ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

Subjective Earnings Risk

Authors	Andrew Caplin, Victoria Gregory, Eungik Lee, Soren Leth-Petersen, and Johan Saeverud
Working Paper Number	2023-003E
Revision Date	August 2025
Citable Link	https://doi.org/10.20955/wp.2023.003
Suggested Citation	Caplin, A., Gregory, V., Lee, E., Leth-Petersen, S., Saeverud, J., 2025; Subjective Earnings Risk, Federal Reserve Bank of St. Louis Working Paper 2023-003. URL https://doi.org/10.20955/wp.2023.003

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

Subjective Earnings Risk

Andrew Caplin
New York University

Victoria Gregory
FRB St. Louis

Eungik Lee
Indiana University

Søren Leth-Petersen
University of Copenhagen

Johan Sæverud
University of Copenhagen

April 25, 2025

Abstract

We introduce a survey instrument to measure earnings risk allowing for the possibility of quitting or being fired from the current job. We find these transitions to be the key drivers of subjective risk. A link with administrative data provides multiple credibility checks and reveals that subjective earnings risk varies systematically across the population. It is also many times smaller than traditional estimates even when conditioning richly on demographics and job history. We show that subjective earnings risk can help explain why many hold limited liquid assets.

Keywords: earnings risk, job transitions, subjective expectations

JEL classification: D31, D84, E24, J31

1 Introduction

Subjective earnings risk impacts search and labor market decisions, savings/insurance choices, and inequality. Yet it has proven challenging to measure. We introduce a survey instrument to measure this risk, which we define as the spread of an individual's probability distribution over future earnings growth.[1] We make explicit allowance for job transitions: both the possibility of quitting and the risk of being fired from the current job. We find these transitions to be the key drivers of subjective earnings risk. A link with administrative data provides multiple credibility checks, but also reveals that subjective earnings risk is very heterogeneous across the population and that average subjective earnings risk is many times smaller than traditional estimates based on the cross sectional distribution of realized earnings growth imply. We show that the significant heterogeneity and generally low level of subjective earnings risk help explain why so many hold limited liquid assets. This indicates the importance of administratively-linked subjective measures such as ours.

One key finding from the survey is the high level of heterogeneity in peoples' beliefs about their future earnings. Respondents vary in terms of their probabilities of job separation, how long they expect to be out of work if they separate from their jobs, and in terms of their probability distribution over future earnings growth conditional on job separation. The link to the administrative data reveals that much of this heterogeneity survives even after accounting for demographics and earnings histories. The match with the administrative data also provides many broad credibility checks when correspondingly aggregated up. Average job separation rates calculated from the administrative data are very similar to the average job separation probabilities in the survey, even when stratified by age groups. Similarly, average expected time out of work following job separation compares nicely with actual time spent out of work as it is recorded in administrative data. Furthermore, based on the answers to the survey instrument, we construct a measure of the overall subjective probability distribution over one-year ahead earnings growth and show that the mean of this lines up tightly with realized earnings growth for the period that the expectations concern. Finally, we pool the subjective distributions over one-year ahead earnings growth for the entire survey sample and compare with the cross sectional distribution of realized earnings growth that we observe in the administrative data. Also

[1] In the literature estimating earnings processes, year-to-year earnings volatility, including earnings fluctuations that may have been the result of a choice and that cannot be explained by observable characteristics, is often interpreted as risk. This amounts to assuming that workers receive a risky but exogenous flow of earnings in each period. The well-known permanent-transitory earnings process fitted to administrative data falls within this category.

here we find that the distribution constructed from the survey data aligns closely with the distribution of realized earnings growth in the administrative data. In sum, the survey answers agree closely with third-party reported administrative data in a sequence of validation exercises, thus giving credence to the quality of the survey.

Despite these matches when the survey data are averaged, we find that the subjective earnings risk at the individual level, measured as the interdecile range of the subjective probability distribution over future earnings growth, is two to six times smaller than would be estimated in the administrative data. This is the case even when taking into account observable characteristics in a very detailed nonparametric fashion. The reason for this is that the cross sectional distribution of earnings growth confounds heterogeneity in expected subjective earnings growth and actual subjective earnings risk. This suggests that individuals have more information than an econometrician who only observes administrative data on realized earnings growth across the population and therefore cannot separate risk from heterogeneity. Our unique combination of survey and administrative data allows us to confirm this and to quantify how much administratively inferred earnings risk overstates subjective risk. By documenting that respondents' expected earnings growth align with their actual subsequent earnings growth, as shown above, we confirm that individuals have superior information.

The survey design allows us to examine the role of job transitions for subjective probability distributions over one-year ahead earnings growth. We show that higher order moments are important for characterizing these distributions. In particular, the typical subjective distribution is negatively skewed and displays excess kurtosis that is increasing in age. When we consider only the probability distributions over the contingency where people stay in the current job then the typical subjective spread is reduced and there is no skewness nor excess kurtosis. In other words, we find that job transitions are responsible for most of the spread that we observe in subjective probability distributions over future earnings including higher order moments.

We present two key insights regarding subjective earnings risk. First, we find that subjective earnings risk is several times lower than risk estimates based on recorded realized earnings growth from administrative data. Second, we observe that subjective risk is highly heterogeneous across the population. We explore the implications of these findings for limited liquid asset accumulation within the framework of a standard incomplete markets savings model (Aiyagari, 1994, Huggett, 1993). When such models are calibrated using risk estimates derived from administrative data on earnings growth, they struggle to explain why so many individuals in the real world hold limited liquid assets. These traditional calibrations tend to overestimate the precautionary savings motive, leading to predictions of higher liquid asset holdings than observed.[2] If a significant proportion of the population perceives their earnings risk to be lower than traditional estimates based on administrative data suggest, they would naturally accumulate fewer liquid assets due to reduced precautionary savings needs. When we calibrate the incomplete markets model to match the distribution of subjective earnings risk from our survey, we find that the model more accurately reflects the observed fraction of the population holding low levels of assets. Finally, using our preferred model to simulate the cross-sectional distribution of earnings growth, we find that it resembles the empirical distribution, capturing both its peakedness and tail thickness.

Our paper builds on recent advances in our understanding of earnings dynamics and earnings risk measured in large administrative data (e.g., Guvenen et al., 2021).[3] This literature shows the importance of higher-order moments in the distribution of realized earnings growth, heterogeneity in earnings dynamics across ages and levels of earnings, and the critical role of job transitions and periods out of the labor force for earnings risk, all of which are mirrored in our subjective data.[4] The other key literature on which we build is the pioneering research of Dominitz and Manski (1997), measuring probabilistic beliefs about one-year ahead earnings, with important subsequent work by De Bruin et al. (2011), Dominitz (1998, 2001), Guiso et al. (2002), and Pistaferri (2001, 2003). In recent work complementary to ours, Koşar and van der Klaauw (2023) and Wang (2023) study wage expectations related to staying in the current job as elicited in the Survey of Consumer Expectations conducted by the New York Fed. Wang finds that subjective wage risk associated with staying in the current job is lower than wage risk inferred from wage realization for job stayers and finds that this can help explain a high concentration of US households with little liquid wealth. Hartmann and Leth-Petersen (2024) find that earnings risk is related to subjective unemployment expectations. Our study follow recent

[2] The literature has addressed this issue through various approaches. One approach introduces preference heterogeneity (e.g., Krusell and Smith, 1998; Carroll et al., 2017), which leads to varying saving behaviors and consequently, a larger fraction of the population with low asset holdings. Another approach incorporates an illiquid asset that offers a higher return than the liquid asset, encouraging individuals to allocate resources into illiquid assets, thereby becoming wealthy-hand-to-mouth (Kaplan and Violante, 2014).

[3] See also Busch et al. (2022), Druedahl and Munk-Nielsen (2020), Halvorsen et al. (2023), Guvenen et al. (2022) and other papers published as part of the Global Repository of Income Dynamics (GRID) project, `https://www.grid-database.org`.

[4] The prominent role of job transitions for earnings risk, and in particular for higher order moments in the distribution of realized earnings growth, is documented earlier by Low et al. (2010). The importance of job transitions for the level of earnings is also revealed in separate studies of layoffs and quits (Topel and Ward, 1992; Jacobson et al., 1993; Von Wachter et al., 2009).

research that combine at the individual level subjective information collected from surveys with objective information from administrative data facilitating direct comparison of subjective and objective information (e.g., Andersen and Leth-Petersen, 2021; Caldwell et al., 2024; Epper et al., 2020; Hvidberg et al., 2023;). Another related literature studies subjective labor market expectations (Manski and Straub, 2000; Stephens, 2004; Campbell et al., 2007; Hendren, 2017), and in terms of methodology, our work builds on a branch of the literature that measures conditional expectations (e.g. Arcidiacono et al., 2020, Wiswall and Zafar, 2021). Finally, our work is related to a recent literature using subjective expectations data to inform and discipline structural models of labor market dynamics and savings (Conlon et al., 2018; Balleer et al., 2021; Bick et al., 2021; Mueller et al., 2021; Faberman et al., 2022; Jäger et al., 2022; Stoltenberg and Uhlendorff, 2022; Balleer et al., 2023; Wang, 2023).

The paper is organized as follows. Section 2 introduces the conditional earnings survey instrument. Section 3 compares survey responses with linked administrative data. Section 4 compares subjective earnings risk with its administratively-estimated counterpart. Section 5 links this risk with job transitions. Section 6 presents and calibrates a standard incomplete-markets model and examines the ability of subjective earnings risk to explain the prevalence of low liquid asset holding. Section 7 concludes.

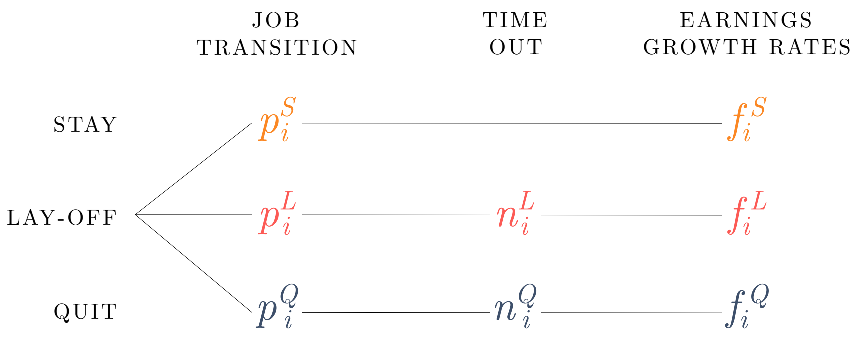
2 The Conditional Earnings Instrument

In this section we introduce the conditional earnings survey instrument through which respondents are asked in January 2021 about their expectations concerning job transitions and earnings throughout 2021. We first present the branching structure and the survey questions.[5] We then introduce the *Copenhagen Life Panel* in which it was implemented and give a branch-by-branch bird's-eye-view of survey responses. We end by explaining how key variables are constructed and providing a high level overview of quantitative findings.

Figure 1 illustrates the branching structure of our survey and our naming convention for each of the components. Starting from the left, we first ask about the probability of job transitions, i.e., the probability of staying in the current job (p_i^S), the probability of being laid off (p_i^L), and the probability of quitting (p_i^Q). For the layoff and quit branches we then ask about the expected time out of work following the separation (n_i^L, n_i^Q). Finally, we elicit the conditional probability distributions over one-year ahead earnings in each

[5] The questionnaire is reported in Online Appendix A.

of the three branches. We subtract last year's earnings, which we also ask about in the survey, from this to arrive at branch-specific distributions of growth rates of earnings, and we denote these distributions (f_i^S, f_i^L, f_i^Q). For each respondent we collect all these eight objects.



Note: The survey instrument consists of three branches, each representing a job transition (Stay, Layoff, Quit), and three domains for each branch: for each individual i we elicit job transition probabilities, p_i^B, time out of work, n_i^B, and distributions of conditional earnings growth rates, f_i^B, where $B \in \{S, L, Q\}$.

Figure 1: Survey instrument overview

2.1 Job transitions

The expectations instrument opens by asking all respondents who report being employed in January 2021 about the likelihood of job transitions during 2021:

- *Please think about your possible relationship with your current employer in 2021. Assign the probability in each possible case. The sum of the probabilities should be 100.*

    1. *Staying with your current employer during 2021*
    2. *Being laid off from your current employer at some point during 2021*
    3. *Quitting from your current employer at some point during 2021*
    4. *Separating from your current employer for some other reason during 2021*

For each individual (i) we denote the branch-specific probability p_i^B, where $B \in \{S, L, Q\}$.

For those who report a positive layoff probability, we follow up by asking about how long they expect to be out of work, and we do this by asking the likelihood of being re-employed within four different horizons: 1, 3, 12, and 24 months:

- *Suppose you were to be laid off from your current employer during 2021. What is the probability that you would start working for pay again within 1/3/12/24 months of termination?*

For those who report a positive probability of quitting during 2021, we ask a similar question, where the probabilities now refer to finding a job within each time horizon after quitting. We use this information to calculate the expected time out of the labor force following a separation (n_i^L and n_i^Q) for each individual. The process is described in detail in Section 2.3.

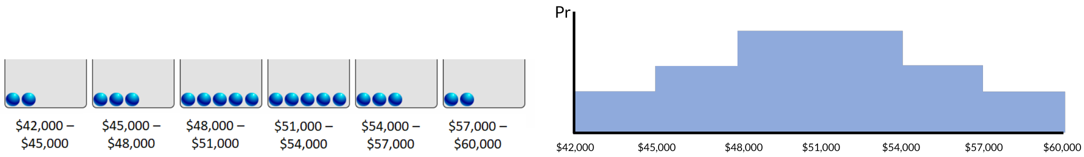
Finally, we ask each respondent their probabilistic beliefs about future earnings for each of the subjectively possible job transitions. This is straightforward for the stay branch as this is just the uninterrupted continuation of the current job. For the layoff and quit branches, we ask about the earnings in the 12 months following the start of the new job, i.e., the annual earnings taking into account that the new job may begin following a period out of work. Here is the basic design for the case of being laid off from the job during 2021.

- *Suppose you were to be laid off from the current employer during 2021 and to start to work for pay at some point in the following 2 years. Think about your possible earnings during the first 12 months in this new job*

In order to elicit the full distribution of future annual earnings in each branch we apply the "balls in bins" method developed by Delavande and Rohwedder (2008), which is intuitive and visually oriented.[6] Respondents are first asked to state the minimum and maximum values for possible future earnings, as in Dominitz and Manski (1997). Then the range between the stated minimum and maximum is divided into six equally sized bins. Respondents are then instructed to move 20 balls into the six bins to reflect how likely their future earnings are to fall in each of the ranges defined by the bins. Figure 2(a) illustrates the "balls in bins" task as it appears in the online survey.

We construct branch-specific subjective distributions of earnings based on the answers to these questions. Since there are 20 balls available, we interpret one ball as representing a probability of 5%. We also assume that probabilities are uniformly distributed within each bin. For example, in Figure 2(a) two balls are placed in the first bin and we interpret this to mean that there is a 10% likelihood of realizing earnings in the interval 42,000 to 45,000 DKK (the first bin). Combining all the bins enable us to characterize the entire

[6] Goldstein and Rothschild (2014) show that bins and balls elicitation increases the accuracy of reported distribution compared to other non-graphical elicitation methods.



(a) Elicitation (b) Distributions

Note: Panel (a) shows the sample screen for the elicitation and Panel (b) shows how we interpret the distribution of the answer in Panel (a) as a mixture of uniform distributions.

Figure 2: Balls in bins

subjective probability distribution and to calculate various moments for each respondent's conditional distribution. For example, Figure 2(b), shows the distribution that the "balls in bins" answers in panel (a) are converted to. The mean of this distribution is 51,000 and the standard deviation is 4,896.

In the survey we ask respondents about last year's earnings.[7] By subtracting this from the conditional distributions collected using the survey instruments outlined above we arrive at branch-specific distributions of growth rates of earnings, which we denote f_i^B, where $B \in \{S, L, Q\}$.

2.2 Copenhagen Life Panel

Our survey instrument is implemented in the newly developed *Copenhagen Life Panel* [8] (CLP) which is an online panel survey implemented in Denmark. We invite a random selection of individuals, who are recorded in the Danish population registry and aged between 20 and 70, to participate in the survey. The population registry is a complete registry of all persons who are born or have ever had an address in Denmark. It contains a personal identifier (CPR-number) applied universally to record any contact an individual has with the public sector. Invitations to participate were sent out using an official email account, called *e-boks*, which all Danes are equipped with. For the purpose of this paper we consider questions about earnings expectations and job transitions that were included

[7] Last years earnings is the earnings realized in 2020. This has a direct counterpart in the administrative data. In Online Appendix C.4 we directly compare the survey responses concerning earnings in 2020 with earnings for 2020 recorded i the income-tax register. The comparison show that survey answers line up very closely with the administrative data.

[8] The *Copenhagen Life Panel* is an ongoing survey that was initiated in 2020 and is issued every year in January.

in CLP issued in January 2021.[9]

Upon survey completion, answers are linked to the administrative records for all individuals who are invited to the survey as well as the rest of the entire Danish population. These data include standard demographic information, such as age, gender, education, household composition, and household wealth, all collected at the annual frequency. All the administrative data are longitudinal by nature and are currently available up to and including 2021.

For this study we include respondents between age 20 and 65 which is the typical working age span. The gross sample includes 14,875 respondents. We restrict the sample to include 10,945 people who are employed at the time of the survey. This is to make sure that we are not dealing with individuals who are permanently or temporarily out of the labor market. In Online Appendix B, we compare the average earnings, age, gender, and educational attainment between the survey sample and the full population belonging to the same age groups. There is wide variation across age and earnings in our sample. In comparison with the larger population, the average survey participant is slightly older, more educated, and has a somewhat higher level of earnings. For the subsequent analysis, we apply population weights that we construct from the administrative data.[10]

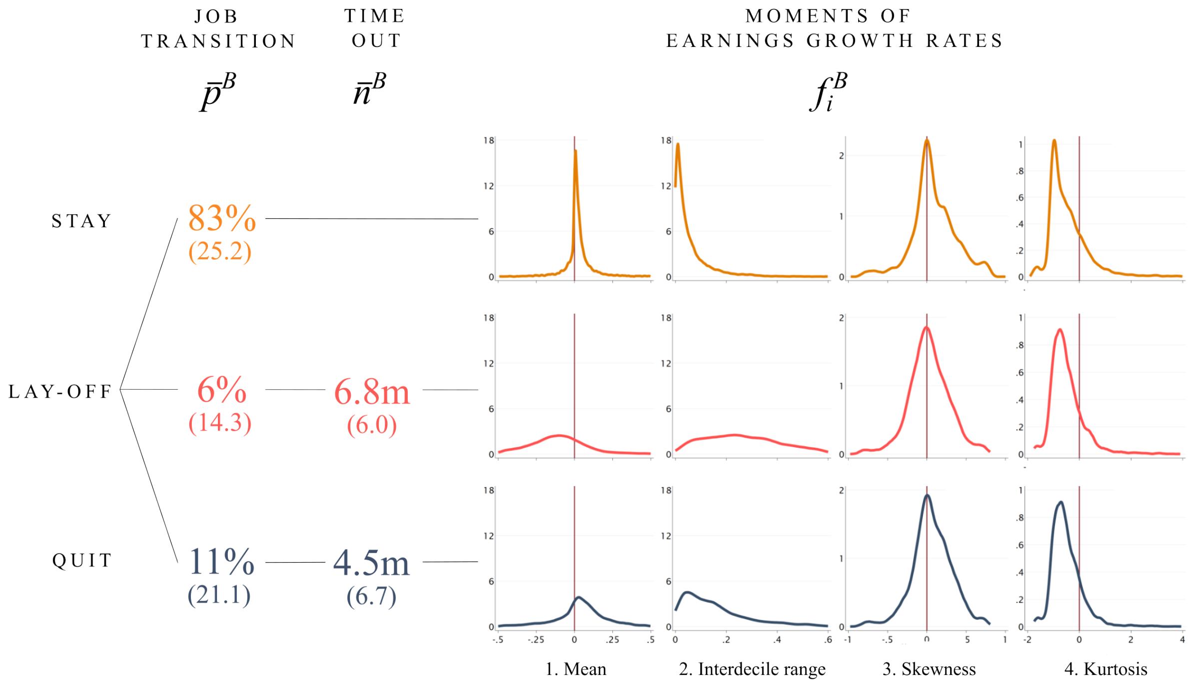
2.3 Job Transitions

In Figure 3 we present an overview of the answers collected. Starting from the left, the probabilities of job transitions $\bar{p}^B$ represent the average job transition probabilities stated by the respondents. With an average likelihood of $\bar{p}^S = 83\%$, the most likely event is remaining with the current employer, followed by quitting, $\bar{p}^Q = 11\%$ and being laid off, $\bar{p}^L = 6\%$. In Online Appendix C.1, we plot the distributions of each job transition probability.

Moving to the right in Figure 3, we report the average expected time out of work upon quitting or being laid off, $\bar{n}^B$. To arrive at a summary measure of the out-of-work duration, we aggregate over the likelihood of being out of work in the 12 months window following the interview . Focusing on time out of work following a layoff, respondents

[9] Those who finished the survey participated in a prize lottery with 50 respondents receiving prizes worth 1,000 DKK (approximately, 140 USD) and one a grand prize of 10,000 DKK (approximately, 1,400 USD).

[10] To construct these weights, we estimate a probability model of survey participation using the 2020 administrative data with information about the demographics of the Danish population who are active in the labor market and use the inverse of the predicted propensity scores as weights. For a detailed description of the construction of the population weights see Online Appendix B.



Note: The figure shows answers to the questions in the conditional survey instrument, where the rows correspond to the branches "Stay", "Layoff", and "Quit". The first column shows the average probabilities of each branch, $\bar{p}^B$. The second column shows the average of the expected reemployment period in each branch, $\bar{n}^B$, in months. The distributions show the cross-sectional distribution of the 1st to 4th moment of the subjective conditional earnings distributions, f_i^B. The parenthesis below each value shows the corresponding standard deviation. We measure the second moment by the interdecile range, $p_{90} - p_{10}$. We measure skewness using Kelley's measure of skewness: $S_K = \frac{(p_{90}-p_{50})-(p_{50}-p_{10})}{(p_{90}-p_{10})}$. We use the Crow-Siddiqui measure of excess kurtosis $K_{CS} = \frac{(p_{97.5}-p_{2.5})}{(p_{75}-p_{25})} - 2.91$. In Online Appendix C.1, we plot the distributions of job transition and time out respectively.

Figure 3: Overview of branch-by-branch survey responses

report $(n^L_{i,1}, n^L_{i,3}, n^L_{i,12}, n^L_{i,24})$ as their reemployment probabilities within 1, 3, 12, and 24 months, respectively. We assume that reemployment occurs in the middle of three time intervals. The expected re-employment period is calculated as follows:

$$n^L_i = 2(n^L_{i,3} - n^L_{i,1}) + 7.5(n^L_{i,12} - n^L_{i,3}) + 18(n^L_{i,24} - n^L_{i,12}) + 24(100 - n^L_{i,24})$$

We use the same procedure for the expected re-employment period after a quit, n^Q_i.

The numbers under "Time out" in Figure 3 are the average periods out of work following a job separation, $\bar{n}^B$. We find that respondents expect to spend 6.8 (4.5) months on average to find a new job after being laid off (quitting).[11] These results imply that the respondents anticipate spending more time out of work following a layoff than following a quit, as might have been expected.

The anticipated time out of the labor force after quitting may reflect either an anticipated break or anticipated job search. In unreported analyses we regress expected time out of work following a quit on liquid assets relative to disposable income, an often used indicator of being liquidity constrained (Leth-Petersen, 2010). We find that workers with less liquid wealth expect to spend less time out of work after quitting, as if pressured back to work more quickly.

2.4 Conditional Earnings Risk

A key innovation in our survey is that we obtain subjective distributions of expected earnings growth **conditional** on job transitions. We are therefore able to calculate the moments of their subjective earnings distributions for each respondent i in each branch B. We simulate the empirical distributions of conditional earnings growth rates for each survey respondent in each branch, $\hat{f}^B_i$, by taking 20,000 random draws from the mixture of uniform distributions of expected earnings, which is illustrated in Figure 2, panel b. We convert expected earnings levels to logs and subtract the log of earnings in 2020 (self-reported) to obtain a distribution of one-year-ahead log earnings growth. This procedure imposes minimal assumptions on the shape of the empirical subjective conditional distributions, $\hat{f}^B_i$.[12]

The last four columns of Figure 3 show the cross sectional distribution of the first four mo-

[11] In Appendix C.1, we plot the distributions of time out labor force.

[12] We also fitted beta distributions to the subjectively reported data from the balls-in-bins answers. This did not change the results in any important way (not reported).

ments of subjective earnings growth distributions.[13] Each row corresponds to a different branch, B. Turning first to the means, the average respondent expects a 3% increase in earnings if staying with their current employer. Following a layoff, individuals on average expect an 11% decrease in annual earnings when they find a new job, and a 7% increase when they find a new job after a quit. These two branches also exhibit considerable heterogeneity in the means relative to the stay branch. Among those who report a positive probability of being laid off, 73% of the respondents expect a decrease in earnings if this state is realized. In contrast, among those who report a positive probability of quitting, 81% of the respondents expect to increase earnings if that state materializes.

The next column shows the interdecile range, $p_{90} - p_{10}$. This measures how much risk people associate with their earnings prospects in each branch. As might be expected, the results show that people tend to associate least risk to their earnings growth in the stay branch and most risk in the layoff branch. There is also less heterogeneity in responses in the stay branch where a considerable amount of the mass is bunched toward zero. In contrast, in the layoff branch different respondents report very different perceptions of earnings risk.

The distribution of skewness is similar across all branches. In all cases, it is clustered around zero and symmetric. This means that the modal respondent is creating symmetric distributions with their bins and balls on all the branches. However, it is noticeable, that there is a lot of heterogeneity and many individuals report distributions that are skewed.

Lastly, we turn to the distribution of excess kurtosis. The final column of Figure 3 shows that also these distributions appear similar across branches, but with a lot of heterogeneity across respondents. On average, excess kurtosis is negative, which means that the average subjective distribution is not as peaked as a normal distribution. This means that most respondents have entered distributions with relatively more mass between the center and the tails than a normal distribution.

[13] Following the practice in the literature we measure the second moment, skewness, and kurtosis for each of the subjective distributions using robust, quantile based measures. We measure the second moment by the interdecile range, $p_{90} - p_{10}$. We measure skewness using Kelley's measure of skewness: $S_K = \frac{(p_{90}-p_{50})-(p_{50}-p_{10})}{(p_{90}-p_{10})}$. S_K measures the relative length of the right and left tails of the distribution. If $S_K > 0 (< 0)$, then it means the right (left) tail is longer and large positive (negative) draws are more likely than large negative (positive) draws. Therefore, this captures the extent to which individuals perceive larger upside or downside risk. Finally, we use the Crow-Siddiqui measure of excess kurtosis $K_{CS} = \frac{(p_{97.5}-p_{2.5})}{(p_{75}-p_{25})} - 2.91$. This measure compares the range of the middle 95% of the distribution to that of the middle 50%. The statistic is normalized by 2.91, such that the Crow-Siddiqui measure of excess kurtosis for a normal distribution is zero. Excess kurtosis is informative about the extent to which expected earnings growth is concentrated in the center of the distribution or in the tails. Large excess kurtosis means larger risk of extreme changes.

Being laid off leads to the worst outcomes, on average, and respondents associate most risk to what may happen here. Staying with the current employer is expected to lead to small increases in earnings and respondents are most certain about the outcome in this state. Quitting leads to the best outcomes, and the level of risk is between that of staying and being laid off. Overall, the data uncover massive heterogeneity in expectations of future earnings, and this is reflected in all four moments and across all labor market transitions.[14]

3 Comparing Survey and Administrative Data

The Danish research data infrastructure allows us to directly compare measures elicited in the survey with their realized counterpart in the administrative data and thereby to assess the credibility of the survey instrument. To that end, the survey data is combined at the individual level with administrative data made available by Statistics Denmark from different sources with third-party reported records from various sources. The Danish administrative data are known to be of high quality (Kleven et al., 2011) and have been used extensively in previous studies, see for example, Browning et al. (2013), Leth-Petersen (2010) and Chetty et al. (2014). The data are made available with a time lag, with data through 2021 currently available for research thus allowing us to directly compare expectations with their realized counterparts. Data gathered in this manner includes earnings from work and job transitions as well a host of other administrative data providing background information about each respondent. For our comparison between survey and administrative earnings, we use monthly data about employer matches and earnings to identify job transitions, time spent out of work, and annual earnings. We also use a standard battery of administrative data compiled by Statistics Denmark.

We start out by comparing job transitions and time out of work following a job separation. We then introduce a method of aggregating our conditional survey responses to arrive at a holistic measure of subjective earnings risk that we then pool to compare with the corresponding numbers in the register data. Our baseline comparison will be based on administrative data for 2021. To allay one possible concern, note that the COVID-19 pandemic hit the Danish economy lightly and respondents seemed to recognize that this would be the case. Massive furloughing schemes were set in place very quickly by the Danish government. As a result, the lowest employment level during 2020 was only 40,000 below the baseline pre-pandemic level of 2,768,766 (February 2020), and by the end of the year this small loss had been recovered. During 2021 employment accelerated and by

[14] In Online Appendix C.2, we reproduce Figure 3 using standard measures of the moments.

end of 2021 total employment was 2,916,139, about 5 percent above the pre-covid level. To verify whether this is consistent with the data that that we collect, we will also verify that subjective expected earnings growth aligns with the realized earnings growth during the period that expectations concern

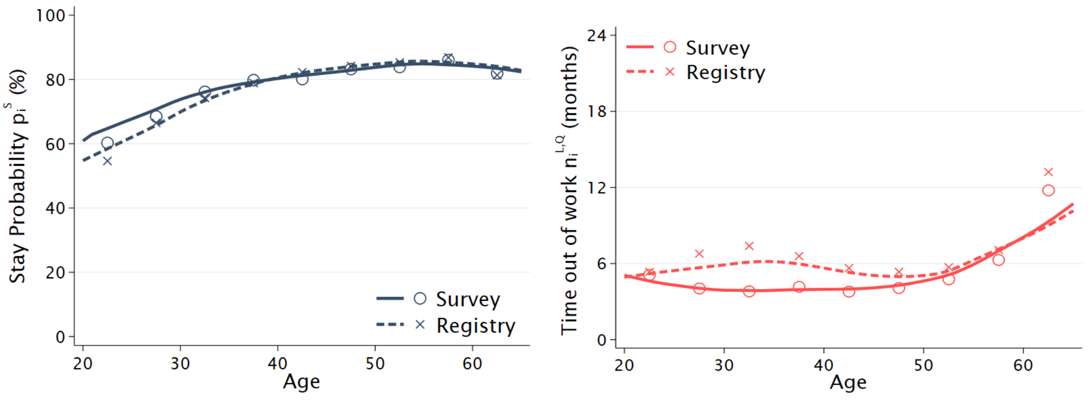
3.1 Job Transitions

In the survey, we ask about the probability of staying with the current employer, the probability of being laid off and the probability of quitting. In the administrative data we only observe job separations but not the reason for the job separation. Consequently, as a first-order check on the survey answers to these items, we compare the probability of staying with the same employer, since this object can be computed directly from the administrative data.

From the survey we consider the average reported probability of staying with the same employer, $\bar{p}^S$. In the administrative data we observe employer-employee matches at the monthly frequency and obtain a direct counterpart to $\bar{p}^S$ by calculating the share of employees who stay with the same employer throughout the calendar year. Figure 4, panel (a) shows the average stated probability (solid line) of staying with the same employer throughout 2021 and the fraction of stable job-matches (dashed line) throughout 2021 in the administrative data, both summarized by age. Generally, the likelihood of remaining in the same job throughout the year is lower among the young and there is only a low likelihood that workers aged 40+ separate from their job. The alignment between survey and registry is striking.

3.2 Time Out of the Labor Force

In the survey we ask about time spent out of work following a potential quit and lay-off. As before, we cannot confront these objects directly with administrative data as we do not know the reason for an observed separation. We therefore compare time spent out of work following any type of separation between the survey and the administrative data. In the administrative data we observe employer-employee matches at the monthly frequency and are able to track the number of months spent out of work following a separation. As we currently only have administrative data until 2022, we consider time spent out of work following job separations that took place in 2020 such that we can follow periods out of work that extend into 2022. From the survey, the expected time out of work in



(a) Stay probability

(b) Time out of work

Note: The figure shows how two components from the survey (solid line) and the administrative data (dashed line) compare over the life-cycle. Panel (a): In the survey, we directly ask about the probability of staying with the same employer for the next 12 months. In the registry, we compute the proportion of workers who stayed from Dec 2020 to Dec 2021 with the same employer. Panel (b): We calculate the expected time out of work after a separation in the survey using Equation (1). In the administrative data we consider job separations that took place during 2020 and follow time spent until reemployment occurs, possibly extending into 2022. In the registry, 1) we exclude the workers going back to the same employer (possibly seasonal work) and 2) we consider the first separation and do not take into account further separations within a year. "o" and "x" represent the empirical mean across 5-years age bins for the survey and the administrative data, respectively. The lines are local linear polynomials, calculated using a bandwidth of 4 years. C.3 shows the corresponding figures for 2019.

Figure 4: Job separations and time out of work in the survey and in the administrative data

the survey is calculated as a combination of quits and layoffs:

$$n_i^{L,Q} = \frac{p_i^Q n_i^Q + p_i^L n_i^L}{p_i^Q + p_i^L} \tag{1}$$

The result of the comparison is shown in Figure 4, panel (b). According to the survey (solid line), the average expected time spent out of work following a job separation is about 5 months for people aged less than 50 and the expected time out of work increases dramatically for workers aged 50+. The pattern is similar when the corresponding measure from the administrative data is plotted (dashed line).

3.3 Overall Subjective Earnings Expectations

We now turn to compare the overall subjective earnings growth expectations to the realized earnings growth for 2021 that we observe in the administrative data.

To do this, we first aggregate the conditional answers for each respondent into one distribution that characterizes the respondent's overall subjective expectations concerning one-year ahead earnings growth and we denote this subjective expected *holistic* earnings growth. This object summarizes the overall subjective probability distribution over one-year ahead earnings growth taking into account all the different contingencies that we asked about.

To construct the subjective holistic expected earnings growth distribution we weight together each of the branches, $B = \{S, Q, L\}$, for individual i,

$$g_i = p_i^S f_i^S + p_i^Q f_i^Q (1 - n_i^Q) + p_i^L f_i^L (1 - n_i^L), \tag{2}$$

where p_i^S, p_i^Q, and p_i^L are the probabilities of staying, quitting, and being laid off, n_i^Q and n_i^L are time out of work following a quit and a layoff. f_i^S, f_i^Q, and f_i^L are the subjective probability distributions over one-year ahead earnings growth rates for each of the three branches, staying, quitting and being laid off. The subjective holistic probability distribution over one-year ahead earnings growth, which we denote g_i, captures the total earnings growth risk, as perceived by individual i.

In practice, we simulate the empirical distribution of $\hat{g}_i$ by making a large number of random draws for each respondent based on the stated transition probabilities, p_i^B, and the individual empirical distributions of $\hat{f}_i^B$ and $\hat{n}_i^B$,[15] which are then weighted together according to Equation (2). The goal is to construct a measure of overall one-year ahead earnings expectations. Some of the contingent questions pertain to a period that goes beyond the 12 months following the interview. For example, the probability distributions over earnings growth in the case of quit and lay-off, f_i^Q and f_i^L concern the 12-months period following separation from the previous job and potential time out of work. When weighting together all the components in equation (2) we take this into account such that the overall probability distribution over one-year ahead earnings, g_i, exactly matches the 12 months following the interview. In this way, our measure of g_i pertains to the exact same period that the administrative data measure, i.e., the calendar year.

We simulate the empirical distribution $\hat{g}_i$ by drawing 20,000 job transition events for each individual based on the stated job transition probabilities. From each of these simulated job transitions, we simulate time out of work and the conditional earnings distribution

[15] In addition to the individual point estimates, n_i^L and n_i^Q, we also construct individual empirical distributions of time out of work following a separation $\hat{n}_i^L$ and $\hat{n}_i^Q$. These distributions are simulated out of stated probabilities of being reemployed within a certain time frame. We refer to Online Appendix D for specifics.

for the relevant branch based on the empirical distributions of $\hat{n}_i^B$ and $\hat{f}_i^B$. In this way, we simulate 20,000 synthetic realizations for each respondent based on the reported survey answers. We give a complete account of the simulation protocol in Online Appendix D.[16] In the following we refer to g_i but in practice we use $\hat{g}_i$.

g_i is the subjective probability distribution over earnings growth during 2021. The mean of $\mathbb{E}_i[g_i]$ has a direct counterpart in the administrative data, namely the realized earnings growth, Δy_i, for 2021. In Figure 5 we show a direct comparison of $\mathbb{E}_i[g_i]$ and Δy_i. The figure displays a binned scatter plot of the $\mathbb{E}_i[g_i]$ against realized earnings in 2021 as observed in the administrative data. The figure is constructed by ranking $\mathbb{E}_i[g_i]$ for all respondents and dividing them into 20 equally sized groups. Within each of the 20 groups the average realized earnings growth in 2021 is calculated. The figure shows that there is a remarkably close alignment between subjective earnings growth expectations and the subsequent realizations for the vast majority of the sample. In other words, respondents are well-informed about their own future earnings growth.[17] Finally, note that the survey was fielded in January 2021, before the COVID-19 lockdown. The close alignment of expectations with realizations confirms that the COVID-19 pandemic did not significantly misalign people's expectations with the actual outcomes.

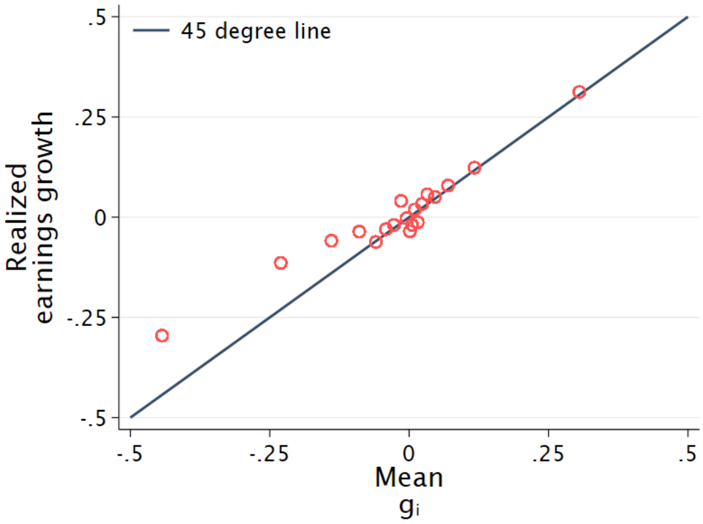
3.4 The Cross-Sectional Distribution of Earnings Growth

Finally, we turn to compare the cross sectional distribution of expected earnings growth in the survey to the cross sectional distribution of realized earnings growth for 2021 in the administrative data. The first characterizes the *expected* earnings growth variability in the population and the second characterizes the corresponding *realized* earnings growth variability in the population.

In order to arrive at a cross sectional distribution of expected earnings growth, we pool the distributions of subjective expected holistic earnings growth, g_i, for all N individuals in our sample to arrive at a cross sectional distribution that describes the expected earnings growth variability in the population that is conceptually comparable to the cross sectional distribution of realized earnings growth that we observe in the administrative data. In doing this, we pool within cells in which individuals share the same observable

[16] Equation (2) implicitly assumes that job separations take place at the beginning of the period. In practice, expected earnings in case of job separation may be a convex combination of earnings in the old job and earnings in the new job following time out of work. We have a more detailed discussion in Online Appendix D.

[17] Also Conlon et al. (2018) find that people's earnings growth expectations are on average consistent with their realizations and Balleer et al. (2023) finds the same for employed workers.



Note: The figure shows a binned scatter plot of the mean of subjective expected earnings growth, $\mathbb{E}_i[g_i]$ and the corresponding realized earnings growth, Δy_i. We first divide the sample into 20 equal-sized bins and plot the empirical mean of the X and Y-axis. The regression coefficient of mean g_i on realized earnings growth is 1.03. For the plot, we focus on the range of mean g_i in between -0.5 and 0.5.

Figure 5: Comparing mean g_i and realized earnings growth

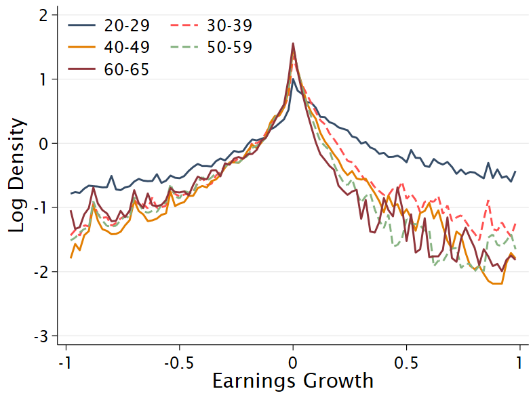
characteristics X:

$$h_S^X = \frac{1}{N^X} \sum_{i=1}^{N^X} g_i^X \tag{3}$$

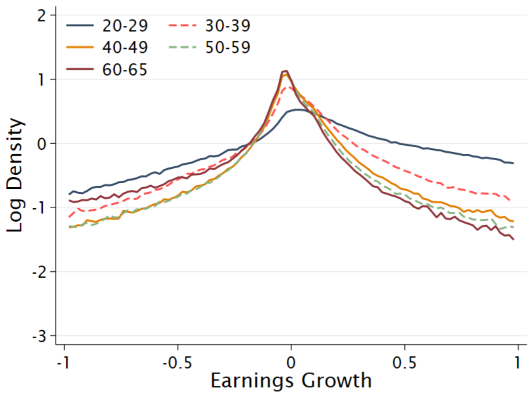
The pooled distribution, h_S^X, reflects the total variability of expected earnings growth in the population and it is thus directly comparable to the distribution of realizations earnings growth observed in the administrative data, which we denote h_A^X, for individuals who have similar observable characteristics, X.

We start out by plotting the distribution of pooled holistic expected earnings growth from the survey, h_S^X, cf. equation (3), within broad age groups and compare it to the corresponding distributions of realized earnings growth from the administrative data, h_A^X.[18]

[18] As a further check we compare the stated level of earnings in 2020 from the survey to the level of earnings recorded in the administrative data. In the survey, which is conducted in January 2021, we ask about total earnings throughout 2020. Information about total earnings for 2020 is also reported directly from employers to the tax agency and is made available in the administrative data. In Online Appendix C.4 we compare these two different measures of earnings in 2020, and it turns out that survey answers line up accurately with the administrative data. This is an indication that respondents are well-informed about their level of earnings.



(a) Survey, h_S^X



(b) Registry, h_A^X

Note: Panel (a) plots log density for the pooled distribution of expected holistic earnings growth rates from the survey, h_S^X, where X indicates partitions by age groups. (b) plots the distribution of annual earnings growth from 2020 to 2021 as observed in the administrative data for the full population, h_A^X. For constructing the distribution of earnings growth in the administrative data we dropped observations where the level of annual earnings is less than 24,000 DKK in 2020. Online Appendix C.5 shows the corresponding figures using administrative data for 2019.

Figure 6: Pooled earnings risk and registry earnings risk

Figure 6 panels (a) and (b) show the distributions of earnings growth in the survey and the registry. The two distributions are similar and have similar life cycle patterns. Generally, the distributions based on the survey data and the administrative data both have thicker and longer tails than a normal distribution. This is analogous to the patterns documented by Guvenen et al. (2021) for the US.[19] In the survey as in the administrative data, younger workers (age 20-29) tend to have a higher density of positive earnings growth and the log density is right tilted, arguably reflecting career progress for individuals in this age group. At older ages, the density of positive earnings growth decreases and the density of negative earnings growth increases. It is also notable that the peak around 0 earnings growth increases. The log density level for each age group is well aligned between the survey and the registry. Panel (b) is based on the full population. In Online Appendix C.6, we present the distribution of realized earnings growth observed in administrative data but only including the individuals from the survey and note that this too looks similar to the distributions shown in Figure 6.

In Figure 7 we examine how the moments of the distributions h_S^X and h_A^X evolve over the life cycle. The graph shows that life cycle patterns are broadly similar between the

[19] Leth-Petersen and Sæverud (2022) document that the distribution of realized earnings growth in Denmark share many of the features that are also observed in the US data.

survey data and the administrative data for all moments. Mean earnings growth, panel (a), decreases with age. Young workers, on average, expect and realize positive earnings growth while the oldest workers expect and realize negative earnings growth. Next, the interdecile range of the distribution of earnings growth, panel (b), is especially high for people in the 20s but is relatively stable after age 30. This means that young workers tend to associate their earnings growth with relatively more risk. Skewness, panel (c), is decreasing with age. One divergence between the survey and the registry is that for those of ages 30-50, skewness is negative while in the registry it is close to zero in the survey. Lastly, note that excess kurtosis, panel (b) increases in age. This means that the earnings growth distribution becomes more peaked and develops fatter tails as age increases.[20]

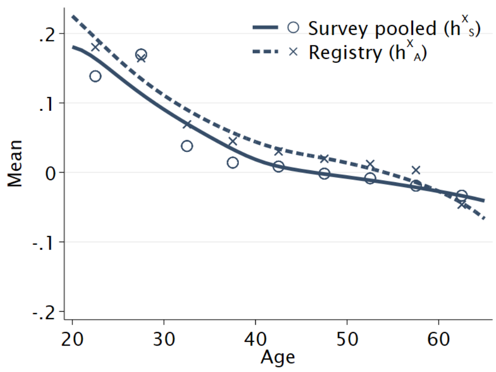
Overall, we find high coherence between the pooled distribution of expected earnings growth based on the survey data and the distribution of actual earnings growth recorded in the administrative data.[21]

4 Administrative versus Survey-Based Earnings Risk

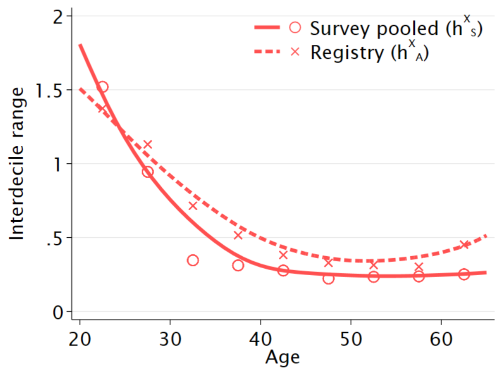
In this section we compare the distribution of realized earnings growth, often used to infer earnings risk, with subjective earnings risk directly measured in survey data. Guvenen et al. (2021), for example, group the population into three broad age groups and percentiles of earnings levels and examines the characteristics of the distribution of earnings growth within these cells. Obviously this method of inferring earnings risk from the moments of the cross sectional distribution of realized earnings growth comes with assumptions about worker homogeneity, i.e. that groups of workers draw earnings realizations from the same underlying distribution which can be characterized by the cross sectional distribution of realized earnings. To explore the validity of such homogeneity conditions, we analyze whether moments calculated from the distribution of earnings growth in the administrative data within these detailed partitions, h_A^X, are able to mimic the moments of the subjective distributions of holistic earnings growth within the same cells, g_i^X.

[20] In Online Appendix C.7, we document that also the standard moment measures (standard deviation, skewness, and kurtosis) show similar patterns in the registry and the survey.

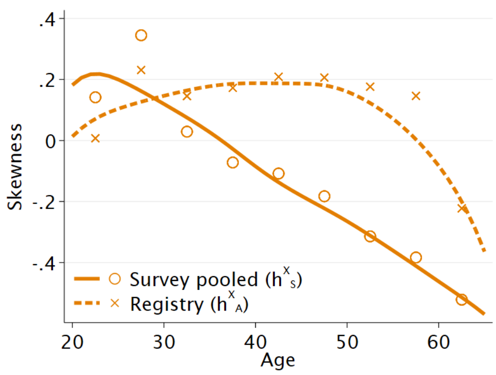
[21] The comparison is based on administrative data for 2021. In Online Appendix C.5 we repeat the exercise using administrative data for 2019, and find that the pattern is practically identical. This confirms that the survey year is not particular.



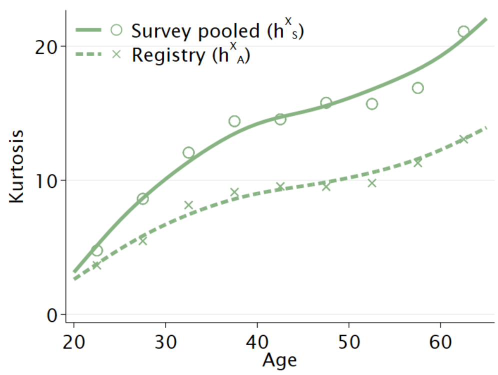
(a) Mean



(b) Interdecile range



(c) Skewness



(d) Kurtosis

Note: The figure shows the average value of the 1st to 4th quantile based moments (See notes to Figure 3) over the life cycle of the pooled earnings distribution in the survey, h_S^X, and in the administrative data, h_A^X. "o" and "x" represent the empirical mean across 5-years age bins for the survey and the administrative data for 2021, respectively. The lines are local linear polynomials, calculated using a bandwidth of 4 years. Online Appendix C.7 shows the corresponding figures using standard moments.

Figure 7: Higher-order moments of h_S^X and h_A^X over the life cycle

4.1 Coarse Stratification

We start out by illustrating the main insight based on a coarse partition of the administrative data which allows us to summarize the main insight graphically. We then implement a more detailed partition that is close to the most granular researchers could achieve with administrative data.

In the coarse stratification we divide the administrative data for the Danish population into six cells based on three age groups (20-34, 35-49, 50-65) and the earnings level being *High* or *Low* (above/below the median). Within each of these cells we calculate the moments of the cross sectional distribution of realized earnings growth from the administrative data, h_A^X. Each moment within a cell will be a unique number. Next, we calculate the corresponding moments for the pooled distribution of subjective holistic earnings growth expectations, h_S^X, cf. equation (3). Also in this case, each moment within a cell will be a number. If these two objects are similar then the survey data and the registry data are consistent with each other. We then compare with the cross sectional *distribution* of moments of the subjective distributions of holistic earnings growth, g_i^X, cf. equation (2), for individuals from the survey belonging to the cell.

Figure 8 presents estimates of mean earnings growth and interdecile range (IDR) for two of the groups in the coarse stratification described above. The first row is for the cell 20-34:*Low* and the second row is for the cell 50-65:*High*. The first column shows the estimates for the mean and the second column shows estimates of the interdecile range. Panel (a) shows that for the 20-34:*Low* group, the mean calculated from the administrative data, $\mathbb{E}[h_A^X]$, and the mean calculated from the pooled subjective distribution, $\mathbb{E}[h_S^X]$, are practically identical while there is considerable heterogeneity in the means of the subjective distributions, $\mathbb{E}_i[g_i^X]$. Panel (b) shows that the interdecile range estimated from the administrative data, $\text{IDR}[h_A^X]$, and the pooled survey data, $\text{IDR}[h_S^X]$, are also very close to each other, but that the subjective interdecile ranges, $\text{IDR}_i[g_i^X]$, are very heterogeneous and centered at much lower values than the interdecile range calculated from the administrative data and the pooled subjective data.

The second row of Figure 8 shows the corresponding figures for the 50-65:*High* group. The estimate of the mean, panel (c), and the interdecile range, panel (d), based on the administrative data and the pooled subjective data are also very similar. The modal point of the distribution of subjective interdecile ranges is also positioned lower than the estimate of the interdecile range based on the administrative data. However, this group displays less heterogeneity in subjective means and the distance between the estimate of the interdecile range based on the administrative data and the modal point of the distribution of subjective interdecile ranges is smaller than for the 20-34:*Low* group.

The main insight from Figure 8 is that risk inferred from the administrative data tends to be larger than risk inferred from the subjective data and that the degree of overshooting tends to be linked to how much dispersion there is in the distribution of subjective means. This is consistent with the view that the pooled distribution of expected earnings growth, h, is a mixture of underlying subjective distributions, g_i, cf. Equation (3). The theoretical variance of a mixture distribution of N equally weighted subjective distributions with individual means and variances μ_i, σ_i^2 is:

$$\text{Var}(h) = \frac{1}{N}\sum_i^N \sigma_i^2 + \frac{1}{N}\sum_i^N \mu_i^2 - \left(\frac{1}{N}\sum_i^N \mu_i\right)^2 \tag{4}$$

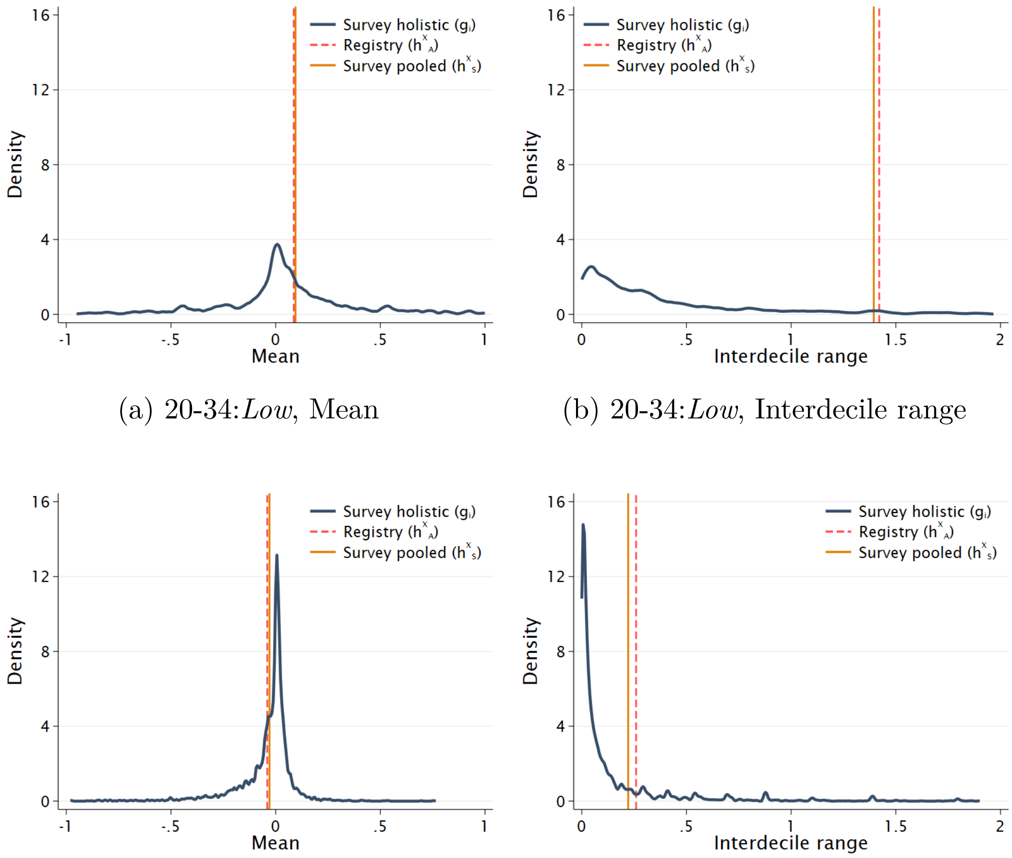
The variance of the mixture distribution is the mixture of the variances of the subjective distributions plus a non-negative term reflecting the differences in means between the subjective distributions. By Jensen's Inequality the average squared mean is weakly greater than the squared average mean, implying that the sum of the last two terms is non-negative and hence that the variance of the mixture distribution is weakly larger than the average variance of the subjective distributions, $\text{Var}(h) \geq \frac{1}{N}\sum_i^N \sigma_i^2$. Put differently, over-dispersion in the pooled holistic distribution, h_S^X, and by extension the distribution from which the registry based variance is calculated from, h_A^X, occurs when the underlying subjective holistic distributions, g_i, have heterogeneous means, and, as a result of this, risk and heterogeneity are confounded.[22]

4.2 Refining the Stratification

The logic above suggests that the gap between subjective and administratively estimated risk will be lower the more we refine the stratification of the population. To pursue this we now consider a finer stratification. Specifically we make use of the administrative data and partition the distribution of realized earnings growth in the population data into 300 cells by three age groups and earnings percentiles following Guvenen et al. (2021). For each of these cells we perform the same calculations as in the illustration above: We calculate the interdecile range of the distribution of realized earnings growth within each cell, $\text{IDR}[h_A^X]$, and the average of the subjective interdecile ranges within each of these cells, $\frac{1}{N^X}\sum_i^{N^X} \text{IDR}_i[g_i^X]$. The result is shown in Figure 9.[23]

[22] How skewness and kurtosis of the pooled distribution are related to skewness and kurtosis of the underlying subjective distributions is ambiguous. We refer to Online Appendix E.2 for derivations.

[23] in Online Appendix E.3, we report results for skewness and kurtosis.

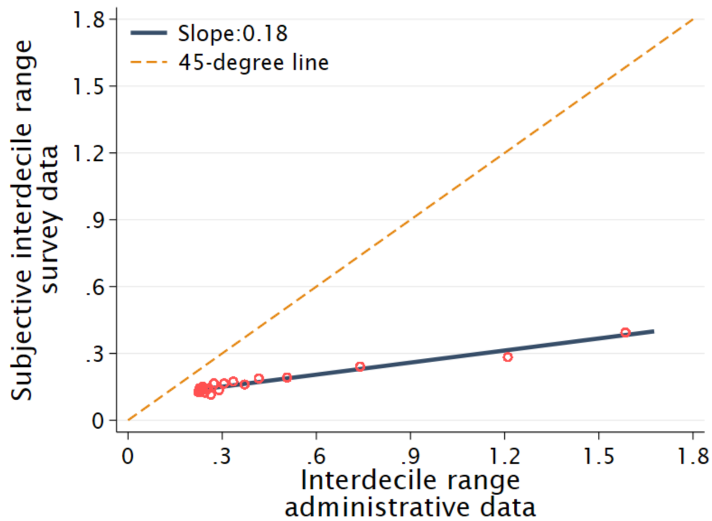


(a) 20-34:*Low*, Mean (b) 20-34:*Low*, Interdecile range

(c) 50-65:*High*, Mean (d) 50-65:*High*, Interdecile range

Note: The figure shows estimates of the mean and interdecile range for h_S^X, h_A^X, and the distribution of g_i for two subgroups in the data. The top row shows these statistics for individuals aged 20-34 and with below median earnings (20-34:*Low*), and the bottom row shows the corresponding statistics for individuals aged 50-65 and with above median earnings (50-65:*High*). Online Appendix E.1 show the corresponding figures for the remaining subgroups.

Figure 8: Mean and interdecile range of h_S^X and h_A^X, and the distributions of individual means and interdecile range of g_i for two selected subgroups



Note: The figure compares average interdecile ranges of subjective holistic earnings expectations, $\frac{1}{N^X}\sum_i^{N^X} \mathrm{IDR}_i[g_i^X]$, to interdecile ranges calculated from administrative data, $\mathrm{IDR}[h_A^X]$, within 300 cells divided by age groups (20-34, 35-49, 50-65) and earnings percentiles. The panel shows a binned scatterplot (red circles) of $\frac{1}{N^X}\sum_i^{N^X} \mathrm{IDR}_i[g_i^X]$ by vigintiles of $\mathrm{IDR}[h_A^X]$. A regression line based on the 300 data points is overlaid.

Figure 9: Comparing interdecile ranges calculated from subjective expectations and from administrative data

We find that the average of the subjective interdecile ranges, $\frac{1}{N^X}\sum_i^{N^X} \text{IDR}_i[g_i^X]$, within each cell is much smaller than the interdecile range calculated from the administrative data within the same cell, $\text{IDR}[h_A^X]$. We find that within each cell there is a lot of heterogeneity in the subjective mean growth rates (not reported). Consistent with the idea that the pooled distribution of earnings growth rates is a mixture of individual distributions of expected earnings growth rates, this finding suggests that heterogeneity is assigned to chance when earnings risk is inferred from the distribution of realized earnings growth and, as a consequence, that risk is systematically overstated compared to how the majority of individuals experience it.

Assigning heterogeneity to risk could potentially be the result of not applying a sufficiently fine partition by observable characteristics. In Online Appendix E.4 we present results for an even finer grid with 1,800 cells for age, earnings deciles, gender, and university education and find results that are practically identical. Furthermore, we also try a version where we include the individual growth rate of earnings in the covariate set. A branch of the literature assumes that individual earnings grow deterministically at an unobserved rate. This is known as the heterogeneous income profiles model (HIP, e.g., Guvenen, 2009; Browning et al., 2013). In order to account for this possible type of heterogeneity we construct an alternative version of Figure 9 where we expand the covariate set to include also the average growth rate of earnings within the past five years. This essentially allows for an individual fixed effect in the growth rate of earnings. The resulting figure is practically identical to Figure 9. These results are reported in Online Appendix E.5.

4.3 The Role of Observable Characteristics

The role of observable characteristics is to control for the otherwise unobservable heterogeneity in the means of the underlying subjective distributions, $\mathbb{E}_i[g_i]$, such that subjective heterogeneity and risk are not confounded when assessing risk from the distribution of realized earnings growth, cf. equation (4). In this section we examine how well observable characteristics predict heterogeneity, i.e., $\mathbb{E}_i[g_i]$, and how using observable characteristics as a way to control for heterogeneity affects the estimate of the interdecile range of the distribution of realized earnings growth, $\text{IDR}[h_A^X]$.

To do this we regress subjective means, $\mathbb{E}_i[g_i]$, on observable characteristics, X_i. We do this for a range of different covariate sets and for each of the covariate sets, we calculate the average of interdecile ranges of the cross sectional distribution of realized earnings growth from the administrative data within X cells, $\text{IDR}[h_A^X]$. To assess how well

Table 1: Regressions of $\mathbb{E}_i[g_i]$ on observable characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable : $\mathbb{E}_i[g_i]$						
Age group indicators	✓	✓	✓	✓	✓	✓
Earnings percentile indicators		✓	✓	✓	✓	✓
Female			✓	✓	✓	✓
Education				✓	✓	✓
Past earnings growth quintile indicators					✓	✓
Unemployment						✓
Industry						✓
Observation	10,746	10,746	10,746	10,746	10,746	10,746
R-squared	0.015	0.028	0.028	0.028	0.028	0.029
Average $\mathrm{IDR}_i[g_i]$			0.179			
$\mathrm{IDR}[h_A]$			0.630			
Average $\mathrm{IDR}[h_A^X]$	0.621	0.602	0.599	0.593	0.591	0.573

Note: The table shows regressions of $\mathbb{E}_i[g_i]$ on observable characterstics. Age group indicators are dummy variables for 5 years age bin. Earnings percentiles include dummies for earning level percentiles. The percentiles are calculated within age groups. Education is a dummy for having completed a college degree. For past earnings growth, we calculate average earnings growth from 2016-2020 and construct quintiles within age group. Unemployment is a dummy variable, taking the value 1 if workers were without an employer for at least one month during 2020. Additionally, the industry classification captures private and public sector employment. $\mathrm{IDR}[h_A]$ is the interdecile range of h_A, i.e. the pooled cross section of earnings growth. Average $\mathrm{IDR}[h_A^X]$ is the average of interdecile ranges calculated from the cross sectional distributions of realized earnings growth in the administrative data within X cells. A weighted average is used to compute the Average $\mathrm{IDR}[h_A^X]$, taking into account the different sample sizes across each cell of X.

observable characteristics control for heterogeneity, we compare with the interdecile range of the pooled cross sectional distribution of realized earnings, i.e., constructed without controlling for observable characteristics, $\mathrm{IDR}[h_A]$, and with the average of the subjective interdecile ranges, $\mathrm{IDR}_i[g_i]$.

The result is shown in Table 1. In column (1), $\mathbb{E}_i[g_i]$ is regressed on a full set of age group dummies. The R^2 is 0.015 meaning that age group dummies are able to explain only 1.5% of the variation in $\mathbb{E}_i[g_i]$. As a result, $\mathrm{IDR}[h_A^X] = 0.621$ is only slightly lower than $\mathrm{IDR}[h_A] = 0.630$ and much bigger than the average subjective interdecile range, $\mathrm{IDR}_i[g_i^X] = 0.179$. In column (2) we add percentile dummies for the level of earnings. This increases R^2 to 0.028 and reduces $\mathrm{IDR}[h_A^X] = 0.602$. In column (3)-(6) we sequentially add a dummy for being female, a dummy variable for having a college degree, quintile dummies for the average earnings growth during 2016-2020, i.e., the past five years, and in the last column, we incorporate a dummy variable for unemployment, which indicates if workers experienced at least one month of unemployment during 2020, along with industry classification differentiating between private and public sectors. In all cases are the covariates significant in explaining the variation in $\mathbb{E}_i[g_i]$, but collectively the covariates explain only a small fraction of the variation in $\mathbb{E}_i[g_i]$. As a result, for the richest specification in column (6), $\text{IDR}[h_A^X] = 0.573$ which is still only slightly lower than $[h_A] = 0.630$ and much bigger than the average subjective interdecile range, $\text{IDR}_i[g_i^X] = 0.179$. In sum, observable characteristics have a significant but limited ability to capture the heterogeneity in the means of the underlying subjective distributions, and this explains our key finding that risk inferred from the cross-sectional distribution of earnings growth systematically overstates the level of subjective risk that the majority of individuals experience, even when taking into account heterogeneity along observable characteristics. As individuals have approximately mean correct expectations about their earnings growth, cf. Figure 5, this indicates that individuals have more information than the econometrician who only observes administrative data on realized earnings growth across the population and therefore cannot separate risk from heterogeneity.

5 Job Transitions and Subjective Earnings Risk

In this section, we take advantage of our conditional survey instrument to decompose the subjective holistic earnings growth distributions, g_i, according to job transitions and show that such transitions are key in explaining the level and heterogeneity of higher-order moments. To illustrate this we compute not only the average life cycle patterns of the four moments of the subjective holistic earnings growth distributions, g_i, but also the subjective risk arising from staying in the current job, f_i^S. Figure 10 illustrates the results, which confirm the great importance of job transitions for earnings risk.

Panel (a) shows average mean earnings growth across the life cycle. Generally, mean earnings growth decreases as the life cycle progresses and this is the case for both holistic earnings growth and for earnings growth conditional on staying. Holistic earnings growth is, on average, positive up to about age 50 and then turns negative. Fixing earnings risk to the stay branch increases expected earnings growth for all ages, and this happens to a degree where also the oldest workers expect positive earnings growth, i.e., the net contribution of job transitions is to reduce expected earnings growth.

Panel (b) shows how the average subjective risk, which we measure as the average interdecile range, pertaining to earnings growth over the life cycle. Considering risk based on subjective holistic earnings growth expectations we find that risk is highest for young people. Fixing earnings growth risk to come only from the stay branch generates a big drop in risk at all ages, but most dramatically for the young. This shows that risk pertaining to one-year ahead earnings growth is intimately tied to job transitions.

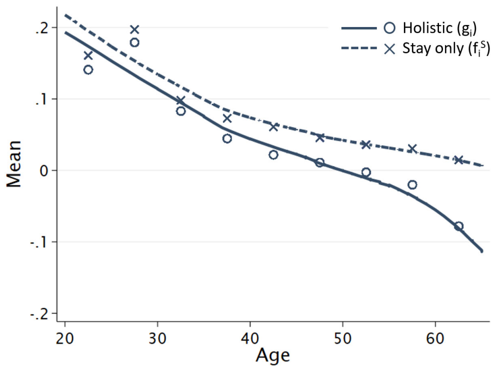
In Panel (c) we consider skewness. For all ages there is, on average, negative skewness in the subjective holistic distributions. Yet when quantifying skewness only from the stay branch, it is close to zero. Negative skewness appears when people expect to disproportionately draw large negative shocks and it indicates that job transitions are, in expectation, responsible for the downside risk that people face.

Finally, in panel (d) we consider kurtosis. According to the holistic measure of subjective earnings growth, kurtosis is significant at a level of about 10-20 and it is increasing in age. When removing risk stemming from job transitions, kurtosis is practically removed. This is consistent with the notion that extreme earnings growth derives from job transitions.

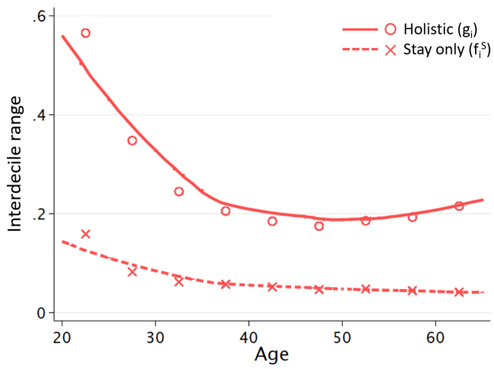
Thus, we find that job transitions are essential in determining life-cycle patterns of higher order moments of subjective holistic one-year ahead earnings risk. This finding is the analog to the differences between job switchers' and job stayers' earnings growth in realizations from administrative data (Guvenen et al., 2021). Our takeaway is that this factor, whether someone expects to make a job transition, also translates over to beliefs about earnings growth.

6 Subjective Earnings Risk and Low Liquid Savings

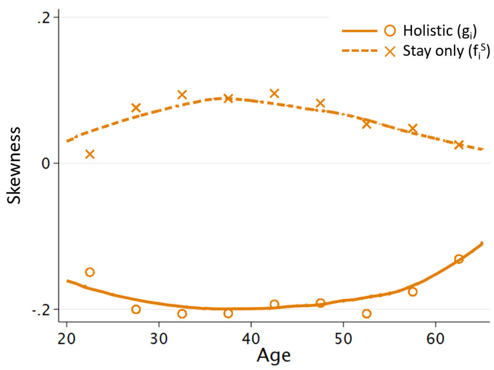
Our new survey data provides two key insights regarding subjective earnings risk. First, average subjective earnings risk in the population is several times lower than risk estimates based on recorded realized earnings growth from administrative data. Second, we observe that subjective risk is highly heterogeneous across the population. In this section, we illustrate the implications of these findings for limited liquid asset accumulation within the framework of a standard incomplete markets savings model (Aiyagari, 1994; Huggett, 1993). In such models, agents face un-insurable idiosyncratic earnings risk and have a precautionary savings motive, and as a result endogenously accumulate assets. When calibrated using risk estimates derived from administrative data on earnings growth, these models struggle to explain why so many individuals in the real world hold limited liquid assets. Our results suggest that models based on these traditional calibrations tend to overestimate the precautionary savings motive, thereby generating higher liquid asset



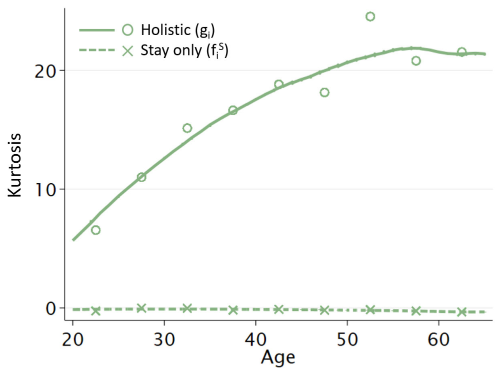
(a) Mean



(b) Interdecile range



(c) Skewness



(d) Kurtosis

Note: The figure shows the average value of the 1st to 4th quantile based moments (See notes to Figure 3) over the life cycle of holistic earnings risk, g_i, and risk conditional on staying, f_i^S. “o” and “x” represent the empirical mean across 5-years age bins. The lines are local linear polynomials, calculated using a bandwidth of 4 years. Online Appendix C.8 show the corresponding figure using standard moments instead of quantile based moments.

Figure 10: Moments of holistic earnings risk, g_i, vs. risk conditional on staying, f_i^S, over the life cycle

holdings than observed in the data.[24]

For simplicity, and in line with most of the literature, we assume that overall earnings risk is governed by a standard earnings process with permanent and transitory shocks. This is the work-horse model of earnings risk that allow us to illustrate the consequences of our findings for precautionary savings in the simplest possible manner. We first calibrate the model to an earnings process that is quantified solely from administrative data as is standard practice in the literature. We then re-calibrate the earnings process to include the features of the cross sectional distribution of subjective earnings risk, $var(g_i)$, which summarizes all sources of risk from our survey instrument including risk from job transitions.

6.1 The Model

In this section, we describe the consumption-savings model, how we calibrate it including how we introduce subjective earnings risks into the model, and finally we present the results.

6.1.1 Consumption-savings problem

The impact of earnings risk is governed by the earnings process. We characterize earnings risk using a standard earnings process where the log earnings (y_t^i) for individual i at age t is represented by a deterministic component and a stochastic component.

$$y_t^i = \underbrace{\gamma_0 + \gamma_1 t + \gamma_2 t^2}_{\text{life cycle}} + \underbrace{(z_t^i + \epsilon_t^i)}_{\text{stochastic}} \tag{5}$$

The first life cycle component in equation (5) captures the average log earnings dynamics across life stages. The second component is stochastic and is assumed to have a permanent, z_t^i, and a transitory component, ϵ_t^i, where $z_t^i = z_{t-1}^i + \eta_t^i$ with $z_0^i = 0$ and where $\eta_t^i \sim \mathcal{N}(\mu_\eta, \sigma_\eta^2)$ and $\epsilon_t^i \sim \mathcal{N}(\mu_\epsilon, \sigma_\epsilon^2)$, i.e. both are independent and identically distributed over time and across individuals. An individual works for the first T years of her life and lives until L ($> T$) years. From year $T+1$ to L, after retirement, workers receive retirement income as a fraction of their earnings before retirement with a certain guaranteed

[24] Other modelling approaches have been applied to reduce the accumulation of liquid assets in incomplete markets models. One approach introduces preference heterogeneity (e.g., Krusell and Smith, 1998; Carroll et al., 2017). This leads to varying saving behaviors and consequently, a larger fraction of the population with low asset holdings. Another introduces an illiquid asset that offers a higher return than the liquid asset, thereby encouraging agents in the model to allocate resources into the illiquid asset to a degree where they become wealthy-hand-to-mouth (Kaplan and Violante, 2014).

minimum level and no uncertainty.

Preferences over consumption follow the constant relative risk aversion (CRRA) utility function with the parameter ϕ. The constant and exogenous interest rate of the asset is r and δ denotes a time discount factor. In each period of time t, an individual chooses consumption c_t^i and savings for the next periods a_{t+1}^i. The cash-on-hand (savings and earnings) is represented as x_t^i. The value function representation is as follows.

$$V_t^i(x_t^i, z_t^i) = \max_{c_t^i, a_{t+1}^i} \left\{ \frac{c_t^{i\,1-\phi}}{1-\phi} + \delta \mathbb{E}_t[V_{t+1}^i(x_{t+1}^i, z_{t+1}^i)] \right\} \tag{6}$$

$$\text{s.t.} \quad c_t^i + a_{t+1}^i = x_t^i \tag{7}$$

$$x_t^i = (1+r)a_t^i + exp(y_t^i) \tag{8}$$

$$a_{t+1}^i \geq k \tag{9}$$

for $t = 1, ..., T-1$, where V_t^i is the value function of individual with age t. We assume that people face a borrowing constraint, $a_t^i \geq k$. Online Appendix F.2 provides details about retirement income and the value function for retired workers from year $T+1$ to J.

6.1.2 Calibration

In order to give the model quantitative content, we calibrate the preference parameters to values commonly used and then solve the model using different calibrations of the earnings process. We first calibrate it to estimates based purely on administrative data. We then calibrate it to the subjective earnings risk from our survey.

Life cycle. The model frequency is annual. The working age spans 20 to 67 ($T = 47$) and the agent dies with certainty at the age of 75 ($L = 55$).

Preference. The coefficient of the relative risk aversion parameter, ϕ, is set to 2.0 which is a commonly used value in the literature. The discount factor, δ, is set to 0.96, which is also a typical value applied in the literature.

The earnings process calibrated to administrative data. We first estimate average life cycle parameters by regressing log earnings on age and age-squared. The regression coefficients are reported as $(\gamma_0, \gamma_1, \gamma_2)$ in Table 2. Using the variance and covariance structure of residualized log earnings, we estimate the parameters of the earnings process, $(\sigma_\eta^2, \sigma_\epsilon^2)$. The details of this estimation step and data are reported in Online Appendix F.1. We find $(\hat{\sigma}_\eta^2, \hat{\sigma}_\epsilon^2) = (0.043, 0.018)$. In the baseline, the parameters for the mean of permanent and transitory shocks are set to 0: $(\hat{\mu}_\eta, \hat{\mu}_\epsilon) = (0, 0)$, since log earnings

Table 2: Baseline parameterization

Block	Parameter	Values	Source
Earnings process	γ_0	11.815	Registry estimated
Earnings process	γ_1	0.082	Registry estimated
Earnings process	γ_2	-0.0014	Registry estimated
Earnings process	σ_ϵ^2	0.018	Registry estimated
Earnings process	σ_η^2	0.043	Registry estimated
Earnings process	μ_ϵ	0.0	Registry estimated
Earnings process	μ_η	0.0	Registry estimated
Life cycle	T	47	Retirement age at 67
Life cycle	L	55	Live up until 75
Preference	ϕ	2	Standard assumption
Preference	δ	0.96	Standard assumption
Asset return	r	0.03	Standard assumption
Borrowing constraint	k	0	No borrowing constraint

Note: These are the parameters we use for the simulation. In Online Appendix F.2, we report parameters for the retirement phase.

have been residualized.

Table 2 summarizes the parameters used in the calibration of the baseline model.

Subjective earnings risk. In our survey, we measure total subjective earnings risk, $\sigma_{g,i}^2$. Similar to Wang (2023), we assume that this corresponds to the sum of the permanent and transitory component, $\sigma_\eta^2 + \sigma_\epsilon^2$, from the earnings process specified above. In our survey, we only observe total risk and we therefore assume that total subjective risk can be divided into a transitory and a permanent component based on the ratio of σ_η^2 and σ_ϵ^2 as estimated from the administrative data. Suppose $\sigma_{g,i}^2$ represents the survey-measured variance of g for individual i. We then decompose the total subjective earnings risk, $\sigma_{g,i}^2$ into a permanent and transitory component using the scaling parameter α_i and the registry-measured permanent/transitory variance parameters, σ_η^2 and σ_ϵ^2:

$$\sigma_{g,i}^2 = \alpha_i(\sigma_\eta^2 + \sigma_\epsilon^2) \tag{10}$$

We interpret $\alpha_i \times \sigma_\eta^2$ as subjective permanent shocks parameter and $\alpha_i \times \sigma_\epsilon^2$ as subjective transitory shocks parameters.

Figure 11 shows the distribution of subjective earnings risk, $var(g_i)$. Most workers' earnings risk is very low and much lower that what is inferred from the administrative register data which is indicated by the vertical yellow line. Equation (10) simply scales the earnings process estimated from administrative data to match the average level of risk that we observe in the survey. This level is illustrated in Figure 11 by the vertical dashed line. The average of $var(g_i)$ is around 0.022, which is approximately 36% of the registry benchmark. Consequently, the scaling parameter α, is set to 0.36 for this version of the earnings process which we call "Subjective 1-G".

However, this specification still overstates risk for the majority and it assumes that all agents face the same level of earnings risk, i.e., it does not exploit the vast heterogeneity in subjective earnings risk that is present in the survey data. To address this, we also calibrate the model to earnings processes that are designed to capture the heterogeneity in $var(g_i)$. To do this, we divide the distribution of $var(g_i)$ into K equally sized groups and calculate scaling factors, α_i^K to match the average level of subjective risk within each group. We then simulate the model for K groups where each group have a level of earnings risk that matches the average in the Kth group. We do this for $K = (3, 10)$ which we denote "Subjective 3-G" and "Subjective 10-G". In Online Appendix F.3, we report the values of α_i^K for Subjective 3-G and 10-G, respectively.

We further allow heterogeneity in the mean of permanent and transitory shocks in each group. We first compute the average of the subjective mean of g ($\mu_{g,i}^K$) in each group, K. We then decompose this into the mean of permanent shocks with ratio $\frac{\sigma_\eta^2}{\sigma_\epsilon^2+\sigma_\eta^2}$ and the mean of transitory shocks according to the ratio $\frac{\sigma_\epsilon^2}{\sigma_\epsilon^2+\sigma_\eta^2}$. We apply the same approach to decompose the subjective variance. In Online Appendix F.3, we report the values of $\mu_{g,i}^K$ for Subjective 1-G, 3-G, and 10-G, respectively.

In summary, we simulate 100,000 workers for each of four different earnings processes.

1. Registry

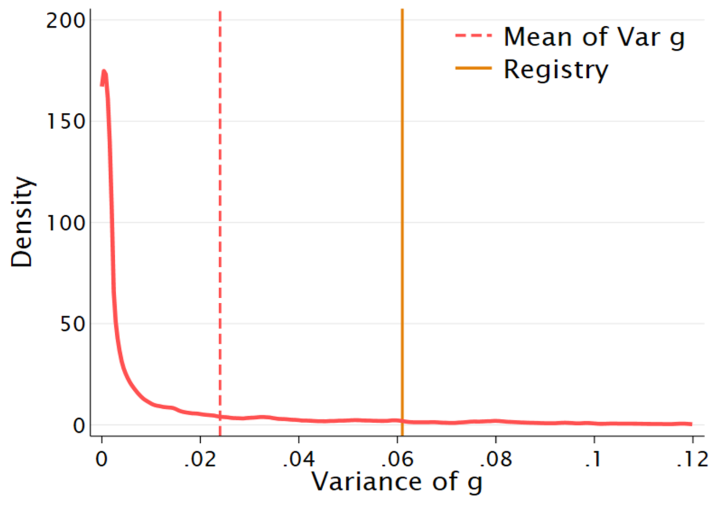
2. Subjective 1-G: Mean of α_i and $\mu_{g,i}$

3. Subjective 3-G: Mean of α_i and $\mu_{g,i}$ across terciles of $\sigma_{g,i}^2$

4. Subjective 10-G: Mean of α_i and $\mu_{g,i}$ across deciles of $\sigma_{g,i}^2$

6.1.3 Results

Figure 11 shows that subjective earnings risk is far smaller than what is suggested by an earnings process estimated to match administrative data on earnings growth. This suggest that most people need much less precautionary savings than what is implied from a model calibrated to a registry-matched earnings process. We start out from the observation that about half of the population hold limited liquid assets. Specifically, we identify workers whose liquid assets compare to less than 30% of their annual earnings and refer to them



Note: The figure presents the distribution of the variance of subjective earnings risk, $var(g_i)$ from the survey (solid red line). The vertical dashed line shows the average value of $var(g_i)$, which is used for calibrating the 1-G model of earnings risk. The solid yellow vertical line shows the level of earnings risk estimated from administrative data only.

Figure 11: The distribution of subjective earnings risk $var(g_i)$

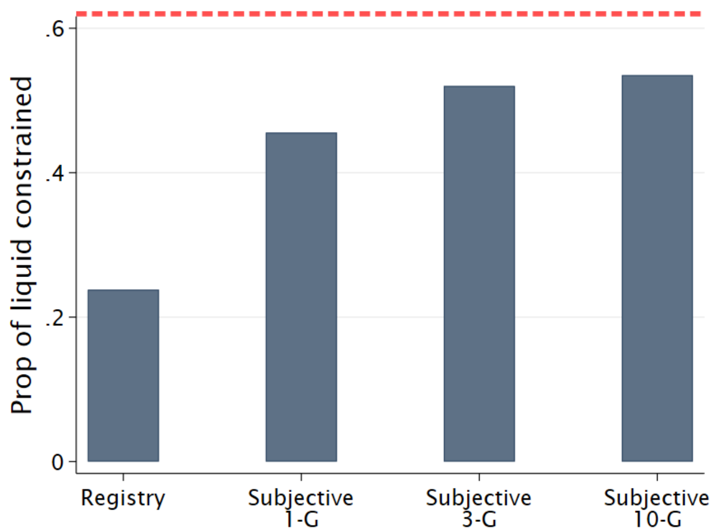
as low liquid asset holders. According to the administrative data, approximately 62% of workers are low liquid asset holders. We then investigate the model's ability to rationalize this fact when calibrated to the four different earnings processes outlined above.

The result is shown in Figure 12. The vertical bars show the fraction of individuals who are low liquid asset holders when the model is calibrated to the four different earnings processes. As a reference, the figure displays the fraction of low liquid asset holders observed in the data by the horizontal line. The first bar shows the fraction of low liquid asset holders resulting from the model where the earnings process is estimated to match administrative data. As shown in Figure 11, this estimate of earnings risk overstates risk when compared to the level of subjective earnings risk that we document in our survey. As a result, only 23% of workers are low liquid asset holders in the model. The second bar shows the fraction of low liquid asset holders when the model is calibrated to match the average level of earnings risk in the survey. In this version 45% are low liquid asset holders. This indicates that introducing subjective parameters helps reduce the gap between the empirical moment and the simulations. However, this version ignores the pervasive heterogeneity in earnings risk that we document in the survey. When we introduce heterogeneity in subjective earnings risk, the proportion of low liquid asset holders increases further. In the Subjective 3-G and 10-G simulations, 49% and 53% are low liquid asset holders. This happens because we allow more and more agents to have low and lower levels of risk as we add more groups. These findings suggests that calibrating the earnings process to match the level and heterogeneity in subjective earnings risk can go far in rationalizing why a large fraction of the population are low liquid asset holders: many hold limited liquid assets because they face very limited earnings risk.

In order to verify that the model featuring heterogeneity is consistent with the data, we investigate whether it is able to replicate the cross-sectional distribution of earnings growth observed in the administrative data. To do this, we simulate the distribution from all versions of the model and compare them to the empirical distribution. The results are shown in Figure 13.

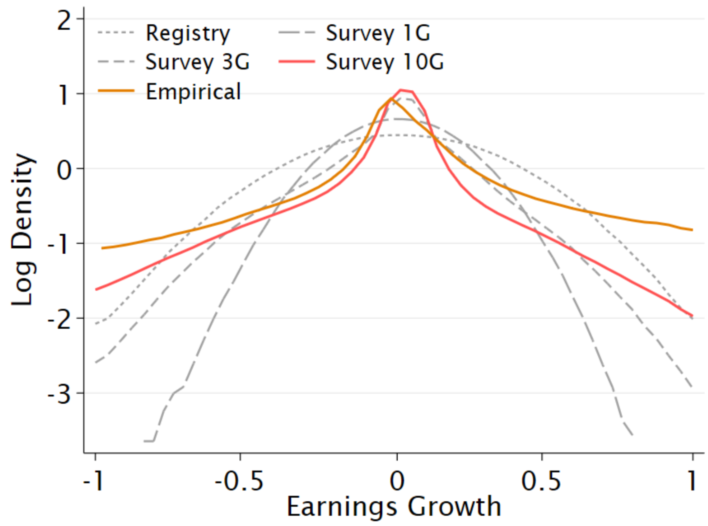
The empirical distribution, shown in orange, is sharply peaked and has fat tails. In contrast, the distribution simulated from the baseline model—where a standard permanent-transitory earnings process is calibrated to administrative data—exhibits a smooth, inverted U-shape and fails to replicate the peakedness and tail thickness of the empirical distribution. Similarly, the distribution resulting from the model calibrated to match the mean and variance of subjective earnings growth (1-G) also yields a log-density that is too smooth. However, as we introduce more heterogeneity into the model, the simulated distributions begin to resemble the empirical one. With three groups (3-G), the distribution becomes more peaked and shows heavier tails. With ten groups (10-G), the simulated distribution resembles the empirical distribution even more closely, capturing both its peakedness and tail thickness.

In a companion paper, Caplin et al. (2024), we show how the distribution of subjective earnings risk in Figure 11 can be simulated from administrative data on job transitions, earnings levels and tenure. This result means that the level and heterogeneity of subjective earnings risk can be inferred even without access to survey data about subjective earnings risk. Using this insight, we also simulate the distribution of earnings risk for the US based on data from the Survey of Income and Program Participation (SIPP) and find that it looks quite similar to distribution in Denmark. Moreover, as our estimates of the permanent-transitory earnings processes based on administrative data on earnings growth are quite similar to estimates based on US data (e.g., Meghir and Pistaferri, 2004), this suggests that the scaling factors that we present here may, as a first order approximation, also be applied to the US context.



Note: The horizontal dashed line represents the empirical data. Each bar shows the proportion of liquid-constrained workers in each simulation.

Figure 12: The proportion of liquid constrained samples



Note: The figure plots with orange the log density of annual earnings growth from 2020 to 2021 as observed in the administrative data for the full population, h_A. It also plots the log density of annual earnings growth as simulated from the model with four different versions of the earnings process: 1. Registry calibrated (gray short dash), 2. Subjective 1-G (gray long dash), 3. Subjective 3-G (gray dash), and 4. Subjective 10-G (red).

Figure 13: The distribution of earnings growth in administrative data and in the model

7 Conclusion

We introduce a survey instrument that measures earnings risk. A key feature of our instrument is that it conditions on possible job transitions, i.e., whether people stay in their current job, quit or are laid off. A link with administrative data provides many credibility checks. We show evidence that subjective expectations about earnings and job transitions are consistent with earnings realizations when appropriately aggregated. Strikingly, this is true across the age distribution, even though job transitions and earnings growth vary significantly across age.

The linked survey and administrative data reveals *subjective* earnings risk to be heterogeneous across the population and two to six times lower than its counterpart estimated from administrative data on realized earnings growth alone. This is because expected earnings growth is heterogeneous and is confounded with risk when inferred from realized earnings growth even when calculated within narrow cells defined by observable characteristics. This finding shows that while observable characteristics, including among other things age, education, the earnings level and past earnings growth, are significant in predicting heterogeneity in earnings growth, such characteristics do not account adequately for the relevant heterogeneity.

Our findings highlight the value of using survey-based measures of subjective earnings expectations to understand the nature of labor market and earnings risk. Among other things, this has implications for understanding and modelling savings behavior. We show that people who face a low level of risk need limited precautionary savings and our findings thus help explain why many households hold limited liquid assets.

Acknowledgements: We thank Anmol Bhandari, Andres Blanco, Martin Browning, Charles Manski, Guido Menzio, Christian Moser, Andreas Mueller, Stijn Van Nieuwerburgh, Serdar Ozkan, Todd Schoellman, Johannes Stroebel, Mirko Wiederholt, and Basit Zafar, as well as seminar/conference participants at Dartmouth, the St. Louis Fed, University of Copenhagen, University of Michigan, LAEF Labor Markets and Macroeconomic Outcomes Conference, Society for Economic Dynamics, NBER Summer Institute 2023, the 7th Workshop on Subjective Expectations, the 3D-In-Macro workshop at Oslo BI, USC, CESifo Venice Summer Institute 2024, and the Minneapolis Fed for comments. The work presented in this paper benefited from financial support from Center for Economic Behavior and Inequality, CEBI, a center of excellence at the University of Copenhagen financed by grant DNRF134 from the Danish National Research Foundation. Leth-Petersen is affiliated with the Danish Finance Institute which financially supports researchers in finance. Caplin thanks the Alfred P. Sloan Foundation and the NOMIS Foundation for their support for the research on the cognitive foundations of economic behavior and acknowledges support for this research from Sloan Foundation Grant 2015-14209. The views expressed herein are solely those of the authors and do not necessarily reflect those of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

References

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics 109*(3), 659–684.

Andersen, H. Y. and S. Leth-Petersen (2021). Housing wealth or collateral: How home value shocks drive home equity extraction and spending. *Journal of the European Economic Association 19*(1), 403–440.

Arcidiacono, P., V. J. Hotz, A. Maurel, and T. Romano (2020). Ex ante returns and occupational choice. *Journal of Political Economy 128*(12), 4475–4522.

Balleer, A., G. Duernecker, S. Forstner, and J. Goensch (2021). The effects of biased labor market expectations on consumption, wealth inequality, and welfare.

Balleer, A., G. Duernecker, S. Forstner, and J. Goensch (2023). Wage bargaining and labor market policy with biased expectations.

Bick, A., A. Blandin, and K. Mertens (2021). Work from home before and after the covid-19 outbreak. *Available at SSRN 3786142*.

Browning, M., M. Gørtz, and S. Leth-Petersen (2013). Housing wealth and consumption: a micro panel study. *The Economic Journal 123*(568), 401–428.

Busch, C., D. Domeij, F. Guvenen, and R. Madera (2022). Skewed idiosyncratic income risk over the business cycle: Sources and insurance. *American Economic Journal: Macroeconomics 14*(2), 207–42.

Caldwell, S., I. Haegele, and J. Heining (2024). Firm pay and worker search. Available at: https://www.sydneecaldwell.com/research.

Campbell, D., A. Carruth, A. Dickerson, and F. Green (2007). Job Insecurity and Wages. *The Economic Journal 117*(518), 544–566.

Caplin, A., V. Gregory, E. Lee, S. Leth-Petersen, and J. Sæverud (2024). Inferring subjective earnings risk from labor market outcomes: A life-cycle search model approach. Available at: https://sites.google.com/view/sorenleth/working-papers.

Carroll, C., J. Slacalek, K. Tokuoka, and M. N. White (2017). The distribution of wealth and the marginal propensity to consume. *Quantitative Economics 8*(3), 977–1020.

Chetty, R., J. N. Friedman, S. Leth-Petersen, T. H. Nielsen, and T. Olsen (2014). Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from denmark. *The Quarterly Journal of Economics 129*(3), 1141–1219.

Conlon, J. J., L. Pilossoph, M. Wiswall, and B. Zafar (2018). Labor market search with imperfect information and learning. Technical report, National Bureau of Economic Research.

De Bruin, W. B., C. F. Manski, G. Topa, and W. van der Klaauw (2011). Measuring consumer uncertainty about future inflation. *Journal of Applied Econometrics 26*(3), 454–478.

Delavande, A. and S. Rohwedder (2008). Eliciting subjective probabilities in internet surveys. *Public Opinion Quarterly 72*(5), 866–891.

Dominitz, J. (1998). Earnings expectations, revisions, and realizations. *Review of Economics and statistics 80*(3), 374–388.

Dominitz, J. (2001). Estimation of income expectations models using expectations and realization data. *Journal of Econometrics 102*(2), 165–195.

Dominitz, J. and C. F. Manski (1997). Using expectations data to study subjective income expectations. *Journal of the American statistical Association 92*(439), 855–867.

Druedahl, J. and A. Munk-Nielsen (2020). Higher-order income dynamics with linked regression trees. *The Econometrics Journal 23*(3), S25–S58.

Epper, T., E. Fehr, H. Fehr-Duda, C. T. Kreiner, D. D. Lassen, S. Leth-Petersen, and G. N. Rasmussen (2020, April). Time discounting and wealth inequality. *American Economic Review 110*(4), 1177–1205.

Faberman, R. J., A. I. Mueller, A. Şahin, and G. Topa (2022). Job search behavior among the employed and non-employed. *Econometrica 90*(4), 1743–1779.

Goldstein, D. G. and D. Rothschild (2014). Lay understanding of probability distributions. *Judgment & Decision Making 9*(1).

Guiso, L., T. Jappelli, and L. Pistaferri (2002). An empirical analysis of earnings and employment risk. *Journal of Business & Economic Statistics 20*(2), 241–253.

Guvenen, F. (2009). An empirical investigation of labor income processes. *Review of Economic dynamics 12*(1), 58–79.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2021). What do data on millions of us workers reveal about lifecycle earnings dynamics? *Econometrica 89*(5), 2303–2339.

Guvenen, F., L. Pistaferri, and G. L. Violante (2022). Global trends in income inequality and income dynamics: New insights from grid. *Quantitative Economics 13*(4), 1321–1360.

Halvorsen, E., H. A. Holter, S. Ozkan, and K. Storesletten (2023). Dissecting idiosyncratic earnings risk. *Journal of the European Economic Association.*

Hartmann, I. M. and S. Leth-Petersen (2024). Subjective unemployment expectations. *Labour Economics.*

Hendren, N. (2017). Knowledge of Future Job Loss and Implications for Unemployment Insurance. *American Economic Review 107*(7), 1778–1823.

Huggett, M. (1993). The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control 17*(5), 953–969.

Hvidberg, K. B., C. T. Kreiner, and S. Stantcheva (2023). Social positions and fairness views on inequality. *The Review of Economic Studies 86*(1), 253–269.

Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan (1993). Earnings losses of displaced workers. *The American Economic Review*, 685–709.

Jäger, S., C. Roth, N. Roussille, and B. Schoefer (2022). Worker beliefs about outside options. Technical report, National Bureau of Economic Research.

Kaplan, G. and G. L. Violante (2014). A model of the consumption response to fiscal stimulus payments. *Econometrica 82*(4), 1199–1239.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica 79*(3), 651–692.

Koşar, G. and H. W. van der Klaauw (2023). Workers' perceptions of earnings growth and employment risk. Working Paper 1056, Federal Reserve Bank of New York.

Krusell, P. and A. A. Smith (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy 106*(5), 867–896.

Leth-Petersen, S. (2010). Intertemporal consumption and credit constraints: Does total expenditure respond to an exogenous shock to credit? *American Economic Review 100*(3), 1080–1103.

Leth-Petersen, S. and J. Sæverud (2022). Inequality and dynamics of earnings and disposable income in denmark 1987–2016. *Quantitative Economics 13*(4), 1493–1526.

Low, H., C. Meghir, and L. Pistaferri (2010). Wage risk and employment risk over the life cycle. *American Economic Review 100*(4), 1432–67.

Manski, C. F. and J. D. Straub (2000). Worker perceptions of job insecurity in the mid-1990s. *The Journal of Human Resources 35*(3), 447–479.

Meghir, C. and L. Pistaferri (2004). Income variance dynamics and heterogeneity. *Econometrica 72*(1), 1–32.

Mueller, A. I., J. Spinnewijn, and G. Topa (2021). Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence, and bias. *American Economic Review 111*(1), 324–63.

Pistaferri, L. (2001). Superior information, income shocks, and the permanent income hypothesis. *Review of Economics and Statistics 83*(3), 465–476.

Pistaferri, L. (2003). Anticipated and unanticipated wage changes, wage risk, and intertemporal labor supply. *Journal of Labor Economics 21*(3), 729–754.

Stephens, M. (2004). Job Loss Expectations , Realizations , and Household Consumption Behavior. *The Review of Economics and Statistics 86*(1), 253–269.

Stoltenberg, C. A. and A. Uhlendorff (2022). Consumption choices and earnings expectations: Empirical evidence and structural estimation. Working Paper 15443, IZA.

Topel, R. H. and M. P. Ward (1992). Job mobility and the careers of young men. *The Quarterly Journal of Economics 107*(2), 439–479.

Von Wachter, T., J. Song, and J. Manchester (2009). Long-term earnings losses due to mass layoffs during the 1982 recession: An analysis using us administrative data from 1974 to 2004. *unpublished paper, Columbia University.*

Wang, T. (2023). Perceived versus calibrated income risks in heterogeneous-agent consumption models.

Wiswall, M. and B. Zafar (2021). Human capital investments and expectations about career and family. *Journal of Political Economy 129*(5), 1361–1424.