Occupational Choice, Income Risk and Preference Heterogeneity

Sebastian Findeisen*
University of Zurich

April 5, 2010

Preliminary and incomplete, please do not cite without permission

Abstract

This paper presents a framework to empirically study how uninsurable income risk affects occupational choices. I use a rich German panel data set to estimate income risk-parameters, transitory and permanent, across occupations. The data also includes information on individuals' attitudes to risk, implying substantial heterogeneity. I find evidence for self-selection into occupations on risk-preferences, as more risk-tolerant work in occupation, with more permanent and transitory risk. A dynamic incomplete markets model featuring partial insurance is used to derive welfare expressions in terms of previously estimated individual-occupation specific wage growth rates and occupation specific income risk. I estimate a random utility mixed-logit model, allowing for heterogeneous non-pecuniary preferences for occupations, using the welfare expressions. Eliminating uninsurable income risk in a counterfactual experiment is found to have a negligible impact on occupational choices. Calibrating the welfare expression with imputed heterogeneous coefficients on risk-aversion fits the data better than the same model assuming homogeneous risk-preferences.

Keywords: Occupational Choice, Preference Heterogeneity, Income Risk, Incomplete Insurance

JEL Classification: J24; J31; D81

---

*Contact: sebastian.findeisen@uzh.ch. I am indebted to Fabrizio Zilibotti for ongoing support and advice. I also thank Matthias Doepke, Christoph Winter and seminar participants in Zurich for helpful comments and suggestions. The usual disclaimer applies.

1 Introduction

The estimation and consequences of (US) idiosyncratic labor income risk have been researched extensively in the last years. This includes both, a large empirical literature using various methods to estimate income risk mostly using data from the Panel Study of Income Dynamics (PSID), and an expanding (macro) theoretical literature using dynamic models featuring incomplete insurance markets to asses and predict how agents react to these shocks and the implications for aggregates and welfare.[1] In this paper, rather than trying to address the whole set of consequences of labor income risk using theory, I focus on one particular margin, namely how income risk affects occupational choices. To this end, I present and implement an empirical framework, using data from the German Socioeconomic Panel (SOEP).

Important for the empirical strategy is to allow for heterogeneity in multiple dimensions. First, income risk parameters, decomposed into a transitory and a permanent part, are estimated across occupations. Second, differences in abilities are incorporated by predicting income streams of individuals across occupations. Third, individuals are allowed to have heterogeneous preferences for unobserved attributes of occupations, not related to economic fundamentals (i.e. the distribution of the income stream).[2] Fourth, in parts of the analysis I will allow for heterogeneous risk preferences. As an input for the structural parameter (the coefficient of constant relative risk aversion), I will use survey and wealth information, directly available in the data set. The analysis has two main focuses, I aim to shed light on: a) As long as labor income risk remains uninsurable, how does this market incompleteness affect the occupational margin? If all risk were insurable, would sorting patterns change? b) When risk-preferences are heterogeneous people will self-select into occupations, at least to some extent, on risk tolerance. I test, if preference data elicited by surveys can help to predict selection patterns, and if a model using this data can improve upon a model assuming homogeneous preferences (the standard approach).

To asses the effect of incomplete markets on occupational choices, a model is needed, imposing structure on the degree of insurance. I adopt a dynamic incomplete markets model from Krebs (2004) and Krebs, Krishna, and Maloney (forthcoming). Importantly, the model predicts partial insurance: transitory shocks are perfectly smoothed

[1] A highly selective list of recent empirical papers includes Meghir and Pistaferri (2004), Carrol and Samwick (1997), Guvenen (2009) and Storesletten, Telmer, and Yaron (2004). For a comprehensive and impressive overview on theoretical developments see Heathcote, Storesletten, and Violante (2009).

[2] These unobservable attributes include the availability of jobs, status, working time, or just how salient a certain occupation is, for an individual making his occupational choice.

out, whereas persistent shocks translate into consumption fluctuations.[3] The model delivers a formula for the utility of individuals across occupations, as a function of income uncertainty and average wage growth rates. I use this expression together with occupation dummies to estimate a random utility (mixed-logit) model, explaining occupational choices. Given the results, I can counterfactually ask, if individuals had chosen different occupations under complete markets and insurable persistent shocks.

The available data on individual risk-preferences show big heterogeneity across individuals. Assuming constant relative risk-aversion (CRRA) utility, I can infer individual risk-aversion coefficients. I then test by simple linear regressions and an ordinary ordered probit model, if higher relative risk-aversion translates into working in riskier occupation. Additionally, I use the imputed CRRA coefficient as an input for the "structural" formula from the theoretical model, and test whether the fit of the discrete choice model improves vis-a-vis the standard procedure of assuming homogeneity in preferences (a representative agent).[4]

I find that only a small number of individuals (about 3%) are predicted to choose occupations differntly under insurable income risk. Most of the movements is towards similar (as defined by belonging to the same broad categoty according to ISCO 88 classification) but higher paying occupations. In line with previous reserach (see below) people with a higher risk tolerance have a higher probability to work in a riskier occupation. This is true for both, permanent *and* transitroy income risk. Fitting the discrete choice model with the imputed CRRA parameters improves the fit against assuming identical risk preferences- in a nutshell, heterogeneity matters.

Previous work has addressed the self-selection of individuals on risk attitudes into occupations using quite different approaches. Fuchs-Schündeln and Schündeln (2005) use a difference-in-difference design to exploit that occupational choice was determined by largely political reasons and, therefore, orthogonal to risk-preferences in the Eastern Part, but not so in the West of Germany. Their contribution delivers and clean and credible identification by a quasi-natural experiment. However, the self-selection is only broken down to working in the private or public sector, and their approach does not distinguish between occupations, which are characterized by substantial heterogeneity in income risk as shown in this paper, at a finer level. Schulhofer-Wohl (2010) splits respon-

[3]This follows the consensus view in the literature. See Jappelli and Pistaferri (2010) for a survey of empirical papers.

[4]Dohmen, Falk, Huffmann, Sunde, Schupp, and Wagner (2005, forthcoming) conduct a field experiment validating the behavioral implications of answers to the risk-questions in the SOEP. The influential paper by Barsky, Juster, Kimball, and Shapiro (1997) pioneers the use of hypothetical gambles to elicit risk-preferences, using the Health and Retirement Study. Guiso and Paiella (2008) use household survey data from the Bank of Italy to construct a measure of absolute risk-aversion.

dents of the Health and Retirement Study into two groups of risk-tolerant and risk-averse workers. The former group sees their (total) income fluctuate more with GDP, which can be interpreted as evidence in favor of the sorting hypothesis. Finally, Bonin, Falk, Huffman, and Sunde (2007) use a cross-section (also from the SOEP) and show that men declaring a higher propensity to take risks, work in occupations with a larger residual in Mincerian wage regressions. My method, chosen in this paper, adds to previous contributions by estimating a panel model of income risk, allowing for unobserved effects across individuals and occupations. This is important since a dynamic model allowing for unobserved heterogeneity gives a better picture of income risk across occupation than a simple cross-section. Furthermore, connecting to the large empirical literature on idiosyncratic income risk, I carefully distinguish between transitory and permanent shocks across occupations. This matters, since theory predicts different effect for persistent and non-persistent innovations. I then investigate, if risk-preferences are important for self-selection using both, a more structural approach imposing assumption on the data guided by theory esp. with respect to the degree of insurance, and reduced-form estimates. The structure imposed enables me to investigate to what degree market incompleteness affects occupational choices.

The structure of the paper is as follows. Section 2 describes how occupational income risk is defined, identified, and estimated. In Section 3, I present evidence of the self-selection on risk attitudes and how income uncertainty and expected wages are related across occupations. Section 4 then estimated the selection into occupations, guided by a simple theory. Section 5 concludes, finally.

2 Occupational Income Risk

A whole series of papers, what might be called an own literature, has used various approaches to estimate US labor income risk, using the PSID with data ranging from the late sixties to the early nineties (among them Carrol and Samwick 1997, Meghir and Pistaferri and Storesletten, Telmer, and Yaron 2004, Guvenen 2009). Building on those, I define and estimate income risk as the variance of unpredictable changes in individual income. Like in previous studies, I will disentangle permanent from transitory innovations to income. Later, when the occupational choice margin is considered, the focus will be on the influence of permanent shocks. Blundell, Pistaferri, and Preston (2008) find that transitory income shocks don't transmit to consumption, concluding there exists partial insurance against temporary fluctuations. Consistent with this evidence I will allow only permanent shocks to affect occupational choices. In the econometric specification, moreover, transitory shocks will pick up measurement error in income, biasing the estimates of true temporary income risk across occupations.

2.1 Specification

Income Definition. The data provide information on an individual's occupation and yearly labor income. Importantly, I observe the labor income that has been earned in the main occupation of an individual. Total labor income might differ from labor income in the main occupation, when, for example, individuals hold secondary, part-time jobs. Since later income shocks will be attributed to the major occupations held, the distinction becomes important. Because only people working full-time are included, the difference is likely to be small, however.

About 30% of people report at least one unemployment spell in their working history, defined by receiving unemployment benefits or assistance, or subsistence (welfare) payments in at least one year. This indicates that any riskiness measure for occupations should explicitly take unemplyoment risk into account. I, therefore, attribute all government transfers explicitly linked to the unemplyoment status to the labor income of the individual in the given year. During an unemployment spell, the last occupation held is assigned to the individual, motivated by the fact that benefits are, in general, linked to past wages.

Income Process. The log of an individual i's labor income working in occupation j at time t is assumed to be governed by the following process:

$$w_{ijt} = \lambda_i + \alpha_{jt} + \beta_j \cdot x_{ijt} + u_{ijt}, \tag{1}$$

where λ_i is occupation and time-independent personal fixed-effect, α_{jt} is an occupation-time trend effect, and x_{ijt} captures time-varying individual effects. Since the fixed effect only allows for variation within individuals over time, x_{ijt} contains age and age squared. Note that β_j is free to vary across occupations, implying that returns to seniority (experience) are heterogeneous.[5] Income variation in the data caused by different life-cycle patterns of income across occupations are, hence, not contained in the error u_{ijt}.

Recently, Guvenen (2007, 2009) has revived interest in the distinction between econometric specifications using *heterogeneous income profiles* (HIP) and *restricted income*

[5] After controlling for individual and occupation-time fixed effects, the variance which is left to identify β_j comes from variation of given individuals across occupations (occupation switchers) and variation within individuals overtime. The group of switchers is small, however.

profiles (RIP). In the first view, individuals are characterized by heterogeneous income profiles over their life-cycle, implying a relatively low-persistence of shocks. In the second view, growth rates are restricted to be homogeneous across individuals, conditional on the educational group. Heterogeneity comes from the large persistence of shocks, which have been estimated to very closely resemble an unit-root process.[6] In the current context of *occupation specific* income risk, my approach can be understood as taking a middle ground. Estimating (1) allows for heterogeneous wage growth across occupations, but not explicitly across individuals.

Income Risk: Identification. I decompose the error into the sum of two components:

(2) $$u_{ijt} = \omega_{ijt} + \epsilon_{ijt},$$

where ω_{ijt} is the permanent component. ω_{ijt} is referred to as permanent, since innovations to its process have lasting effects on labor income, as captured by the random walk:

$$\omega_{ijt} = \omega_{ijt-1} + \eta_{ijt},$$

where innovations η_{ijt} are iid normal over time, occupations, and individuals, so $\eta_{ijt} \sim N(0, \sigma^2_{\eta_j})$. For different occupations these innovations are, hence, drawn from different distributions, as indicated by $\sigma^2_{\eta_j}$. Relatedly, the transitory shocks have no persistence and are drawn from a normal, so $\epsilon_{ijt} \sim N(0, \sigma^2_{\epsilon_j})$ with a occupation-specific variance. My specification and the decomposition into permanent and transitory components follows the previous work estimating income risk using the PSID, adopted to the purpose of distinguishing wage risk across occupations.

2.2 Estimation

Taking the n-th difference between two residuals u_{ijt} in (2) yields:

(3) $$\Delta_n u_{ijt} = u_{ijt+n} - u_{ijt} = \eta_{ijt+1} + \ldots + \eta_{ijt+n} + \epsilon_{ijt+n} - \epsilon_{ijt}.$$

Applying the variance operator:

(4) $$var\left[\Delta_n u_{ijt}\right] = n\sigma^2_{\eta_j} + 2\sigma^2_{\epsilon_j}.$$

[6] See Guvenen (2009) for an enlightening discussion and a reassessment of the competing specifications.

In their influential paper Carrol and Samwick (1997) propose to estimate a similar specification like (4) by OLS. The LHS variable is conveniently obtained by taking the (estimated) squared differences $(\Delta_n \hat{u}_{ijt})^2$. $\sigma^2_{\eta_j}$ and $\sigma^2_{\epsilon_j}$ are then estimated as coefficients on the regressors n and 2. Measurement error in the LHS will add noise and inflate standard errors, but will not attenuate the estimates. To account for occupational switching, I estimate a more general version of (4):

$$\begin{aligned} var\left[\Delta_n u_{ijt}\right] &= \sigma^2_{\eta_1}\sum_{k=1}^{n} I_{1k=1} + \cdots + \sigma^2_{\eta_J}\sum_{k=1}^{n} I_{Jk=1} + \sigma^2_{\epsilon_1}\sum_{m=1}^{2} I_{1m=1} + \cdots + \sigma^2_{\epsilon_J}\sum_{m=1}^{2} I_{Jm=1} \\ &= \sum_{j=1}^{J}\left[\sigma^2_{\eta_j}\sum_{k=1}^{n} I_{jk=1} + \sigma^2_{\epsilon_j}\sum_{m=1}^{2} I_{jm=1}\right], \end{aligned} \tag{5}$$

where $I_{jk=1}$ is an indicator function, picking up the number of times a permanent innovation from distribution j is drawn, in between two periods with distance n, and $I_{jm=1}$ picks up the number of transitory shocks; note that for any difference $\Delta_n u_{ijt}$ only two such shocks enter.

Potential Caveats. Implicit in my identification strategy is that individual labor income uncertainty as captured by the second moments of the permanent and transitory shocks is attributed to the occupation(s) held by a particular individual. If some workers see the unexplainable part of their earnings fluctuate for reasons not related to their occupation, this may bias my estimates since the variance is falsely attributed to their occupation rather than to the individual. Suppose two types of workers with occupation independent income risk exist- high risk workers and low risk workers. If there is systematic selection of one group into one occupation, my estimation strategy would pick up these shocks and wrongly connect them to the occupation held. The presence of both (significant) occupation-independent income risk and self-selection could, hence, be dangerous for my empirical strategy. As long as workers' idiosyncratic risk is (reasonably) independent of other factors determining the self-selection into occupations, however, this will not affect the results.

In accordance with the bulk of the literature, but in contrast with some rare contributions (Meghir and Pistaferri 2004, Krebs, Krishna, and Maloney forthcoming), I do not estimate a flexible model with time-varying income risk parameters. That is I assume that the distribution of innovations for a given occupation is constant over time. Since I estimate income risk parameters over a comparably big set of 28 occupations, for some

Table 1: Occupations

Occupation Code	Description
11	Legislators and Senior Officials (inc. senior gov. and party officials)
12	Corporate Managers (incl. responsibilities for certain divisions/departments)
13	General Mangers (incl. responsibilities for whole companies)
21	Science Professionals
22	Life Science and Health Professionals
23	Teaching Professionals (includes professors, school teaching which requires educational background
24	Other Professionals (inc. accountants, lawyers and others)
31	Physical and Engineering Professionals
32	Life Science and Health Associate Professionals (includes medical assistance and nursing)
33	Teaching Associate Professionals (incl. teaching not requiring higher educational background)
34	Other Associate Professionals (includes estate agents and travel consultants)
41	Office Clerks
42	Customer Service Clerks
51	Personal Services Workers (includes personal care workers and barbers)
52	Shop Salespersons/Models
61	Agricultural and Fishery Workers
71	Extraction and Building Trade Workers
72	Metal, Machinery, and Related Trades Workers
73	Precision, Handicraft, Printing and Related Trades Workers
74	Other Craft and Related Trades Workers (includes textile, wood and processing)
81	Stationary-Plant and Related Operators
82	Machine Operators and Assemblers
83	Drivers and Mobile-Plant Operators
91	Sales and Services Elementary Occupations (incl. street vendors and door-to-door salesmen)
92	Laborers in Agriculture and Fishery
93	Laborers in Mining, Construction, Manufacturing and Transport
99	Laborers without a Defined Role
110	Armed Forces

periods and occupations, only a small number of observation are available, which makes it difficult to obtain precise estimates of time varying income risk.

2.3 Data

The (G)SOEP is a representative panel survey of the adult population in Germany. The SOEP came into life 1984, and for this study all waves from 1984 to 2008 are used. It surveys all members of a household on wide range on economic and non-economic topics. Importantly, information on the occupation held by an individual is included, according to the International Standard Classification of Occupations (ISCO88) by the International Labour Organization. Throughout the analysis, the two digit occupations code (plus the armed forces) will be used, differentiating between 28 occupations. Table 2 summarizes and describes them. The earnings variable is the annual labor income, derived from the main occupation of an individual, which includes all compensation received, i.e. wages, bonuses, commissions etc. from that source. Like motivated above, if during any year an individual goes through a spell of unemployment, earnings include all benefits and government transfers linked to the labor market status of the individual.

The sample is restricted to males working full-time, for which I observe a labor market history of at least 15 years (the maximum being 25). In particular, only those observations are used for which the individual is either working full-time or unemployed. The analysis is, moreover, restricted to Western Germans, since occupational in the former German Democratic Republic (GDR) was severely restricted. Also Fuchs-Schündeln and Schündeln (2005) document that labor income risk seemed to completely absent in the GDR. This creates a non-balanced panel of 1838 individuals, with complete information for 19 years on average, and 34831 observations in total.

A part of the later analysis will combine survey and wealth data to impute individual risk-aversion. Details of the procedure are provided in the Appendix. Much more complete and detailed information on the survey questions and, very importantly, their usefulness to deliver input for empirical research, is provided in Dohmen et al. (2005, forthcoming).

2.4 Results

Table 2 presents the estimated income risk across occupations by fitting a linear regression to (5), where the data is pooled across persons, occupations, and time periods, which provides 322677 observations used for identification. Across occupations transitory shocks are estimated to be much larger in magnitude than persistent innovations. This finding is consistent with the literature dealing with US labor income risk. On average, the ratio of transitory to permanent shocks is about 13, where one should keep in mind that transitory innovation are likely to be *upward biased* in the presence of measurement error in labor income.[7]

Since the existing literature has almost exclusively focused on the US with data drawn from the PSID, it is instructive to compare the evidence for German labor income risk with previous results. Pooling across occupation, the average permanent component is estimated around 0.0062. In contrast, using different methods, most of the other studies have found a value in the ballpark of 0.020 (among them Meghir and Pistaferri 2004, Caroll and Samwick 1997 and Storesletten, Telmer, and Yaron 2004).[8] Finally, Krebs, Krishna, and Maloney (forthcoming) use Mexican data, mainly from the nineties, and find an annualized variance of 0.0320. To be clear, the specification of the wage process used

[7]The number is calculated excluding the two occupations, for which permanent income risk is numerically indistinguishable from zero.

[8]Storesletten, Heathcote, and Violante (2008) even use a value of 0.2200 in their welfare calculation. Their specification is different from the rest of the literature, however, since the permanent component is tied to a person fixed-effect, picking up innate differences in abilities.

Table 2: Estimates of Persistent and Transitory Income Risk

Occupation Code	Permanent Shock	Transitory Shock	Occupation Code	Permanent Shock	Transitory Shock
11	0.011	0.0035	52	0.0030	0.0571
12	0.0051	0.0430	61	0.0083	0.0343
13	0.0132	0.2648	71	0.0031	0.0972
21	0.0034	0.0456	72	0.0042	0.0952
22	0.0023	0.0078	73	0.0037	0.0676
23	0.0164	0.0000	74	0.0073	0.0887
24	0.0020	0.0442	81	0.0049	0.0581
31	0.0080	0.0759	82	0.0091	0.0833
32	0.0064	0.0652	83	0.0183	0.0471
33	0.000	0.0571	91	0.0000	0.1145
34	0.0026	0.0665	92	0.0007	0.0237
41	0.0046	0.0620	93	0.0064	0.0896
42	0.0006	0.0071	99	0.0000	0.1149
51	0.0173	0.0278	110	0.0066	0.0097

Based on pooled regression with 322,677 observations.

in this paper controls for a bigger set of covariates than in most other studies. In particular, my empirical strategy includes person-fixed effects, occupation-time-fixed effects and different slope coefficients to capture heterogeneous returns to age across occupations. It is likely that this relatively rich specification reduces the variance of what we label the unexplainable part of earnings. It remains, hence, an open question if my estimates can be directly interpreted as evidence that German workers faced less labor income risk than their American counterparts.

Using the sum of transitory and permanent components as a measure of total labor income risk in a given occupation yields intuitive patterns, at first glance. Managers, salespersons, laborers in construction, mining, extraction, and related trades face great uncertainty; doctors, soldiers, professors, and teachers face safe income streams. Just ranking occupations with respect to their permanent risk component, however, reveals quite different insights. According to this ordering, personal service workers, drivers, and also professors and teachers are subject to relatively large persistent shocks. So, for ex-

Table 3: Income Variance and Expected Wages

All Occupations	Permanent Shocks	Transitory and Permanent Shocks
Expected Wage Growth	-0.456***	-2.7***
	(0.061)	(0.44)
Expected Starting Wage	-0.059	-0.15***
	(0.04)	(0.03)
Observations	37828	37828
Only Occupations chosen	**Permanent Shocks**	**Transitory and Permanent Shocks**
Expected Wage Growth	-1.33***	-6.36***
	(0.32)	(1.94)
Expected Starting Wage	-0.050**	-1.53***
	(0.022)	(0.14)
Observations	1832	1832

Robust std. errors (in parentheses). Constant included, but results omitted. Estimated coefficients multiplied by 100.

Statistically different from zero at *10%, **5% and ***1% level

ample, university professors seem to face minimal unpredictable, temporary fluctuations, but are subject to permanent innovations like being tenured. On the other hand, laborers in agriculture tend so see low persistent uncertainty, but see their wages fluctuate temporarily, for example, caused by seasonal variations.

The next sections take a closer look on how uncertainty and expected wages are related across occupations, and present first evidence on the sorting into occupation on attitudes to risk.

3 Risk-Preference Heterogeneity, Self-Selection and Variation Across Occupations

3.1 Bad Occupations?

Table 3 shows the result of regressing both, permanent and total (permanent + transitory) income risk on expected average wage growth and the expected (log) base wage. Individual-occupation specific average income growth rates are recovered by taking first differences to the predicted values for income:

$$\mu_{ij} = (\hat{\alpha}_{jt} - \hat{\alpha}_{jt-1}) + (\hat{\beta}_{jt} - \hat{\beta}_{jt-1})(x_{ijt} - x_{ijt-1}), \quad (6)$$

and averaging over time. The expected base wage is simply the predicted staring wage in logs for an individual i starting to work in an occupation j at time t. Perhaps surprisingly, the coefficients show a significant negative relationship- high-risk occupations do not seem to compensate with a higher mean! Back of the envelope calculations, however, imply that the magnitudes are rather small- a one standard deviation increase in the expected wage growth rate is associated with 0.04 standard deviations decrease in the variance of permanent shocks, and 0.002 standard deviations in the case of total shocks. Restricting the sample to the occupations actually chosen shows that the relationship is even more negative. So at least from the point of an agent forming expectations over expected income streams across different occupations, there is evidence of strict dominance of certain occupations over others in terms of the first two moments of the wage distribution. Still we observe occupations picked, which are dominated in pecuniary terms, because of non-monetary attributes, which are later taken care off in the estimation procedure.

3.2 Preference Heterogeneity and Self-Selection

Arguably, the simplest way to study the relationship between risk preferences and the self-selection into occupations is a simple regression framework, linking income risk across occupations to preferences. The data allows to impute individuals' CRRA coefficients, by combining wealth and survey data. A detailed account of the procedure can be found in the appendix, as well as in Dohmen et al. (2005, forthcoming).[9]

Table 4 presents the result of two regressions of income risk on an individual's (imputed) coefficient of relative risk-aversion, controlling for the expected growth rates of income.[10] Under the assumption of classical measurement error, this leads to attenuation bias in the estimated coefficients. Individuals of higher risk tolerance go into riskier occupations, as indicated by the results. The effect is even more pronounced for total income risk, despite that theories featuring effective insurance against transitory shocks would dictate that there shouldn't be a difference. Since the LHS is restricted to taking on 27 different values, additionally, an ordered probit model is estimated, in which occupations are ranked by their respective income risk. Agents are predicted to be more likely to be

[9]The imputation relies on a combination of a survey question asking to indicate the willingness to take risk *in general*, and a lottery question. Bonin et al. (2007) use the responses to the general risk question in their analysis. There are good reasons, however, to work with the imputed CRRA coefficients as a measure of risk-aversion instead. First, as the point is made in Dohmen et al. (2005, forthcoming), the general risk question likely picks up *risk-perceptions* and *risk-attitudes*, whereas the lottery focuses on the variable of interest, namely risk-aversion. Additionally, the CRRA coefficients control for individual wealth levels, which matter if absolute risk-aversion is not constant.

[10]The number of individuals for whom a risk-aversion parameter can be imputed is 1435, due to either missing wealth data or non-responses to the risk question(s).

Table 4: Income Risk and Heterogeneous Risk Aversion

OLS	Permanent Shocks	Transitory and Permanent Shocks
CRRA Coefficient	-4.09*	-62.28***
	(2.42)	(14.18)
Observations	1435	1435
Ordered Probit	Permanent Shocks	Transitory and Permanent Shocks
CRRA Coefficient	-0.12**	-0.234***
	(0.05)	(0.05)
Observations	1435	1435

Robust std. errors (in parentheses). Constant and Income Growth included, but results omitted. Estimated coefficients multiplied by 10^5 in OLS specification.

Statistically different from zero at *10%, **5% and ***1% level

active in more risky occupations, the smaller their imputed CRRA coefficient. Looking at the marginal effects evaluated at the sample mean (results available on request), one can infer that the likelihood of choosing a low income risk occupation is increasing in risk-aversion. Up to the 12th “safest”occupation (metal, machinery, and related trades workers) the marginal effect of the CRRA coefficient is positive; starting from rank 13 (office clerks) it gets negative. Using the sum of the variance of both shocks, effects get larger in magnitude, indicating that even transitory occupational income risk, said to well insurable, correlates with risk preferences.[11]

The evidence in the previous section suggested that non-pecuniary aspects play a big part on the occupational choice margin, since some occupations seem to dominate other with regards to the first two moments of the wage distribution. Despite this big “noise”in the sorting process, risk-preferences robustly correlate with occupational labor income risk for both persistent and transitory shocks. The next section, finally, tries to link the influence of economic fundamentals (risk-preferences, income risk) and non-monetary, unobservable attributes in the occupational choice process.

4 Self-Selection Into Occupations

I opt for a parsimonious framework of modeling the self-selection into occupations. I don’t aim to structurally model all the potential margins affecting occupational choice, of which most are unobserved to the econometrician. To make the influence of the economic

[11] A plausible objection to the conclusions concerns the (interesting) possibility that preferences are endogenous; i.e. people with higher labor income risk learn to cope with that risk. To test this, a panel containing repeated observation on risk-aversion would be needed. These data could become available within the next years, but is not yet.

variables of biggest interest (income risk, expected wage growth) and their interaction with risk-aversion salient, I specify the latent utility function of agents to depend on two (additively separable) parts. One captures the expected utility of risk-averse agents associated with the expected *consumption streams* across the occupations in their choice set. This part of the latent utility function is rooted in a simple but tractable dynamic incomplete market model, making predictions how agents respond to uninsurable permanent income shocks across different occupations. The second part will consist of simple occupation dummies, meant to capture all unobservable attributes and the agents' preferences for these attributes. So we can denote by U_{ij} the (latent) utility of individual i from occupation j:

$$U_{ij} = V_{ij}\left(\gamma_i, \mu_{ij}, \sigma^2_{\eta_j}\right) + \delta_{ij} X_j,$$

where X_j is an indicator variable for each occupation in the choice set of an individual. This specification allows for multidimensional preference heterogeneity. β_i captures the effect of all non-pecuniary effects unrelated to the income process of an occupation. Preference heterogeneity in this dimension comes as a natural assumption in the context of occupational choice. Moreover, V_{ij}, the expected lifetime utility from consumption streams in occupation j, allows for different degrees of risk-aversion. That is, I will follow the literature and specify the flow utility function to be of the CRRA form with coefficient γ, but will in parts of the application explicitly allow for differences in γ_i, implying that permanent income risk and income growth rates have heterogeneous effects across individuals.

The specification implies that the model is solved and estimated under the assumption, that individuals decide on occupations only once in their working life. If the data show that occupational mobility is important and switching frequent, this could be too strong an assumption. It turns out that 62% of individuals hold only two or one occupation in their work history I observe (which varies between 15 and 25 years). About 85% hold a total of three or less occupations. Moreover, of these 85% the median time spent working in one's main occupation, defined as the mode, is 80%. To account for the small amount of occupational switching, I, hence, define the chosen occupation of an individual as her mode occupation of her observable work history. Averaging over all individuals and time periods, this definition of the "main" occupation is equal to the actual occupation in 76% of all cases. Additionally, much of the observed occupational mobility is driven by forth-and-back switching; people leaving their mode occupation for a short spell before returning. This indicates that specifying a model, in which agents make occupational choices, acting as if they select only one major occupation seems appropriate.[12]

The next section describes the theoretical foundation, adopted from Krebs (2004) and Krebs, Krishna, and Maloney (forthcoming), to recover V_{ij}. Finally, I explain how the model is identified as well as the details of the estimation.

4.1 Income Risk and Occupational Choice: A Simple Theory

At the beginning of her working life an agent decides on one occupation. After that, she is assumed to remain in this occupation. This makes the occupational choice problem static, whereas optimal consumption/saving decisions are made dynamically.

Conditional on having chosen an occupation, agents face persistent income shocks. Labor income y_{itj} of individual i at time t in occupation j follows the law of motion:

$$y_{ijt} = (1 + \mu_{ij})(1 + \upsilon_{jt})y_{ijt-1}, \tag{7}$$

where $(1+\mu_{ij})$ is the growth rate of worker i's wage in occupation j, and υ_j is an occupation specific shock. Staying close to the literature estimating the welfare consequences of income risk, I assume that $log(1 + \upsilon_j)$ is normally distributed with occupation dependent variance $\sigma^2_{\upsilon_j}$ and mean $-\sigma^2_{\upsilon_j}/2$.[13] Workers are allowed to save at the risk free rate r_t, but have no opportunity to borrow in financial markets, leading to a standard sequential budget constraint:

$$y_{ijt} = c_{ijt} - (1 + r_t)a_{ijt} - a_{ij,t+1}, \tag{8}$$

with $a_{ijt} \geq 0$. For each possible occupation j (of which only is chosen in equilibrium), agents choose the optimal sequence of consumption and assets $\{c_{ijt}, a_{ij,t+1}\}_{t=0}^{\infty}$ over an infinite time span:

$$\max_{\{c_{ijt},a_{ij,t+1}\}_{t=0}^{\infty}} V_{ij} = \sum_{t=0}^{\infty} \beta^t u(c_{ijt}), \tag{9}$$

[12]Kambourov and Manovskii (2009) find strong evidence that human capital is occupation specific, as opposed to industry or employer linked, implying there are strong returns to occupation specific tenure. When agents anticipate this the incentive to occupation switching strongly decrease, one might interpret observed changes as unexpected shocks, giving support to a model trying to explain the selection of individuals into their main occupation.

[13]This ensures $E[\upsilon_j] = 0$, since the moment generating function of a normal with mean μ and variance σ^2 is given by $e^{t*\mu+0.5t^2\sigma^2}$. This is an especially attractive feature, when one is interested in counterfactual experiments changing income risk, since it leaves the mean growth rate unaffected.

where $u(c_{ijt}) = \frac{u(c_{ijt})^{1-\gamma}}{1-\gamma}$.[14] In an equilibrium all markets clear, so $\sum a_{ijt} = 0$ in every t. Conditional on occupational choices, the equilibrium notion is adopted from Krebs (2004) and Krebs, Krishna, and Maloney (forthcoming). Importantly, households will not use asset holdings to smooth out permanent shocks, which will be shown now.

The equilibrium interest rate r_{t+1} obeys:

$$(10) \qquad r_{t+1} = \min_{\{ij\}} \frac{u_c(c_{ijt})}{\beta E[u_c(c_{ijt+1})]} - 1 = \min_{\{ij\}} \left(c_{ijt}^{-\gamma} [\beta(1+\mu_{ij})^{-\gamma} E\left[(1+\upsilon_{jt})^{-\gamma}\right]]^{-1} \right) - 1.$$

The price q_{t+1} of the riskless asset is then just given by the inverse of the gross interest rate:

$$(11) \qquad q_{t+1} = \frac{1}{1+r_{t+1}} = \max_{\{ij\}} \frac{\beta E[u_c(c_{ijt+1})]}{u_c(c_{ijt})} = \max_{\{ij\}} \left(c_{ijt}^{\gamma} [\beta(1+\mu_{ij})^{-\gamma} E\left[(1+\upsilon_{jt})^{-\gamma}\right]] \right).$$

Agents with the highest valuation of the asset like to buy it, bidding up its price and driving down the interest rate. The Euler equation of an agent reads:

$$(12) \qquad c_{ijt}^{-\gamma} \geq \beta(1+r_{t+1}) E[c_{ijt+1}^{-\gamma}],$$

where the usual equality is replaced by a weak inequality because of the short-sale constraint on the asset.[15] A useful way to read the Euler equation is to see the LHS as the cost and right hand side the benefit of saving one unit in the asset. Denote by $(ij)^\star$ the agent i active in occupation j with the highest propensity to save. From (10) we see that this agent is pinning down the equilibrium interest rate. Using the interest rate in the Euler equation of agent $(ij)^\star$, it is clear that the equation holds with equality for all possible consumption levels $c_{(ij)^\star t}$. For all other agents it immediately follows that the Euler condition becomes:

$$\frac{u_c(c_{ijt})}{\beta E[u_c(c_{ijt+1})]} > \frac{u_c(c_{(ij)^\star t})}{\beta E[u_c(c_{(ij)^\star t+1})]},$$

intuitively, agents with a lower valuation of the asset will never want to save when facing the equilibrium interest rate. Since no one strongly prefers to hold assets (the marginal agent just being indifferent), we have found an equilibrium of the economy involving no trades and savings with $y_{ijt} = c_{ijt}$ and $a_{ijt} = 0$.

The main implication of the equilibrium is that agents are not able to smooth out persistent shocks. Because in this stylized economy *persistent* shocks to the growth rate

[14] The assumption of homogeneous risk-preferences is relaxed later, and doesn't affect the equilibrium.
[15] And so (12) holds with equality when $a_{ijt+1} > 0$.

of income are the only source of uncertainty, this also implies that agent don't use any assets to smooth their income at all. Assuming the absence of *transitory* innovations to income comes without loss of generality, however. The simplified model and its no-trade equilibrium are rather a shortcut to follow the consensus view in the literature that transitory shocks, like short-term unemployment, and illness are known to be well-insured through various mechanisms.[16]

Given the above equilibrium in financial and goods markets, and the normality assumption for the shocks it is possible to get an expression for the expected lifetime utility an agents derives from an occupation. Importantly, the expression depends on the occupation specific variance of permanent shocks, the individual-occupation specific growth rate of income, and their interaction with an agent's risk-aversion:

$$V_{ij} = \frac{y_0^{(1-\gamma)}}{(1-\gamma)\left(1-\beta(1+\mu_{ij})^{(1-\gamma)}\exp\left(0.5\gamma(\gamma-1)\sigma^2_{\upsilon_j}\right)\right)}, \tag{13}$$

and intuitively, the expression is increasing in wage growth μ_{ij}, the base wage y_0, and decreasing in the variance of permanent shocks $\sigma^2_{\upsilon_j}$. Allowing for preference heterogeneity, it is, moreover, straightforward to show that more risk-averse agents suffer more from uncertainty in the sense of being willing to give up a larger amount of money to eliminate persistent shocks under some weak parameter assumptions (details on request).

4.2 Empirical Model

Expression (13) allows for a tight link with the parameters estimated in the previous sections. Individual-occupation specific income growth rates are inferred by taking first differences to the predicted values for income in equation (6), and then averaged over time. The variance of the permanent shocks can simply be recovered by taking the occupation specific estimates from the previous section: $\sigma^2_{\upsilon_j} = \hat{\sigma}^2_{\eta_j}$. The discount factor β is set to 0.99, since the one period correspondents to one year.

Moreover, in the empirical analysis I will allow for differences in the risk-aversion parameter γ_i. The input for each individual's CRRA coefficient comes from the same combination of wealth and survey data used to analyze the self-selection into occupation

[16] Levin and Zame (2002) show that self-insurance is a very effective device to smooth transitory shocks in a model of incomplete markets. Heathcote, Storesletten, and Violante (2008) propose a partial insurance framework in which contingent securities are traded to insure transitory fluctuations, but persistent shocks remain uninsured. The welfare expressions obtained in their article hence resemble those obtained in this paper. Empirically, Blundell, Pistaferri, and Preston (2008) find in a recent influential paper that transitory shocks seem to be very well absorbed, and do not cause significant fluctuations in consumption. Relatedly, Guiso, Pistaferri, and Schivardi (2005) present evidence that firms insulate their workers completely against transitory productivity shocks, but only partially against persistent shocks.

in the previous section. The model will be estimated under the assumption of both, homogeneous and heterogeneous risk-preferences. This allows for an evaluation of the usefulness of combining survey and micro data to recover deep parameters.

Given a V_{ij} for every individual, we can specify a random unobserved utility model of the form:

$$U_{ij} = V_{ij}\left(\gamma_i, \mu_{ij}, \sigma^2_{\eta_j}\right) + \delta_{ij}X_j + \epsilon_{ij}, \tag{14}$$

where X_j is a dummy variable for occupation j. I make the standard assumption that ϵ_{ij} is i.i.d. of the extreme value form. Additionally, I assume that the δ_{ij} are drawn from normal distributions, with a separate distribution for each occupation j, so $\delta_{ij} \sim N(\nu_j, \sigma^2_{\delta_j})$. Equation (14) can be, hence, conveniently estimated using a mixed logit model (Revelt and Train 1998 and McFadden and Train 2000). This specification is particularly suited for the problem at hand, since it takes differences in tastes of the decision makers explicitly into account. In the occupational choice context, heterogeneity along the non-pecuniary attributes of an occupation seem particularly important. As is well known, mixed logit specifications don't suffer from the curse of independence from irrelevant alternatives like other more standard discrete choice models. Individual coefficients are then recovered using the method proposed by Revelt and Train (2001).

Because the model is computationally very intensive to fit, in practice only 10 dummy variables are used. For each occupation, I, therefore, assign a dummy for its first digit in the occupational code. This allows for unobserved preferences over broader categories (for example, working in the armed forces or in agricultural/fishing occupations). The set of possible choices for each individual is further endogenously restricted by the level of education an individual obtained in her life. The sample is divided into six degree groups. In the estimation individuals are allowed to choose from those occupations, which have been selected by at least one individual in their respective degree group. E.g., if no one who just attended primary school is observed to be a university professor, the occupation university professor is eliminated from the choice of those individuals, who only attended primary school. This assumption allows to economize on computing time.

4.3 Results of Discrete Choice Model

Section contains preliminary results.

Table 5 presents the estimated population means and standard deviation of the occupation parameters δ_{ij}, corresponding to specification (14) with heterogeneity in risk-preferences.

Table 5: Estimates of non-observable, non-monetary occupation preferences

Occupational Group	Estimated Mean	Standard Error
LEGISLATORS, SENIOR OFFICIALS AND MANAGERS	-2.70	1.35
PROFESSIONALS	1.98	0.32
TECHNICIANS AND ASSOCIATE PROFESSIONALS	-221.96	127.79
CLERKS	1.15	0.80
SERVICE WORKERS AND SHOP AND MARKET SALES WORKERS	-78.21	40.41
SKILLED AGRICULTURAL AND FISHERY WORKERS	-15.34	4.72
CRAFT AND RELATED TRADES WORKERS	2.31	0.32
PLANT AND MACHINE OPERATORS AND ASSEMBLERS	-28.05	17.18
ELEMENTARY OCCUPATIONS	1.29	0.34
Occupational Group	Estimated Standard Dev.	Standard Error
LEGISLATORS, SENIOR OFFICIALS AND MANAGERS	4.37	1.04
PROFESSIONALS	0.01	0.16
TECHNICIANS AND ASSOCIATE PROFESSIONALS	228.33	128.76
CLERKS	1.27	1.01
SERVICE WORKERS AND SHOP AND MARKET SALES WORKERS	59.21	29.36
SKILLED AGRICULTURAL AND FISHERY WORKERS	9.62	2.59
CRAFT AND RELATED TRADES WORKERS	0.08	0.20
PLANT AND MACHINE OPERATORS AND ASSEMBLERS	35.15	17.89
ELEMENTARY OCCUPATIONS	0.14	0.50

Based on random-utility mixed logit model with 29710 observations

A likelihood-ratio test for the joint significance of the standard deviations rejects the null hypothesis that all the standard deviations are equal to zero (with a Wald statistic of 134.15 and 9 degrees of freedom).[17] Although the number of choices is large with 27, the model is able to predict 933/1435 $\approx$ 65% cases correctly. When evaluating the performance by the numer of rightly predicted occupations at the 1-digit level the number increases to 76%.

Perfect Insurance Counterfactual. The model was estimated under the assumption that persistent income shocks are not insurable, and hence directly affect occupational choices through their effect on consumption fluctuations, risk-averse workers would like to avoid. Having identified the distribution of preferences for occupations across individuals, I now compare a counterfactual situation, in which all shocks are insured- in this world agents make their occupational choices by comparing expected wages and their non-pecuniary preferences. Concretely, I evaluate expression (13) at $\sigma^2_{\upsilon_j} = 0$ for each individual and occupation, and use it together with the estimates form Table 5 to get the counterfactual. From the set of individuals for which the estimated model with imperfect insurance predicted occupational choices correctly, I select those who are estimated to derive a higher utility from a different occupation under perfect insurance. This gives an estimated number of 47 individuals, about 3%, who would have made a different occupational selection under complete markets. Of these 47 the majority of 41 individuals would have chosen

[17]Under the null, a conditional logit model would be sufficient for estimation.

an occupation in the same 1-digit class of occupations, which is also a consequence of the definition of preferences over the 1-digit class. It follows that these agents are predicted to choose a different occupation, because their expected earnings are higher in their predicted occupation under complete markets. The absence of uncertainty induces the agents to choose occupations paying them higher expected wages.[18]

The model, hence, implies a relatively small effect of uninsurable income risk on occupational choices. The counterfactual switching is mainly due to the ability of agents to choose occupations offering higher expected individual-occupation specific wages.

Heterogeneous or Homogeneous Risk Preferences. Next, I test whether the model with heterogeneous, imputed CRRA coefficients performs vis-a-vis a standard model, assuming the same risk-attitudes. I re-estimate specification (14) using a coefficient of $\gamma = 1.5$ (Chetty 2006). The model predicts only $374/1435 \approx 26\%$ choices correctly. Additionally, equaility of the estimated log-likelihoods is formally rejected, in favor of the model with heterogeneous preferences. So, although is is certainly not uncontroversial to employ the imputed CRRA coeffcients as a structural input, this approach outperforms a model with homogeneity. This points to the importance of regocnizing that risk affects the welfare and therefore economic decisions of individuals differently.

5 Conclusion

This paper has tried to shed light how income risk, defined as the variance of the unexplainable part of labor earnings, distorts occupational choices. Importantly the empirical analysis takes heterogeneity among individuals seriously. I have demonstrated that less risk-averse agents self-select themselves into more risky occupations. Assuming that transitory shocks are well-insured, but permanent risk is uninsurable, I have estimated a model of the self-selection process. Calculating a counterfactual with perfect insurance, the model suggests small effects of income risk on occupational choices. Moreover, inferring coefficients of risk-aversion from questionnaires (surprisingly) yields a much better model fit than assuming identical risk-attitudes.

It has been abstracted from some potentially important effects. There are no general equilibrium effects considered, when counterfactually changing income risk. Implicitly, educational decisions have been separated from occupational choices- that is the education level was taken exogenously. It is plausible that income risk also affects how much

[18]For the other 6 individuals it can be both higher expected wages or non-monetary preferences driving the counterfactual choice.

education to obtain. Moreover, the paper has nothing to add to the question, how the information set of the econometrician and the individuals are aligned, when estimating income risk. It might be true that something unobservable is known to the individual and vice-versa. Deeper explorations of these issues are left for the future.

References

Barsky, R. B., F. T. Juster, M. S. Kimball, and M. D. Shapiro, "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *Quarterly Journal of Economics*, 112 (1997), 537-580.

Blundell, Richard, L. Pistaferri, and I. Preston, "Consumption Inequality and Partial Insurance," *American Economic Review*, 98(5) (2008), 1887-1921.

Bonin, H., T. Dohmen, Thomas, A. Falk, D. Huffman, and U. Sunde, "Cross-sectional earnings risk and occupational sorting: The role of risk attitudes," *Labour Economics*, 14(6) (2007), 926-937.

Carroll, C. D., and A. A. Samwick, "The Nature of Precautionary Wealth," *Journal of Monetary Economics* 40:1 (1997), 41-71.

Chetty, R., "A New Method of Estimating Risk Aversion," *American Economic Review*, vol. 96(5) (2006), 1821-1834.

Dohmen T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. Wagner, "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey," *IZA Discussion Paper* No. 1730 (2005).

_______________, _______________, "Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences," *Journal of the European Economic Association*, (forthcoming).

Fuchs-Schündeln, N., and M. Schündeln, "Precautionary Savings and Self-SelectionEvidence from the German Reunification Experiment," *Quarterly Journal of Economics*, 120:3 (2005), 1085-1120.

Guiso, L., M. Paiella, "Risk Aversion, Wealth, and Background Risk," *Journal of the European Economic Association*, 6(6) (2008), 1109-1150.

Guiso, L., L. Pistaferri, and F. Schivardi, "Insurance within the Firm," *Journal of Political Economy*, 113(5) (2005), 1054-1087.

Guvenen, F., "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, 97(3) (2007), 687-712.

Guvenen, F., "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, 12(1) (2009), 58-79.

Heathcote, J., Storesletten, K., and G. Violante, "Insurance and opportunities: A welfare analysis of labor market risk," *Journal of Monetary Economics*, 55(3) (2008), 501-525.

________________, ________________, "Quantitative Macroeconomics with Heterogeneous Households", *Annual Review of Economics*, (2009).

Holt, C.A., and S. K. Laury, "Risk Aversion and Incentive Effects," *American Economic Review*, 92(5) (2002), 1644-1655.

Kambourov, G., and I.Manovskii, "Occupational Specificity Of Human Capital, " *International Economic Review*, 50(1) (2009), 63-115.

Jappelli, T., and L. Pistaferri, The consumption response to income changes, *Annual Review of Economics*, (2010).

Krebs, T., "Testable implications of consumption-based asset pricing models with incomplete markets," *Journal of Mathematical Economics*, 40(1-2) (2004), 191-206.

Krebs, T., P. Krishna, and W. Maloney, "Trade Policy, Income Risk, and Welfare," *Review of Economics and Statistics*, (forthcoming).

Levine, D. K., and W. R. Zame, "Does Market Incompleteness Matter?," *Econometrica*, 70(5) (2002), 1805-1839.

McFadden, D.,and K. Train, "Mixed MNL models for discrete response," *Journal of Applied Econometrics*, 15(5) (2000), 447-470.

Meghir, C., and L. Pistaferri, "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1) (2004), 1-32.

Revelt, D., and K. Train, "Mixed Logit With Repeated Choices: Households Choices of Appliance Efficiency Level," *Review of Economics and Statistics*, 80:4 (1998), 647-657.

________________, ________________, "Customer-Specific Taste Parameters and Mixed Logit: Households' Choice of Electricity Supplier,", *Econometrics 0012001 EconWPA*, (2001).

Schulhofer-Wohl, S., "Heterogeneity and Tests of Risk Sharing," *Working Paper*, (2010).

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

Appendices

A Summary Statistics Data

B Recovering Risk Preferences

A more detailed account of how individual CRRA coefficients can be imputed in the SOEP is found in Dohmen et al. (2005). First, responses to a standard hypothetical investment scenario are used. Individuals were asked how much of 100,000 Euros in lottery winnings they would choose to invest in a risky asset; with 50% probability they would double or halve the invested amount. Respndents could select from six different of 0, 20,000, 40,000, 60,000, 80,000, or 100,000 Euros. The SOEP also contains data on individual wealth levels. Following the approaches taken by Barsky et al., (1997), and Holt and Laury (2002), one can now solve for an upper and lower bound on an individual's CRRA coefficient, assuming the respondent selected the investment level most to suitable to her personal risk tolerance. In addition, the SOEP offers researchers with individuals' answers to a question measuring willingness to take risks, in general on an 11-point scale. Importantly, Dohmen et al. (forthcoming) find the question to be a robust predictor of risky choices with real money at stake in a field experiment. Following their suggestion, I calculate average amounts invested for each possible answer to the risk question. Coefficients are then simply imputed by calculating:

$$\gamma_i = \frac{-\ln 2}{\ln\left((100000 + W_i - 0.5 * I_R)/(100000 + W_i + I_R)\right)},$$

where W_i is the personal wealth level, and I_R the average investment of all people with the same response. Figure 1 displays a histogram (capped at 10), with evident mass points between 2 and 4.

Figure 1: Distribution of imputed CRRA coefficients