True and Spurious State Dependence in Earnings Transitions *

Andrea Weber[†]

July 2004

Abstract

This paper focuses on the measurement of earnings mobility under special consideration of individual heterogeneity. We model transitions between quintiles in the wage distribution by a first order Markov process. The model is estimated using a fixed effects dynamic panel procedure that uses a minimum of assumptions. Thus we are able to control for the effects of observed individual heterogeneity by including time-varying exogenous variables as well as for unobserved individual effects. The estimates, derived from a large administrative panel for Austria, indicate that controlling for individual heterogeneity reduces persistence of earnings by at least 27%.

Keywords: Wage mobility, Markov process, dynamic panel estimation, fixed effects estimation

Jel classification: C23, C25, J31, J60

*I would like to thank Helmut Hofer, Bo Honoré and Rudolf Winter-Ebmer, participants at the 10th Internationa Panel Data Conference in Berlin, July 2002, participants at the CAM workshop for dynamic panel data models, Copenhagen, January 2003, and seminar participants in Vienna and Linz for many helpful comments and suggestions.

[†]Institute for Advanced Studies, Department of Economics and Finance, Stumpergasse 56, A-1060 Vienna, phone: +43-1-59991-147, fax: +43-1-59991-163, e-mail: andrea.weber@ihs.ac.at

1 Introduction

The inequality of income and the persistence of low income are important social indicators. From the welfarist point of view the static picture of inequality in income in a single point in time can only be completed by considering the dynamics of the income distribution as well. Individual mobility in the income distribution gives an impression of the equality of opportunities in a society and it also informs about the income risks an individual faces by moving downwards in the distribution.

Individual heterogeneity is an important issue when studying earnings mobility. The persistence in individual earnings is comparable to many other economic situations where we observe that an individual who has experienced an event in the past is more likely to experience that event in the future than an individual who has not experienced that event. **?** discusses two explanations for this phenomenon. The first one is the presence of “true state dependence”, in the sense that the lagged state enters the model in a structural way as an explanatory variable. Industry or country wide wage bargaining and firm specific wage schemes could be reasons for true state dependence. The second explanation is that heterogeneity makes the individuals differ in their propensity to experience the event in all time periods. This would mean that part of the observed earnings persistence is due heterogeneity either in characteristics like sex or education or in unobservable characteristics like motivation or ability. Heckman calls the latter source of serial correlation ”spurious state dependence”.

Earnings mobility is often measured at an aggregate level by calculating mobility indices according to several theoretical concepts (**???**). The indices are helpful for comparisons across countries or social groups. But the results often differ according to the mobility concepts used (**?**). To get a better insight into the underlying dynamics and to allow for heterogeneity among individuals it is preferable to model individual earnings processes directly. In the literature we find two main approaches to this end. The literature on earnings dynamics uses continuous time series models for individual earnings development. Several authors, among them **???**, fit models to the covariance structure of earnings and estimate a permanent and a transitory component.[1] The models constrain the nature of earnings dynamics and assume a great deal of homogeneity in the dynamics across individuals. Recent developments increasingly allow for heterogeneity (**?**) or for more complicated processes(**???**) to achieve a better fit of the actual earnings distribution.

[1] **?** discuss the most prominent papers in a summary.

Whereas the first approach models earning levels, the second strand of literature is concerned with movement between relative positions or ranks in the distribution. Typically Markov processes are used to model transitions between discrete earnings states. This approach explicitely models the dynamics of the process, but the problem is the inclusion of heterogeneity into the model. In a seminal paper **?** investigates transitions into/out of low income and allows different processes for stayers and movers. The advances of econometric technique in the estimation of discrete response panel data models make it feasible to use more detailed models. Many authors still concentrate on transitions between two income states like poverty and non-poverty (**??**). **?** model transitions between earnings quintiles using a copula approach to reduce the number of parameters. All these papers use random effects estimation to account for unobservable individual heterogeneity.

The aim of this paper is to determine the degree of true state dependence in the measurement of earnings mobility. We model the dynamics of transitions between wage quintiles as a first order Markov process, which is heterogeneous among individuals. We adopt a fixed effects multinomial logit estimation procedure designed by **?**, which is based on conditional likelihood maximization.[2] With this approach it is possible to consider explicitly the effects of observed and unobserved individual characteristics on the measure of wage mobility. Modeling quintiles gives an impression of the entire wage distribution and does not focus on the lower part of the distribution alone. The fixed effects approach has the advantage of leaving the distribution of the individual effects and the correlation between the individual effects, the initial state, and exogenous variables completely unspecified which may be crucial for the consistency of the estimated parameters. A major challenge for the fixed effects approach is parameter interpretation. Partial effects on the response probability or conditional mean are not identified. Therefore the economic importance of covariates, or the amount of state dependence are difficult to determine. We take different approaches to learn from the estimation results as much as possible: by comparing them to a model not controlling for individual effects, giving an upper bound for a wage persistency measure and by simulating wage profiles.

We study wage dynamics for a large sample of Austrian employees who are observed between 1986 and 1998. The data consist of a sample drawn from the Austrian social security records. This data source provides most accurate

[2]Similar estimation procedures are used by **?** in a study of transitions between labor market states, **?** analyzing household brand choices, and **?** studying welfare participation and female labor force participation.

wage information over a long time horizon but wages are top coded at the contribution cap.[3]

The analysis of Austria is interesting because in an international comparison **?** find that wage mobility is particularly low in Austria. This is probably a consequence of the highly centralized wage bargaining system in Austria. Centralized wage setting makes aggregate wages move according to the macroeconomic conditions but it leaves the firms little room for individual adjustments. Hence, the rigid wage system would imply that there is a high degree of genuine state dependence in Austrian wages.

The remainder of the paper is organized as follows. In the next section we present the general model and two special cases in which only observed heterogeneity is allowed for or individuals are assumed homogeneous. In Section 3 we discuss the choice of the estimation method and present the estimator. Section 4 describes the data and the variables used in the model. Section 4 contains the estimation results and Section 5 concludes.

2 A model distinguishing true state dependence and heterogeneity

To describe transitions between categories of the wage distribution we adopt the latent propensity framework a la **?**. At each period, the latent variable y^*_{kit} denotes the propensity level to be in state k out of states $0, \ldots, m$ for individual i at time t. In our case states are non-employment $k = 0$ and five wage quintiles $k = 1, \ldots, m$ with $m = 5$. We observe N individuals i at $T + 1$ points in time $t = 0, \ldots, T$. The propensity function is determined by

$$y^*_{kit} = x_{it}\beta_k + \sum_{j=0}^{m} \gamma_{jk}\mathbf{1}\{y_{i(t-1)} = j\} + \alpha_{ki} + \epsilon_{kit} \tag{1}$$

where x_{it} is a vector of observable personal characteristics, $\mathbf{1}$ is the indicator function, $y_{i(t-1)}$ indicates the lagged state, $y_{i(t-1)} = j$ if the individual was in state j at $t - 1$, α_{ki} is an unobservable individual specific effect and ϵ_{kit} is an unobservable error term. The parameters of interest to be estimated are $\beta = (\beta_0, \ldots, \beta_m)$ which give the influence of observed covariates on the propensity of being in each state and γ the coefficient on the lagged endogenous

[3] This makes it complicated to model continuous wage dynamics. One way out could be a censored dynamic panel data model by **?**.

variable. The parameter γ is allowed to depend upon both the lagged state and the current state, so that there are in total m^2 feedback parameters and γ_{jk} is the feedback effect when the state j at $t-1$ is followed by the state k at time t, where $j, k \in 0, \ldots, m$. In the model γ represents true state dependence whereas $\alpha_i = (\alpha_{0i}, \ldots, \alpha_{mi})$ represents the source of spurious state dependence. We model individual heterogeneity depending on the state: each individual has a specific propensity for each alternative.

The link between the latent and the observed variables is given by the device that the observed state has maximal propensity:

$$y_{it} = k \text{ if } y^*_{kit} = \max_l (y^*_{lit})$$

As a consequence, if we assume that the underlying errors ϵ_{kit}, are independent across alternatives and over time conditional on (x_i, α_i, y_{i0}) and identically distributed according to the Type1 extreme value distribution, the probability of individual i of being in state k at time t, is given by

$$P(y_{it} = k | y_{i(t-1)} = j, x_i, \alpha_i) = \frac{\exp(x_{it}\beta_k + \gamma_{jk} + \alpha_{ki})}{1 + \sum_{l=1}^{m} \exp(x_{it}\beta_l + \gamma_{jl} + \alpha_{li})}$$

with $x_i = (x_{i0}, \ldots, x_{it})$. This implies that the transition matrix of this first order Markov process is heterogeneous between individuals.

It is worth noticing some special cases of the general model (1)

- No unobserved heterogeneity $\alpha_{ki} = \alpha_k \;\; \forall\, i = 1, \ldots, N$

$$y^*_{kit} = x_{it}\beta_k + \sum_{j=0}^{m} \gamma_{jk} \mathbf{1}\{y_{i(t-1)} = j\} + \alpha_k + \epsilon_{kit} \tag{2}$$

  This is a model where no unobserved individual heterogeneity is present and hence it is of the form of a standard multinomial logit model. If unobserved heterogeneity is absent, this model yields consistent and efficient estimates of the transition parameters. If unobserved heterogeneity is, however, present in this model, the lagged state variables and the error terms ϵ_{kit} are not independent and the estimates are inconsistent. We use the comparison of the general model (1) with unobserved heterogeneity and the multinomial logit model (2) to perform a test for the presence of unobserved heterogeneity.

- No observed or unobserved heterogeneity $\alpha_{ki} = \alpha_k$ and $x_{it} = 0 \;\; \forall\, i =$

$1, \ldots, N$

$$y^*_{kit} = \sum_{j=0}^{m} \gamma_{jk} \mathbf{1}\{y_{i(t-1)} = j\} + \alpha_k + \epsilon_{kit} \tag{3}$$

This model is the standard first order Markov process in the absence of any heterogeneity. This model is usually applied for the calculation of transition matrices like the one given in Table 2. Also mobility indices based on transition matrices are usually calculated based on this model. Again, estimates of the transition probabilities are inconsistent in presence of heterogeneity. We compare a measure of mobility for all three models (1), (2), (3) to see whether mobility increases when we control for heterogeneity

3 Conditional Maximum Likelihood Estimation

When specifying and estimating (1) one has the choice of whether to take a random effects approach or a fixed effects approach. There is a trade off between these two settings: In the random effects model, one specifies the distribution of the individual specific effects α_i. The main advantage of this approach is that it delivers a completely specified model. As a consequence all probabilities of interest under any "what-if" scenario can be estimated, provided that the model remains true. One has, however, to make assumptions about the interrelation of the distribution of α_i with the time varying explanatory variables in all periods. The usual assumption is that the time-invariant unobserved effects are independent of the observed covariates. A violation of this assumption is a source of inconsistency when estimating the model. In the model with lags of the dependent variable a further problem arises with the initial condition of the process. This requires to specify the distribution of α_i conditional on (y_{0i}, x_i) (**?**) or alternatively the distribution of y_{0i} conditional on (α_i, x_i) (**?**).[4]

The fixed effects approach attempts to estimate the model parameters β and γ without making any assumptions on the distribution of α_i, and on the way they depend on x_i. Only in special cases it is possible to estimate nonlinear models with fixed effects. **?** present a method to identify and consistently estimate panel data discrete choice models in the presence of exogenous variables, the lagged endogenous variable, and unobserved heterogeneity. Their method does also not require modeling of the initial conditions or of their statistical

[4]See **?** for a discussion of these points.

relationship with the unobserved heterogeneity. Drawbacks of this approach are, first, that the semi-parametric nature of fixed effects models may lead to estimates that are much less precise than the corresponding random effects estimates. Second, the parameter estimates by this approach do not allow one to calculate objects such as the average effect of the explanatory variables on the probability that y_{it} equals a certain state, because this will depend on the distribution of α_i.[5]

To assess the practical relevance of the theoretical considerations some studies compare fixed and random effects approaches to estimation dynamic binary response panel data models in empirical applications. **?** evaluate alternative approaches to differentiating state dependence from spurious correlation. **?** compare robustness of estimators across econometric methods and investigate in Monte Carlo simulations how sensitive the methods are to model misspecification. Both papers find that the size of estimated parameters varies considerably across estimation methods and that ignoring the contribution of heterogeneity to the initial sample observations leads to drastically overstated estimates of state dependence.

In this paper we trade the convenience of a fully specified model and straight forward parameter interpretation against the freedom from parametric restrictions on the unobserved heterogeneity and initial conditions and use a fixed effects approach. We implement the extension of the **?** method for the case of multinomial discrete choice variables, which covers our model for wage mobility. [6]

The individual fixed effects parameters α_{ik} in the general model (1) cannot be estimated consistently. Unlike in linear models the problem of incidental variables cannot be overcome by differencing. The idea applied by **?** for fixed effects logit estimation was to derive a set of conditional probabilities that do not depend on the individual effects. **?** follow up on this approach. They regard events where the state variable y switches from say state k to state l or reverse between two points in time, say s and t with $1 \leq t < s \leq T-1$. Conditional on such a switch and on the constancy of the explanatory variables in the following periods $x_{i(t+1)} = x_{i(s+1)}$, the probabilities of the events are independent of the individual effects. As the equality assumption may be too restrictive for

[5]The discussion about fixed effects versus random effects estimation in dynamic discrete choice models has been enriched by recent paper by **?** who propose to estimate bounds parameters in a random effects model in which the correlation of the individual effect and the initial condition is unspecified.

[6]A further argument is that in a multinomial framework random effects specification requires integration over multiple dimensions which is a major computational challenge.

continuous explanatory variables the exact equality condition is replaced by weighting the differences with a kernel function and giving the observations with smallest differences the highest weights. The likelihood function for model (1) is given in the Appendix. The semi-parametric estimator converges to the true value at a speed slower than the standard $\sqrt{N}$ rate. In addition, due to the use of the weighting scheme only a smaller number of observations is used for maximization. Therefore it is crucial that the estimation is based on a large sample.

The method allows strictly exogenous variables x which are time varying but with the support of $x_{it} - x_{is}$ in the neighborhood of 0 for any $t \neq s$. For this reason time dummies are excluded. We model age effects on wage mobility by defining dummy variables for age groups.

For every contribution to the likelihood function the state at two different points in time, the state in the periods before and afterwards and the values of the independent variables at these dates are important. Therefore the method can be interpreted as collecting similar histories of states and covariates, which make similar contributions to the likelihood. In contrast the estimation of the multinomial logit, model (3), corresponds to the estimation of the pooled data, neglecting the panel structure or individual histories. This estimator also assumes that the initial states are exogenously given, which is less of a restriction once we assume there are no individual effects.

4 Data

In Austria the social security authority collects detailed information for all workers, except for self-employed, civil servants and marginal workers. We use a random sample drawn from these administrative records. The data contain information on wages and labor market status of the individuals for every day and cover the years 1986 to 1998.

There are major advantages of using administrative data compared to the analyzes based on surveys. First, there is no outflow apart from death and migration and inflow into the sample is random. Hence sample attrition, which is often considerable in longitudinal surveys, is not an issue in administrative data. Second, the measurement of individual wages is highly reliable, which is extremely important for analyzing longitudinal wage development. Finally, the sample size is very large, which is crucial for the chosen estimation method. The total sample contains daily information on about 73,000 persons, who have been in the labor force at least for one day between 1986 and 1998.

As the data are collected for social security reasons there are several shortcomings for empirical analysis. Earnings data are top censored because of the contribution assessment ceiling in the social security system. The sample we use for the analysis contains at most 15% censored wage observation per year. We avoid problems with top censoring by analyzing wage quintiles.Further, the number of observable worker characteristics is rather scarce. Especially, we have no information on schooling and working time. In our analysis wage mobility is examined in terms of monthly earnings. The lack of information on working time is important mainly for women, as part-time work is quite unusual for men in Austria.[7]

As a measure for income we use the gross monthly wage on May 31st of each year. Wages are categorized according to the quintiles of the yearly wage distribution. Individuals with zero wage income on May 31st fall into the category non-employed.

We exclude all individuals from the sample who have zero earnings throughout the whole period and who are younger than 16 in 1998 and older than 64 in 1986. We are only interested in analyzing movements within the wage distribution. Transitions from education into the labor force or transitions to retirement should therefore be not considered.[8] The reductions leave an unbalanced panel of 43,078 (18,422 female) workers.

To control for macro economic effects we include the unemployment rate at a regional level as an explanatory variable in the model.We use the average unemployment rate during the second quarter of each year for 120 political districts.[9] The choice of the other explanatory variables is motivated by the results in **?**. Young individuals and individuals who changed their employer were found to be the most mobile, whereas other population groups displayed only minor differences in wage mobility. Hence we are interested to study the effects of age (especially young age) and employer changes on wage mobility. We include age effects in the model by defining dummy variables for 3 five-year age groups between the ages of 20 and 35. The effects of employer changes are measured by aggregating the number of different employers over the years.

[7]The share of part-time work 1990 was 20% for women and 1.5% for men; it was rising during the 1990's.

[8]For any individual over the age of 55 we define a series of observations in state non-employment which reaches the end year 1998 as retirement. Analogously for an individual under the age 27 we define a series of non-employment observations which starts in the first year (1986) as education. Those observations are excluded from the estimation.

[9]We thank Alfred Stigelbauer for providing the unemployment rates.

There may be problems with the assumption of strict exogeneity of the number of employers in the dynamic model, however. An unsatisfactory income situation might induce the individual to look for a better job and change the employer. This means there could be feedback effects from the lagged dependent variable on the number of employers. **?** and **?** estimate models allowing for feedback effects using a random effects specification. No such alternative is available for the fixed effects model. We alternatively estimate a specification omitting the number of employers variable. The main results do not change, so we leave the variable in the model for the results presented here.

A list of descriptive statistics is given in Table 1. We notice that the distribution among wage quintiles is different for men and women. Men are rather to be found in the upper part of the wage distribution. In the top quintile we find 24% of all male observations but only 7% females. The picture is reversed at the bottom, where women are dominant. This could also be an effect of the inclusion of part-time working women in the sample. A matrix of yearly transitions between wage categories is given in Table 2. Persistence seems to be highest in the top quintile for both sexes. Men, however, move out of the bottom wage quintiles more quickly than women.

5 Results

In this section we contrast the estimation results from the model controlling for unobserved heterogeneity (1) with the multinomial logit model (2) to see whether there is spurious state dependence. For the apparent differences between the sexes we conduct all estimations separately for men and women and compare the results. Further we analyze the influence of the explanatory variables on wage mobility.

The fixed effects estimation does not allow to estimate marginal effects and to quantify the effects. We can only interprete the signs of the estimated parameters and compare the relative magnitudes of parameters between the models. One way of measuring wage mobility would be to look at persistence, e.g. the probability to stay in the current state. We investigate how the degree of persistence changes from a model not controlling for any heterogeneity (3) to the model controlling for observed heterogeneity only (2), and in the model controlling for both observed and unobserved heterogeneity (1). In the latter model we cannot give an exact measure of persistence but we calculate an upper bound. In this way we are able to quantify the effect of spurious state dependence relative to genuine state dependence on wage mobility.

Another possibility to compare the relative importance of the effects and the differences between women and men is by simulating wage profiles. This is what we do in the last part of this section.

5.1 Estimated parameters

Estimation results of the model controlling for unobserved heterogeneity (1) are given for women and men in Tables 3 and 4. Results from estimation of the multinomial logit model (2) with no unobserved heterogeneity are given in Tables 5 and 6. In each model we choose non-employment as the reference state.

First, let us compare the feedback parameters γ in models (1) and (2). The elements on the diagonal as well as on the upper and lower diagonal are larger in model (2) not controlling for unobserved heterogeneity for both, men and women. On the other hand, the elements in the corners of the matrix are higher in this model whenever they are significantly different from zero. Thus, indeed, mobility is underestimated if we do not take spurious state dependence into account.

A formal test for the hypothesis that there is no unobserved heterogeneity can be constructed as a Hausman test. Tables 5 and 6 present consistent and efficient estimates under the null hypothesis and estimates in Tables 3 and 4 are consistent under the null and the alternative. The null hypothesis is strongly rejected. The test statistic is equal to 56,099 for female results (61,900 for male results) and is under the null hypothesis distributed as $\chi^2(45)$.

After thus rejecting the multinomial model (2) we go on to discuss the difference between women and men in model (1), looking at Tables 3 and 4. In both cases the γ parameters take on higher values above the diagonal, than below. This would imply a tendency to move upwards in the wage distribution. Women have higher probability to move to any position in the wage distribution than to move to non-employment if they are currently in the lower quintiles. This can be seen from the coefficients in the first two rows of the γ matrix, which are higher for women than for men. Women starting in the upper quintiles, however, have a lower probability to stay in paid employment than men.

Next, we turn to the effects of the covariates on wage mobility. As expected higher regional unemployment rate has a negative effect on the probability to move to any quintile relative to moving to non-employment. For men these effects are increasing with wage quintiles, with the more negative effects in the upper quintiles. For women, the unemployment only plays a significant role in the first and second quintiles. An employer change raises the probability of moving to each wage quintile as compared to moving to non-employment. Again the effects are different for men and women. For women the effect of employer change seems to be stable across quintiles, which would mean that changing the employer does not contribute to upward movement in the distribution. For men, the coefficient estimates are increasing and changing the employer helps them to move to higher income states. (Note, problems with strict exogeneity of this variable.)

We also estimate the effects of three age groups ($<$ 25 years, 25 to 29 years and 30 to 34 years) as we assume that higher aged individuals show less mobility than young ones. Indeed the parameter estimates are highest for the youngest group, who are most likely to move to any quintile compared to moving to non-employment both for men and for women. For all age groups it is most likely to move to the bottom wage quintile, but the differences among parameters for the different quintiles are highest for the youngest individuals.

5.2 Persistence measures

The aim is to define a measure of persistency and compare the measure for three models: the model with no heterogeneity (3), the model with only observed heterogeneity (2) and the model with observed and unobserved heterogeneity (1). The measure of persistency will be reduced when we control for all sources of individual heterogeneity. In the third model it should only measure the degree of genuine state dependence.

We define a measure of persistence by the sum of average probabilities of staying in each state

$$M(x, \alpha) = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=0}^{5} P(y_{it} = k | y_{i,t-1} = k, x_i, \alpha_i)$$

In in model (3), the probabilities are constant for all individuals and F is simply given by the sum of the diagonal elements in the transition matrix (Table 2). In model (2), estimated by the pooled multinomial logit, we average the predicted probabilities for staying in each state over all individuals

In model (1) things are more complicated. As we do not have consistent estimates for the α_i we cannot calculate predicted probabilities. We can, however, give an upper bound for the persistence measure F by selecting α such that for every individual the persistence for the predicted probabilities from the fixed effects logit model is maximized.

$$M(x, \alpha^*) = \max_{\alpha} \frac{1}{n} \sum_{i=1}^{n} \sum_{k=0}^{5} P(y_{it} = k | y_{i,t-1} = k, x_i, \alpha_i)$$

The maximum value of the objective function is independent of x so it is sufficient to find α^* for a single value of x. We need not perform the maximization for all possible combinations of x values. Further, the objective function is not strictly concave. It is not possible to give an analytical expression for the maximum. We found the maximum value by numerical methods.

Table 7 gives the persistence measures for women and men in all three models. We see that only controlling for observed heterogeneity does not reduce the measure a lot. This may be, because we only have a limited number of personal characteristics available. The upper bound of the persistence measure of model (1), however is considerably smaller. Controlling for unobserved heterogeneity at least reduces persistence by 27% for women (28% for men). For comparison, in the random effects model **?** find that heterogeneity explains about 50% to 70% of the overall persistence in welfare and labor force participation, respectively.

5.3 Simulations

The interpretation of the transition parameters in the multinomial model is complicated by the large number of parameters and the conditioning on reference state and the unobserved fixed effects. Also as discussed before the magnitude of the single effects cannot be quantified. Here we use a simulation approach: we design artificial individuals with special unobserved and observed characteristics and plot their simulated earnings profiles. This will enable us to see whether the interpretations we gave above to the parameter estimates show up in earnings profiles and which are the more important and less important effects.

We begin with choosing the unobserved propensity of being in each quintile with respect to non-employment $\alpha = (\alpha_1, \ldots, \alpha_5)$. We consider one individual with high propensity to move up in the wage distribution (where α_1 is small compared to α_5), one with high propensity for the lower part of the distribution (where α_1 is large compared to α_5) and one who is indifferent between the wage quintiles ($\alpha_1 = \ldots = \alpha_1$).[10] For these individuals we choose 3 different starting

[10]The constant term estimates in the multinomial logit model give an approximation for

ages (20 years, 30 years and over 35 years) and employer careers (the same employer over the total period employer changes every two years). For each individual female and male profiles are contrasted. The unemployment rate is assumed constant for all cases.

We generate earnings profiles by calculating the probability distribution over states in each year and choosing the state nearest to the expected value as the state of the current year. In period $T = 0$ all individuals start at quintile 1, to display a maximum of movements in the profiles. This process is run iteratively over 15 years and the resulting earnings profiles are displayed graphically in Figures 1 to 3. We use the mean from the unconditional distribution (the equilibrium distribution of the Markov process) as a benchmark state. If the individual starts her earnings career at the bottom of the distribution the persistence parameters will determine if or how fast she is able to reach the unconditional mean state over time.[11]

It should be stressed that this method cannot give evidence for the development of earnings for some representative individual in the sample, as the values for α are taken ad hoc. The only use of the graphical results is in simplifying the interpretation of the estimated parameters.

Figure 1 shows earnings profiles for an individual with low propensity of moving to the upper wage quintile, α is set to $-(1, 1.5, 2, 2.5, 3)$. The pictures on the left show women with different ages in the starting year. Male equivalents are shown on the right hand side. An individual never changing their employer of either sex and age group remains in the first quintile for the whole period. I takes them longer than 15 years to reach their mean income state (quintile 2). The man's upward movement can however be achieved by employer change.[12] For those men we also observe an age effect: older men moving sooner.

Next, Figure 2 with individuals indifferent between quintiles in the wage distribution ($\alpha = -(1, 1, 1, 1, 1)$) shows more mobility already. The unconditional mean positions shift to wage quintiles three to five. Men move quicker and farther in the wage distribution. The youngest man reaches the unconditional

the magnitude of the chosen α values.

[11] For a first order Markov process the equilibrium or unconditional probability distribution can be calculated from the matrix of transition probabilities. It is given by the eigenvector to the eigenvalue one, normalized to the length of one. The equilibrium probability distribution then determines the unconditional mean state for every individual. If we let the individual start at the mean state and do not change its observed characteristics the individual will remain in this state forever.

[12] Employer changes every two years are a very dynamic scenario. The average number of employer changes per individual over twelve years is two in the sample, only one per cent of individuals changes their employer more than five times.

mean after seven years. Again men profit a lot from employer changes.

The most fortunate individuals are shown in Figure 3, with $\alpha = -(3, 2.5, 2, 1.5, 1)$. We now observe everyone to move out of the bottom wage quintile after one year and most of them reach their mean state within a few periods. Men move faster and further even without employer changes.

Overall the simulations confirm the interpretation of a tendency towards upward movements in the wage distribution. Only men are supported by employer changes. Older individuals move higher in the distribution. Men are more mobile than women.

6 Conclusion

This study has investigated wage mobility in terms of transitions between quintiles of the wage distribution. We have modeled individual transitions as a first order Markov process with heterogeneous individuals. In the model we controlled for macroeconomic influences via the regional unemployment rate, age effects, and the effects of employer changes. In addition, we allowed for unobserved individual effects to capture spurious state dependence. The chosen estimation method is a dynamic multinomial logit framework with fixed effects based on conditional likelihood maximization (**?**). With this method we can consistently estimate the model parameters although we leave the distribution of the individual effects as well as the correlation between the individual effects, the initial state, and exogenous variables completely unspecified. For the empirical analysis we have used data from Austrian administrative sources, which have the advantage of providing highly accurate wage information for a large number of individuals over a long time period.

The results confirm that it is important to control for spurious state dependence. A model controlling only for observed heterogeneity is strongly rejected against the more general specification. Thus, wage mobility is underestimated if we do not take spurious state dependence into account. In the case of the Austrian economy with a very rigid wage setting system **?** found that wage mobility is extremely low in international comparison. Our results show that about 70% of the aggregate persistence in wages are due to true state dependence. Hence the centralized wage bargaining scheme seems to contribute to wage persistence to a large amount. On the other hand, a share of at least 30% in the persistence stems from individual heterogeneity. This means that individual effects certainly should not be neglected when measuring wage mobility.

The examination of simulated wage profiles shows that women are less mobile than men. This is especially disturbing as women tend to remain in the lower part of the wage distribution. Our results give the impression that, even conditional on individual heterogeneity, there exist huge barriers for women to move out of the lower part of the wage distribution.

There are several ways in which the research issues in this paper can be extended. There are arguments that the large effects of unobserved heterogeneity on transition parameters may be due to heterogeneity of transitions themselves between individuals. Meaning that the dynamics of the Markov process are higher than one. To the best of our knowledge for the multinomial model there exist neither Monte Carlo simulations for the convergence properties of the estimation method, sensitivity of the method to model misspecification and comparison of alternative estimation methodologies (fixed effects, random effects). This would of course complement our investigation of wage mobility.

References

Abowd, J. M. and Card, D. (1998). On the covariance structure of earnings and hours changes. *Econometrica*, 57:411–445.

Alvarez, J., Browning, M., and Ejrnaes, M. (2002). Modelling income processes with lots of heterogeneity. *Paper presented at the 10th International Conference on Panel Data*. Berlin.

Baker, M. (1997). Growth-rate heterogeneity and covariance structure of lifecycle earnings. *Journal of Labor Economics*, 15:338–375.

Baker, M. and Solon, G. (2003). Earnings dynamics and inequality among Canadian men, 1972-1992: evidence from longitudinal tax records. *Journal of Labor Economics*, 21:389–321.

Biewen, M. (2004). Measuring state dependence in individual poverty status: are there feedback effects to employement decisions and household composition? *IZA Discussion Paper*, (1138). Bonn.

Bonhomme, S. and Robin, J.-M. (2004). Modeling individual earnings trajectories using copulas with an application to the study of earnings inequality. *Paper presented at the 18th Annual Conference of the European Society for Population Economics*. Bergen.

Buchinsky, M., Fields, G. S., Fougére, D., and Kramarz, F. (2003). Francs or ranks? Earnings mobility in France, 1967-1999. *CEPR Discussion Papers*, (3937).

Cappellari, L. and Jenkins, S. P. (2004). Modelling low income transitions. *Journal of Applied Econometrics*. forthcoming.

Chamberlain, G. (1984). Panel data. In Grilliches, Z. and Intriligator, M., editors, *Handbook of Econometrics*, volume 2. Elsevier, Amsterdam.

Chay, K. Y. and Hyslop, D. R. (2004). Identification and estimation of dynamic binary response panel data models: empirical evidence using alternative approaches. *Journal of Econometrics*. forthcoming.

Chintagunta, P., Kyriazidou, E., and Perktold, J. (2001). Panel data analysis of household brand choices. *Journal of Econometrics*, 103:111–153.

Dickens, R. (2000). The evolution of individual earnings in Great Britain: 1975-95. *The Economic Journal*, 110:27–49.

Fields, G. S. and Ok, E. A. (1999). The measurement of income mobility: an introduction to the literature. In Silber, J., editor, *Handbook of income inequality measurement*. Kluwer Academic Publishers, New York.

Geweke, J. and Keane, M. (2000). An empirical analysis of earnings dynamics among men in the PSID: 1968-1989. *Journal of Econometrics*, 96:293–356.

Heckman, J. J. (1981a). The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In Manski, C. F. and McFadden, D., editors, *Structural analysis of discrete data with econometric applications*, chapter 4. MIT Press, Cambridge Massachusetts.

Heckman, J. J. (1981b). Statistical models for discrete panel data. In Manski, C. F. and McFadden, D., editors, *Structural analysis of discrete data with econometric applications*, chapter 3. MIT Press, Cambridge Massachusetts.

Hofer, H. and Weber, A. (2002). Wage mobility in Austria 1986 - 1996. *Labour Economics*, 9:563–577.

Honoré, B. E. (2002). Nonlinear models with panel data. *Portuguese Economic Journal*, 1:163–179.

Honoré, B. E. and Kyriazidou, E. (2000). Panel data discrete choice models with lagged dependent variables. *Econometrica*, 68:839–874.

Honoré, B. E. and Tamer, E. (2004). Bounds on parameters in dynamic discrete choice models. *Paper presented at the 11th International Conference on Panel Data*. Texas A&M.

Hu, L. (2002). Estimation of a dynamic censored panel data model. *Econometrica*, 70:2499–2517.

Magnac, T. (2000). Subsidised training and youth employment: distinguishing unobserved heterogeneity from state dependence in labour market histories. *The Economic Journal*, 110:805–837.

McCall, J. J. (1971). A Markovian model of income dynamics. *Journal of the American Statistical Association*, 66:439–447.

McFadden, D. (1974). Conditional logit anaysis of qualitative choice behaviour. In Zarembka, P., editor, *Frontiers in Econometrics*. Academic Press, New York.

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

Shorrocks, A. F. (1978a). Income inequality and income mobility. *Journal of Economic Theory*, 19:376–393.

Shorrocks, A. F. (1978b). The measurement of mobility. *Econometrica*, 46:1013–1024.

Wooldridge, J. (2000). A framework for estimating dynamic, unobserved effects panel data models with possible feedback to future explanatory variables. *Economics Letters*, 68:245–250.

Wooldridge, J. (2004). Simple solutions to the initial conditions problem in dynamic nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics*. forthcoming.

Appendix: Likelihood function for the dynamic multinomial logit model with fixed effects

Define the binary variable $y_{hit} = 1$ if the individual i is in state $h \in \{0, 1, \ldots, m\}$ in period t and zero otherwise. The estimation of model (1) with person specific individual characteristics x_{it} and fixed individual effects α_{ki} can be based on the maximization of the following likelihood function:

$$
\begin{aligned}
L \;=\; & \sum_{i=1}^{N} \sum_{1 \le t < s \le T-1} \sum_{k \ne l} \mathbf{1}\{y_{kit} + y_{kis} = 1\} \mathbf{1}\{y_{lit} + y_{lis} = 1\} \\
& K\left(\frac{x_{i(t+1)} - x_{i(s+1)}}{\sigma_n}\right) \ln \frac{\exp(D_1)}{1 + \exp(D_1)} \mathbf{1}\{s - t = 1\} \\
+ & \sum_{i=1}^{N} \sum_{1 \le t < s \le T-1} \sum_{k \ne l} \mathbf{1}\{y_{kit} + y_{kis} = 1\} \mathbf{1}\{y_{lit} + y_{lis} = 1\} \\
& K\left(\frac{x_{i(t+1)} - x_{i(s+1)}}{\sigma_n}\right) \ln \frac{\exp(D_2)}{1 + \exp(D_2)} \mathbf{1}\{s - t > 1\}
\end{aligned}
\tag{4}
$$

with

$$
\begin{aligned}
D_1 \;=\; & (x_{it} - x_{is})(\beta_k - \beta_l) + \gamma_{y_{i(t-1)},k} + \gamma_{kl} + \gamma_{l,y_{i(s+1)}} \\
- \; & \gamma_{y_{i(t-1)},l} - \gamma_{lk} - \gamma_{k,y_{i(s+1)}}
\end{aligned}
\tag{5}
$$

$$
\begin{aligned}
D_2 \;=\; & (x_{it} - x_{is})(\beta_k - \beta_l) + \gamma_{y_{i(t-1)},k} + \gamma_{k,y_{i(t+1)}} + \gamma_{y_{i(s-1)},l} \\
+ \; & \gamma_{l,y_{i(s+1)}} - \gamma_{y_{i(t-1)},l} - \gamma_{l,y_{i(t+1)}} - \gamma_{y_{i(s-1)},k} - \gamma_{k,y_{i(s+1)}}
\end{aligned}
\tag{6}
$$

In the objective function $K(.)$ is a kernel and σ_n is a bandwidth which approaches 0 as the number of observations increase to ∞. The model identification of β and γ is based on sequences of states where the individual switches between alternatives at least once during the periods 1 to $T-1$. However, only $(m^2 - (2m - 1))$ feedback parameters are identified. We impose the following identification restrictions:

$$
\begin{aligned}
\beta_0 &= 0 \\
\gamma_0 &= (\gamma_{00}, \ldots, \gamma_{m0}) = 0 \\
\gamma_{0k} &= 0 \;\; \forall\, k = 1, \ldots, m
\end{aligned}
\tag{7}
$$

$$\alpha_{i0} \quad = \quad 0 \;\; \forall\, i = 1, \ldots, N$$

which means that all parameters with respect to the reference state $k = 0$ are equal to zero. In the empirical analysis we choose non-employment as reference group, as transitions from and to this state are different from transitions between wage quintiles. The method requires at least four periods of observations since there must be some variability between the dates 1 and $T - 1$. Stable histories, where the same state is occupied between 1 and $T - 1$ do not contribute to the likelihood.

The estimator defined by maximizing the likelihood function depends on a bandwidth and a kernel to be chosen. The choice of the kernel is usually less critical than the choice of the bandwidth in applications of semi- and nonparametric methods. We choose the Epanichnikov kernel given by

$$K(u) = \max\{0, 1 - u^2\}$$

$K(.)$ has a bounded support which implies that many terms in the objective function are 0. We will experiment with different values of the bandwidth, since the choice of the bandwidth is more important than the choice of the kernel.

Table 1: Descriptive Statistics

	Women Mean	Std. Dev.	Men Mean	Std. Dev.
No income	0.241		0.160	
Quintile 1	0.293		0.055	
Quintile 2	0.184		0.138	
Quintile 3	0.119		0.195	
Quintile 4	0.097		0.214	
Quintile 5	0.067		0.237	
Observations	199,932		273,323	
Number of employers	1.93	1.23	2.15	1.47
Age (years)	31.91	12.74	32.89	12.62
Age < 25	0.34	0.47	0.30	0.46
Age 25 to 29	0.13	0.33	0.15	0.36
Age 30 to 34	0.12	0.33	0.12	0.32
Mean unemployment rates				
1987	5.52	1.91	5.60	1.95
1998	4.24	1.77	4.45	1.69
Individuals	20,897		27,826	

NOTE: Age of individual in 1987
Cumulative number of employers in 1998

Table 2: Estimated transition probabilities, yearly transitions 1986-1998, no heterogeneity

Women

Destination state	No income	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Origin state						
No income	0.760	0.156	0.047	0.020	0.012	0.005
Quintile 1	0.108	0.796	0.080	0.011	0.005	0.001
Quintile 2	0.081	0.068	0.741	0.100	0.008	0.001
Quintile 3	0.064	0.012	0.078	0.728	0.114	0.003
Quintile 4	0.056	0.005	0.008	0.064	0.790	0.076
Quintile 5	0.044	0.002	0.001	0.004	0.048	0.899

Men

Destination State	No Income	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Origin state						
No income	0.728	0.068	0.088	0.057	0.036	0.023
Quintile 1	0.176	0.610	0.155	0.038	0.015	0.006
Quintile 2	0.096	0.039	0.662	0.177	0.023	0.003
Quintile 3	0.053	0.008	0.086	0.704	0.145	0.005
Quintile 4	0.036	0.003	0.010	0.091	0.767	0.093
Quintile 5	0.025	0.002	0.002	0.004	0.055	0.913

NOTE: number of observations 199,932 females; 273,323 males. number of individuals 20,897 females; 27,826 males

Table 3: Estimated parameters from transition model with observed and unobserved heterogeneity, Women, yearly transitions 1986-1998

Destination State	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Origin state					
Quintile 1	3.264	2.73	1.925	1.625	0.998
	(0.085)	(0.106)	(0.181)	(0.293)	(0.574)
Quintile 2	1.8	3.892	3.369	2.323	1.915
	(0.103)	(0.112)	(0.141)	(0.248)	(0.574)
Quintile 3	0.392	2.219	4.051	3.545	2.26
	(0.204)	(0.133)	(0.150)	(0.192)	(0.371)
Quintile 4	-0.681	0.552	2.582	4.51	3.593
	(0.346)	(0.250)	(0.169)	(0.200)	(0.266)
Quintile 5	0.065	-0.825	0.358	2.623	4.454
	(0.642)	(0.723)	(0.519)	(0.258)	(0.305)
Regional unemployment rate	-0.084	-0.115	-0.053	-0.088	-0.072
	(0.033)	(0.035)	(0.050)	(0.061)	(0.085)
Number of employers	3.291	3.322	3.277	3.134	3.32
	(0.120)	(0.132)	(0.174)	(0.211)	(0.299)
Age $<$ 25	2.018	1.77	1.206	-0.124	0.203
	(0.374)	(0.455)	(0.529)	(0.644)	(1.069)
Age 25 to 29	0.551	0.301	0.074	-0.522	-1.212
	(0.286)	(0.372)	(0.439)	(0.490)	(0.777)
Age 30 to 34	0.196	-0.278	-0.191	-0.694	-1.143
	(0.209)	(0.265)	(0.319)	(0.372)	(0.488)
number of cases	41890				
number of individuals	16841				
mean log Likelihood	-0.174				

NOTE: fixed effects logit model estimated with conditional ML, Epanichnikov Kernel with bandwidth 0.4, standard errors are in parentheses.

Table 4: Estimated parameters from transition model with unobserved heterogeneity, Men, yearly transitions 1986-1998

Destination State	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Origin state					
Quintile 1	2.808	1.855	1.264	0.172	0.015
	(0.123)	(0.123)	(0.180)	(0.304)	(0.523)
Quintile 2	1.398	2.921	2.77	1.688	0.946
	(0.127)	(0.097)	(0.105)	(0.161)	(0.415)
Quintile 3	1.109	2.055	3.698	3.325	1.87
	(0.190)	(0.106)	(0.112)	(0.125)	(0.261)
Quintile 4	0.744	1.141	2.94	4.392	4.139
	(0.345)	(0.165)	(0.124)	(0.135)	(0.184)
Quintile 5	0.503	0.693	1.767	3.233	5.079
	(0.458)	(0.493)	(0.235)	(0.168)	(0.209)
Regional unemployment rate	-0.079	-0.107	-0.122	-0.164	-0.123
	(0.040)	(0.036)	(0.035)	(0.038)	(0.048)
Number of employers	2.736	3.231	3.272	3.365	3.601
	(0.138)	(0.118)	(0.123)	(0.133)	(0.162)
Age < 25	2.471	1.757	1.144	1.054	0.101
	(0.516)	(0.442)	(0.441)	(0.483)	(0.604)
Age 25 to 29	0.25	0.407	0.386	0.496	-0.035
	(0.390)	(0.344)	(0.337)	(0.365)	(0.434)
Age 30 to 34	0.111	0.263	0.3	0.498	0.331
	(0.282)	(0.250)	(0.245)	(0.266)	(0.306)
number of cases	47175				
number of individuals	21429				
mean log Likelihood	-0.240				

NOTE: fixed effects logit model estimated with conditional ML, Epanichnikov Kernel with bandwidth 0.4, standard errors are in parentheses.

Table 5: Estimated parameters from pooled transition model (only observed heterogeneity), Women, yearly transitions 1986-1998

Destination State	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Origin state					
Quintile 1	3.590	2.436	1.298	0.966	0.238
	(0.019)	(0.030)	(0.053)	(0.074)	(0.149)
Quintile 2	1.409	4.956	3.803	1.749	0.887
	(0.030)	(0.029)	(0.042)	(0.074)	(0.154)
Quintile 3	-0.110	2.942	6.043	4.677	1.971
	(0.066)	(0.041)	(0.043)	(0.053)	(0.131)
Quintile 4	-0.854	0.780	3.729	6.720	5.137
	(0.105)	(0.089)	(0.054)	(0.053)	(0.075)
Quintile 5	-1.475	-0.864	1.151	4.138	7.762
	(0.182)	(0.247)	(0.150)	(0.073)	(0.077)
Regional unemployment rate	-0.025	-0.065	-0.081	-0.100	-0.082
	(0.004)	(0.004)	(0.005)	(0.007)	(0.010)
Number of employers	0.235	0.282	0.227	0.233	0.211
	(0.009)	(0.011)	(0.013)	(0.016)	(0.023)
Age < 25	-0.807	-0.277	0.019	-0.051	-1.249
	(0.026)	(0.030)	(0.038)	(0.056)	(0.163)
Age 25 to 29	-0.804	-0.602	-0.335	-0.256	-0.581
	(0.024)	(0.028)	(0.033)	(0.040)	(0.069)
Age 30 to 34	-0.431	-0.486	-0.333	-0.286	-0.449
	(0.023)	(0.028)	(0.034)	(0.040)	(0.058)
Constant	-1.508	-2.626	-3.413	-3.825	-4.539
	(0.028)	(0.037)	(0.049)	(0.061)	(0.089)
number of cases	199,932				
number of individuals	20,897				
mean log Likelihood	-0.541				

NOTE: multinomial logit estimation, T=12, the reference state is non-employment, standard errors are in parentheses.

Table 6: Estimated parameters from pooled transition model (only observed heterogeneity), Men, yearly transitions 1986-1998

Destination State	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Origin state					
Quintile 1	3.641	1.950	0.991	0.514	0.171
	(0.030)	(0.033)	(0.049)	(0.072)	(0.106)
Quintile 2	1.479	4.008	3.143	1.568	0.045
	(0.036)	(0.025)	(0.030)	(0.045)	(0.098)
Quintile 3	0.462	2.600	5.143	4.016	1.148
	(0.056)	(0.030)	(0.029)	(0.034)	(0.071)
Quintile 4	-0.180	0.898	3.505	6.094	4.414
	(0.083)	(0.050)	(0.034)	(0.035)	(0.042)
Quintile 5	-0.481	-0.613	0.681	3.825	7.037
	(0.107)	(0.104)	(0.072)	(0.041)	(0.042)
Regional unemployment rate	0.001	-0.005	-0.020	-0.040	-0.087
	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)
Number of employers	0.217	0.205	0.138	0.084	-0.003
	(0.010)	(0.008)	(0.008)	(0.009)	(0.011)
Age < 25	-0.468	0.081	0.023	0.049	-0.400
	(0.037)	(0.029)	(0.030)	(0.037)	(0.067)
Age 25 to 29	-0.206	0.116	0.101	0.243	0.199
	(0.033)	(0.025)	(0.024)	(0.027)	(0.036)
Age 30 to 34	-0.176	-0.019	-0.048	0.036	0.049
	(0.031)	(0.024)	(0.024)	(0.025)	(0.031)
Constant	-2.651	-2.463	-2.685	-2.981	-2.998
	(0.040)	(0.033)	(0.034)	(0.039)	(0.048)
number of cases	273,323				
number of individuals	27,826				
mean log Likelihood	-0.523				

NOTE: multinomial logit estimation ,T=12, the reference state is non-employment, standard errors are in parentheses.

Table 7: Persistency measures according to different models

Model	Women	Men
A: no heterogeneity	4.767	4.451
B: observed heterogeneity	4.672	4.376
C: observed and unobserved heterogeneity	3.488	3.212

NOTE: Persistency is measured by the trace of the transition matrix resulting from the esitmation of each model.
For the model with oberved and unobserved heterogeneity an upper bound persistency measure is used.

Figure 1: Simulated wage profiles $\alpha = -(1, 1.5, 2, 2.5, 3)$

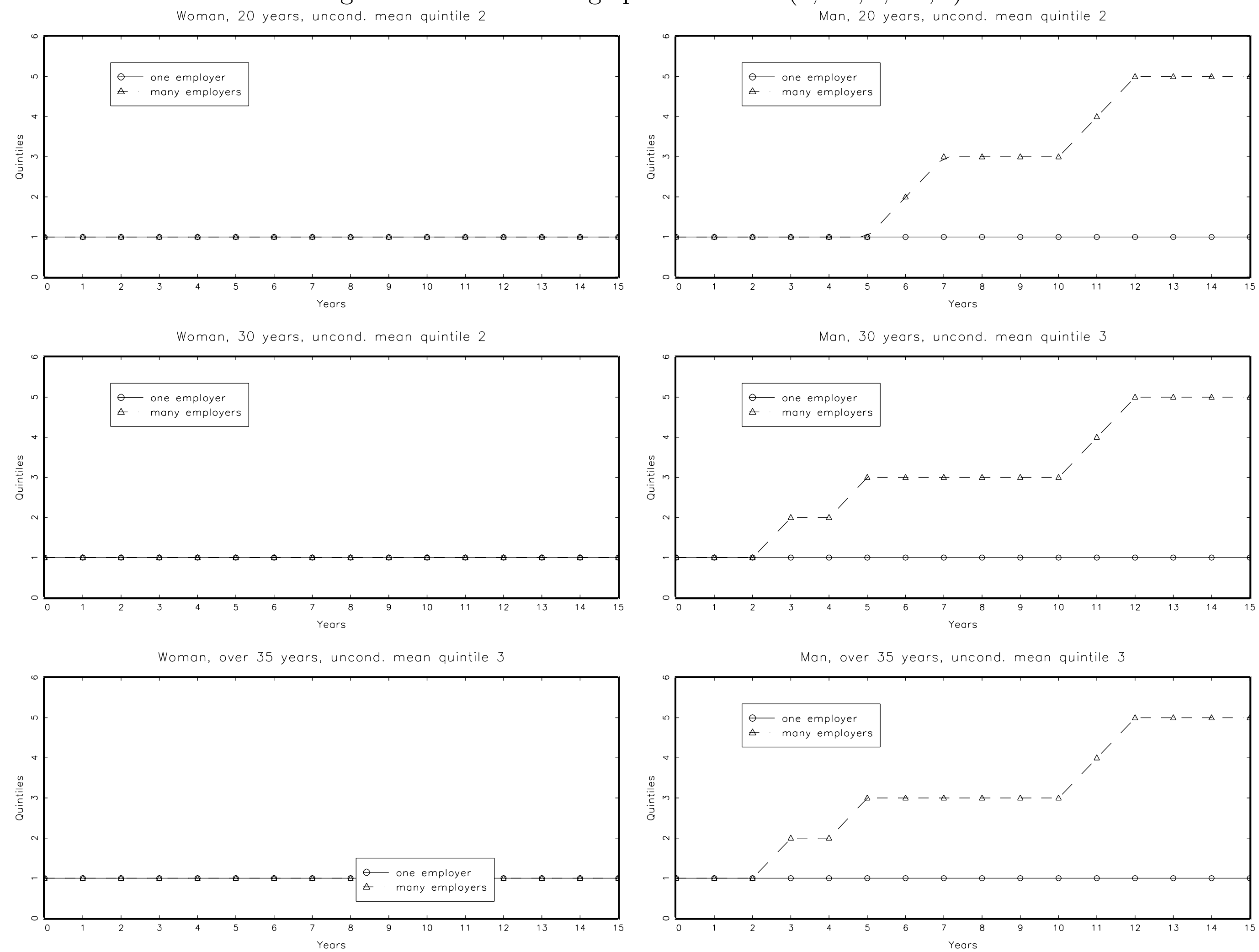

Figure 2: Simulated wage profiles $\alpha = -(1, 1, 1, 1, 1)$

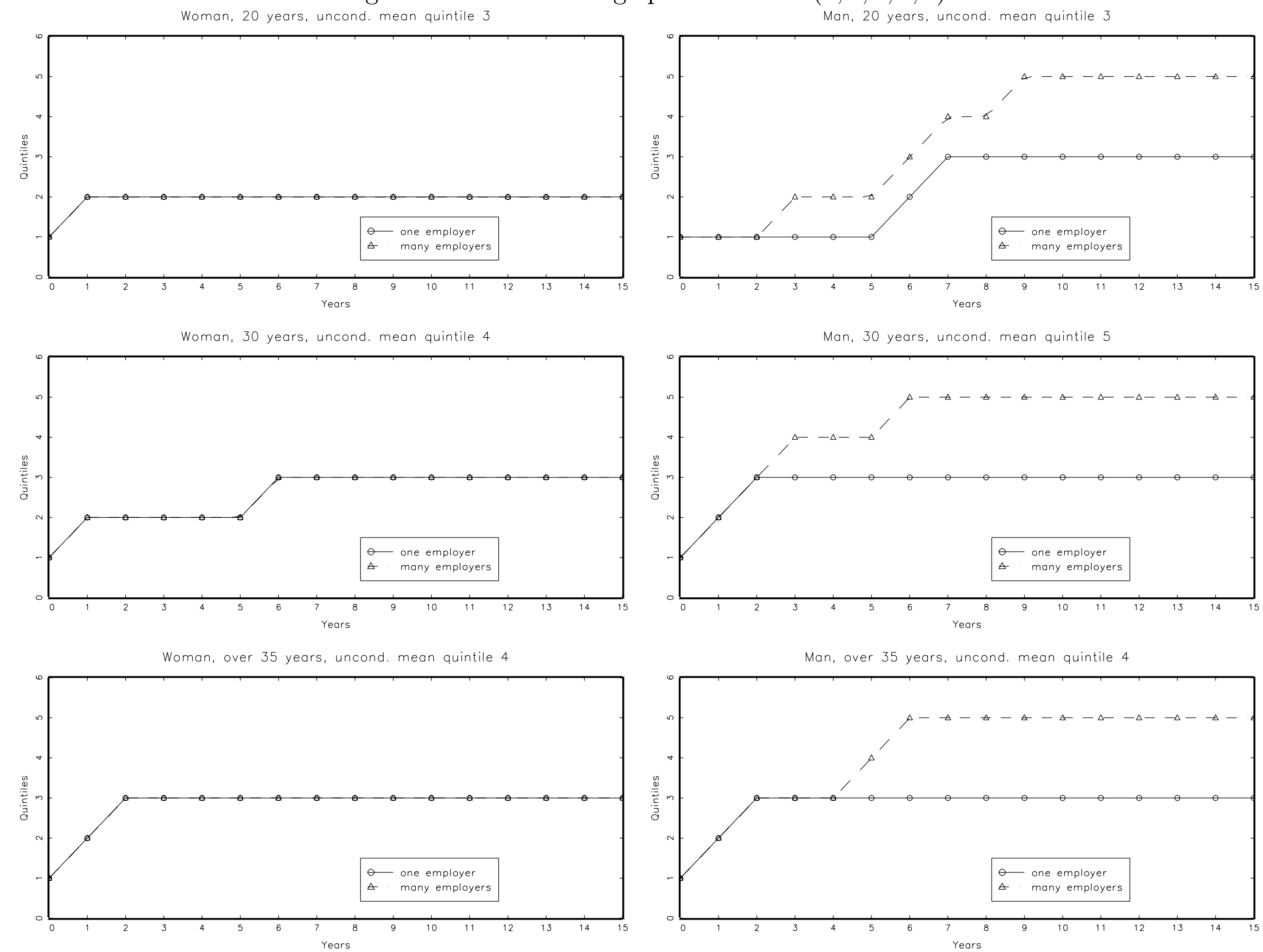

Figure 3: Simulated wage profiles $\alpha = -(3, 2.5, 2, 1.5, 1)$