# The simple life? Heterogeneity in income risk and household portfolios

EEA-ESEM 2024: Consumption and Saving, Rotterdam

Sebastian Hildebrand[1], with Christian Bayer[2], Thomas Hintermaier[3], Moritz Kuhn[4], and Gašper Ploj[5]

August 29, 2024

[1]Univ. of Bonn [2]Univ. of Bonn, CEPR, IZA [3]Univ. of Bonn [4]Univ. of Mannheim, CEPR, IZA, CESifo [5]Banka Slovenije

# The paper in one slide: double-dividend of persistently low income risk

Literature

Income risk is the largest economic risk faced by workers

- How is it distributed *across* workers? How persistent is it *within* worker?
- How is it correlated with income, unemployment, and health?
- How is it correlated with wealth and portfolio choices?

Document large and persistent heterogeneity in income risk empirically

- Concentration of risk in a small set of high-risk workers
- Estimation of a regime-switching model of workers' earnings
- Double-dividend: low-risk workers face higher earnings, less unemployment, and better health, accumulate more wealth, and invest in less liquid, higher-return assets

Rationalize in life-cycle model: persistent risk types & frictional asset markets [not today]

# The paper in one slide: double-dividend of persistently low income risk

Literature

Income risk is the largest economic risk faced by workers

- How is it distributed *across* workers? How persistent is it *within* worker?
- How is it correlated with income, unemployment, and health?
- How is it correlated with wealth and portfolio choices?

Document large and persistent heterogeneity in income risk empirically

- Concentration of risk in a small set of high-risk workers
- Estimation of a regime-switching model of workers' earnings
- Double-dividend: low-risk workers face higher earnings, less unemployment, and better health, accumulate more wealth, and invest in less liquid, higher-return assets

Rationalize in life-cycle model: persistent risk types & frictional asset markets [not today]

# Data

# Data – German data for earnings and portfolios

Details

**Earnings data:** Representative, administrative data from social security records (VSKT)

- Worker-year panel of complete employment histories – 129,808 workers
- Sample: male salaried workers with employment records in West Germany
- Main variables: gross labor earnings, employment status, and health

**Wealth data:** Match with survey data from the Socio-Economic Panel (SOEP)

- Worker-year panel, information on labor market experiences and financial situations
- Focus on level and liquidity of an individual's net worth
- Predictive mean matching algorithm (and some linked observations)

**Outcome:** Detailed matched earnings histories and wealth and portfolio information

# Data – German data for earnings and portfolios

Details

**Earnings data:** Representative, administrative data from social security records (VSKT)

- Worker-year panel of complete employment histories – 129,808 workers
- Sample: male salaried workers with employment records in West Germany
- Main variables: gross labor earnings, employment status, and health

**Wealth data:** Match with survey data from the Socio-Economic Panel (SOEP)

- Worker-year panel, information on labor market experiences and financial situations
- Focus on level and liquidity of an individual's net worth
- Predictive mean matching algorithm (and some linked observations)

**Outcome:** Detailed matched earnings histories and wealth and portfolio information

# Data – German data for earnings and portfolios

Details

**Earnings data:** Representative, administrative data from social security records (VSKT)

- Worker-year panel of complete employment histories – 129,808 workers
- Sample: male salaried workers with employment records in West Germany
- Main variables: gross labor earnings, employment status, and health

**Wealth data:** Match with survey data from the Socio-Economic Panel (SOEP)

- Worker-year panel, information on labor market experiences and financial situations
- Focus on level and liquidity of an individual's net worth
- Predictive mean matching algorithm (and some linked observations)

**Outcome:** Detailed matched earnings histories and wealth and portfolio information

# Empirics

# Empirics – birds-eye view of the empirics

Using the **administrative data** from social security records, we:

1. Document the German earnings growth distribution (over the life cycle)
2. Show large, significant, and persistent heterogeneity in earnings risk across workers
3. Estimate a structural regime-switching model of workers' earnings
4. Document relationship b/w earnings risk and earnings, unemployment, and sickness

Using the **combined data** (RV + SOEP), we:

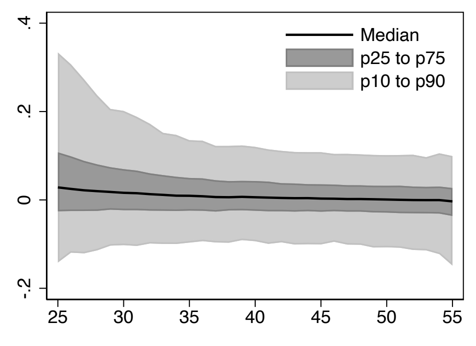
4. Show relationship b/w earnings risk and portfolio outcomes

# Empirics – earnings growth dispersion over a worker's life

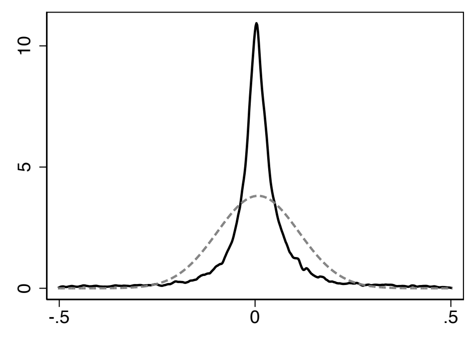


Figure 1: Earnings growth rate distribution

(a) Earnings growth rate by age




(b) Density of earnings growth rate at age 40




Notes (lhs): Percentiles of the earnings growth rate distribution by age in terms of pension points, i.e., relative to aggregate earnings growth. Notes (rhs): Earnings growth rate distribution at age 40. The solid line shows a kernel density estimate and the dashed line an estimate of the normal distribution.

**Main takeaway:** Large dispersion of growth rates across workers, clear age patterns, still overdispersion at age 40; similar for wage growth rates [appendix]

# Empirics – taking stock: what drives the overdispersion of earnings growth?

**Standard way to model overdispersion:** Mixing of risk types (Guvenen et al., 2021)

- Some workers draw from a high-variance, others from a low-variance distribution

**Crucial question:** Are risk types transitory, persistent, or permanent?

**Unique data feature:** Entire working histories, not revolving panel

- Compute within-worker earnings growth standard deviations
- Study (i) dispersion and (ii) persistence of earnings growth volatilities
- Step 1: Compare to simulated models to understand data-generating process Details
- Step 2: Estimate a regime-switching model of workers' earnings

# Empirics – taking stock: what drives the overdispersion of earnings growth?

**Standard way to model overdispersion:** Mixing of risk types (Guvenen et al., 2021)

- Some workers draw from a high-variance, others from a low-variance distribution

**Crucial question:** Are risk types transitory, persistent, or permanent?

**Unique data feature:** Entire working histories, not revolving panel

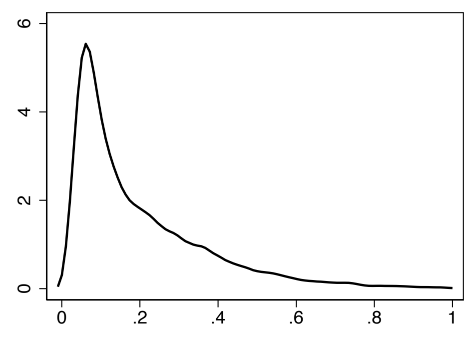
- Compute within-worker earnings growth standard deviations
- Study (i) dispersion and (ii) persistence of earnings growth volatilities
- Step 1: Compare to simulated models to understand data-generating process Details
- Step 2: Estimate a regime-switching model of workers' earnings

# Empirics – dispersion of earnings growth volatilities

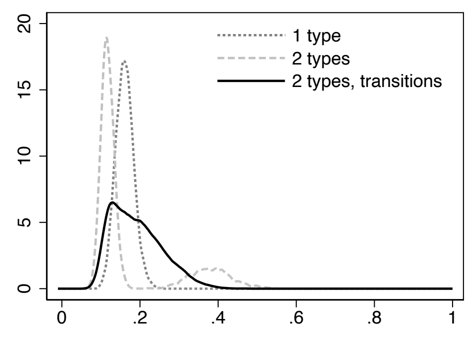
Details

Figure 2: Distribution of earnings growth rate volatility

(a) Density of empirical lifetime volatility




(b) Density of simulated lifetime volatility




Notes: Empirical and simulated distributions of lifetime earnings growth volatility (std. dev. b/w age 25-55). Kernel density estimates.

**Main takeaway:** Large dispersion of lifetime earnings growth volatilities across workers; model with two persistent, but not permanent, risk types fits the shape

# Empirics – persistence of earnings volatility

Details

**Suggestive evidence:** Persistence of earnings volatility

- Dispersion decreases over lifetime $\longrightarrow$ mean reversion in risk
- Large dispersion of lifetime volatility $\longrightarrow$ not very transitory

**Predictive regression:** Moving windows of worker-specific volatilities

$$\text{SD(Log growth rate)}_{i,j+6,t+6} = \alpha + \beta \times \text{SD(Log growth rate)}_{i,j,t} + \gamma_j + \delta_t + \varepsilon_{i,j+6,t+6} \quad (1)$$

- $\text{SD}(-)_{i,j,t}$: 5-year rolling std. dev. of earn. growth for worker $i$, age $j$, time $t$
- $\beta$: persistence of worker-level earnings volatility
- $\gamma_j$ and $\delta_t$: age and year fixed effects

# Empirics – persistence of earnings volatility

Details

**Suggestive evidence:** Persistence of earnings volatility

- Dispersion decreases over lifetime $\longrightarrow$ mean reversion in risk
- Large dispersion of lifetime volatility $\longrightarrow$ not very transitory

**Predictive regression:** Moving windows of worker-specific volatilities

$$\text{SD(Log growth rate)}_{i,j+6,t+6} = \alpha + \beta \times \text{SD(Log growth rate)}_{i,j,t} + \gamma_j + \delta_t + \varepsilon_{i,j+6,t+6} \quad (1)$$

- $\text{SD}(-)_{i,j,t}$: 5-year rolling std. dev. of earn. growth for worker $i$, age $j$, time $t$
- $\beta$: persistence of worker-level earnings volatility
- $\gamma_j$ and $\delta_t$: age and year fixed effects

# Empirics – persistence of earnings volatility

**Table 1:** Predictability of future earnings volatility, 5-year windows

| | SD(Log growth rate)$_{j+6,t+6}$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Data | 1 type (simulation) | 2 types (simulation) | 2 types with transitions (simulation) |
| SD(Log growth rate)$_{j,t}$ | 0.1216*** | -0.0024 | 0.6612*** | 0.1214*** |
| | (95.27) | (-1.06) | (245.15) | (48.50) |
| Observations | 1422719 | 190000 | 190000 | 190000 |

Notes: OLS regression estimates for Equation (1). Regressions on empirical data include age and year fixed effects, while regressions on simulated data include age fixed effects. Robust standard errors used. $t$ statistics in parentheses, $^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

**Main takeaway:** significant, but not permanent, degree of persistence of earnings growth volatility, matched by transitions between risk types (remain high-risk for ≈ four years)

# Empirics – structural evidence on risk types in earnings dynamics

Details

**Simple model of earnings dynamics:** homogeneous parameters, incl. risk

| | | |
|---|---|---|
| a worker's income: | $e_{i,t} = \eta_{i,t} + \tau_{i,t} + \mu_0$ | |
| persistent component: | $\eta_{i,t} = \rho\,\eta_{i,t-1} + \varepsilon_{i,t}$ | with $\varepsilon_{i,t} \sim \mathcal{N}(0, \sigma^2_{\varepsilon,0})$ |
| transitory component: | $\tau_{i,t} = \xi_{i,t}$ | with $\xi_{i,t} \sim \mathcal{N}(0, \sigma^2_{\xi,0})$ |

**Regime-switching model:** heterogeneous parameters, incl. risk, and transitions

| | | |
|---|---|---|
| a worker's income: | $e_{i,t} = \eta_{i,t} + \tau_{i,t} + \mu_{s_{i,t}}$ | |
| persistent component: | $\eta_{i,t} = \rho\,\eta_{i,t-1} + \varepsilon_{i,t}$ | with $\varepsilon_{i,t} \sim \mathcal{N}(0, \sigma^2_{\varepsilon,s_{i,t}})$ |
| transitory component: | $\tau_{i,t} = \xi_{i,t}$ | with $\xi_{i,t} \sim \mathcal{N}(0, \sigma^2_{\xi,s_{i,t}})$ |

with $s_{i,t} \in \{0,1\}$ and Markov transition matrix $\Pi$.

# Empirics – structural evidence on risk types in earnings dynamics

Details

**Simple model of earnings dynamics:** homogeneous parameters, incl. risk

| | | |
|---|---|---|
| a worker's income: | $e_{i,t} = \eta_{i,t} + \tau_{i,t} + \mu_0$ | |
| persistent component: | $\eta_{i,t} = \rho\,\eta_{i,t-1} + \varepsilon_{i,t}$ | with $\varepsilon_{i,t} \sim \mathcal{N}(0, \sigma^2_{\varepsilon,0})$ |
| transitory component: | $\tau_{i,t} = \xi_{i,t}$ | with $\xi_{i,t} \sim \mathcal{N}(0, \sigma^2_{\xi,0})$ |

**Regime-switching model:** heterogeneous parameters, incl. risk, and transitions

| | | |
|---|---|---|
| a worker's income: | $e_{i,t} = \eta_{i,t} + \tau_{i,t} + \mu_{s_{i,t}}$ | |
| persistent component: | $\eta_{i,t} = \rho\,\eta_{i,t-1} + \varepsilon_{i,t}$ | with $\varepsilon_{i,t} \sim \mathcal{N}(0, \sigma^2_{\varepsilon,\,s_{i,t}})$ |
| transitory component: | $\tau_{i,t} = \xi_{i,t}$ | with $\xi_{i,t} \sim \mathcal{N}(0, \sigma^2_{\xi,\,s_{i,t}})$ |

with $s_{i,t} \in \{0, 1\}$ and Markov transition matrix $\Pi$.

# Empirics – comparison of parameter estimates

**Table 2:** Estimation results for the earnings process of workers

| | Parameter estimates | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | $\mu_0$ | $\mu_1$ | $\rho$ | $\sigma_{\varepsilon,0}$ | $\sigma_{\varepsilon,1}$ | $\sigma_{\xi,0}$ | $\sigma_{\xi,1}$ | $\pi_{00}$ | $\pi_{11}$ |
| homogeneous | 0.00 | – | 0.95 | 0.08 | – | 0.06 | – | – | – |
| heterogeneous, switching | 0.02 | -0.04 | 0.98 | 0.04 | 0.13 | 0.00 | 0.15 | 0.88 | 0.52 |

Notes: Estimates from maximum likelihood estimation of the simple and regime-switching model of earnings dynamics, likelihood computed using the Kalman filter and Kim filter, respectively. Initial values based on raw data moments, robust to different initial values.

## Main takeaways:

- High-risk workers ($s = 1$) have lower but more volatile income
- Approx. 20 % of workers are high-risk (per year)
- High-risk workers stay high-risk for approximately two years (on average)

# Empirics – taking stock: large, persistent, and heterogeneous earnings risk

**What did we learn?**

- Large dispersion of earnings growth rates across workers
- Evidence of persistence in individual workers' earnings volatility
- Model with two risk types and transitions fits the data

**Next step:** Link earnings risk to other labor market outcomes

- Identify (ex-post) risk types based on lifetime earnings growth volatility
    - robust to using the smoothed type-probabilities from the estimation
- Strength of our data: observe unemployment and sickness spells
- Document lower income growth, higher unemployment and health risks

# Empirics – taking stock: large, persistent, and heterogeneous earnings risk

What did we learn?

- Large dispersion of earnings growth rates across workers
- Evidence of persistence in individual workers' earnings volatility
- Model with two risk types and transitions fits the data

**Next step:** Link earnings risk to other labor market outcomes

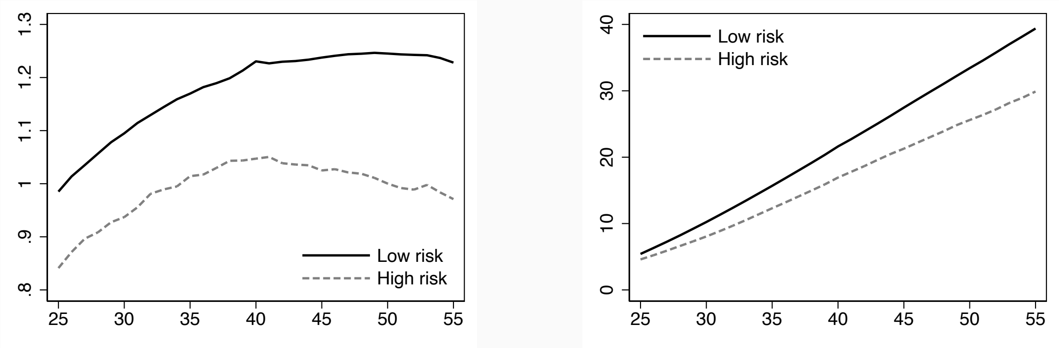
- Identify (ex-post) risk types based on lifetime earnings growth volatility
  - robust to using the smoothed type-probabilities from the estimation
- Strength of our data: observe unemployment and sickness spells
- Document lower income growth, higher unemployment and health risks

Details

Figure 3: Lifetime income by risk group

(a) Average earnings in pension points by age

(b) Cumulative earnings in pension points by age




Notes: Risk types are defined based on lifetime earnings volatility where the high-risk type is mapped to the top 20% of the distribution whereas the remaining 80% are low-risk.

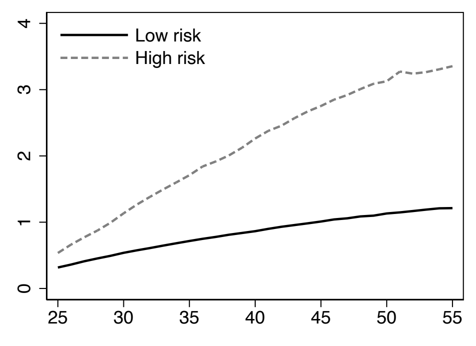
**Main takeaway:** workers with more volatile earnings histories have lower average earnings per year, cumulating up to sizable differences at retirement

# Empirics – strong correlation with unemployment

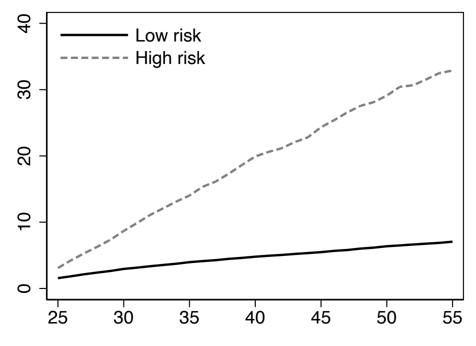
Details

Figure 4: Lifetime unemployment by risk group

(a) Unemployment spells by age




(b) Unemployment duration in months by age




Notes: Life cycle profile of the number of independent unemployment spells and the total number of months spent as unemployed. Risk types are defined based on lifetime earnings volatility where the high-risk type is mapped to the top 20% of the distribution whereas the remaining 80% are low-risk.

**Main takeaway:** strong positive correlation between unemployment risk and risk types

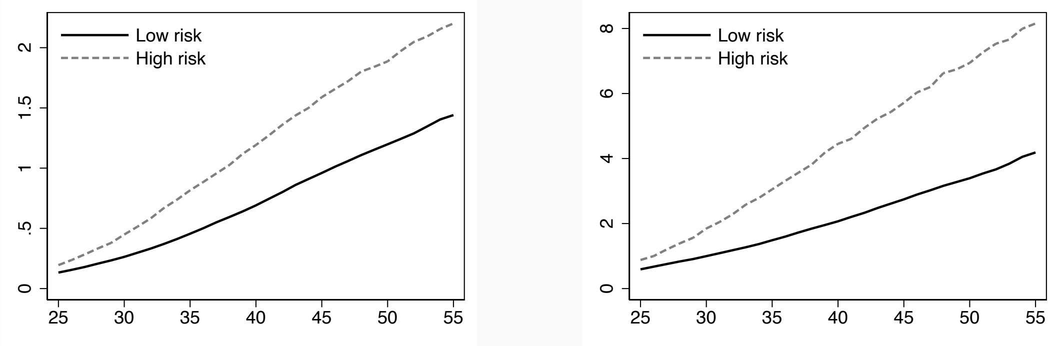
# Empirics – strong correlation with sickness

Details

Figure 5: Lifetime sickness by risk group

(a) Sickness spells by age

(b) Sickness duration in months by age




Notes: Life cycle profile of the number of independent sickness spells and the total number of months spent sick. Risk types are defined based on lifetime earnings volatility where the high-risk type is mapped to the top 20% of the distribution whereas the remaining 80% are low-risk.

**Main takeaway:** strong positive correlation between health risk and risk types

# Empirics – taking stock: simple vs. hard life

Details

**What did we learn?**

- Large heterogeneity in earnings risk, highly concentrated
- Correlates with lower income, higher unemployment and more sickness
- Risk-types are not permanent: turbulence eventually calms down

**Next step:** Link earnings risk to workers' portfolios

- Economic theory: workers respond to labor market risk in their savings decisions
- Test relationship between labor market risk and workers' portfolios

# Empirics – taking stock: simple vs. hard life

Details

**What did we learn?**

- Large heterogeneity in earnings risk, highly concentrated
- Correlates with lower income, higher unemployment and more sickness
- Risk-types are not permanent: turbulence eventually calms down

**Next step:** Link earnings risk to workers' portfolios

- Economic theory: workers respond to labor market risk in their savings decisions
- Test relationship between labor market risk and workers' portfolios

# Empirics – risk heterogeneity and workers' portfolios

Details

**Table 3:** Risk types and workers' portfolios

Panel (a): OLS

| | (1) Earnings | (2) Net wealth | (3) Net wealth to earnings | (4) Illiquid share | (5) Average portfolio return |
|---|---|---|---|---|---|
| Risk type | -0.2466*** | -0.7834*** | -0.2292*** | -0.0584*** | -0.3769*** |
| | (-249.11) | (-77.97) | (-22.78) | (-50.49) | (-65.80) |
| Observations | 938130 | 929170 | 929170 | 929170 | 929170 |

Notes: OLS regression estimates of the outcome variables (each column) on risk type and a set of controls of year, age and state fixed effects as well as age-year and state-year interactions. Robust standard errors used. $t$ statistics in parentheses, $^{+}\ p < 0.10$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

**Main takeaway:** higher labor market risk comes with lower earnings, less net wealth, less net wealth relative to earnings, more liquid assets, and lower portfolio returns

$\Longrightarrow$ double dividend: higher earnings growth and higher returns on financial wealth

# Model

# Model – birds-eye view of the model

Details

**Objective:** Rationalize the empirical patterns through life-cycle model

- Large heterogeneity in earnings risk, highly concentrated
- Correlates with lower income, higher unemployment and more sickness
- Risk-types are not permanent: turbulence eventually calms down
- High-risk have lower wealth, more liquid portfolios, and lower returns

**Core features:** Persistent risk types & frictional asset markets

- Income risk, and second-order risk of switching between persistent risk types
- Portfolio choice: low-return liquid asset vs. illiquid savings plan

# Conclusion

# Conclusion – empirics and theory of double-dividend of low income risk

**Novel data:** Combine largely unexplored German administrative labor market data with workers' balance sheets, document within-worker facts

**Empirical patterns:** Large heterogeneity in earnings risk, highly concentrated

- Persistent differences in earnings risk, but turbulence eventually calms down
- Correlates with lower income, higher unemployment and more sickness
- Strong effects on wealth accumulation and portfolio choice

**Life-cycle model:** Qualitatively and quantitatively rationalizes the empirical patterns

- Persistent risk types with second-order risk of switching
- Frictional asset market with systematic return differences

# Appendix

# Related literature & our contribution

Back

**Heterogeneity in labor market risk:**

- Extensive heterogeneity in labor market outcomes, challenging the adequacy of a representative earnings process [e.g., Arellano et al., 2017; Guvenen et al., 2015; Hall, 1982; Jung and Kuhn, 2018; Karahan et al., 2019; Kuhn and Ploj, 2020; Morchio, 2020]
- Labor income risk varies across age, education, skill groups, and with the business cycle [e.g., Bayer et al., 2019; Blundell et al., 2015; Jung and Kuhn, 2018; Karahan and Ozkan, 2013; Kuhn and Ploj, 2020; Mukoyama and Şahin, 2006; Storesletten et al., 2004]
- Detailed administrative data allows a more nuanced analysis of labor market risk [e.g., Guvenen et al., 2015; Manuel et al., 2017; Morchio, 2020; Schmillen and Möller, 2012]

**Our contribution:** Leverage detailed data and focus on within-worker persistence

- Additional evidence: workers systematically vary in terms of lifetime earnings risk
- Difference largely explained by unemployment and health risk

# Related literature & our contribution

Back

**Implications for household savings and portfolio choices:**

- Income risk reduces the demand for risky assets [e.g., Angerer and Lam, 2009; Chang et al., 2020; Fagereng et al., 2018; Guiso et al., 1996; Heaton and Lucas, 2000, 2001; Palia et al., 2014]
- Effects are often small and insignificant
- Macroeconomic perspective on variations in income risk [e.g., Bayer et al., 2019; Larkin, 2023; McKay, 2017]

**Our contribution:** Additional evidence for significant effect of risk heterogeneity on

- Wealth accumulation, portfolio returns, and asset participation
- Life-cycle patterns of heterogeneity in portfolio liquidity and wealth accumulation

# Appendix: Data

# Appendix: Data – German pension data for wages and earnings

Back Details

**First data source:** Representative, administrative data from social security records

- Entire employment histories of individuals starting at age 14 on a monthly basis
- Gender, age, earnings, labor market status, unemployment, and health

**Sample selection:** Male salaried workers with employment records in West Germany

- Cohorts 1937-1960, complete labor market histories up to age 60 – 129,808 workers
- Exlude: GDR, women, civil servants, self-employed

**Earnings measure:** Labor market income subject to social security taxes

- Monthly pension points – relative to macroeconomic annual average, capped
- Aggregate to annual panel: worker-year

# Appendix: Data – German pension data for wages and earnings

Back Details

**First data source:** Representative, administrative data from social security records

- Entire employment histories of individuals starting at age 14 on a monthly basis
- Gender, age, earnings, labor market status, unemployment, and health

**Sample selection:** Male salaried workers with employment records in West Germany

- Cohorts 1937-1960, complete labor market histories up to age 60 – 129,808 workers
- Exlude: GDR, women, civil servants, self-employed

**Earnings measure:** Labor market income subject to social security taxes

- Monthly pension points – relative to macroeconomic annual average, capped
- Aggregate to annual panel: worker-year

# Appendix: Data – German pension data for wages and earnings

Back Details

**First data source:** Representative, administrative data from social security records

- Entire employment histories of individuals starting at age 14 on a monthly basis
- Gender, age, earnings, labor market status, unemployment, and health

**Sample selection:** Male salaried workers with employment records in West Germany

- Cohorts 1937-1960, complete labor market histories up to age 60 – 129,808 workers
- Exlude: GDR, women, civil servants, self-employed

**Earnings measure:** Labor market income subject to social security taxes

- Monthly pension points – relative to macroeconomic annual average, capped
- Aggregate to annual panel: worker-year

# Appendix: Data – German SOEP for wealth and portfolios

Back Details

**Second data source:** Survey data from the Socio-Economic Panel (SOEP)

- Longitudinal panel dataset of the German population
- Labor market experiences and financial situations of individuals

**Sample selection:** Choose individuals that are most comparable to the pension data

**Wealth data:** Available in the SOEP for every fifth year since 2002

- Focus on level and liquidity of an individual's net worth
- Liquid-illiquid split of assets
- Historical portfolio returns

# Appendix: Data – German SOEP for wealth and portfolios

Back Details

**Second data source:** Survey data from the Socio-Economic Panel (SOEP)

- Longitudinal panel dataset of the German population
- Labor market experiences and financial situations of individuals

**Sample selection:** Choose individuals that are most comparable to the pension data

**Wealth data:** Available in the SOEP for every fifth year since 2002

- Focus on level and liquidity of an individual's net worth
- Liquid-illiquid split of assets
- Historical portfolio returns

# Appendix: Data – German SOEP for wealth and portfolios

Back Details

**Second data source:** Survey data from the Socio-Economic Panel (SOEP)

- Longitudinal panel dataset of the German population
- Labor market experiences and financial situations of individuals

**Sample selection:** Choose individuals that are most comparable to the pension data

**Wealth data:** Available in the SOEP for every fifth year since 2002

- Focus on level and liquidity of an individual's net worth
- Liquid-illiquid split of assets
- Historical portfolio returns

# Appendix: Data – combining labor market and portfolio data

Back Details

**Imputation:** Impute wealth to the pension data based on a set of common observables

- Variables of interest: net worth, share of illiquid assets, average portfolio return
- Regressors: everything available in both data sources
  - Year, age, state, (cum.) unemployment, (cum.) sickness indicators (and interactions)
  - Levels, means, volatilities, interactions over different lag horizons of workers' earnings
- Predictive mean matching algorithm to capture nonlinearities:
  - Regress the variables of interest on these regressors in the SOEP
  - Predict values in the pension data, find the neighbors of predicted value & average
  - Compute multiple implicates by drawing from the error term distribution

**Outcome:** Detailed earnings histories and detailed wealth and portfolio information

# Appendix: Data – combining labor market and portfolio data

Back Details

**Imputation:** Impute wealth to the pension data based on a set of common observables

- Variables of interest: net worth, share of illiquid assets, average portfolio return
- Regressors: everything available in both data sources
  - Year, age, state, (cum.) unemployment, (cum.) sickness indicators (and interactions)
  - Levels, means, volatilities, interactions over different lag horizons of workers' earnings
- Predictive mean matching algorithm to capture nonlinearities:
  - Regress the variables of interest on these regressors in the SOEP
  - Predict values in the pension data, find the neighbors of predicted value & average
  - Compute multiple implicates by drawing from the error term distribution

**Outcome:** Detailed earnings histories and detailed wealth and portfolio information

# Appendix: Data: pension data – sample selection

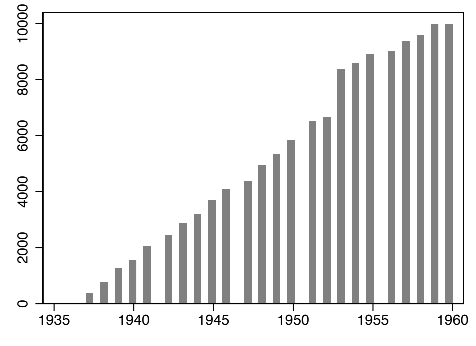
Back

Details on sample selection:

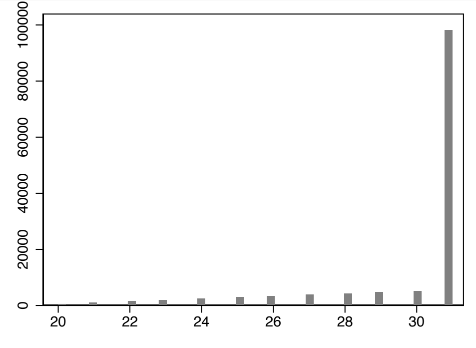
- West Germany: exlude if employment records from the GDR period, if place of residence in East Germany, or if "Fremdrenten" points
- (Sworn-in) civil servants: exclude if all observations after age 30 are missing (no contributions to public pension system), implicit assumption that civil servants do not return to the public pension system at a later stage
- Self-employed: if recorded as mainly self-employed (SES == 11) over their lifetime

Figure 6: Histograms on the sample

(a) Workers per cohort




(b) Workers per number of observations




Notes: The left panel shows the number of workers per birth year. The right panel shows the number of workers per number of ID-year-observations, between age 25-55.

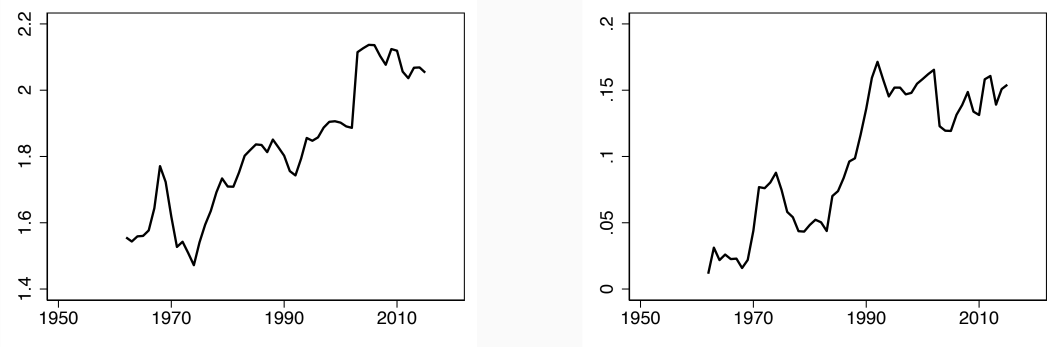
# Appendix: Data: pension data – sample selection

Back

Figure 7: Social security contribution limit and the share of capped observations

(a) Contribution limit in pension points

(b) Share of earnings observations




Notes: The left panel shows the social security contribution limit expressed in the number of pension points. The right panel shows the share of earnings observations that are affected by the social security contribution limit.

# Appendix: Data: wealth data – sample selection

Back

Details on sample selection:

- Male, West Germany, age 20-60, no self-employed, civil servants, or refugees
- Self-employed: exclude if income from self-employment greater than other income
- Civil servants: exclude based on identifier measured in the SOEP

# Appendix: Data: wealth data – portfolios

Back

**Net wealth in a category:** value of all assets minus value of all debt in that category

- Net real estate wealth (primary residence plus other real estate net of mortgages)
- Business wealth
- Tangible assets
- Net building loan and insurance wealth
- Net liquid financial wealth (savings accounts, bonds, stocks, investments net of financial liabilities other than mortgages and building loans, i.e. consumer credit)

**Liquid-illiquid split:** three measures (for robustness)

- Narrow: net real estate and insurance assets are illiquid
- Medium: additionally, building loan contracts are illiquid
- Wide: additionally, business assets are illiquid

# Appendix: Data: wealth data – portfolios

Back

**Home real estate:** “Are you personally the owner of the house or apartment in which you live?”
“If you were to sell today, how much would you receive for your house/apartment including land?”
“If you are still paying off a loan on your house/apartment, how much is left to repay (excluding interest)?”

**Other real estate:** “Apart from the home you live in, do you own any other homes or more land?”
“If you were to sell your property today (excluding the one where you live in), how much would you receive?”
“If you are still paying off a loan on your property, how much is left to repay (excluding interest)?”

**Business wealth:** “Are you the owner of a commercial enterprise, [...], or are you co-owner of one of these types of enterprises?”
“What do you estimate the current value of your enterprise or your share thereof to be? This is the price before taxes that you would receive on the sale of your enterprise or your share thereof, after deducting any remaining debts.”

**Tangible assets:** “Do you own any tangible assets in the form of gold, jewelry, coins, or valuable collections?”
“If you could sell these assets, what would their total value be?”

**Net building loan and insurance wealth:** “Do you have a building loan agreement?”
“What would you estimate the building loan credit balance to be, including interest and dividends?”
“Do you have a life insurance policy or private retirement plan, with contributions paid either by you or by your employer?”
“What do you estimate the current cash surrender value or balance of these plans to be?”

**Net liquid financial wealth:** “Do you own financial assets in the form of a savings account, bonds, shares, or investments?”
“What do you estimate to be the value of your financial assets?”
“Leaving aside any mortgages on homes and property or building loans: Do you currently owe money on loans [...]?”
“What are your outstanding debts?”

# Appendix: Data: wealth data – returns

Back

**Table 4:** Average historical returns, real

| | (1) Equity | (2) Housing | (3) Bonds | (4) Bills | (5) Deposits |
|---|---|---|---|---|---|
| Constant | 5.8030 | 4.7334 | 3.6345 | 1.3852 | -0.2123 |
| Observations | 50 | 50 | 50 | 50 | 50 |

Notes: The table reports average historical returns based on German data between 1970 and 2020. The first four columns use data from Jordà et al. (2019), the last column uses data by the Bundesbank.

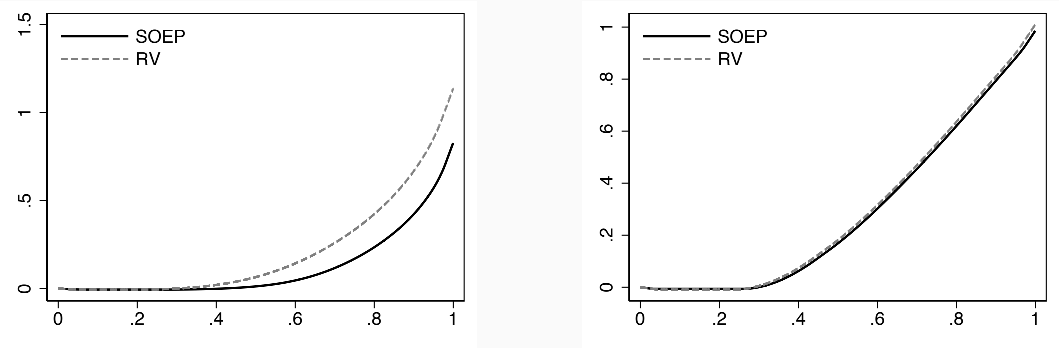
# Appendix: Data: wealth data – imputation quality

Back

Figure 8: Comparison of wealth distributions

(a) Net wealth in pension points

(b) Illiquid share (narrow)

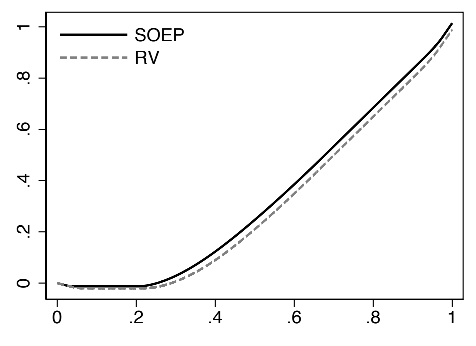



Notes: Lorenz curves comparing the wealth distribution in the SOEP with the imputed wealth distribution in the pension data (RV).

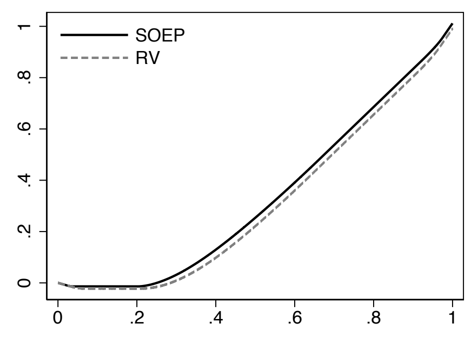
Back

Figure 8: Comparison of wealth distributions

(c) Illiquid share (medium)




(d) Illiquid share (wide)



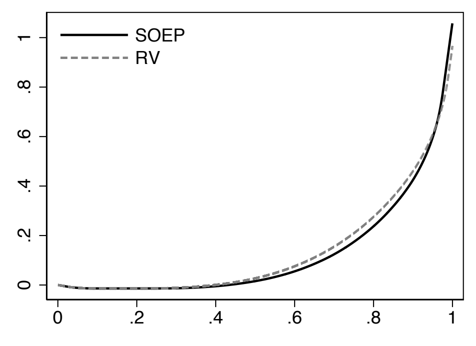

Notes: Lorenz curves comparing the wealth distribution in the SOEP with the imputed wealth distribution in the pension data (RV).

# Appendix: Data: wealth data – imputation quality

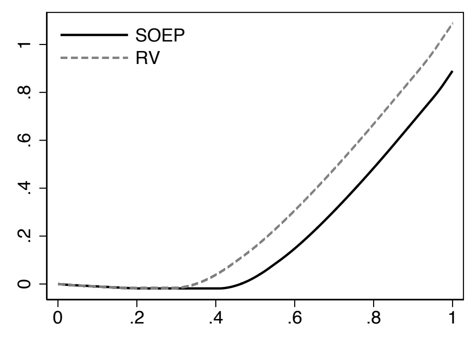
Back

Figure 8: Comparison of wealth distributions

(e) Wealth-to-income ratio




(f) Portfolio returns




Notes: Lorenz curves comparing the wealth distribution in the SOEP with the imputed wealth distribution in the pension data (RV).

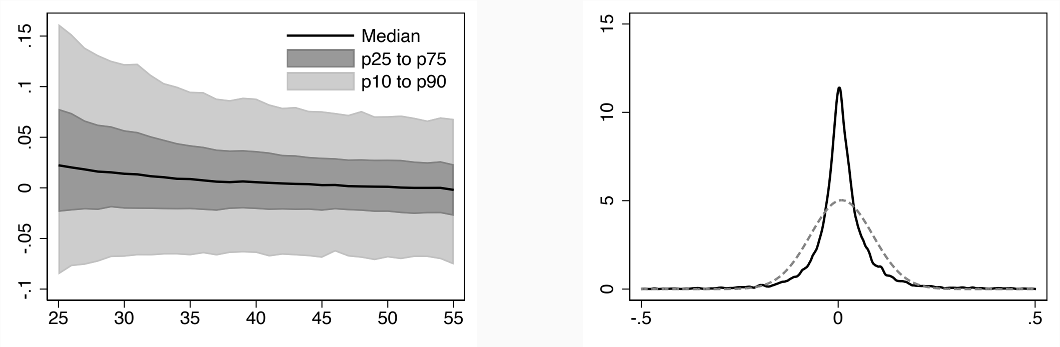
# Appendix: Empirics – wage growth dispersion over a worker's life

Back

Figure 9: Wage growth rate distribution

(a) Wage growth rate by age

(b) Density of wage growth rate at age 40




Notes (lhs): Percentiles of the wage growth rate distribution by age in terms of pension points, i.e., relative to aggregate wage growth. Notes (rhs): Wage growth rate distribution at age 40. The solid line shows a kernel density estimate and the dashed line an estimate of the normal distribution.

**Main takeaway:** similar to earnings growth, earnings growth more dispersed than wage growth (through hours worked)

# Appendix: Empirics – simulated earnings histories of workers

Back

**Standard earnings process:** log earnings of worker $i$ at age $j$, $e_{i,j}$, follow

$$e_{i,j} = p_{i,j} + \eta_{i,j} \quad \text{with} \quad p_{i,j} = p_{i,j-1} + \varepsilon_{i,j}$$

- $p_{i,j}$: permanent component, with permanent innovation $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$
- $\eta_{i,j}$: transitory component, $\eta_{i,j} \sim \mathcal{N}(0, \sigma_\eta^2)$

**Sample split:** Sort workers by lifetime earnings growth volatility

- Assume: bottom 80% are low-risk, top 20% are high-risk
- Estimate earnings process on pooled and split samples
- Obtain: standard deviations of permanent and transitory components

# Appendix: Empirics – simulated earnings histories of workers

Back

**Standard earnings process:** log earnings of worker $i$ at age $j$, $e_{i,j}$, follow

$$e_{i,j} = p_{i,j} + \eta_{i,j} \quad \text{with} \quad p_{i,j} = p_{i,j-1} + \varepsilon_{i,j}$$

- $p_{i,j}$: permanent component, with permanent innovation $\varepsilon_{i,j} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$
- $\eta_{i,j}$: transitory component, $\eta_{i,j} \sim \mathcal{N}(0, \sigma_\eta^2)$

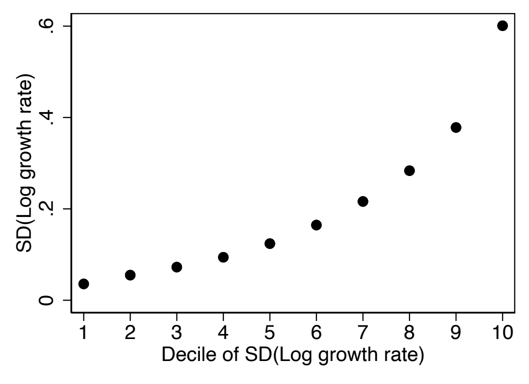
**Sample split:** Sort workers by lifetime earnings growth volatility

- Assume: bottom 80% are low-risk, top 20% are high-risk
- Estimate earnings process on pooled and split samples
- Obtain: standard deviations of permanent and transitory components

# Appendix: Empirics – simulated earnings histories of workers

Back

**Figure 10:** Lifetime earnings growth rate volatility




Notes: Percentiles of std. dev. of lifetime earnings growth rates.

**Table 5:** Estimated earnings processes

| SD | Pooled | Low risk | High risk |
|---|---|---|---|
| Permanent | 0.15 | 0.10 | 0.21 |
| Transitory | 0.09 | 0.07 | 0.18 |

Simulate 3 cases:

- homogeneous risk
- permanent high-low type
- transitions between types

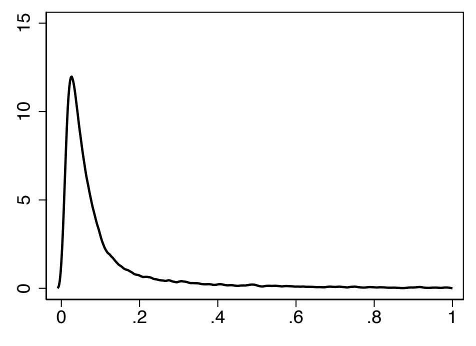
**Next step:** Compare empirical realized earnings with simulated earnings processes

# Appendix: Empirics – dispersion of earnings growth volatilities

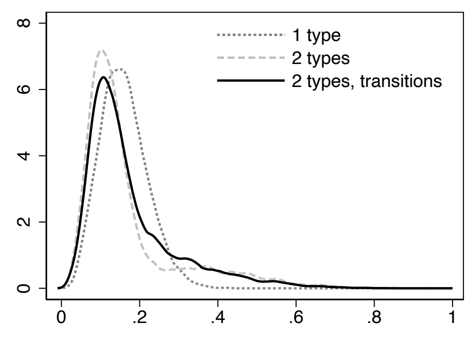
Back

Figure 2: Distribution of earnings growth rate volatility

(c) Density of empirical 5-year volatility at age 40




(d) Density of simulated 5-year volatility at age 40




Notes: Empirical and simulated distributions of earnings growth rate volatility. Kernel density estimates.

**Main takeaway:** model with two persistent, but not permanent, risk types fits the shape of both the lifetime and 5-year volatility distributions

# Appendix: Empirics – the same model(s) in state-space form

Back

Let $y_{i,t} = [e_{i,t}]$, $\beta_{i,t} = [\eta_{i,t}]$, $z_{i,t} = [1]$, $\widetilde{\xi}_{i,t} = [\xi_{i,t}]$, $\widetilde{\varepsilon}_{i,t} = [\varepsilon_{i,t}]$, then

$$y_{i,t} = H_{s_{i,t}} \beta_{i,t} + A_{s_{i,t}} z_{i,t} + \widetilde{\xi}_{i,t} \qquad \text{with } \widetilde{\xi}_{i,t} \sim \mathcal{N}(0, R_{s_{i,t}})$$

$$\beta_{i,t} = \widetilde{\mu}_{s_{i,t}} + F_{s_{i,t}} \beta_{i,t-1} + \widetilde{\varepsilon}_{i,t} \qquad \text{with } \widetilde{\varepsilon}_{i,t} \sim \mathcal{N}(0, Q_{s_{i,t}})$$

with the regimes $s_{i,t} \in \{0, 1\}$, the transition matrix $\Pi$ defined as

$$\Pi = \begin{bmatrix} \pi_{00} & 1 - \pi_{00} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix}$$

with $\pi_{00}, \pi_{11} \in [0, 1]$, and

$$H_0 = H_1 = [1], \qquad \widetilde{\mu}_0 = \widetilde{\mu}_1 = [0],$$

$$A_0 = [\mu_0] \text{ and } A_1 = [\mu_1], \qquad F_0 = F_1 = [\rho],$$

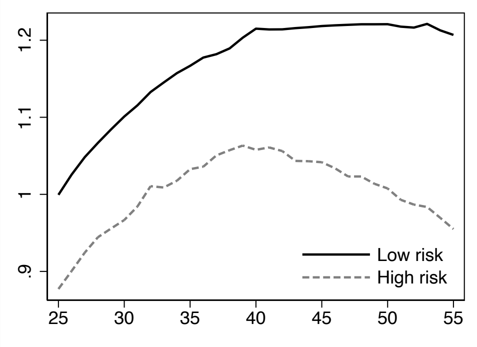
$$R_0 = [\sigma^2_{\xi,0}] \text{ and } R_1 = [\sigma^2_{\xi,1}], \qquad Q_0 = [\sigma^2_{\varepsilon,0}] \text{ and } Q_1 = [\sigma^2_{\varepsilon,1}].$$

# Appendix: Empirics – earnings by risk type, median

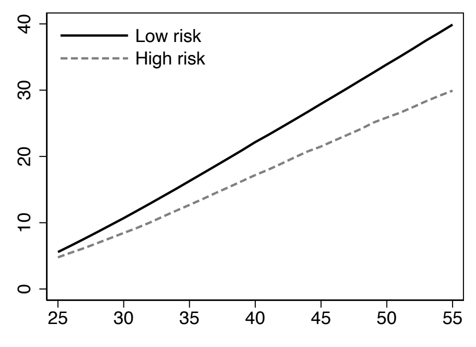
Back

Figure 11: Lifetime income by risk group

(a) Earnings in pension points by age




(b) Cumulative earnings in pension points by age



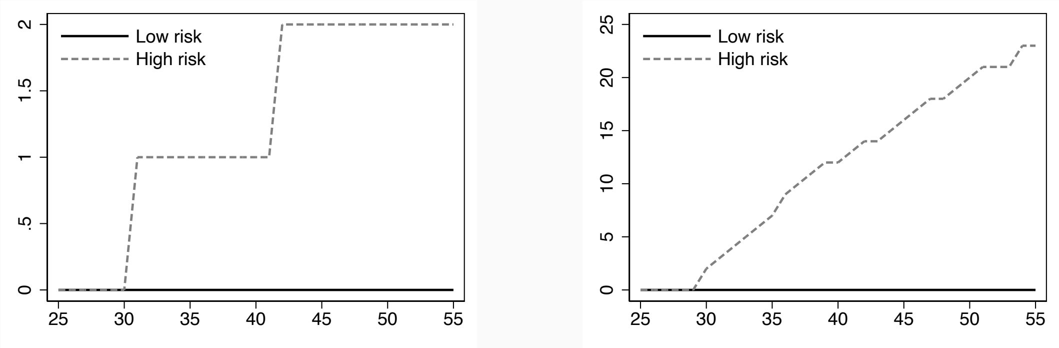

Notes: Risk types are defined based on lifetime earnings volatility where the high-risk type is mapped to the top 20% of the distribution whereas the remaining 80% are low-risk.

Back

Figure 12: Lifetime unemployment by risk group

(a) Unemployment spells by age

(b) Unemployment duration in months by age



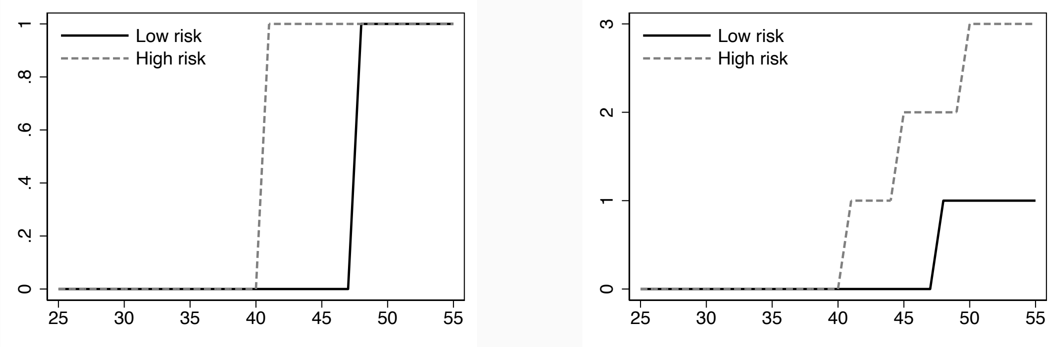

Notes: Life cycle profile of the number of independent unemployment spells and the total number of months spent as unemployed. Risk types are defined based on lifetime earnings volatility where the high-risk type is mapped to the top 20% of the distribution whereas the remaining 80% are low-risk.

# Appendix: Empirics – sickness by risk type, median

Back

Figure 13: Lifetime sickness by risk group

(a) Sickness spells by age (b) Sickness duration in months by age




Notes: Life cycle profile of the number of independent sickness spells and the total number of months spent as sick. Risk types are defined based on lifetime earnings volatility where the high-risk type is mapped to the top 20% of the distribution whereas the remaining 80% are low-risk.

# Appendix: Empirics – regressions for life-cycle plots

Back

**Table 6:** Risk types and income variables

Panel (a): OLS, first set of variables

| | (1) Earnings | (2) Cumulative earnings | (3) Unemployment | (4) Sickness |
|---|---|---|---|---|
| Risk type | -0.2166*** | -4.1647*** | 0.8727*** | 0.1379*** |
| | (-371.32) | (-519.09) | (322.19) | (102.87) |
| Observations | 3275531 | 3565948 | 3275531 | 3275531 |

Notes: OLS regression estimates of the outcome variables (each column) on risk type and a set of controls of year, age and state fixed effects as well as age-year and state-year interactions. Robust standard errors used. $t$ statistics in parentheses, $^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

Back

**Table 6:** Risk types and income variables

Panel (b): OLS, second set of variables

| | (1) Cumulative unemployment | (2) Cumulative sickness | (3) Cumulative spells unemployment | (4) Cumulative spells sickness |
|---|---|---|---|---|
| Risk type | 13.1707*** | 2.2103*** | 1.1174*** | 0.4127*** |
| | (597.62) | (273.33) | (401.71) | (215.88) |
| Observations | 3565948 | 3565948 | 3565948 | 3565948 |

Notes: OLS regression estimates of the outcome variables (each column) on risk type and a set of controls of year, age and state fixed effects as well as age-year and state-year interactions. Robust standard errors used. $t$ statistics in parentheses, $^{+}$ $p < 0.10$, $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$

# Appendix: Empirics – risk heterogeneity and workers' portfolios

Back

**Table 3:** Risk types and workers' portfolios

Panel (b): Quantile Regression: Median

| | (1) Earnings | (2) Net wealth | (3) Net wealth to earnings | (4) Illiquid share | (5) Average portfolio return |
|---|---|---|---|---|---|
| Risk type | -0.2182*** | -0.6103*** | -0.1013*** | -0.0909*** | -0.7827*** |
| | (-150.91) | (-93.77) | (-17.47) | (-41.52) | (-43.29) |
| Observations | 938130 | 929170 | 929170 | 929170 | 929170 |

Notes: OLS regression estimates of the outcome variables (each column) on risk type and a set of controls of year, age and state fixed effects as well as age-year and state-year interactions. Robust standard errors used. $t$ statistics in parentheses, $^{+}\ p < 0.10$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

**Main takeaway:** at median, results are amplified

# Appendix: Model – worker's problem & financial market friction

Back

## Incomplete markets life-cycle model with portfolio choice:

- Life-cycle of $J$ periods: working and retirement
  - working $j = 1, ..., J_{ret} - 1$: stochastic labor income
  - retirement $j = J_{ret}, ..., J$: retirement benefits
- Heterogeneous earnings processes by risk type
- Risk-averse agents: Consumption-smoothing and self-insurance

## Portfolio choice between two assets:

- Safe and liquid asset $a$: return $\pi$, traded without frictions
- Illiquid asset $k$: risky return $r$, trading friction
  - Commitment to saving plan $d$
  - Changes in $d$ at utility cost $\kappa \longrightarrow$ discrete choice
- Trade-off between return and liquidity: $\pi < r$

# Appendix: Model – worker's discrete choice

Back

**States:** $x_j = (a_j, k_j, d_j, y_{p,j}, y_{\tau,j})$: liquid & illiquid assets, saving plan, earnings states

**Adjusting or non-adjusting the saving plan:**

$$V_{X,j}(x_j) = \max_{(c_j, a_{j+1})} \Big\{ u(c_j) + \tilde{\beta}_j \mathbb{E}_{p,\tau} V_{j+1}(x_{j+1}) \;\Big|\; d_{j+1} = d_j, \quad k_{j+1} = (k_j + d_{j+1})\xi_{j+1} \geq 0,$$
$$c_j + a_{j+1} + d_{j+1} = (1+\pi)a_j + rk_j + y(y_{p,j}, y_{\tau,j}, j), \quad a_{j+1} \geq 0 \Big\},$$

where $V_{X,j}$ is either $V_{N,j}$ (non-adjusting) or $V_{A,j}$ (adjusting).

**Value function:** maximize over both branches, with taste shocks and utility cost

$$V_j(x_j) = \max \Big\{ V_{N,j}(x_j) + \zeta_{N,j}, \quad V_{A,j}(x_j) - \kappa + \zeta_{A,j} \Big\}$$

# References

# References


Angerer, X. and Lam, P.-S. (2009). **Income Risk and Portfolio Choice: An Empirical Study.** *The Journal of Finance* **64** (2), 1037–1055 (cit. on p. 34).

Arellano, M., Blundell, R. and Bonhomme, S. (2017). **Earnings and consumption dynamics: a nonlinear panel data framework.** *Econometrica* **85** (3), 693–734 (cit. on p. 33).

Bayer, C. and Kuhn, M. (2019). **Which ladder to climb? Decomposing life cycle wage dynamics.** Tech. rep. IZA Discussion Paper.

Bayer, C., Luetticke, R., Pham-Dao, L. and Tjaden, V. (2019). **Precautionary Savings, Illiquid Assets, and the Aggregate Consequences of Shocks to Household Income Risk.** *Econometrica* **87** (1), 255–290 (cit. on pp. 33, 34).

Blundell, R., Graber, M. and Mogstad, M. (2015). **Labor income dynamics and the insurance from taxes, transfers, and the family.** *Journal of Public Economics* **127**, 58–73 (cit. on p. 33).

Busch, C., Domeij, D., Guvenen, F. and Madera, R. (2022). **Skewed Idiosyncratic Income Risk over the Business Cycle: Sources and Insurance.** *American Economic Journal: Macroeconomics* **14** (2), 207–42.

Chang, Y., Hong, J. H., Karabarbounis, M., Wang, Y. and Zhang, T. (2020). **Income Volatility and Portfolio Choices.** Tech. rep. FRB Richmond (cit. on p. 34).

Drechsel-Grau, M., Peichl, A., Schmid, K. D., Schmieder, J. F., Walz, H. and Wolter, S. (2022). **Inequality and income dynamics in Germany.** *Quantitative Economics* **13** (4), 1593–1635.

# References


Fagereng, A., Guiso, L. and Pistaferri, L. (2018). **Portfolio Choices, Firm Shocks, and Uninsurable Wage Risk.** *The Review of Economic Studies* 85 (1), 437–474 (cit. on p. 34).

Guiso, L., Jappelli, T. and Terlizzese, D. (1996). **Income Risk, Borrowing Constraints, and Portfolio Choice.** *The American Economic Review* 86 (1), 158–172 (cit. on p. 34).

Guvenen, F., Karahan, F., Ozkan, S. and Song, J. (2015). **What do data on millions of US workers reveal about life-cycle earnings risk?** Tech. rep. National Bureau of Economic Research (cit. on p. 33).

— (2021). **What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics?** *Econometrica* 89 (5), 2303–2339 (cit. on pp. 11, 12).

Hall, R. E. (1982). **The Importance of Lifetime Jobs in the U.S. Economy.** *The American Economic Review* 72 (4), 716–724 (cit. on p. 33).

Hamilton, J. D. (1988). **Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates.** *Journal of Economic Dynamics and Control* 12 (2), 385–423.

— (1989). **A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle.** *Econometrica* 57 (2), 357–384.

Heaton, J. and Lucas, D. (2000). **Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk.** *The Journal of Finance* 55 (3), 1163–1198 (cit. on p. 34).

— (2001). **Portfolio Choice in the Presence of Background Risk.** *The Economic Journal* 110 (460), 1–26 (cit. on p. 34).

# References


Jordà, Ò., Knoll, K., Kuvshinov, D., Schularick, M. and Taylor, A. M. (2019). **The Rate of Return on Everything, 1870–2015*.** *The Quarterly Journal of Economics* **134** (3), 1225–1298 (cit. on p. 50).

Jung, P. and Kuhn, M. (2018). **Earnings Losses and Labor Mobility Over the Life Cycle.** *Journal of the European Economic Association* **17** (3), 678–724 (cit. on p. 33).

Karahan, F. and Ozkan, S. (2013). **On the persistence of income shocks over the life cycle: Evidence, theory, and implications.** *Review of Economic Dynamics* **16** (3), 452–476 (cit. on p. 33).

Karahan, F., Ozkan, S. and Song, J. (2019). **Anatomy of lifetime earnings inequality: Heterogeneity in job ladder risk vs. human capital.** *FRB of New York Staff Report* ( 908) (cit. on p. 33).

Kim, C.-J. (1994). **Dynamic linear models with Markov-switching.** *Journal of Econometrics* **60** (1), 1–22.

Kim, C.-J. and Halbert, D. C. R. (1999). ***State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications.*** MIT press.

Kuhn, M. and Ploj, G. (2020). **Job stability, earnings dynamics, and life cycle savings.** Tech. rep. mimeo (cit. on p. 33).

Larkin, K. P. (2023). **Job risk, separation shocks and household asset allocation.** Tech. rep. (cit. on p. 34).

# References


Low, H. and Pistaferri, L. (2010). **Disability risk, disability insurance and life cycle behavior.** Tech. rep. National Bureau of Economic Research.

Manuel, A., Blundell, R. and Stéphane, B. (2017). **Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework.** *Econometrica* **85** (3), 693–734 (cit. on p. 33).

McKay, A. (2017). **Time-varying idiosyncratic risk and aggregate consumption dynamics.** *Journal of Monetary Economics* **88**, 1–14 (cit. on p. 34).

Morchio, I. (2020). **Work Histories and Lifetime Unemployment.** *International Economic Review* **61** (1), 321–350 (cit. on p. 33).

Mukoyama, T. and Şahin, A. (2006). **Costs of business cycles for unskilled workers.** *Journal of Monetary Economics* **53** (8), 2179–2193 (cit. on p. 33).

Palia, D., Qi, Y. and Wu, Y. (2014). **Heterogeneous Background Risks and Portfolio Choice: Evidence from Micro-level Data.** *Journal of Money, Credit and Banking* **46** (8), 1687–1720 (cit. on p. 34).

Schmillen, A. and Möller, J. (2012). **Distribution and determinants of lifetime unemployment.** *Labour Economics* **19** (1), 33–47 (cit. on p. 33).

Storesletten, K., Telmer, C. I. and Yaron, A. (2004). **Cyclical Dynamics in Idiosyncratic Labor Market Risk.** *Journal of Political Economy* **112** (3), 695–717 (cit. on p. 33).