irrespective of their skills, which also includes the risk premium. Conversely, approximately 30 percent of the baseline earnings gap is attributed to risk-averse workers accumulating different skill sets compared to risk-tolerant workers.

In summary, the decomposition analysis reveals that the effects of risk preferences on earnings, through their interaction with worker characteristics and skill accumulation, are comparable to the effects stemming from differential returns to occupation choices. Ignoring unobserved heterogeneity and its evolution may lead to underestimation of the effects of risk preferences on labor market outcomes, or an overestimation of the occupational risk premium.

8 Risk Aversion and Routine-biased Technological Changes

The decomposition analysis underscores the pivotal role of occupation choices in understanding earnings disparities among workers with varying risk preferences. This section delves into using model estimates to assess the disproportionate impact of changes in occupational structure induced by routine-biased technological changes (RBTCs). As per Cortes (2016), recent technological advancements have led to an increased wage premium for cognitive occupations and a decreased premium for routine occupations. Given that risk-averse workers typically opt for more routine occupations to mitigate higher risks, these technological shifts can have a distributional impact across workers with different risk preferences along two fronts. First, since risk-averse workers are inclined to work in more routine occupations on average, they may bear greater costs from technological progress given their careers. Second, their distinct preferences in occupation choices can exacerbate these impacts. With a heightened return to cognitive tasks and a diminished return to routine tasks, the average worker would gravitate towards more cognitive and fewer routine careers. However, moving away from routine occupations implies exposure to higher earnings risks, deterring risk-averse workers from transitioning towards cognitive careers in response to RBTCs.

To investigate the disproportionate impact of RBTCs, I simulate data with an increased premium for cognitive tasks and a decreased premium for routine tasks by setting $\pi^{New}_{1,Cog} = \hat{\pi}_{1,Cog} + 0.1$ and $\pi^{New}_{1,Rou} = \hat{\pi}_{1,Rou} - 0.1$. In this counterfactual scenario, the most cognitive occupation ($y_{Cog} = 1$) commands a 10 percent higher return, with the impact diminishing with cognitive task intensity. Similarly, the return to the most routine occupation ($y_{Rou} = 1$) decreases by 10 percent, with less routine occupations experiencing a proportionally smaller impact based on their task intensities. Using the simulated data, I compare the gaps in the lifetime average earnings and task intensities between each risk attitude group and the most risk-averse group to the baseline values.

The results are detailed in Table 9. In the counterfactual scenario, the most risk-tolerant workers exhibit approximately 14 percent higher earnings compared to the most risk-averse, marking an 8 percent larger earnings gap compared to the baseline. This widened earnings gap coincides with stronger occupational segregation. The most risk-tolerant workers, on average, show a 15 percentile increase in cognitive task intensities in the counterfactual, approximately 10 percent higher than the baseline model. Simultaneously, the average routine task intensity for the most risk-tolerant workers diminishes even further relative to the most risk-averse workers compared to the baseline, changing from -0.122 to -0.132. This emphasizes that the widened earnings gap arises not only from reduced returns to baseline careers but also from less adjustment in careers among the more risk-averse workers.

9 Conclusion

This paper investigates the influence of risk preferences on occupation choices and earnings dynamics, recognizing the profound impact of individual risk attitudes on labor market outcomes in the face of pervasive risks inherent in labor market decisions. Occupations are one of the most important sources of earnings and are likely varied in their earnings risks. Self-selection based on risk preferences does affect earnings through on-the-job learning as well as the compensation for risks. This paper provides an empirical framework where risk preference can have a persistent influence on earnings through occupation choices with differential accumulation of skills.

I first provide descriptive evidence of the significant relationship between risk preferences, earnings dynamics, and occupation choices. Using the survey on the hypothetical job offer choices, I document that individuals with more risk-loving attitudes tend to have higher earnings with faster growth rates over their lives. This pattern is associated with the fact that more risk-averse people on average hold more routine less cognitive occupations, which can be characterized by lower returns and slower growth of earnings.

Building upon the suggestive evidence, I develop a dynamic occupation choice model that allows for heterogeneity in individual risk preferences. Beyond self-selection on risk preferences, an essential mechanism underpinning the effects of risk preferences lies in the interplay between skill accumulation across occupations and occupation-specific returns to skills. Once sorting into different careers, risk-averse workers begin to accrue different sets of skills and the impact of diverging skills propagates throughout career trajectories via dynamic selection. The model is estimated using the EM algorithm modified to accommodate unobserved heterogeneity and its dynamics.

In the decomposition analysis, the estimated model highlights the importance of considering both concurrent compensation to riskier occupations and the dynamic interplay between risk preferences, occupation choices, and unobserved skills. The structural model enables me to examine the distributional implications of routine-biased technological changes on workers with varied risk preferences. More risk-averse workers experience more adverse impacts on their labor market outcomes due to technological changes, as their risk aversion impedes flexible behavioral adjustments compared to their risk-tolerant counterparts.

The empirical framework presented in this paper offers comprehensive insights into the impact of risk preferences on labor market outcomes, while also providing a behavioral perspective on the influence of economic changes on the labor market. The strong association between risk preferences and labor market outcomes suggests that labor market risks are not entirely insurable, prompting workers to leverage their occupation choices as a form of insurance. Future research may extend the model outlined in this paper to incorporate other financial and labor market decisions, such as savings and marriage, and explore the relative importance of occupation choices as a risk mitigation strategy compared to other decision-making contexts.

References

Ahn, T. (2010). Attitudes toward risk and self-employment of young workers. *Labour Economics*, 17:434–442.

Arcidiacono, P. and Miller, R. A. (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, 79:1823–1867.

Barseghyan, L., Molinari, F., O'Donoghue, T., and Teitelbaum, J. C. (2018). Estimating risk preferences in the field. *Journal of Economic Literature*, 56:501–564.

Bonin, H., Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2007). Cross-sectional earnings risk and occupational sorting: The role of risk attitudes. *Labour Economics*, 14:926–937.

Buurman, M., Delfgaauw, J., Dur, R., and den Bossche, S. V. (2012). Public sector employees: Risk averse and altruistic? *Journal of Economic Behavior and Organization*, 83:279–291.

Cortes, G. M. (2016). Where have the middle-wage workers gone? a study of polarization using panel data. *Journal of Labor Economics*, 34:63–105.

DeLeire, T. and Levy, H. (2004). Worker sorting and the risk of death on the job. *Journal of Labor Economics*, 22:925–953.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *Quarterly Journal of Economics*, 132:1593–1640.

Dohmen, T. and Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *American Economic Review*, 101:556–590.

Fouarge, D., Kriechel, B., and Dohmen, T. (2014). Occupational sorting of school graduates: The role of economic preferences. *Journal of Economic Behavior and Organization*, 106:335–351.

Gathmann, C. and Schönberg, U. (2010). How general is human capital? a task-based approach. *Journal of Labor Economics*, 28:1–49.

Gould, E. D. (2002). Rising wage inequality, comparative advantage, and the growing im-portance of general skills in the united states. *Journal of Labor Economics*, 20.

Grazier, S. and Sloane, P. J. (2008). Accident risk, gender, family status and occupational choice in the uk. *Labour Economics*, 15:938–957.

Guvenen, F., Kuruscu, B., Tanaka, S., and Wiczer, D. (2020). Multidimensional skill mismatch. *American Economic Journal: Macroeconomics*, 12:210–244.

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60:497.

Kambourov, G. and Manovskii, I. (2009). Occupational specificity of human capital. *International Economic Review*, 50:63–115.

Kasahara, H. and Shimotsu, K. (2009). Nonparametric identification of finite mixture models of dynamic discrete choices. *Econometrica*, 77:135–175.

Keane, M. P. and Wolpin, K. I. (1994). The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte carlo evidence. *The Review of Economics and Statistics*, 76:648.

Keane, M. P. and Wolpin, K. I. (1997). The career decisions of young men. *Journal of Political Economy*, 105:473–522.

Kimball, M. S., Sahm, C. R., and Shapiro, M. D. (2008). Imputing risk tolerance from survey responses. *Journal of the American Statistical Association*, 103:1028–1038.

King, A. G. (1974). Occupational choice, risk aversion, and wealth. *Industrial and Labor Relations Review*, 27:586.

Kitagawa, G. (1987). Non-gaussian state—space modeling of nonstationary time series. *Journal of the American Statistical Association*, 82:1032–1041.

Lazear, E. P. (2009). Firm-specific human capital: A skill-weights approach. *Journal of Political Economy*, 117:914–940.

Lindenlaub, I. (2017). Sorting multidimensional types: Theory and application. *The Review of Economic Studies*, page rdw063.

Lise, J. and Postel-Vinay, F. (2020). Multidimensional skills, sorting, and human capital accumulation. *American Economic Review*, 110:2328–2376.

Meeuwis, M. (2020). Wealth fluctuations and risk preferences: Evidence from u.s. investor portfolios. *SSRN Electronic Journal*.

Patnaik, A., Venator, J., Wiswall, M., and Zafar, B. (2020). The role of heterogeneous risk preferences, discount rates, and earnings expectations in college major choice. *Journal of Econometrics*.

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

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3:135–146.

Sahm, C. R. (2012). How much does risk tolerance change? *Quarterly Journal of Finance*, 2.

Shaw, K. L. (1996). An empirical analysis of risk aversion and income growth. *Journal of Labor Economics*, 14:626–653.

Sullivan, P. (2010). A dynamic analysis of educational attainment, occupational choices, and job search. *International Economic Review*, 51:289–317.

Yamaguchi, S. (2010). The effect of match quality and specific experience on career decisions and wage growth. *Labour Economics*, 17:407–423.

Yamaguchi, S. (2012). Tasks and heterogeneous human capital. *Journal of Labor Economics*, 30.

Figures and Tables

Figure 1: Risk Preferences and Labor Market Outcomes

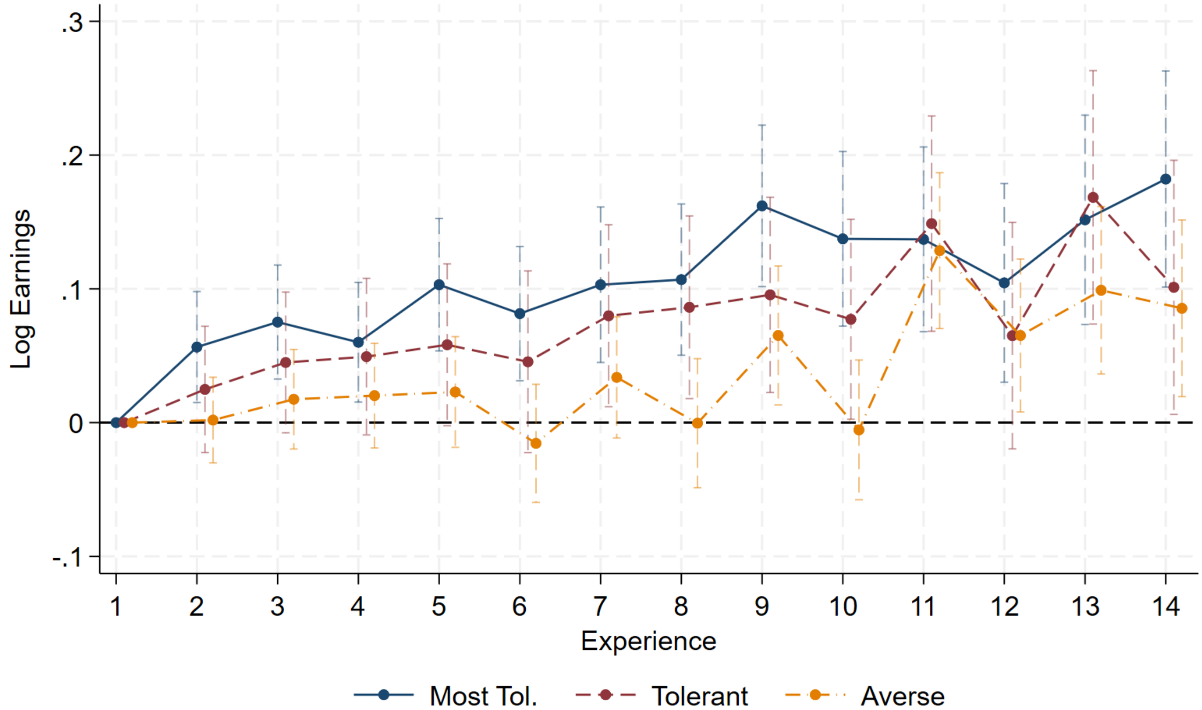


Figure 2: Risk Preferences and Relative Earnings Growth

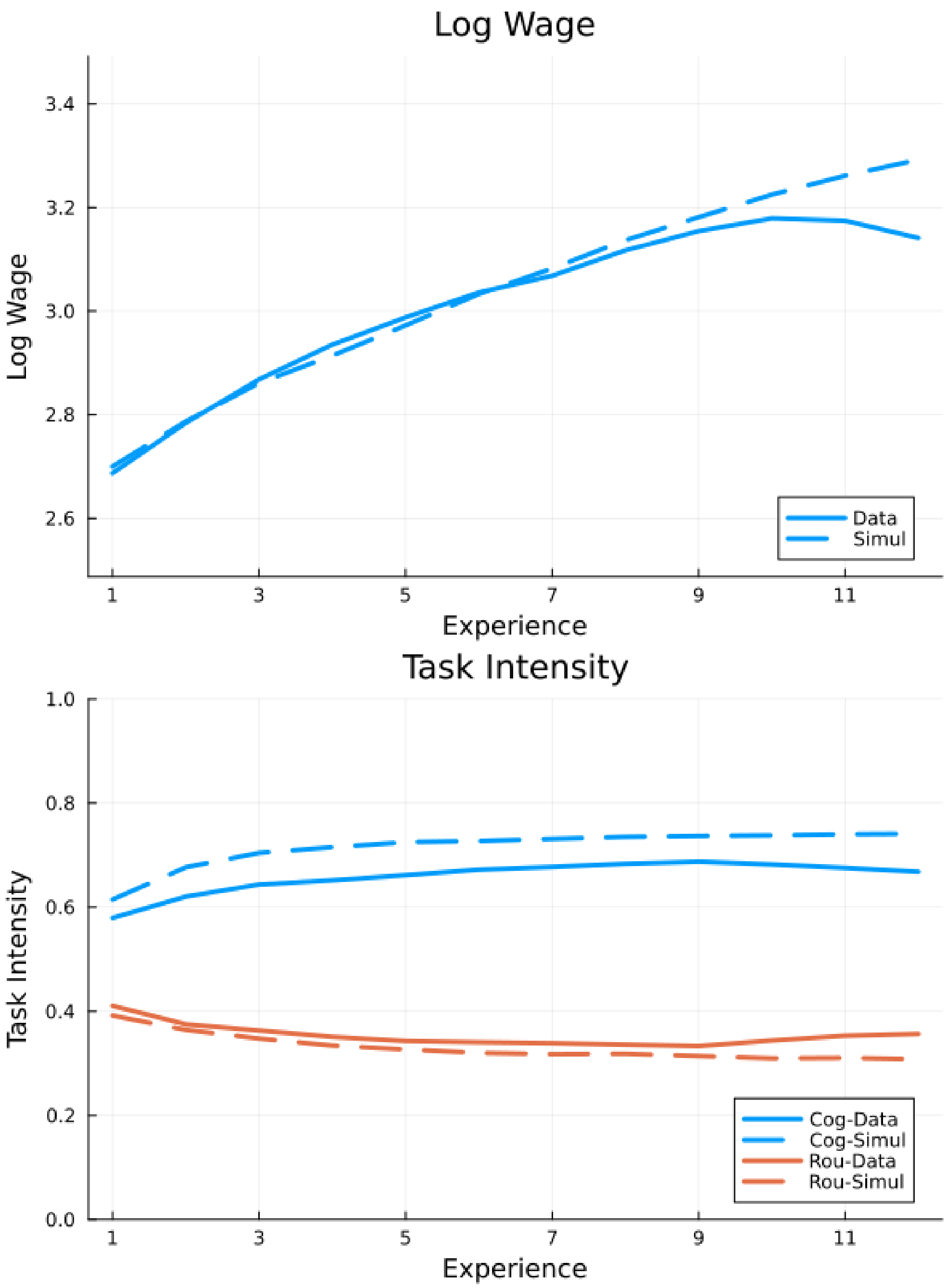


Figure 3: Model Fit

Table 1: Summary Statistics

	Risk Aversion Group				Total
g	1	2	3	4	
	Most Tol.	Tolerant	Averse	Most Aver.	
	Job Lottery				
Choice	Risky-Risky	Risky-Safe	Safe-Risky	Safe-Safe	
N_{indiv}	762	587	1,143	2,702	5,195
%	15	11	22	52	100
N_{obs}	8,121	6,057	11,978	27,975	54,131
	Demographics				
Black	0.214	0.227	0.234	0.291	0.26
Men	0.621	0.543	0.514	0.463	0.506
Education	13.91	14.08	13.9	13.19	13.55
AFQT	0.184	0.261	0.153	-0.115	0.031
	Labor Market Outcomes				
Log Wage	2.873	2.843	2.795	2.686	2.755
Weeks Worked	44.92	44.98	44.69	43.65	44.22
Cognitive	0.59	0.596	0.581	0.534	0.559
Routine	0.445	0.445	0.446	0.491	0.469

Source.–National Longitudinal Survey of Youth 1997.

Table 2: Risk Preference and Lifetime Earnings (Base: Most Averse)

	Log Average Earnings			Log SD of Log Earnings		
	(1)	(2)	(3)	(4)	(5)	(6)
Most Tolerant	0.107***	0.167***	0.166***	0.146***	0.161***	0.155***
	(0.015)	(0.031)	(0.031)	(0.027)	(0.052)	(0.051)
Tolerant	0.059***	0.074*	0.069*	0.111***	0.150**	0.147**
	(0.017)	(0.040)	(0.039)	(0.032)	(0.065)	(0.065)
Averse	0.033***	0.051*	0.049*	0.073***	0.071	0.069
	(0.013)	(0.027)	(0.027)	(0.024)	(0.048)	(0.048)
Demographics	✓	✓	✓	✓	✓	✓
Non-cognitive			✓			✓
Parent income 97			✓			✓
N	5,187	1,082	1,082	5,162	1,079	1,079
R^2	0.352	0.382	0.388	0.045	0.073	0.082

Source.–National Longitudinal Survey of Youth 1997

Notes.–The estimates are from the regression of lifetime average and volatility of earnings on risk attitude group indicators. Control variables include race, gender, age-adjusted AFQT scores and education level. The second column restricts the sample to individuals whose non-cognitive skill measures and parent income in 1997 are available. Non-cognitive skill measures include social skill measures and noncognitive skill measures constructed using Goldberg's Big Five personal factor survey. Robust standard errors in parenthesis.

Table 3: Earnings Risks and Task Intensity

	Earnings Risk (SD)		
	(1)	(2)	(3)
Cognitive (SD)	-0.411***		0.006
	(0.153)		(0.126)
Routine (SD)	-0.698***	-0.431***	
	(0.171)	(0.131)	
Constant	-0.103	-0.054	0.000
	(0.099)	(0.105)	(0.111)
Observations	81	81	81
R-squared	0.187	0.116	0.000

Source.–National Longitudinal Survey of Youth 1997

Note.–The estimates are from the regression of standardized residual wage variances on task intensity indices at the 3-digit occupation level. Residual wages are calculated from wage regression controlling individual-occupation fixed effects and year fixed effects. Robust standard errors in parentheses.

Table 4: Self-selection on Risk Preference

	(1) Routine	(2) Cognitive	(3) Risk	(4) Wage
Most Tolerant	-0.078***	0.100***	0.121***	0.032***
	(0.021)	(0.023)	(0.027)	(0.007)
Tolerant	-0.023	0.062**	0.102***	0.015*
	(0.024)	(0.027)	(0.033)	(0.008)
Averse	-0.046**	0.025	0.078***	0.007
	(0.018)	(0.019)	(0.025)	(0.007)
Demographics	✓	✓	✓	✓
N	54,131	54,131	54,131	54,131
R^2	0.289	0.208	0.029	0.261

Source.–National Longitudinal Survey of Youth 1997

Note.–The estimates are from the regression of average task intensity, earnings risk measures, and occupational average of wages on risk attitude group indicators. Control variables include race, gender, age-adjusted AFQT scores and education level indicators. Standard errors clustered at individual-occupation level.

Table 5: Regression of Log Earnings on History of Task Intensity

Log Earnings, t	All	Hs	Co
Cognitive (SD), $t-1$	0.022***	0.022***	0.015**
	(0.003)	(0.004)	(0.006)
Routine (SD), $t-1$	0.001	0.003	-0.001
	(0.003)	(0.004)	(0.007)
Cognitive (SD), $t \times t-1$	0.012***	0.010***	0.017***
	(0.003)	(0.003)	(0.005)
Routine (SD), $t \times t-1$	0.010***	0.006*	0.017***
	(0.003)	(0.004)	(0.006)
Current Occ FEs	✓	✓	✓
Indiv FEs & Year Fes	✓	✓	✓
Covariates	✓	✓	✓
N	44,736	27,864	16,871
R^2	0.694	0.617	0.717

Source.–National Longitudinal Survey of Youth 1997

Note.–The estimates are from the regression of log earnings on task intensity indices at the last period. Control variables include cubic experience profiles, current occupation, individual, and year fixed effects. Standard errors clustered at the individual-occupation level.

Table 6: Uncertainty and Wage Parameters

	Wage			Risk	
	Cognitive	Routine		Cognitive	Routine
π_0	2.43		a_0	0.28	
π_1	0.40	0.07	a_1	-0.084	-0.089
q_0	0.36	0.13	c_0	0.75	0.91
Q_1	0.043	-0.01	$C_1(\cdot, Cog)$	-0.15	-0.06
			$C_1(\cdot, Rou)$	0.06	0.16

Source.–National Longitudinal Survey of Youth 1997

Note.–The estimates are for wage and risk parameters. The wage equation is $\ln w(s_t, j) = \pi_0 + \pi_1' y_j + q_0' s_t + y_j' Q_1 s_t + \sigma(y_j) e_t$ where $e_t \sim N(0,1)$. $\sigma(y) = a_0 + a_1' y$ and $[\zeta_1(y), \zeta_2(y)] = c_0 + C_1' y$ where $\sigma(y)$ is transitory risks and ζs are persistent risks. a_{1k} is the element of a_1 corresponding to task k. $C_1(\cdot, k)$ refers to the coefficients on each task in the function of k-task persistent risks.

Table 7: Skill and Preference Parameters

	Skill			Preference	
	Cognitive	Routine		Cognitive	Routine
d_0	-0.07	0.60	f_0	-1.29	-0.93
D_1	0.60	0.97	$F_{1,Black}$	-0.13	-0.14
D_2	0.23	-0.54	$F_{1,Man}$	0.83	1.07
$D_{3,Black}$	0.00	-0.17	$F_{1,AFQT}$	0.6	0.32
$D_{3,Man}$	0.05	0.19	F_2	0.17	-0.01
$D_{3,AFQT}$	-0.03	0.16	F_3	7.55	7.54
$H_{1,Black}$	0.15	-0.11	$\overline{y}_0$	0.46	0.41
$H_{1,Man}$	0.03	0.11	$Y_{1,Black}$	-0.02	0.04
$H_{1,AFQT}$	0.23	0.27	$Y_{1,Man}$	-0.07	0.1
$H_{1,\gamma}$	-0.02	-0.04	$Y_{1,AFQT}$	0.05	-0.04

Source.–National Longitudinal Survey of Youth 1997

Note.–The estimates are for skill formation and preference parameters. The skill transition equation is $s_{t+1}(x, s_t, j) = d_0 + D_1 s_t + D_2 y_j + D_3 x + \eta_t(y_j)$, where $\eta(y_j) \sim N(0, S(y_j))$ and $S(y_j)$ is a diagonal matrix with diagonal elements $[\zeta_1(y_j)^2, \zeta_2(y_j)^2]$. Initial skill equation is $s_1 = h_0 + H_1[x, \gamma] + \xi$. The preference equation is $C(j; s_t, x, j_{t-1}) = (f_0 + F_1 x + F_2 s_t)' y_j - (y_{j_{t-1}} - y_j)' F_3 (y_{j_{t-1}} - y_j)$. Initial task equation is $y_0 = \overline{y}_0 + Y_1 x$.

Table 8: Earnings Gap Decomposition - Log Lifetime Earnings Relative to Most Averse

Earnings Gap Relative to Most Averse	Data	Baseline	No Init	No Learn
Most Tolerant	0.158	0.129	0.096	0.055
% of Baseline		(1.00)	(0.74)	(0.43)
Tolerant	0.075	0.068	0.042	0.028
% of Baseline		(1.00)	(0.62)	(0.41)
Averse	0.069	0.023	0.010	0.009
% of Baseline		(1.00)	(0.43)	(0.39)

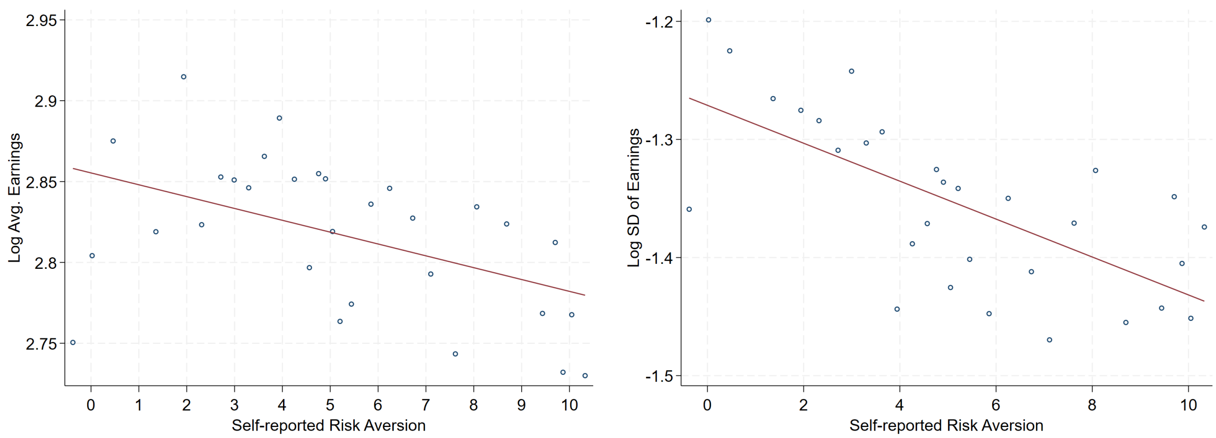
Note.–The table reports the estimates of the regression of log lifetime average earnings on risk attitude group indicators using the real data and the data simulated with the estimated parameters in Table 6 and Table 7 and their variations. The simulated data are 5 times the size of the real data. The percent of the earnings gap relative to the baseline is reported in the parenthesis. A detailed description of the simulations is provided in Section 7.

Table 9: Counterfactuals Analysis - Routine-biased Technological Changes
Higher return to cognitive tasks and lower return to routine tasks

	Log Earnings		Cognitive		Routine	
Gap Relative to Most Averse	Baseline	Counter	Baseline	Counter	Baseline	Counter
Most Tolerant	0.129	0.141	0.136	0.149	-0.122	-0.132
Tolerant	0.068	0.056	0.037	0.040	-0.030	-0.033
Averse	0.023	0.019	0.002	0.005	-0.002	-0.003

Note.–The table reports the estimates of the regression of dependent variables on risk attitude group indicators using the simulated data. Dependent variables include log lifetime average earnings, lifetime average cognitive task intensities, and lifetime average of routine task intensities. The baseline model is simulated with the estimated parameters in Table 6 and Table 7. The counterfactual model is simulated with 10% higher return to cognitive tasks ($\pi_{1,Cog}$) and 10% lower return to routine tasks ($\pi_{1,Rou}$). A detailed description of the simulations is provided in Section 8.

A Additional Figures and Tables



(a) Lifetime Average Earnings

(b) Lifetime Earnings Volatility

Figure A1: Self-reported Risk Attitudes and Labor Market Outcomes

Table A1: Risk Preference, Lifetime Earnings and Occupations (Base: Most Averse)

	Log Average Earnings			Log SD of Log Earnings		
	(1)	(2)	(3)	(4)	(5)	(6)
Most Tolerant	0.107***	0.074***	0.069***	0.125***	0.098***	0.094***
	(0.015)	(0.014)	(0.014)	(0.026)	(0.026)	(0.027)
Tolerant	0.052***	0.042***	0.042***	0.099***	0.090***	0.091***
	(0.018)	(0.016)	(0.016)	(0.030)	(0.029)	(0.029)
Averse	0.045***	0.034***	0.037***	0.038*	0.021	0.018
	(0.013)	(0.012)	(0.012)	(0.023)	(0.023)	(0.023)
Demographics	✓	✓	✓	✓	✓	✓
Major Occ.		✓	✓		✓	✓
2nd Major Occ.			✓			✓
N	4,509	4,509	4,509	4,501	4,501	4,501
R^2	0.345	0.472	0.513	0.051	0.113	0.134

Source.–National Longitudinal Survey of Youth 1997

Notes.–The estimates are from the regression of lifetime average and volatility of earnings on risk attitude group indicators. Control variables include race, gender, and AFQT scores. Robust standard errors in parenthesis.

Table A2: Task Intensity and Productivity Uncertainty Proxy by 2-digit Occupations

	Cognitive		Routine		Wage. Disp.	
	Mean	SD	Mean	SD	Mean	SD
Management	0.933	0.061	0.205	0.143	0.241	0.056
Buss/Finance	0.852	0.078	0.229	0.129	0.233	0.02
Computer/Math	0.907	0.054	0.232	0.084	0.243	0.099
Engineer	0.928	0.101	0.25	0.186	0.208	0.017
Physical/Social Science	0.905	0.092	0.293	0.279	0.209	0.056
Social Services	0.93	0.042	0.051	0.069	0.313	0.161
Legal	0.794	0.19	0.298	0.125	0.241	0.086
Educ	0.772	0.143	0.108	0.091	0.231	0.075
Arts	0.69	0.137	0.35	0.19	0.342	0.05
Healthcare	0.866	0.087	0.445	0.29	0.221	0.045
Health Supp	0.577	0.172	0.41	0.238	0.199	0.022
Protective	0.732	0.25	0.372	0.203	0.189	0.04
Food Service	0.282	0.273	0.639	0.217	0.215	0.062
Clean	0.376	0.347	0.511	0.103	0.302	0.045
Personal	0.466	0.239	0.298	0.231	0.329	0.098
Sales	0.655	0.264	0.234	0.221	0.252	0.077
Administrative	0.499	0.235	0.608	0.243	0.164	0.014
Farm	0.385	0.351	0.562	0.293	0.313	
Construct	0.393	0.226	0.741	0.226	0.238	0.031
Repair	0.517	0.181	0.629	0.23	0.226	0.082
Produce	0.269	0.191	0.903	0.117	0.163	0.054
Transportation	0.363	0.257	0.771	0.199	0.233	0.034
All	0.578	0.305	0.519	0.321	0.229	0.066

Source.–Occupational Information Network, National Longitudinal Survey of Youth 1997

Table A3

SOC3	Title	SOC3 New	Title
111	Top Executives	111	Top Executives
112	Advertising and Sales Managers	112	Advertising and Sales Managers
113	Operations Specialties Managers	113	Operations Specialties Managers
119	Other Management Occupations	119	Other Management Occupations
131	Business Operations Specialists	131	Business Operations Specialists
132	**Financial Specialists**	**131**	**Business Operations Specialists**
151	Computer Occupations	151	Computer Occupations
152	**Mathematical Science Occupations**	**192**	**Physical Scientists**
172	Engineers	172	Engineers
173	Drafters, Engineering Technicians, and Mapping Technicians	173	Drafters, Engineering Technicians, and Mapping Technicians
191	**Life Scientists**	**192**	**Physical Scientists**
192	Physical Scientists	192	Physical Scientists
193	Social Scientists and Related Workers	193	Social Scientists and Related Workers
194	Life, Physical, and Social Science Technicians	194	Life, Physical, and Social Science Technicians
211	Other Community and Social Service Specialists	211	Other Community and Social Service Specialists
212	**Religious Workers**	**211**	**Other Community and Social Service Specialists**
231	Lawyers, Judges, and Related Workers	231	Lawyers, Judges, and Related Workers
232	Legal Support Workers	232	Legal Support Workers
251	**Postsecondary Teachers**	**253**	**Other Teachers and Instructors**
252	**Preschool, Elementary, Middle Education Teachers**	**253**	**Other Teachers and Instructors**
253	Other Teachers and Instructors	253	Other Teachers and Instructors
254	**Librarians, Curators, and Archivists**	**259**	**Other Educational Instruction and Library Occupations**
259	Other Educational Instruction and Library Occupations	259	Other Educational Instruction and Library Occupations
271	Art and Design Workers	271	Art and Design Workers
272	Entertainers and Performers, Sports and Related Workers	272	Entertainers and Performers, Sports and Related Workers
273	Media and Communication Workers	273	Media and Communication Workers
274	Media and Communication Equipment Workers	274	Media and Communication Equipment Workers
291	Healthcare Diagnosing or Treating Practitioners	291	Healthcare Diagnosing or Treating Practitioners
292	Health Technologists and Technicians	292	Health Technologists and Technicians
311	**Health and Personal Care Aides**	**312**	**Occupational and Physical Therapist Assistants**
312	Occupational and Physical Therapist Assistants	312	Occupational and Physical Therapist Assistants
331	Supervisors of Protective Service Workers	331	Supervisors of Protective Service Workers
332	**Firefighting and Prevention Workers**	**333**	**Law Enforcement Workers**
333	Law Enforcement Workers	333	Law Enforcement Workers
339	Other Protective Service Workers	339	Other Protective Service Workers
351	Supervisors of Food Preparation and Serving Workers	351	Supervisors of Food Preparation and Serving Workers
352	Cooks and Food Preparation Workers	352	Cooks and Food Preparation Workers
353	**Food and Beverage Serving Workers**	**359**	**Other Food Preparation and Serving Related Workers**
359	Other Food Preparation and Serving Related Workers	359	Other Food Preparation and Serving Related Workers
371	Supervisors of Cleaning and Maintenance Workers	371	Supervisors of Cleaning and Maintenance Workers
372	Building Cleaning and Pest Control Workers	372	Building Cleaning and Pest Control Workers
373	Grounds Maintenance Workers	373	Grounds Maintenance Workers
392	**Animal Care and Service Workers**	**399**	**Other Personal Care and Service Workers**
393	Entertainment Attendants and Related Workers	393	Entertainment Attendants and Related Workers
395	Personal Appearance Workers	395	Personal Appearance Workers
396	**Baggage Porters, Bellhops, and Concierges**	**399**	**Other Personal Care and Service Workers**
399	Other Personal Care and Service Workers	399	Other Personal Care and Service Workers
411	Supervisors of Sales Workers	411	Supervisors of Sales Workers
412	Retail Sales Workers	412	Retail Sales Workers
413	Sales Representatives, Services	413	Sales Representatives, Services
414	**Sales Representatives, Wholesale and Manufacturing**	**413**	**Sales Representatives, Wholesale and Manufacturing**
419	Other Sales and Related Workers	419	Other Sales and Related Workers

Table A3 continued from previous page

SOC3	Title	SOC3 New	Title
431	Supervisors of Office and Administrative Support Workers	431	Supervisors of Office and Administrative Support Workers
432	Communications Equipment Operators	432	Communications Equipment Operators
433	**Financial Clerks**	**434**	**Information and Record Clerks**
434	Information and Record Clerks	434	Information and Record Clerks
435	Material Recording and Distributing Workers	435	Material Recording and Distributing Workers
436	**Secretaries and Administrative Assistants**	**439**	**Other Office and Administrative Support Workers**
439	Other Office and Administrative Support Workers	439	Other Office and Administrative Support Workers
452	Agricultural Workers	452	Agricultural Workers
471	**Supervisors of Construction and Extraction Workers**	**511**	**Supervisors of Production Workers**
472	Construction Trades Workers	472	Construction Trades Workers
473	**Helpers, Construction Trades**	**472**	**Construction Trades Workers**
474	**Other Construction and Related Workers**	**472**	**Construction Trades Workers**
475	**Extraction Workers**	**472**	**Construction Trades Workers**
491	**Supervisors of Installation, Maintenance, and Repair Workers**	**511**	**Supervisors of Production Workers**
492	**Electrical and Electronic Equipment Mechanics, Installers, and Repairers**	**499**	**Other Installation, Maintenance, and Repair Occupations**
493	**Vehicle and Mobile Equipment Mechanics, Installers, and Repairers**	**499**	**Other Installation, Maintenance, and Repair Occupations**
499	Other Installation, Maintenance, and Repair Occupations	499	Other Installation, Maintenance, and Repair Occupations
511	Supervisors of Production Workers	511	Supervisors of Production Workers
512	Assemblers and Fabricators	512	Assemblers and Fabricators
513	Food Processing Workers	513	Food Processing Workers
514	Metal Workers and Plastic Workers	514	Metal Workers and Plastic Workers
515	Printing Workers	519	Other Production Occupations
516	Textile, Apparel, and Furnishings Workers	516	Textile, Apparel, and Furnishings Workers
517	**Woodworkers**	**516**	**Textile, Apparel, and Furnishings Workers**
519	Other Production Occupations	519	Other Production Occupations
531	**Supervisors of Transportation and Material Moving Workers**	**511**	**Supervisors of Production Workers**
533	Motor Vehicle Operators	533	Motor Vehicle Operators
536	Other Transportation Workers	536	Other Transportation Workers
537	Material Moving Workers	537	Material Moving Workers