NORFACE MIGRATION Discussion Paper No. 2012-39

# Eliciting earnings risk from labor and capital income

Sarah Meyer and Mark Trede

www.norface-migration.org

# Eliciting earnings risk from labor and capital income

Sarah Meyer and Mark Trede
Center for Quantitative Economics
Department of Economics
Westfälische Wilhelms-Universität Münster

May 2012

**Abstract**

Earnings risk is an inherently subjective concept. Observing volatile earnings paths does not necessarily imply that the perceived earnings risk is large. If a drastic change in earnings is known well in advance, there is no additional risk involved. Individuals are likely to have more information about their earnings prospects than the observing econometrician. Since it is the subjectively perceived earnings risk that influences economic decisions like consumption, we need to develop methods that allow to elicit the perceived risk from observable variables. This paper suggests an estimation method based on variables that are not only observable in principle, but can be observed in fact in many panel data sets.

# 1 Introduction

Earnings risk is hard to measure since it is a subjective concept. If we observe large swings in earnings, that does not necessarily mean that the perceived earnings risk is high. If drastic changes in earnings are known in advance, there is no risk involved. Naturally, the information set of individuals about their own earnings prospects is larger than the information set of an observing econometrician and it is hard to distinguish expected earnings changes from unexpected ones. This paper suggests a method that allows to elicit the perceived earnings risk from observable variables. To make the method applicable, the variables need not only be observable in principle, but must be observed in fact.

Most of the literature on earnings risk has ignored problems concerning the superior information of individuals about their own prospects, see e.g. Moffitt and Gottschalk (2008). Guiso, Jappelli and Terlizzese (1992) used a self-reported measure of earnings uncertainty from an Italian cross-sectional survey. This approach is problematic as there are hardly any actual datasets, providing these information. Recently, Guvenen and Smith (2010) have suggested an estimation method for earnings risk that takes into account economic choices of the individuals. Obviously, the most important economic choice is the decision about the consumption level. If individuals feel insecure about their future earnings, i.e. if their perceived earnings risk is high, they tend to consume less and save more to build up a buffer against negative future shocks, see e.g. Carroll (2004) on the theory of buffer saving. However, from a practical point of view, estimation methods that need earnings *and* consumption panel data are unattractive since there are hardly any panels providing information about both earnings and consumption on an individual level over longer time spans.

To overcome these problems, our paper suggests a method to measure perceived earnings risk from individual earnings and capital income data, whereas consumption data are not required. If available, the method can also take into account the different sources of capital income, for instance divided into income from risk-free assets and from risky assets.

The estimation method is close to the suggested procedure by Guvenen and

Smith (2010). In particular, it is based on the classical dynamic stochastic optimization described in Samuelson (1969). While the Samuelson model ignores labor income, we assume that earnings are an (exogenous) stochastic process which makes the dynamic optimization problem more complex. We apply indirect inference to estimate the parameters of the model.

The structure of the paper is as follows. Section 2 describes the economic model and the stochastic dynamic optimization framework. In section 3 we explain the estimation procedure. Section 4 presents simulation studies to gauge the gain of taking capital income into account. Section 5 concludes.

# 2 The economic model

Reconsider the model of Samuelson (1969) with instantaneous CRRA utility function

$$U(C) = \frac{C^{1-\gamma}}{1-\gamma}$$

where $\gamma$ is the measure of relative risk aversion. Assume that the individual has initial wealth $W_0$ and initial labor income $Y_0$. Labor income evolves stochastically according to a geometric random walk

$$Y_{t+1} = Y_t \exp\left(\mu_\varepsilon + \theta_\varepsilon \varepsilon_{t+1}\right) \tag{1}$$

with parameters $\mu_\varepsilon$ and $\theta_\varepsilon$ and a random innovation $\varepsilon_t \sim N(0,1)$. Each period, the individual decides about two control variables: the level of consumption $C_t$, and the portfolio composition $\alpha_t$ (the proportion of wealth invested into the risky asset). The budget constraint is

$$W_{t+1} = \left(\left(1-\alpha_t\right)\left(1+r\right) + \alpha_t Z_{t+1}\right)\left(W_t + Y_t - C_t\right). \tag{2}$$

In addition we constrain the proportion of wealth invested in the risky asset to $0 \leq \alpha_t \leq 1$. Further, our model contains no bequests. From this may be deduced that $W_{T+1} = 0$ and $W_T = C_T$. The random asset return shocks $Z_t \sim LN\left(\mu_Z, \theta_Z\right)$ and $\varepsilon_t$ can be dependent; however, for ease of illustration we assume that they are uncorrelated. The timing of decisions and random shocks is as follows:

$$\ldots \rightarrow \underset{\text{observe the state}}{\begin{pmatrix} W_t \\ Y_t \end{pmatrix}} \rightarrow \underset{\text{decide the control}}{\begin{pmatrix} \alpha_t \\ C_t \end{pmatrix}} \rightarrow \underset{\text{shocks happen}}{\begin{pmatrix} \varepsilon_{t+1} \\ Z_{t+1} \end{pmatrix}} \rightarrow \underset{\text{new state}}{\begin{pmatrix} W_{t+1} \\ Y_{t+1} \end{pmatrix}} \rightarrow \ldots$$

Since earnings $Y_t$ are modelled as an exogenous stochastic process, simply observing $Y_1, \ldots, Y_T$ would, of course, already allow the econometrician to estimate the earnings risk parameter $\theta_\varepsilon$. However, taking into account additional information about capital income from risk-free and risky assets may increase the precision of the estimator. This is particularly important if the observation period is relatively short.

If individuals feel insecure their saving will increase to build up a buffer. An important practical problem is the fact that savings or, equivalently, consumption or wealth, are not reliably observable for the econometrician in most panels (of course, wealth and consumption levels are observable for the individuals themselves). In contrast, many panels provide accurate and detailed information about different sources of income. No matter if savings are invested in the risky asset or in the risk-free asset, capital income will tend to increase in the following periods. Hence, in an indirect way, an increase in capital income indicates larger perceived earnings risk.

Regardless whether consumption or wealth data are available or not, the intertemporal optimization problem

$$\max_{\alpha_1, C_1, \ldots, \alpha_T, C_T} \left( \frac{1}{1+\rho} \right)^t \frac{C_t^{1-\gamma}}{1-\gamma}$$

subject to (2) and (1) has to be solved prior to implementing any estimation strategy. The parameter $\rho$ is a subjective discount rate. Of course, the corresponding value function is

$$V_t \left( W_t, Y_t \right) = \max_{\alpha_t, C_t} \left\{ \frac{C_t^{1-\gamma}}{1-\gamma} + \frac{1}{1+\rho} E \left( V_{t+1} \left( W_{t+1}, Y_{t+1} \right) \right) \right\}. \tag{3}$$

While the model of Samuelson (1969) without labor income can be solved analytically, adding a stochastic labor income process makes the problem much more complex and analytically untractable. Numerical optimization methods are required, and therefore, it is important to keep the number of state variables as low as possible. In the following, we show that the two state variables $W_t$ and $Y_t$ can be collapsed into a single state variable $w_t = W_t/Y_t$.

As the amount of consumption equals the remaining wealth in the last period, its value function can be written as

$$V_T\left(W_T, Y_T\right) = Y_T^{1-\gamma} \frac{\left(w_T + 1\right)^{1-\gamma}}{1-\gamma} \tag{4}$$

where $w_T = W_T/Y_T$. The transition equation for $w_t$ is

$$w_{t+1} = \frac{\left(1-\alpha_t\right)\left(1+r\right) + \alpha_t Z_t}{\exp\left(\mu + \theta\varepsilon_t\right)}\left(w_t + 1 - c_t\right) \tag{5}$$

where $c_t = C_t/Y_t$. Define $\phi_t = \exp\left(\left(\mu+\theta\varepsilon_t\right)\left(1-\gamma\right)\right)$. In the last period but one,

$$\begin{aligned} V_{T-1}\left(W_{T-1}, Y_{T-1}\right) &= \max_{\alpha_{T-1}, C_{T-1}} \left\{ \frac{C_{T-1}^{1-\gamma}}{1-\gamma} + \frac{1}{1+\rho} E\left(V_T\left(W_T, Y_T\right)\right) \right\} \\ &= Y_{T-1}^{1-\gamma} \max_{\alpha_{T-1}, c_{T-1}} \left\{ \frac{c_{T-1}^{1-\gamma}}{1-\gamma} + \frac{1}{1+\rho} E\left( \phi_{T-1} \frac{\left(w_T+1\right)^{1-\gamma}}{1-\gamma} \right) \right\}. \end{aligned}$$

Define a new optimization problem with single state variable $w_t$

$$v_t\left(w_t\right) = \max_{\alpha_t, c_t} \left\{ U\left(c_t\right) + \frac{1}{1+\rho} E\left(\phi_t v_{t+1}\left(w_{t+1}\right)\right) \right\} \tag{6}$$

with transition equation (5) and terminal value function (4). The relation between the value functions for one and two state variables is

$$V_t\left(W_t, P_t\right) = Y_t^{1-\gamma} v_t\left(W_t/Y_t\right)$$

for all $t = 1, \ldots, T$. The policy function $c_t^*(w_t)$ can be re-transformed into

$$C_t^*(W_t, Y_t) = Y_t c_t^*(W_t/Y_t).$$

The policy function $\alpha_t^*(w_t)$, i.e. the share of wealth invested in the risky asset, needs not be re-transformed.

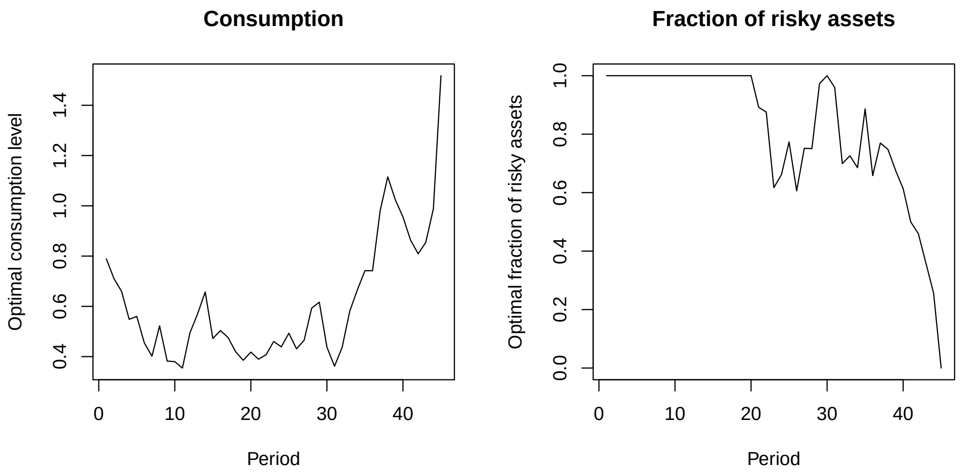


Figure 1: Typical optimal policy functions

A grid based solution of the dynamic optimization problem (6) reveals that the value function can be very well approximated by the functional form

$$v_t(w_t) = \frac{(w_t + a_t)^{b_t}}{c_t}.$$

Figure 1 shows typical policy functions $c_t^*(W_t/Y_t)$ and $\alpha_t^*(W_t/Y_t)$ for various periods $t$.[1] In this case, consumption only faces slight changes during the first 30 years of the work life, but then rises and ends up with its maximum at the end of work life. The rise at the end is what we could have expected since our model contains no bequests. The second policy function reveals that the optimal fraction of risky assets starts at 100%. This could be due to the fact that the expected value of the risky asset is slightly larger than the expected value of the riskfree asset and the amount of risk aversion is moderate.[2] Regardless of the larger expected value of risky assets, the figure shows a shift towards the riskfree investments over time. One possible explanation for this observation is the nature of the CRRA utility function that implies constant relative risk aversion. One has to bear in mind that

[1] The model parameters are set to $r = \rho = 0.05$, $\mu_\varepsilon = 0$, $\theta_\varepsilon = 0.2$, $\mu_Z = 0.05$, $\theta_Z = 0.25$, $\gamma = 2$.

[2] $Z_t$ is lognormal distributed with $E(Z_t) = \exp(\mu_Z + \frac{\theta_Z^2}{2}) = 1.084$

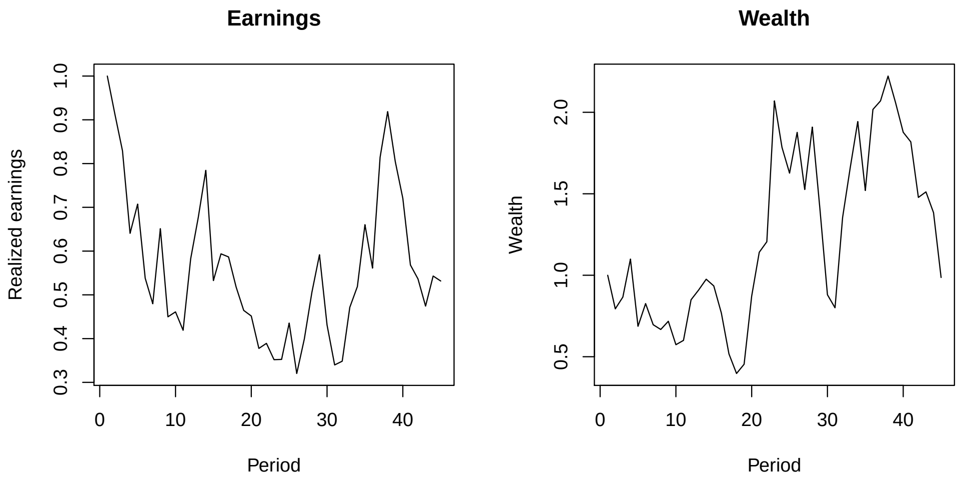


Figure 2: Earnings and wealth paths

in last period $T$ the remaining wealth is liquidated as there are no bequests and our model provides no pension system. With $\gamma = 2$ the CRRA utility function implies a decreasing marginal utility, hence liquidating a large amount of wealth yields not sufficient utility to justify large risk.

Figure 2 depicts the corresponding earnings and wealth paths. As assumed, earnings are exogenous and follow a geometric random walk. The wealth path indicates that the individual accumulates wealth in the middle of work life. However, wealth is reduced at the end in a large amount because it yields no further utility after work life.

The next section implements an indirect inference estimation procedure in order to determine the underlying parameters of the income process and if required the risk aversion parameter $\gamma$. We suppose the estimators to be more precisely than maximum likelihood estimators, because in addition to the earnings process, we now take capital income into account.

# 3 Estimation procedure

The three parameters we are interested in are $\mu_\varepsilon$ and $\theta_\varepsilon$ for the stochastic earnings process (1) and optional the risk aversion parameter $\gamma$ of the utility function. We assume that the other parameters (i.e. interest rate $r$, subjective discount rate $\rho$, expected stock return $\mu_Z$, and volatility $\theta_Z$) are known or have been estimated outside our model.

Let $Y_{it}$ denote earnings of individual $i = 1, \ldots, N$ in period $t = 1, \ldots, T$, the other variables are denoted in the same way, e.g. consumption $C_{it}$. For simplicity we assume that the panel is balanced and that all individuals have the same parameters and known starting values. These restrictions can be relaxed, but simplify the simulation studies below.

The parameters are estimated by indirect inference (Gourieroux, Monfort and Renault (1993)). The auxiliary model is a composition of

- the mean growth rate of earnings,
- the standard deviation of the growth rate of earnings,
- and additional information about capital income, e.g. a linear regression of interest payments on a constant, age, and age squared

The mean and standard deviation of earnings growth are already sufficient to estimate the parameters of interest $\mu_\varepsilon$ and $\theta_\varepsilon$. However, adding information about capital income may increase the precision of the estimates and also allows to estimate the coefficient of risk aversion $\gamma$. The parameters of interest are collected into a vector $\beta = (\mu_\varepsilon, \theta_\varepsilon, \gamma)$.

How much information there is about capital income depends on which variables are observable. In the most realistic setting, the econometrician only observes capital income from the risk-free asset in period $t$, i.e. $(1 - \alpha_t)\, r\, (W_t + Y_t - C_t)$, but not capital income due to changing stock prices. Observing interest payments is, of course, much less informative than observing wealth or consumption, but more common in typical panel data sets. The estimation method can be very easily adapted to more informative settings.

The five auxiliary model's parameters are estimated for the observed earnings and interest payments data of each individual $i = 1, \ldots, N$ and then averaged over all individuals. We denote the set of auxiliary parameters by $\psi$, and their estimates (from the observed data) by $\hat{\psi}(Z)$ where $Z$ represents all observed variables.

Let $b$ be an arbitrary vector of the parameters of interest. It is straightforward to compute the corresponding policy functions $\alpha_t^*, C_t^*$, and to simulate artificial data for a large number $H$ of individuals. $H$ can be larger than $N$ to mitigate the influence of sampling errors. Denote the artificial data set by $Z^*(b)$. In the next step, the auxiliary model's parameters are estimated for $Z^*(b)$. Denote the estimates by $\hat{\psi}(Z^*(b))$.

The indirect inference estimator of $\beta$ is defined by

$$\hat{\beta} = \arg\min_b \left(\hat{\psi}(Z) - \hat{\psi}(Z^*(b))\right)' W \left(\hat{\psi}(Z) - \hat{\psi}(Z^*(b))\right)$$

where $W$ is any positive definite weighting matrix. No matter how $W$ is chosen, the indirect inference estimator is consistent and asymptotically normal. As the asymptotically efficient weighting matrix can be far from optimal in finite samples, we set $W$ to the identity matrix.

# 4 Simulation studies

In order to gain accuracy, the estimation procedure, that takes capital income into account, is repeated 500 times. Within each repetition, both the income process and capital income of $N = 50$ individuals is simulated in a dynamic optimal way. For each group of 50 individuals, the stochastic income parameters $\mu_\varepsilon$, $\theta_\varepsilon$ are estimated by indirect inference, as described in the previous section. Table (1) presents the average results of the repeated estimation. Compared with the true values, the mean estimators seem to be quite accurate, as the mean squared error of all two estimates is near 0. After considering the histograms in figure (3), one can conclude that the estimates are also unbiased.

As mentioned in the first section, the stochastic income parameters $\mu_\varepsilon$ and $\theta_\varepsilon$ may also be estimated by just observing income data, since earnings are modelled

| | $\mu_\varepsilon$ | $\theta_\varepsilon$ |
|---|---|---|
| True values | 0 | 0.2 |
| Mean | -0.000825 | 0.19151 |
| MSE | 0.000198 | 0.00492 |

Table 1: Mean estimators and MSE

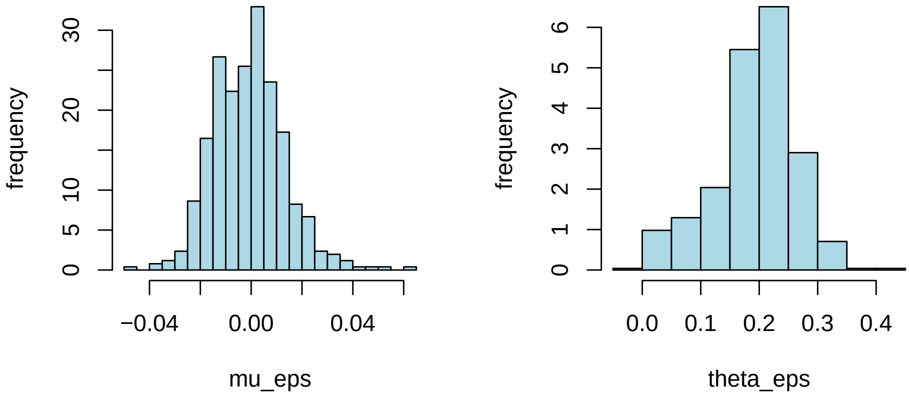


Figure 3: Histograms of the estimators

as an exogenous stochastic process. In this case, we expect a deterioration of our results compared with the estimation before, because we do not take into account capital income data, that (as we assume) contain additional information about consumption and portfolio decisions and thus about perceived income risk.

As we only estimate the given income process (1), one can carry out the estimation by maximum likelihood and then obtains the results, reported in table (2). Contrary to our expectation, table (2) indicates a slight improvement of income parameter estimates, regarding the mean squared errors as a measure of precision. Hence, including capital income data and a dynamic framework can not improve the estimation results. To verify the reliability of our first estimation routine, we now focus on the accuracy of the optimization algorithm in the statistical computer software R, that we used for our programming. We expand a 2-dimensional

| | $\mu_\varepsilon$ | $\theta_\varepsilon$ |
|---|---|---|
| True values | 0 | 0.2 |
| Mean | 0.000136 | 0.198838 |
| MSE | 0.000016 | 0.000010 |

Table 2: Mean estimators and MSE

grid on a small range around our estimators in order to examine whether we have found local or global optima. The results indicate, that the optimization routine does not work reliable as we did not find the global optimum in each case.

Nevertheless, our method provides the possibility of not only estimating income parameters but also the risk aversion parameter $\gamma$. The results of this extended estimation are shown in table (3). Basically, the results for $\mu_\varepsilon$ and $\theta_\varepsilon$ are similar

| | $\mu_\varepsilon$ | $\theta_\varepsilon$ | $\gamma$ |
|---|---|---|---|
| True values | 0 | 0.2 | 2 |
| Mean | 0.003270 | 0.203002 | 2.072478 |
| MSE | 0.000154 | 0.004710 | 0.062578 |

Table 3: Mean estimators and MSE

to the first estimation, presented in table (1). Although the direct estimation that uses only income data still provides better results, we can here obtain an estimator for the risk aversion parameter $\gamma$. Since the estimators seem to be unbiased (see histograms in figure (4)), one can summarize that our estimation procedure, that includes capital income data and mostly unobservable economic decisions, provides precise estimators and a flexible estimation, but can not improve the results of an estimation by maximum likelihood.

# 5 Conclusion

In general, the information set of individuals about their own future income differs with respect to the information set of an econometrician. Therefore, large swings in earnings do not necessarily mean that the individuals face large income risk.

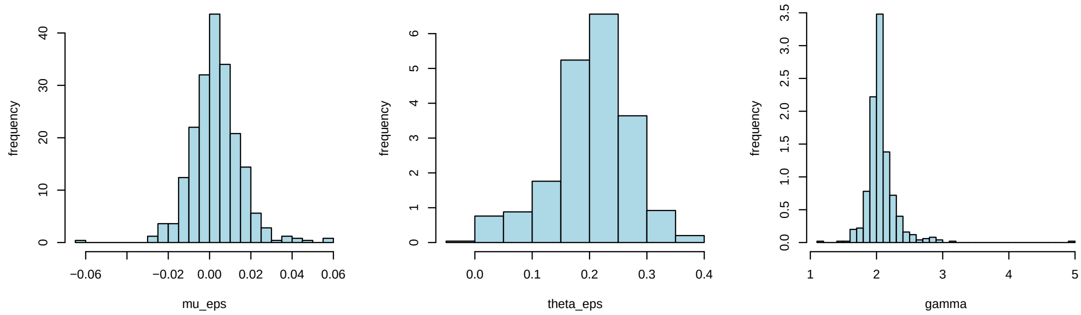


Figure 4: Histograms of the estimators

This paper suggests a method that uses those information indirectly contained in capital income data. Since capital income reflects the amount of savings, it contains important information about consumption and savings behavior and thus about individual risk perception.

We estimate two labor income parameters of a geometric random walk by indirect inference and assume that the true model is formulized in a stochastic dynamic optimal framework. The results indicate that the estimates, obtained from indirect inference, seem to be unbiased. However, they have a larger MSE than the maximum likelihood estimators, but on the other side they provide the additional possibility of estimating the individuals' risk aversion. The latter would not be feasible by maximum likelihood. This and the possibility of changing any part of the model framework (e.g. the income equation or any other assumption) very easily, lead us to state that we obtain a more flexible estimation procedure.

# References

Carroll, C. (2004). Theoretical foundations of buffer stock saving, NBER Working Paper 10867.

Gourieroux, C., Monfort, A. and Renault, E. (1993). Indirect inference, *Journal of Applied Econometrics* **8**: S85–S118.

Guiso, L., Jappelli, T. and Terlizzese, D. (1992). Earnings uncertainty and precautionary saving, *Journal of Monetary Economics* **30**(2): 307–337.

Guvenen, F. and Smith, A. (2010). Inferring labor income risk from economic choices: An indirect inference approach, working paper.

Moffitt, R. and Gottschalk, P. (2008). Trends in the transitory variance of male earnings in the U.S., 1970-2004, working paper.

Samuelson, P. A. (1969). Lifetime portfolio selection by dynamic stochastic programming, *Review of Economics and Statistics* **51**(3): 239–246.

Viceira, L. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income, *The Journal of Finance* **56**(2): 433–470.