RELATIVE WAGE MOBILITY: A NEW SEMI-NONPARAMETRIC ESTIMATION METHOD

Costanza Naguib

A thesis submitted for the degree of Ph.D. in Economics

Faculty of Economics
Università della Svizzera italiana

March 2018

Members of the jury:
Prof. F. Mazzonna, President of the jury
Prof. B. E. Honoré, external member
Prof. R. Maggi, thesis director
Prof. P. Gagliardini, thesis codirector

To my family

Aknowledgements

I wish to deeply thank my thesis director, Prof. Rico Maggi and my thesis co-director, Prof. Patrick Gagliardini, for their invaluable help and support during my doctoral studies. I thank them for all that they have taught me, both at a scientific and at a human level. I also would like to thank Prof. Bo Honoré and Prof. Fabrizio Mazzonna for their insightful comments and suggestions, it is an honor for me that they have accepted to be members of the jury.

I wish to thank my colleagues and friends at the Institute for Economic Research, the Institute for Political Economy and at the Institute of Finance; I really appreciated the fruitful discussions that we had from time to time. I also wish to express my gratitude to the researchers of the Laboratoire de Finance et Assurance at CREST, in particular Prof. Christian Gourieroux and Prof. Jean-Michel Zakoian, for their kind hospitality during my visiting period in Paris.

I would like to thank all the persons at Università della Svizzera italiana who have been close to me during these years, for their positive attitude and for all the enjoyable moments that we have shared. Finally, thanks to my family, for their loving support.

Contents

1 Introduction

This dissertation contains three chapters in the area of empirical labour economics. All the chapters deal with the concept of wage mobility and how to adequately model it. The first chapter focuses on the specification of a new functional copula model aimed at providing a better measure of relative wage mobility than those used in the previous literature. The second chapter applies this novel model to two countries, namely Germany and the United Kingdom and specifically investigates whether wage mobility changed after the financial crisis of 2007-2009. The third chapter more specifically concentrates on the influence of relative wage mobility on individual job satisfaction.

There is general agreement among scholars on the fact that measuring wage inequality at every single year does not allow to really understand wage dynamics over time within a country (Friedman 1962, Shorrocks 1978, Atkinson et al. 1988, Formby et al. 2004, Bonhomme and Robin 2009). This motivates the approach of the present thesis, which aims at studying the dynamics of individual wages over time. In the previous literature, many studies focused on inter-generational wage mobility (see for example, among the most recent studies, Corak 2013 or Chetty et al. 2014). However, labor markets currently appear to be characterized by increasingly flexible and discontinuous career patterns, so that the interest in factors at the roots of intra-generational wage mobility appears to be justified. There are two main approaches to analyse earnings mobility (Dickens 2000, Jantti and Jenkins 2013). The first one considers the distance between different individuals in term of their wages (absolute mobility), whereas the second one analyzes changes in the ranking of individuals from one period to another (relative mobility). The focus of the present thesis is on the latter definition. More specifically, the present work is about micro-mobility, i.e. the investigation of which individuals experience larger changes in their wage ranks. In Chapter 3 of the present thesis, we show that the impact of relative wage mobility on job satisfaction is far larger than that of absolute wage mobility on the same variable. Moreover, we find that the individual cares about her position in the entire wage distribution, not only about her position in a restricted group of peers. These findings justify the approach followed in Chapters 1 and 2. In general, we expect positional mobility to be rather high in the US, given that this country has been often defined as ”the land of opportunities” (Ferrie 2005). Indeed, Schiller (1977) finds a relevant degree of relative earnings mobility among employed males in the US. A consistent proportion of low-paid workers is found to be able to move out of their past position and obtain better-paid jobs, thus providing evidence against the existence of a so-called low-wage trap (Schiller 1977). Hence, the study of the degree of wage mobility of different groups of workers has relevant implications for the evolution of long-term earnings inequality. However, the estimation of the degree of relative or positional wage mobility poses a number of methodological challenges.

Many empirical studies tried to assess the degree of wage mobility within an economy by means of transition matrices derived from a Markov chain. However, this approach rests on some restrictive assumptions. Transition matrices, in fact, result in loss of information at the individual level, since they neglect income variations that take place within the same class. The model also assumes individual homogeneity, which may lead to a significant underestimation of wage mobility (Fields and Ok 1999, Fachinger and Himmelreicher 2012). Moreover, partition of income into quantile intervals is not operated on the basis of some economic criteria, but it is simply determined by the empirical income distribution (Schulter 1997). Bonhomme and Robin (2009) apply a copula model to the ranks of the transitory wage component. The joint distribution of the present and the past transitory component is modeled by the authors via the one-parameter Plackett copula. The copula parameter represents individual positional persistence and is a function of calendar time and individual explanatory variables (experience, experience squared and five dummies, one for each education level). However, in Bonhomme and Robin (2009), the dependence of the mobility parameter on the past rank is rigidly defined by the type of copula chosen by the authors. The goal of Chapter 1 of the present thesis, instead, consists in defining a model in which there is a function controlling the copula and mobility is defined as a functional measure, e.g. as a function of the past position in the wage distribution. This approach, which is intermediate between the standard parametric specification (which usually results too restrictive), and the fully unrestricted approach (which suffers from the curse of dimensionality), has been suggested by Gagliardini and Gourieroux (2007).

In Chapter 2, we aim at applying the new flexible model developed in Chapter 1 to a policy-relevant economic question. In this Chapter, indeed, we want to analyze and compare the degree of relative wage mobility before and after the latest financial crisis (2007-2009) in Germany and in the United Kingdom. In the previous literature, only a few studies tried to perform a cross-country comparison of mobility, in order to assess the role played by the institutional framework on the degree of positional persistence. At a first glance, one may argue that the lower the degree of Government intervention in the labor market, the higher the degree of wage mobility. It is reasonable to expect that higher levels of labor markets regulation are associated with a lower turnover rate among firms. Since job changes have been found to be positively related to both upward and downward wage mobility (respectively in the case of voluntary and in that of involuntary job changes), then high regulation should be associated, in turn, with a lower degree of wage mobility (Siebert 1997).

In particular, we expect labor market flexibility to be associated with a higher degree of earnings mobility. In a flexible labor market, people are free to fluidly move from one job to another, as a consequence of changes in demand or in technology (Atkinson and Bourguignon 1992). However, this hypothesis does not find confirmation in empirical analysis (Abraham and Housman 1993, Cardoso 2004, Sologon and O'Donoghue 2011). Labor market features such as the presence and the level of a minimum wage, the unionization rate and the degree of collective wage bargaining are not unambiguously correlated with a higher or a lower degree of positional mobility (Cardoso 2004). Indeed, it has been found that wage mobility is remarkably similar in Germany and in the United States, despite the relevant differences in the degree of labor market regulation in these two countries (Burkhauser et al. 1997). Germany is the typical "Continental" European welfare state, where employment protection is strict and welfare benefits are relatively high. On the other hand, the UK is characterized by weak employment protection and low unemployment benefits. The UK has a low coverage of collective bargaining agreements and an intermediate level of union membership. A legal minimum wage now exists in the UK as well as in Germany. However, minimum wage has been introduced in Germany only in 2015. In the UK, on the other hand, the Low Pay Commission was created in 1997 and its work culminated in 1999 with the introduction of a national minimum wage (Metcalf 1999).

In a 2016 paper, Cockx and Ghirelli (2016) find evidence that the financial crisis that took place in 2007- 2009 seriously worsened the career prospects of the young graduates. One can expect that the effects of a recession are more long-lasting in a rigid labor market such as the German one rather than in a more flexible context (like the British one). However, empirical evidence on the impact of the financial crisis on wage mobility is still scarce. According to Rinne and Zimmermann (2012 and 2013) and to Caliendo and Hogenacker (2012), the German model showed remarkable resilience during the Great Recession also thanks to the major labor market reforms that had been introduced in Germany in 2003-2005. Implemented in four waves, the so-called Hartz reforms targeted important areas that broadly affect the functioning of labor markets; unemployment benefits and social assistance schemes were restructured and generally downsized and fixed-term contracts and agency work were massively deregulated. All these measures lead to an improved performance of the German labor market in the following years. Indeed, German employment remained almost unaffected during the Great Recession. Chapter 2 sheds further light on the impact of different labor market institutions on the economic performance of an economy, by focusing on a single dependent variable, i.e. relative wage mobility.

Chapter 3 takes a detour towards the field of behavioral economics. This detour is important to explain why the focus of the first two chapters has been on relative wage mobility, defined as the change in ranks or positions from one year to the following one. There is a large number of studies that look into the role of income in determining the degree of job satisfaction. Judge et al. (2010), for example, reports that evidence on this theme is mixed; some studies found a strong positive relationship between the two variables, whereas others reported no influence at all of the level of pay on satisfaction (Judge et al. 2010). Some researchers also tried to investigate the impact of changes in the absolute wage level on individual job satisfaction (for example see Diener et al. (1993), Clark and Oswald (1996), Clark (1999), Leontaridi and Sloane (2004), Di Tella et al. (2010)). However, in these studies wage mobility has always been defined uniquely as the change in the absolute level of pay, i.e. they only considered absolute wage mobility.

Rank is usually regarded as a fundamental variable in fields such as sport economics (see for example Macmillan and Smith, 2007) or education economics (e.g. for the university rankings, see Marginson and Van der Wende 2007). The idea of positional goods, i.e. goods whose utility depend on how much of them is consumed by our neighbors, has gained attention in the economic literature in the recent decades (e.g. Frank (1991), Easterlin (1995), Stutzer (2004)). The importance of rank for individual well-being has been proved in the field of cognitive psychology (Brown et al. 2008, Boyce et al. 2010). However, to the best of our knowledge, rank (or relative) wage mobility has never been included among the determinants of job satisfaction in the past literature. This is precisely the purpose of the third and last chapter of the present thesis.

According to the range frequency theory, developed by Parducci in 1965 in the field of psychology, satisfaction will be predicted partly by the ordinal position of one's wage within a comparison set, i.e. by the individual rank. Therefore, we expect that there is a significant association between relative wage mobility and job satisfaction. On the other side, we expect absolute wage mobility to have only a limited role in determining the degree of individual satisfaction. Wage mobility may affect job satisfaction in two ways: first, since people are generally risk-averse, they prefer to earn a stable income over time. Therefore, the higher wage mobility is, the lower reported job satisfaction will be. On the other hand, wage mobility is often linked to the concept of equality of opportunity. If wage mobility in the lowest ranks of the distribution is high, low-paid workers will be able to improve their positions from one year to the following one (Friedman 1962, Clark 2003, Clark et al. 2009, Bjornskov et al. 2013). Hence, in principle, the sign of the impact of wage mobility on job satisfaction may be either positive or negative.

In what follows, it is therefore presented at first the development of a new functional copula model, to be estimated with semi-nonparametric econometric methods (Chapter 1). Then, the analysis of how relative wage mobility changed in Germany and in the UK after the financial crisis is presented (Chapter 2). Finally, a contribution on the relevance of our main variable, i.e. relative wage mobility, in determining the degree of individual job satisfaction, finds place (Chapter 3). A conclusion section draws results and summary of the present dissertation.

2 Wage Mobility: A Functional Copula Approach

2.1 Introduction

[1] In recent years, the study of inequality and its evolution over time has attracted much attention both from the side of scholars and from that of policy-makers. However, there is general agreement on the fact that measuring earnings[2] inequality in a single year does not allow the economist to draw general conclusions about earnings inequality within a country (Friedman (1962), Shorrocks (1978), Atkinson et al. (1988), Formby et al. (2004), Bonhomme and Robin (2009)). This is why, in this work, we analyse the dynamics of individual earnings over time. The contributions of the present paper are manifold. On the one side, we specify and study a new copula model, in which mobility is represented by a function instead than by a single parameter, so that the degree of mobility is allowed to change across the wage distribution, without being rigidly defined by the specific copula imposed on the data. This model feature allows us to assess whether a so-called ”low-wage trap” exists or not, i.e. whether the degree of wage mobility is lower at the bottom end of the distribution than in the middle or at the top.

In the previous literature, many studies focused on inter-generational wage mobility (see, for example, Behrman and Taubman (1985), Solon (1992), Bratsberg et al. (2007), Corak (2013), Chetty et al. (2014)). However, at present, labor markets appear to be characterized by more and more flexible and discontinous careers, so that the interest in the factors at the root of intra-generational wage mobility seems to be justified. Indeed, there are only a few studies on lifetime's wage mobility in the US. Therefore, our aim is to obtain further insights on the

[1]We would like to thank Richard Blundell, Martin Browning, Robert Gary-Bobo, Christian Gourieroux, Bo Honoré, Thierry Kamionka, Rico Maggi, Blaise Melly, Alain Monfort, Jean-Marc Robin, Sébastien Roux, Mirco Rubin, Stefan Sperlich, Mark Watson, the participants to the Brown Bag seminar at USI Lugano on 9^{th} December 2015, the participants to the Annual Conference of the Swiss Society of Economics and Statistics on $9^{th} - 10^{th}$ June 2016 in Lugano and the participants to the junior seminar at CREST on 11^{th} April 2017 for their insightful comments.

[2]In this paper the terms earnings and wages are used interchangeably; earnings stand for labor earnings.

wage dynamics in the US labor markets, controlling for both the individual's past position in the distribution and for her characteristics.

Our aim is to assess the impact of education and that of the past position in the distribution on the current wage rank. Indeed, if we were to find empirical evidence that highly educated individuals previously found at the bottom of the distribution, have a higher likelihood to experience upward positional mobility in the following year than their less-educated colleagues, then the policy implication would be that encouraging higher education helps low-paid workers to escape the low-wage trap. In general, the finding of a higher degree of mobility at the lower end of the wage distribution would suggest that, in the US labor market in the period considered, there is no such thing as a low wage trap.

Piketty and Saez (2003) as well as Gabaix et al. (2016) point out that income inequality in the US started to rise again from the '60s after a period of decline in correspondence to WWII. According to Autor et al. (2006), to Autor and Dorn (2013) and to Foote and Ryan (2015), since the'70s in the US wage inequality began to increase and, starting from the '80s, the US labor market witnessed a growing polarization between low-paid and high-paid jobs. Moreover, according to Goos et al. (2009), starting from the early '90s a similar polarization process also took place extensively in the European countries. This suggests that, in the period that we consider, a low-wage trap actually existed in the US, preventing workers in low-paid or secondary jobs to access the high-paid (primary) positions. This is consistent with the theory of dual or segmented labor market (Reich et al. (1973), Piore (1975), Dickens and Lang (1988)).

There are two main approaches to analyse earnings mobility (Dickens (2000), Jantti and Jenkins (2013)). The first one considers the distance between different individuals in term of their wages (absolute mobility), whereas the second one analyzes changes in the ranking of individuals from one period to another (relative mobility). The focus of the present paper is on the latter definition. More specifically, the present work is about micro-mobility, i.e. the investigation of which individuals experience larger changes in their wage ranks. In a companion paper (Naguib and Maggi (2017)) we show that the impact of relative wage mobility on job satisfaction is far larger than that of absolute wage mobility on the same variable. Moreover, we find that the individual cares about her position in the entire wage distribution, not only about her position in a restricted group of peers. These findings justify the approach followed in the present paper. Our methodology allows us to model the present rank as a nonlinear function of both the past rank and of an index which is constructed as a weighted sum of individual explanatory variables. In this way, it is possible to virtually analyze the transition from the past to the present rank of each individual in the sample and to assess which are the characteristics associated with a higher or a lower degree of positional mobility. Differently from what has been done by Bonhomme and Robin (2009), dependence of positional mobility on the past rank is here explicitly taken into account nonparametrically in the model, rather than being determined by the chosen parametric copula function.

In a recent paper, Arellano and Bonhomme (2017) develop a novel quantile selection model, with a method to correct for sample selection, and apply it to the estimation of wage percentiles in the UK for the period 1978-2000. The focus of the authors is rather different from ours, since they aim at estimating wage quantiles and their evolution over time, rather than at modelling the individual dynamics of wage ranks. Further, Arellano et al. (2017) propose relevant advances in the modellization of income processes. In particular, they study nonlinear persistence of the earnings process. The focus of the authors lies in macro-persistence of income, i.e. evaluated at different (aggregate) percentiles. Our focus, instead lies in the analysis of individual dynamics of wage ranks.

2.1.1 Literature review

Previous studies often found evidence of a lower degree of wage mobility at the extremes of the distribution and in particular at the lowest decile, for example in the case of the US (Hungerford (1993), Gottschalk (1997)). Indeed, Auten et al. (2013) find evidence of a rather high positional persistence in the top 1% of taxpayers in the US. In principle, we expect upward wage mobility to be increasing in education and downward wage mobility to be decreasing in the same variable. Aggregating these two countervailing effects, Gittleman and Joyce (1996) claim that stability of earnings increases with education. Tansel et al. (2014), too, find a negative relationship between the education level and the degree of mobility. This is mainly due to the fact that high-educated individuals have substantially lower probabilities of experiencing downward mobility. As for the US, Auten and Gee (2009) find evidence that the most important explanatory variables of wage mobility are age and the past position in the distribution. Hence, these variables are included in our model. The authors find upward wage mobility to be decreasing both in age and in the past position, i.e. upward mobility is higher for workers being at the bottom percentiles in the previous period. In general, we expect positional mobility to be rather high in the US, given that the country has been often defined as "the land of opportunities" (Ferrie (2005)). Indeed, Schiller (1977) finds a relevant degree of relative earnings mobility among employed males in the US. A consistent proportion of low-paid workers is found to be able to move out from their past position and obtain better-paid jobs, thus providing evidence against the existence of a so-called low-wage trap (Schiller (1977)).

Buchinsky and Hunt (1999) find that wage mobility contributes to lower wage inequality in the US. More specifically, the authors find evidence that the highest probability of remaining in the same quantile in the next period are recorded at the top of the distribution. However, the staying probabilities at the bottom of the wage scale are also rather high. Burkhauser et al. (1997), too, finds a large amount of rank persistence in the US labor market. To conclude, Kopczuk et al.

(2010) show that the degree of earnings mobility in the US has been quite stable over the past decades.

2.1.2 Methodological considerations

Many empirical studies tried to assess the degree of wage mobility within an economy by means of transition matrices derived from a Markov chain. However, this approach rests on some unrealistic assumptions. Transition matrices, in fact, result in loss of information at the individual level, since they neglect income variations that take place within the same class. The model also assumes individual homogeneity, which may lead to a significant underestimation of wage mobility (Fields and Ok (1999), Fachinger and Himmelreicher (2012)). Moreover, partition of income into quantile intervals is not operated on the basis of some economic criteria, but it is simply determined by the empirical income distribution (Schulter (1997)). To consistently estimate the transition probabilities, it is necessary to adequately take into account past individual histories. Bonhomme and Robin (2009) apply a copula model to the ranks of the transitory wage component. The joint distribution of the present and the past transitory component is modeled by the authors via the one-parameter Plackett copula. The copula parameter represents individual positional persistence and is a function of calendar time and individual explanatory variables (experience, experience squared and five dummies, one for each education level). The authors find a significant degree of wage immobility in the early '90s in France. Relative mobility appears to be decreasing in both experience and the level of education, a feature that is likely to represent the decreased downward mobility risk for the older and high-educated workers. However, in Bonhomme and Robin (2009), the dependence of the mobility parameter on the past rank is rigidly defined by the type of copula chosen by the authors. Our goal, instead, consists in defining a model in which dependence of the mobility measure on the past position derives from the data and not from the copula structure imposed over them. Indeed, as it emerges from the past literature, wage mobility may take an inverted U-shaped pattern, being low at the extremes of the distribution and high in the middle of it (de Coulon and Zuercher (2001), Cardoso (2004), Pavlopoulos et al. (2007), Germandt (2009). In particular, Cardoso (2004) finds that mobility is lower in the bottom and the upper tails of the distribution, both in the UK and in the US.

The aim of the present work is to estimate a copula model, in which mobility is a function of both the past rank, and of some individual explanatory variables. To the best of our knowledge, the analysis of a copula density with a functional parameter, the arguments of the copula being the current value and the first-order lag of an underlying non-linear autoregressive process, has never been performed in the past literature. A copula is a function that joins a multivariate distribution function to its one-dimensional marginal distribution functions (Joe (1997), Nelsen (1999), see Appendix D in the Supplementary Material). As explained by Patton (2012), copula-based models allow the researcher to specify the models for the marginal distributions separately from the dependence structure that links these distributions to form a joint distribution. This feature of copulas allows for a large degree of flexibility in model specification and estimation. Indeed, in the past literature the marginal distributions have been often estimated by nonparametric methods (usually resorting to the empirical cumulative distribution function), whereas the unknown parameter characterizing the copula model is estimated via a Maximum Likelihood approach (Patton (2012)). Chen and Fan (2006) and Otto (2005) study the efficient estimation of a model with fully parametric copula function and non-parametric marginal distributions. Chen et al. (2009) examine the asymptotic properties of a copula-based quantile autoregressive model; however, their autoregressive model is fully parametric. Indeed, in the conclusions the authors mention semi-parametric modelling of the copula itself via the method of Sieves as a feasible strategy to expand the menu of the currently available parametric copula. Such semi-parametric estimation will be performed in the present work. The focus of our paper is on a model in which the parameter of the Gaussian copula is replaced by a bivariate function to be estimated by means of nonparametric methods. We use this copula with the present rank and the arguments of the bivariate autoregressive function, i.e. the past rank and the weighted sum of individual characteristics, to model the dynamic of the ranks so that the marginal law of the present rank remains standard normal. In this setting, the challenge arises from the introduction of explanatory variables in the model.

Our approach is intermediate between the standard parametric specification, and the fully unrestricted approach and builds on Gagliardini and Gourieroux (2007). The authors examine the case in which the joint density is constrained and depends on a small number of one-dimensional functional parameters. Such a constrained nonparametric approach has some advantages with respect to both extremes presented above in order to model nonlinear dependence. First, by using functional parameters instead of scalar ones, it is possible to achieve greater flexibility and a better fit of the model to the data, compared to fully parametric specifications. Second, as shown by Gagliardini and Gourieroux (2007), the rate of convergence of the appropriate estimators for the functional parameters and for the joint density is equal to the standard one-dimensional nonparametric rate, thus avoiding the curse of dimensionality of the fully nonparametric approach.

The remaining of the present paper is structured as follows. As a descriptive analysis of the data, Section 2 presents some kernel estimates of the relationship between the present and the past ranks as well as preliminary unconstrained estimates. Section 3 introduces the full model used to define the joint dynamics of the ranks. Section 4 is devoted to the description and comment of the results of the constrained estimates and Section 5 concludes.

2.2 Data and exploratory analysis

2.2.1 Data description

In this section we perform a descriptive analysis of the data, in order to get preliminary insights on the shape of the relationship between present and past wage ranks. From the Panel Study of Income Dynamics we collect an unbalanced sample that contains 35'378 observations (nT) and covers the period 1975-1996[3]. Since our aim is to extend Bonhomme and Robin (2009) work and hence to obtain comparable results, we follow the same procedure used by the authors in order to construct our sample. More precisely, we drop observations for students, retirees and self-employed workers and we only include observations relative to full-time male employees, in order to limit the role of variations in the intensive margin of labor supply on wage dynamics (Bonhomme and Robin (2009), Bachmann et al. (2014)). We consider individuals aged between 15 and 64, since these are the usual thresholds for the definition of the active population. Moreover, as usual in the wage mobility literature (Buchinsky and Hunt (1999)) we exclude observations with wage equal to zero. This means that, for each couple of years considered, we exclude from our sample observations of female and part-time workers, of those who are beyond the age thresholds of the active population and of those who earn a zero wage. We are aware that there may be differences in the functioning of the labor market in different states of the US. However, due to feasibility constraints and in the wake of the previous literature (Burkhauser et al. (1997), Buchinsky and Hunt (1999), Auten et al. (2013), Kim (2013)) we will consider the US labor market as a whole.

Similarily to what has been done by Bonhomme and Robin (2009) we use as individual explanatory variables age, age squared, and five education dummies. We assume that these explanatory variables are exogenous. The education dummies are constructed on the basis of the varia-

[3]The choice of the time span analyzed here is due to the features of the PSID data. Indeed, data on years of education completed, our main explanatory variable of wage mobility, are only collected starting from 1975. Moreover, the structure of the survey changed in 1997, becoming bi-annual instead of annual.

ble ”years spent in education”, which is recorded in the PSID. According to the US education system, the first dummy correponds to 1-8 years of education, thus including elementary and middle school; the second dummy stands for junior high school (9-11 years), the third one represents senior high school (12-15 years), the fourth is for college completion (16 years) and the last one stands for graduate education (17 years of education or more).

Table 2 in Appendix A reports the descriptive statistics for our main variables of interests, i.e. wage and education. Note that in Table 2 education is expressed by the number of year of education completed by the individual. However, as it has been explained above, for the estimation of our model we transform this variable, by replacing it with five education dummies, one for each education level. Each education level corresponds to a certain number of years spent in school, as explained above. We then divide our sample into 6 sub-samples, one for each age cohort (from 17 to 24 years, from 25 to 32, from 33 to 40, from 41 to 48, from 49 to 56 and from 57 to 64). Descriptive statistics by age cohort for our main variables of interest are reported in Table 2 in Appendix A. These summary statistics about education and wages are almost constant across different age groups; this suggest that our estimation results will not be biased by a cohort effect. In Figure 15 (Appendix A), we also plot the empirical cumulative distribution functions (cdf) of real annual wage (i.e. adjusted for inflation), in order to assess whether there is a significant cohort effect. These empirical distribution functions are rather close to each other (they almost fully overlap) and they are also close to the global cdf (i.e. that computed on the real annual wage of all age groups). From Figure 13, reported in Appendix A, we notice that the empirical cumulative distributions of real wages in 1976 and in 1996 almost overlap, once we correct for inflation in both years considered. This suggests that the wage distribution in real terms in the US did not witness dramatic changes in the period considered in our sample. In this context of a rather stable real wage distribution, we want to assess which groups experienced the highest degree of positional mobility and which ones, instead, had the lowest mobility. As usual in the education economics literature, we find evidence of a positive relationship between wage and education. Indeed, if we regress log wage on the years of education completed, the estimated coefficient is positive and and statistically significant (at 99% confidence level, result not reported for brevity).

2.2.2 Exploratory estimates

In this Section, we start to specify some simple models of the relationship between the present and the past rank, before studying some nonparametric model specification. First, we specify a linear OLS model, then we perform a polynomial regression and finally we resort to an unconstrained nonparametric model for the rank dynamics. In our first OLS preliminary regression, presented in equation (2.1), the dependent variable is gross annual wage. We define individual annual wage as the sum of two components, a permanent one and a transitory one (Bonhomme and Robin (2009)). We first regress log earnings on age and the ratio of experience to age after 15 (the threshold for active population) in order to purify earnings from the permanent component, and then we study the evolution over time of the transitory component alone. This preliminary estimation is perfomed via the panel data fixed-effect technique, in order to adequately take into account the potential presence of unobserved heterogeneity across workers[4]. We also include in this regression a time fixed-effect (λ_t) in order to take into account all the macroeconomic shocks on wages, among them also the impact of inflation on wages. The model reads:

[4]Experience is recorded in the PSID as the number of years that the individual worked full time. For the period of interest (1975-1996) and with reference to our sample of full-time male workers, the empirical correlation between age and experience is rather high (0.89). Hence, in order to take into account the effects of both age and experience on log wages, we introduce the ratio of years of experience to years after 15, in order to solve the problem of potential collinearity between our explanatory variables. There are many possible procedures to disentangle the permanent and the transitory wage components, we chose here a rather simple one, which is already established in the literature. A throughout analysis of the relationship between the permanent and the transitory individual earnings component, indeed, lies beyond the scope of the present paper. For detailed information on a more elaborate procedure to disentagle the permanent and the transitory earnings components, see Bonhomme and Robin (2010).

$$Wage_{i,t} = \beta_1 Age_{it} + \beta_2 \frac{Experience_{it}}{Age_{it} - 15} + \eta_i + \lambda_t + \epsilon_{i,t} \tag{2.1}$$

where $Wage_{i.t}$ stands for log earnings, $\frac{Experience_{it}}{Age_{it}-15}$ is the proportion of years of activity in labor market to age in excess of 15, η_i represents the individual fixed effect and $\epsilon_{i,t}$ the transitory wage component[5]. From the residuals of this preliminary regression, we obtain the Gaussian wage ranks via the following formula:

$$Z_{i,t} = \Phi^{-1}(\hat{F}_t(\hat{\epsilon}_{i,t})).$$

We compute earnings ranks using the empirical cdf of earnings residuals, $\hat{F}_t$, and we apply the quantile function of the standard normal distribution, Φ^{-1}, to impose standard gaussianity (Gottschalk (1982), Moffitt and Gottschalk (2002), Kalwij and Alessie (2007), Bonhomme and Robin (2009)). We perform some unit-root tests on $Z_{i,t}$, to ensure that these n variables are stationary. The results of the Levin-Lin-Chu test, of the Harris-Tzavalis test and of the Breitung test allow us to reject at a 99% confidence level the null hypothesis that some panels contain unit roots (results reported in Table 3 of Appendix A). We assume for simplicity that the conditional rank variance is constant across the wage distribution. A kernel estimation of the conditional rank variance seems to confirm this hypothesis. This result is reported in Appendix A (Figure 17). Moreover, from Figure 14 in Appendix A we notice that the relationship between present and past Gaussian ranks does not seem to be significantly influenced by the age cohort. We introduce now a slightly more complex model, in which quadratic and cubic terms are also present:

$$Z_{i,t} = a_0 + aZ_{i,t-1} + bZ^2_{i,t-1} + cZ^3_{i,t-1} + e_{i,t} \tag{2.2}$$

[5]We tried different model specifications for this preliminary regressions, e.g. also including a quadratic term, such as experience squared or age squared, among the explanatory variables. The results did not significantly change across different model specifications.

We choose to estimate a polynomial rank regression of third degree in order to allow for asymmetry in the relationship between the present and the past ranks, given that such asymmetry emerges from the data. From the polynomial regression reported in Table 1, we deduce that the coefficient of the linear term in equation (2.2) is rather high (0.97), thus suggesting a substantial degree of positional immobility from one year to the following one. The fit of this regression is represented by the dashed line in Figure 1. The coefficients of the quadratic term and that of the cubic past rank term are statistically significant at a 99% confidence level, thus providing further support to the hypothesis of a nonlinear relationship between the present and the past Gaussian ranks. The statistical significance of these coefficients is preserved when we run separate polynomial regressions for each copule of years in the sample (results are not reported for brevity). Indeed, this finding is consistent with Altonji et al. (2013) and Arellano et al. (2017), who claim that linear approximation of earnings dynamics may not always be accurate.

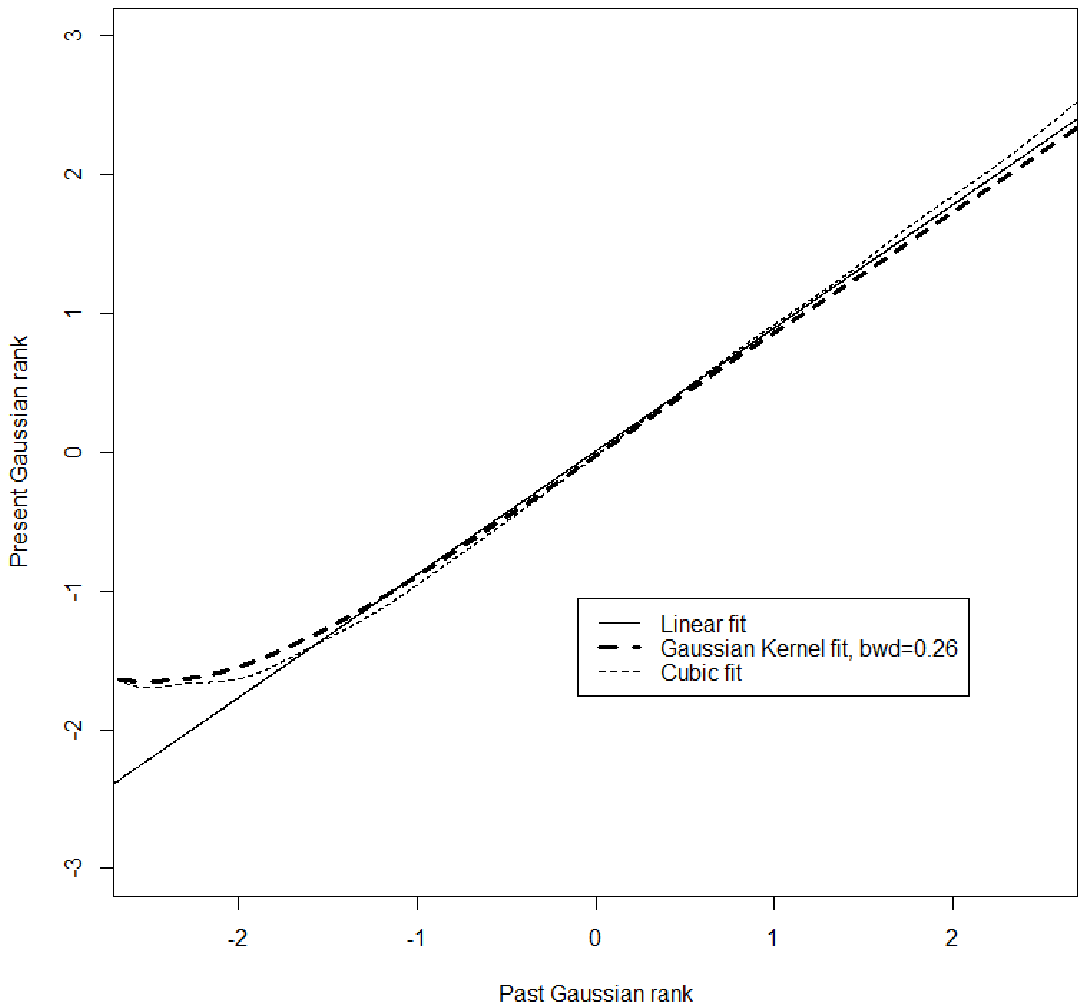
Table 1: Third-degree polynomial regression

	Full sample (1975-1996)
constant	-.0454***
	(.0030)
$Z_{i,t-1}$	.9717***
	(.0040)
$Z_{i,t-1}^2$	.0456***
	(.0017)
$Z_{i,t-1}^3$	-.0290***
	(.0011)
n. of obs.	35'346

This table reports the estimated coefficients of the third-degree polynomial regression in (2.3). This regression has been run on pooled PSID data for the period 1975-1996.

As an additional exploratory data analysis, we run a kernel regression of $Z_{i,t}$ on $Z_{i,t-1}$, in

Figure 1: Gaussian ranks, full sample (1975-1995), nT=35'346



In this figure the results of three preliminary estimations are reported. The black line stands for the fit of a linear regression of the present position on the past one; the dashed line represents the fit of a polynomial regression of degree 3 in the past rank, where the dependent variable is again the present rank. The bold dashed line stands for the result of a Gaussian kernel regression of the present rank on the past one.

order to get a first idea of the shape of the relationship between present and past ranks. The fit of this kernel regression is represented by the bold dashed line in Figure 1. Both from the kernel regression and the polynomial rank regression (Figure 1), we get some intuition that the relationship between the present and the past Gaussian ranks is nonlinear in nature. In particular, the slope of the function linking the present and the past ranks appears lower at the extremes of the wage distribution. This finding suggests that the degree of wage mobility is higher both at the bottom end and at the upper end of the distribution than in the middle of it.

In unreported results, we regress the cross-sectional correlation between the present and the past ranks over the annual GDP growth rate[6], in order to check for the stability over time of the relationship between the present and the past rank, and we find the the estimated coefficient is very small and not statistically significant. Hence, we deduce that the correlation between the present and the past ranks exhibits no cyclical behavior. We also tried to regress our correlation on the annual unemployment rate in the US[7] and on the share of immigrants on the total population[8]. In both cases, the estimated association between the two variables was not statistically significant (results not reported for brevity). This convinced us not to include macroeconomic factors among the explanatory variables that we deem at the origin of the rank dynamics[9].

We now want to assess how the relationship between present and past ranks changes, depending on individual explanatory variables such as age and education. We are interested in performing a preliminary unconditional estimate of the model:

$$Z_{i,t} = \rho(Z_{i,t-1}, W_{i,t}) + \varepsilon_{i,t} \tag{2.3}$$

where $\rho(Z_{i,t-1}, W_{i,t})$ is the autoregressive function which represents the present individual posi-

[6]Data source: World Bank Indicators

[7]Data source: OECD

[8]Data source: OECD

[9]Note that, in all the three OLS regression described here, the dependent variable, i.e. the correlation between the present and the past rank, which is always positive, has been transformed by means of the inverse logistic function, $y = \log(\rho/(1-\rho))$, so that it has an unbounded support.

tion in the distribution as a function of his position in the previous period and of some individual explanatory variables (here age, age squared and education)[10]. We define the variable $W_{i,t}$ as $W_{i,t} = X'_{i,t}\beta$; this variable is called a score, or index. It is the weighted sum of the individual characteristics and allows to avoid the curse of dimensionality when numerous variables are included in the model. This index is then standardized so that its distribution is (approximately) standard normal. The details of the standardization are reported in Appendix B.

Single index models have been largely studied by the previous literature. Ichimura (1987) proposed a semi-parametric estimator for a class of single index models. Powell et al. (1989) solved the problem of estimating coefficients of index models via the estimation of the density-weighted average derivative of a general regression function, without making restrictive assumptions on the regression function or on the distribution of the data. Hardle et al. (1993) tackled the problem of optimal smoothing in the context of single index models. Li and Racine (2007) offer a summary of the most diffused methods for estimation and review the asymptotic properties of these estimators.

In order for model (2.3) to be identified, some conditions must be satisfied (Li and Racine (2007)). First, the vector of explanatory variables, X_{it}, cannot include a constant term and at least one of the variables must be continuous. Moreover, the coefficients in the β vector have to be normalized. Here, we adopt the common practice of setting the coefficient of one of the variables (here: the dummy for the first education level) equal to one. The statistical significance of this coefficient is confirmed by some preliminary OLS regressions that we run. The other variables are age at the beginning of the period considered, age squared (multiplied by a factor 10^{-2}, measured at the beginning of the sample period) and other four education dummies.

In model (2.3), we do not know neither the bivariate function $\rho(\cdot,\cdot))$, nor the vector of coefficients β in the index $W_{i,t}$, so we need to estimate them. We assume that the explanatory

[10]This unconstrained model does not take into account the marginally Gaussian law of Z_{it}. A complete model, which includes both the marginal and the joint laws of Z_{it} and W_{it} is presented in Section 3.

variables X_{it} included in the score are exogenous. This makes sense from an economic viewpoint. Indeed, it is reasonable to assume that the education level achieved before labor market entry is not influenced by the rank assumed in the wage distribution once the individual actually entered the job market. Hence, we exclude on-the-job training from our explanatory variables due to the risk of potential endogeneity. We estimate the model by an iterative procedure, the details are reported in Appendix B for brevity[11].

In the remaining of the present section, we present some unconstrained estimation results obtained from model (2.3) on the US data from the Panel Study of Income Dynamics. From our unconstrained estimate of the autoregressive function by means of a Hermite polynomial of degree 2, we find that the score is always increasing in age and it is increasing in the education level until college ($Edu4_{it}$):

$$\begin{aligned}\hat{W}_{it} = -0.0004Age_{it} + 0.0048Age_{it}^2 + Edu1_{it} + 1.0155Edu2_{it} + 1.4691Edu3_{it}\\ +1.6757Edu4_{it} + 1.3786Edu5_{it}.\end{aligned}$$

Hence, people with a low value of the score may be young or low educated or both. The estimated autoregressive function is given by the following expression:

$$\begin{aligned}\rho(Z_{i,t-1}, \hat{W}_{i,t}) = -0.0068 + 0.0263\hat{W}_{i,t} + 0.8816Z_{i,t-1} + 0.0383\hat{W}_{i,t}Z_{i,t-1}+\\ -0.0153(\hat{W}_{i,t}^2 - 1) + 0.0417(Z_{i,t-1}^2 - 1) - 0.0164(\hat{W}_{i,t}^2 - 1)Z_{i,t-1}+\\ -0.0008(Z_{i,t-1}^2 - 1)\hat{W}_{i,t} + 0.0003(\hat{W}_{i,t}^2 - 1)(Z_{i,t-1}^2 - 1)\end{aligned}$$

On the basis of the previous equation, we compute the partial derivative of the autoregressive function with respect to the past position. This is our summary mobility measure; a high value of this derivative means that the present rank moves close together with the past rank, therefore

[11]It is worth mentioning that, at each estimation step, the score is centered and standardized, so that it is approximately standard normally distributed. More details on this are provided in Appendix B.

there is a high degree of positional immobility in the labor market. On the other side, a low value of this partial derivative would indicate that present and past rank are not closely associated; therefore, the labor market would be characterized by a high degree of rank mobility. Of course, this partial derivative is an incomplete and summary mobility measure. The full information about rank dynamics lies in the bivariate autoregressive function ρ. However, our summary mobility measure has the merit of allowing us to investigate mobility pattern in a more readily interpretable and comparable way.

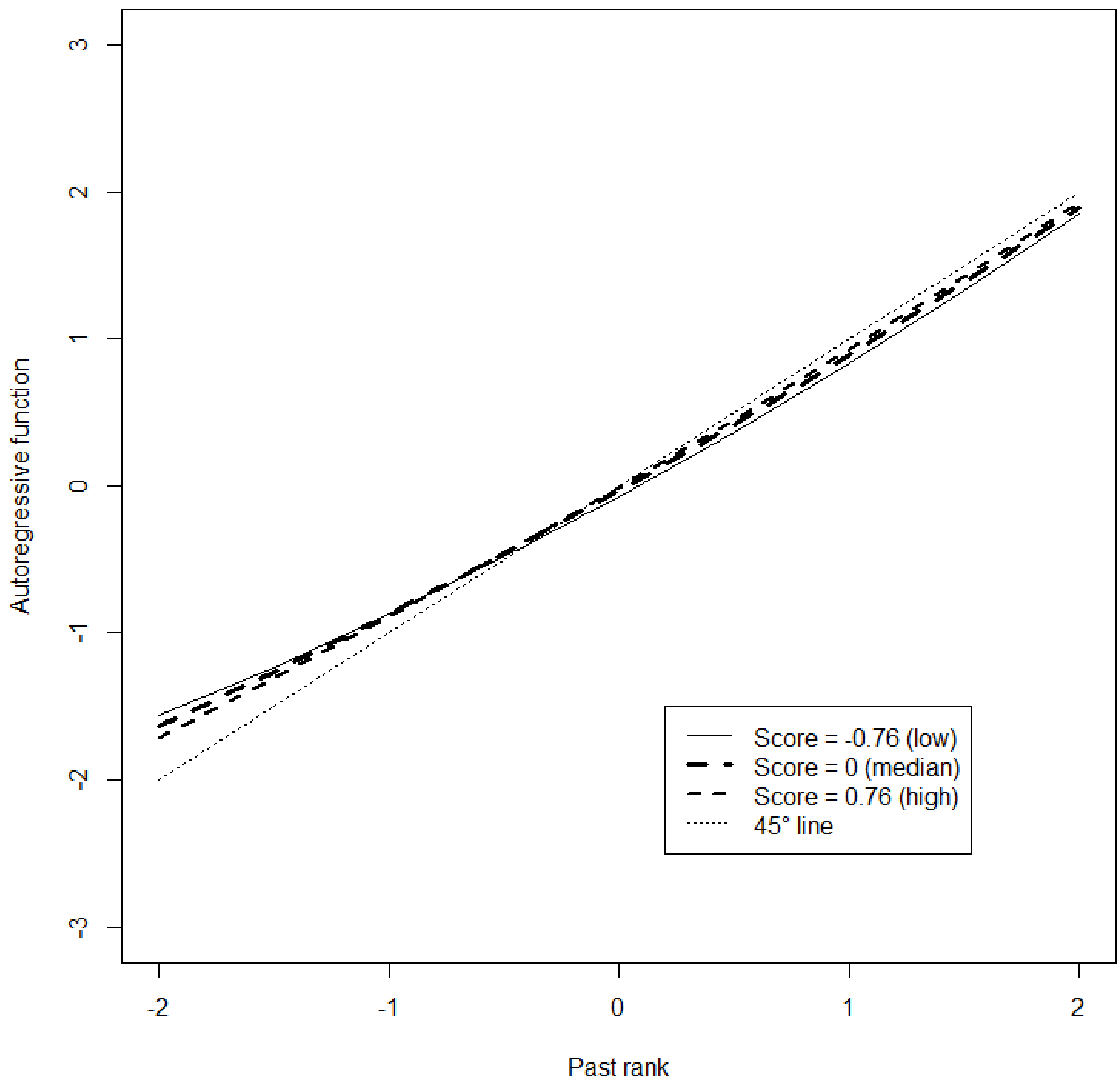
$$\begin{aligned}\frac{\partial\rho(Z_{i,t-1},\hat{W}_{i,t})}{\partial Z_{i,t-1}} = 0.8980 + 0.0383\hat{W}_{i,t} + 0.0322Z_{i,t-1} - 0.0164\hat{W}_{i,t}^{2} \\ -0.0016Z_{i,t-1}\hat{W}_{i,t} + 0.0006Z_{i,t-1}\hat{W}_{i,t}^{2}\end{aligned} \tag{2.4}$$

Of course, given that we have approximated the autoregressive function $\rho(\cdot,\cdot)$ with a polynomial basis of degree two, this partial derivative, for a given value of W_{it}, is necessarily a straight line. As reported by Jantti and Jenkins (2013), many empirical studies as well as our descriptive analysis of the data show that the relationship between the present and the past rank is positive, i.e. individuals who are initially in the highest positions tend to end up in the upper part of the distribution in the following period and vice versa. Consistently with this observation, we find that the values of the partial derivative of the estimated autoregressive function are always greater than zero, thus reflecting the existence of the above-mentioned positive relationship between present and past rank. From Figures 2-4, we notice that the estimated autoregressive function is increasing in the past rank, for any value of the score. This means that the highest present ranks are associated with the highest past ranks, as it could be reasonably expected. From Figure 2 we deduce that, almost regardless of the value of the score, individuals whose past rank is from low to middle tend to improve their position from time $t-1$ to time t. Indeed, the autoregressive function lies above the 45 degree line in all the bottom and middle part of the wage distribution. This effect counterbalances the low-wage trap.

On the other hand, this autoregressive function is decreasing in the score for low values of the past ranks and increasing in the score otherwise (see Figure 3 and 4). This suggests that the effect of age and education level on the probability of improving or worsening one's relative position is heavily influenced by the past position in the distribution, i.e. the dynamics of the Gaussian ranks exhibit a rather high degree of state dependence. From Figure 3 we notice that the score seems to have only a slight influence on the shape of the autoregressive function. However, the role of the score in determining variations in the shape of the autoregressive function is somewhat larger at the top of the wage distribution.

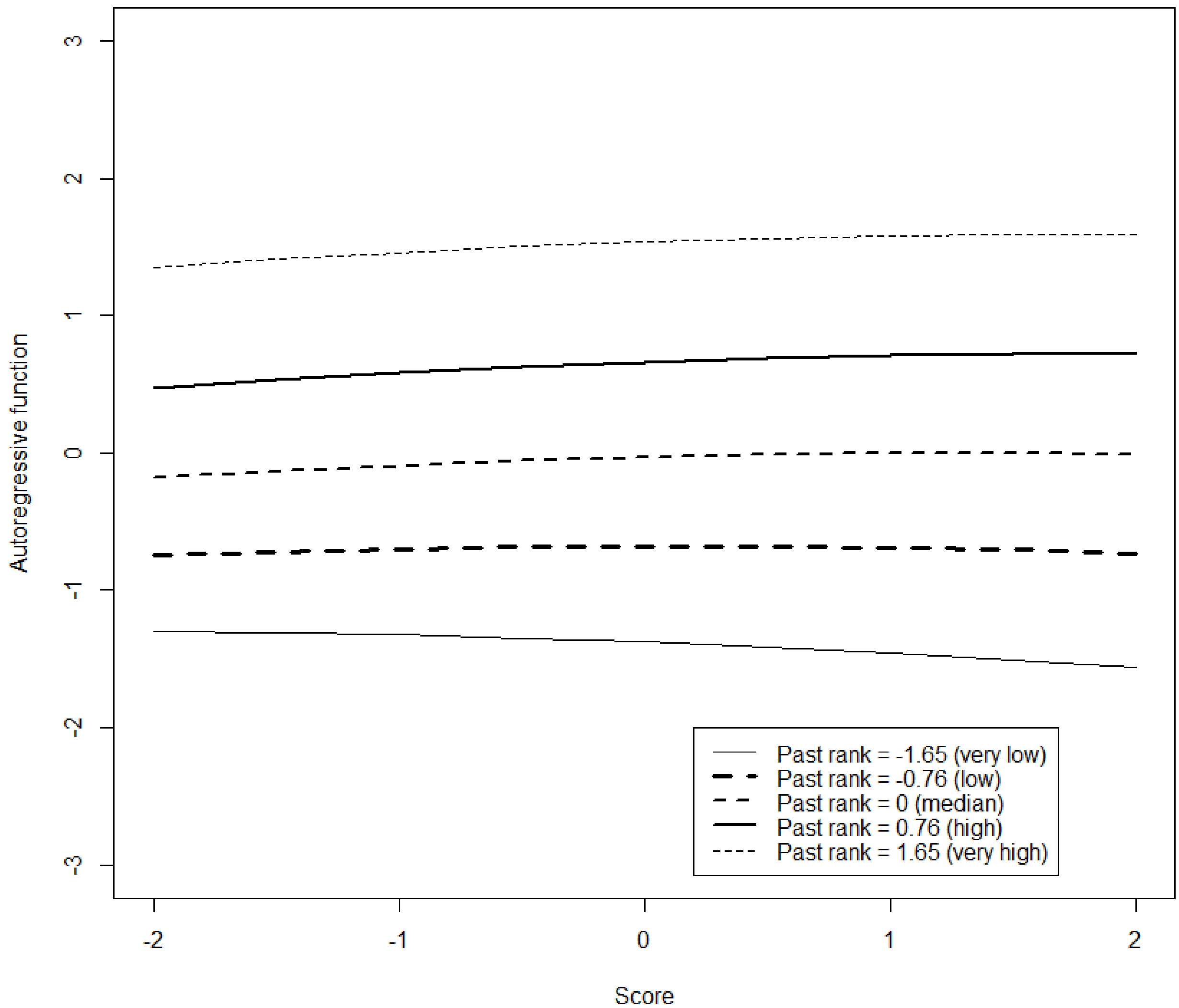
However, age and education may have a countervailing effect on the shape of the autoregressive function and hence, by aggregating these variables in a single index, it is difficult to disentangle the effect of each of them. In the following, we compute the partial derivatives of this function for different groups of individuals, in order to identify more precisely the effects of age and those of education on wage mobility.

Figure 2: Estimated unconstrained autoregressive function for different values of the score



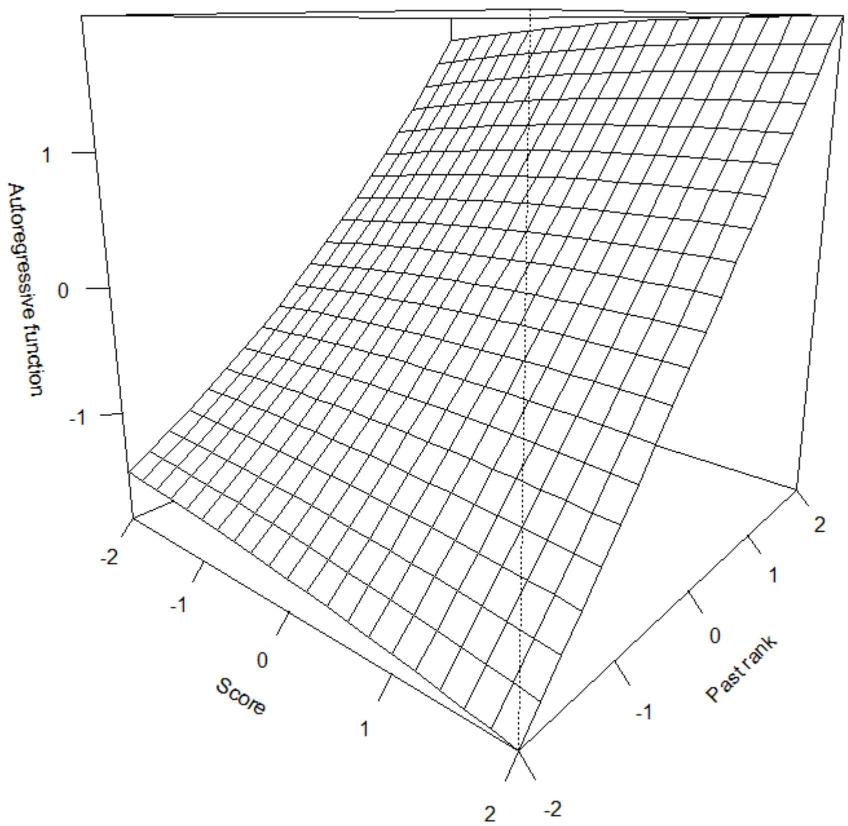
In this figure we display the value of the estimated autoregressive function $\rho(Z_{i,t-1}, W_{it})$, as a function of past rank $Z_{i,t-1}$,for different values of the score W_{it}, which correspond to quantile 25%, 50% and 75% of the score distribution.

Figure 3: Estimated unconstrained autoregressive function for different values of the initial rank



In this figure we display the value of the estimated autoregressive function $\rho(Z_{i,t-1}, W_{it})$ as a function of the score W_{it}, for different values of the past rank, which correspond to percentile 5%, 25%, 50%, 75% and 95%.

Figure 4: Unconstrained estimate of the function $\rho(Z_{i,t-1}, W_{i,t})$



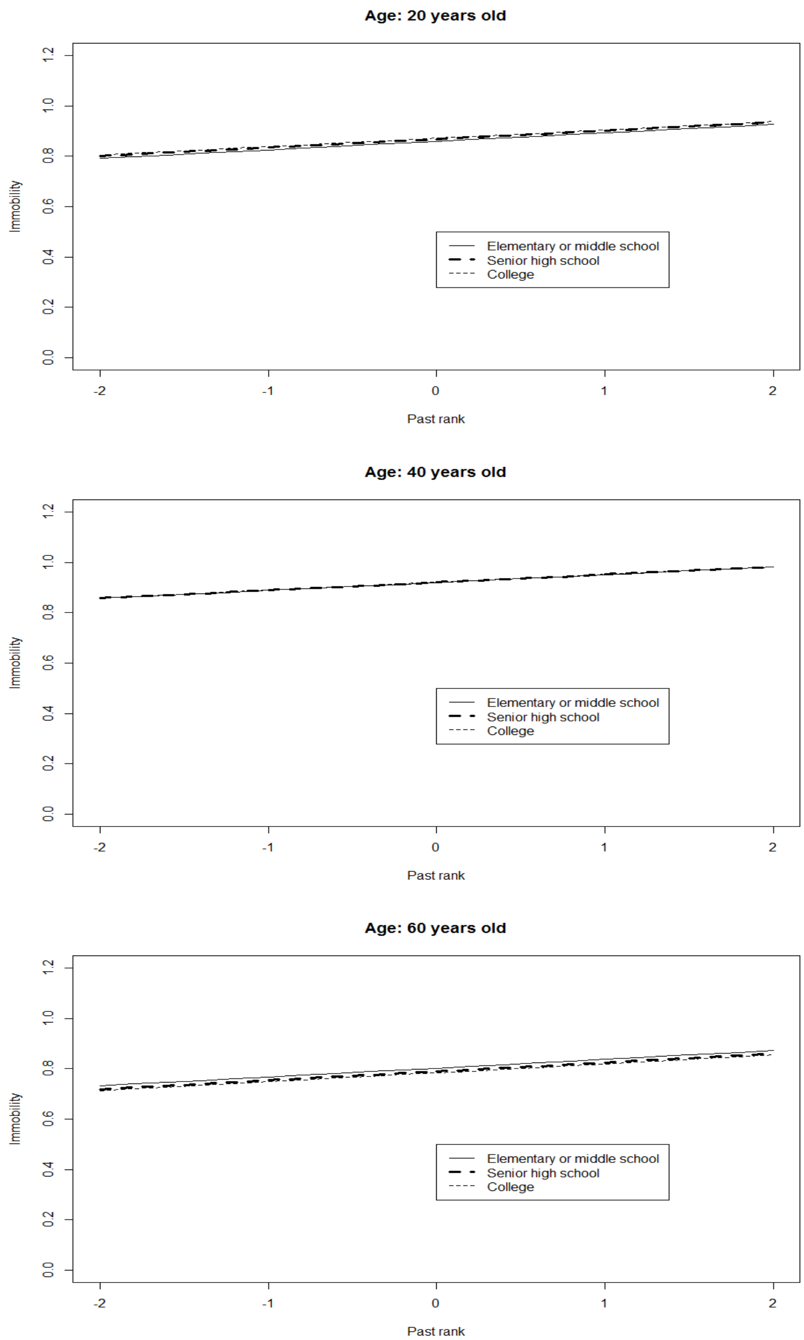
This figure shows the three-dimensional representation of the autoregressive function. This function has been approximated with a Hermite basis of degree 2.

As mentioned above, in order to be able to draw more precise conclusions on the different degrees of wage mobility for people of different ages and education levels, we resort to our mobility measure, i.e. the partial derivative of the autoregressive function with respect to the past position in the distribution. As it can be observed from Figure 5, for any education level this partial derivative is increasing in the past rank for people in any age group (from 20 to 60). Given that this partial derivative stands for positional persistence, i.e. immobility, this means that the degree of mobility is lower at the top of the distribution than at the bottom of it for workers belonging to any age group in our sample. This constitutes a hint that in our sample there is no low-wage trap, i.e. workers being at the bottom of the distribution exhibit lower positional persistence than those at the top and hence should be able to improve their relative position over time. In general, we find that the mobility pattern for people in their 20s and for those in their 60s are rather close, people in their 60s being only slightly more mobile at the bottom of the wage scale with respect to their 20-years-old colleagues. On the other hand, middle-aged workers (in their 40s) appear less mobile than both their younger and their older colleagues in all parts of the distribution. This suggest that they are more prone to the risk of being stuck in a low-pay trap at the bottom of the wage distribution.

Note that for all the age groups and education levels, the degree of rank persistence is rather high, with the value of the partial derivative of our autoregressive function ranging from around 0.7 to around 1. This reflect a rigid labor market, in which the current position in the wage distribution is largely determined by the past position of the individual and education only plays a small role. Indeed, in general, we observe that there is no clear monotonically increasing or decreasing relationship between the highest education level achieved and the degree of positional mobility. We find that the mobility patterns for people of the same age who completed different education levels almost fully overlap. Hence, it seems that age and the past position in the rank distribution play a far more important role than education in determining the degree of individual positional persistence. We find that the upper ranks are generally characterized by a higher level of positional immobility, i.e. once reached a high position, it is likely that the worker will maintain it. To summarize, our unconstrained estimates provide indication that education does not play a clear role in determining the degree of positional mobility. On the other side, we find some indication that people who are in their 20s or in their 60s are more likely to escape the low-wage trap than their middle-aged colleagues.

Note that no constraint has been imposed on estimation. As a consequence, the model estimated in this section (Equation (2.3)) does not necessarily respect the constraint that the present rank must be marginally distributed as a standard normal: $Z_{i,t} \sim N(0,1)$. The aim of this section, indeed, is just to get some preliminary insights on the relationship between the present and the past rank. In the following section, instead, we will provide a complete framework for the estimation of a constrained autoregressive function.

Figure 5: Partial derivatives of the unconstrained autoregressive function w.r.t. past rank, for different education levels



In each panel, we represent the pattern of the partial derivative $\frac{\partial \rho}{\partial Z_{i,t-1}}$ as a function of the past rank $Z_{i,t-1}$, for individuals in a given age class, across three different education levels. The partial derivative $\frac{\partial \rho}{\partial Z_{i,t-1}}$ is a measure of positional immobility, i.e. larger values of this partial derivative imply more positional persistence.

2.3 The model

We are interested in the analysis of the dynamics of the joint process $(Z_{i,t}, W_{i,t})$, where $Z_{i,t}$ stands for the cross-sectional Gaussian individual rank in the wage distribution and $W_{i,t}$ is the (univariate) score, i.e. a weighted sum of the individual explanatory variables, X_{it}.

2.3.1 The general framework

In the following, N stands for the total number of individuals in the sample.

Assumption 1. *The processes* $\{(Z_{it}, W_{it}), t \in \mathbb{N}\}, i = 1, ..., N$ *are i.i.d. across individuals.*

Let us define $Z_t = (Z_{1,t}, ..., Z_{N,t})$. We want to specify a (semi-nonparametric) model for the joint distribution of the Gaussian rank Z_{it} and of the (standardized) score W_{it}. The sample density is:

$$l(\underline{Z_T}, \underline{W_T}) = \prod_{i=1}^{N} l(\underline{Z_{i,T}}, \underline{W_{i,T}}) = \prod_{i=1}^{N} \prod_{t=1}^{T} l(Z_{i,t}, W_{i,t} | \underline{Z_{i,t-1}}, \underline{W_{i,t-1}})$$

where $\underline{Z_{i,t-1}} = (Z_{i,t-1}, Z_{i,t-2}, Z_{i,t-3}, ...)$ and $\underline{Z_T} = (Z_T, Z_{T-1}, Z_{T-2}, ...)$.
In the remaining of this section, we focus on modelling the conditional density $l(Z_{i,t}, W_{i,t} | \underline{Z_{i,t-1}}, \underline{W_{i,t-1}})$. Let us now consider the following decomposition:

$$l(Z_{it}, W_{it} | \underline{Z_{i,t-1}}, \underline{W_{i,t-1}}) = l(W_{it} | \underline{Z_{i,t-1}}, \underline{W_{i,t-1}}) \cdot l(Z_{it} | \underline{Z_{i,t-1}}, \underline{W_{it}}) \tag{2.5}$$

The distribution of the process (W_{it}, Z_{it}) is thus characterized by the two conditional densities, which are the transition density of the rank given the score history:

$$l(Z_{it} | \underline{W_{it}}, \underline{Z_{i,t-1}})$$

and the transition density of the score given the past ranks:

$$l(W_{it} | \underline{W_{i,t-1}}, \underline{Z_{i,t-1}}) \tag{2.6}$$

We deploy the following assumptions on these conditional densities:

Assumption 2. *Process* (Z_{it}) *does not Granger cause process* (W_{it})*, for any* i*.*

Granger non-causality implies Sims non-causality (see e.g. Gourieroux and Monfort (1995) Property 1.2).

Assumption 2 implies that the conditional density in Equation (2.6) is such that:

$$l(W_{it}|\underline{W_{i,t-1}}, \underline{Z_{i,t-1}}) = l(W_{it}|\underline{W_{i,t-1}}) \tag{2.7}$$

Hence, under Assumption 2, the past ranks do not affect the score dynamics, i.e. the individual explanatory variables included in the score are exogenous.

Assumption 3. *Process* (W_{it}) *is Markovian of first-order with transition density* $l(W_{it}|W_{i,t-1})$ *and is strictly stationary.*

The first-order Markov property implies that Equation (2.7) can be further rewritten as:

$$l(W_{it}|\underline{W_{i,t-1}}, \underline{Z_{i,t-1}}) = l(W_{it}|W_{i,t-1}). \tag{2.8}$$

To ensure stationarity of the score, only strictly stationary explanatory variables X_{it} are included in it.

Assumption 4. *The rank dynamics is such that:* $l(Z_{it}|\underline{W_{it}}, \underline{Z_{i,t-1}}) = l(Z_{it}|W_{it}, W_{i,t-1}, Z_{i,t-1})$*.*

This means that information about rank and explanatory variables occurring before time $t-1$ is not relevant in determining the present rank Z_{it}.

Moreover, from Equations (2.5), (2.8) and Assumption 4 we get:

$$l(Z_{it}, W_{it}|\underline{Z_{i,t-1}}, \underline{W_{i,t-1}}) = l(W_{it}|W_{i,t-1}) \cdot l(Z_{it}|Z_{i,t-1}, W_{it}, W_{i,t-1}). \tag{2.9}$$

As a consequence, (W_{it}, Z_{it}) is first-order Markov. Equation (2.9) provides functional restrictions on the specifications of the transition density in our model[12]. The two transition densities

[12] Note that we allow the conditional law of the present rank in Assumption 4 to depend also on the past value of the score, $W_{i,t-1}$ and not only on its present value W_{it}. However, in the following we will assume that the autoregressive function which links the present and the past rank only depends on the present value of the score, W_{it}. Otherwise, the framework would become far too complicated in order to accommodate our copula model that will be presented in the following Section.

of interest are the following:

$$l(Z_{it}|Z_{i,t-1}, W_{it}, W_{i,t-1}) \tag{2.10}$$

and

$$l(W_{it}|W_{i,t-1}).$$

The second of these two transition laws is exogenously given, hence we will exclusively focus on the first one (equation (2.10)). These two laws define a stationary process for (Z_{it}, W_{it}), which will be characterized by a certain joint stationary law (implied by the model), whose density we denote by $l(Z_{it}, W_{it})$. In order for the model to be consistent with the interpretation of Z_{it} as a Gaussian rank, we need to ensure that the stationary density of Z_{it} is standard normal:

$$Z_{it} \sim N(0, 1). \tag{2.11}$$

In the following we prove that, under Assumptions 1-4 and additional constraints on transition density (2.10), equation (2.11) holds.

2.3.2 A copula model

In this subsection we introduce a nonparametric specification for transition density (2.10) based on a copula model. Given Assumption 1, for explanatory purpose, we omit the subscript i in the following. Let $c(\cdot, \cdot, \rho)$ be a copula function that is indexed by the possibly functional parameter ρ. Let us suppose that $g(\cdot; \theta), \theta \in \Theta$, is a parametric family of probability density functions (pdf), and that $G(\cdot; \theta), \theta \in \Theta$ is the corresponding family of cdf. Let $l(W_t)$ be the stationary distribution of W_t. We make parameter depend on variable W as a function $\theta(W)$ such that:

$$\int g(Z;\theta(W))l(W)dW = \phi(Z), \tag{2.12}$$

where ϕ is the standard normal pdf. Let function c be a copula pdf, and define:

$$l(Z_t|Z_{t-1}, W_t, W_{t-1}) = g(Z_t;\theta(W_t))c[G(Z_t;\theta(W_t)), G(Z_{t-1};\theta(W_{t-1}));\rho(\cdot, W_t)], \tag{2.13}$$

where $c(\cdot;\cdot;\rho)$ is a generic copula density, where $\rho(\cdot, W_t)$ is the function indexing the copula. This is a valid density function since it integrates to 1 w.r.t. argument Z_t and $\int c(u,v)du = 1$. To prove our main result for this Section we use the next instrumental lemma.

Lemma 1 Under Assumptions 1-4, if:

$$l(Z_{t-1}|W_{t-1}) = g(Z_{t-1};\theta(W_{t-1})) \tag{2.14}$$

then the following holds:

$$l(Z_t|W_t) = g(Z_t;\theta(W_t))$$

Proof. We have:

$$\begin{aligned} l(Z_t|W_t) &= \int\int l(Z_t|Z_{t-1}, W_{t-1}, W_t)l(Z_{t-1}, W_{t-1}|W_t)dZ_{t-1}dW_{t-1} \\ &= \int\int l(Z_t|Z_{t-1}, W_{t-1}, W_t)l(Z_{t-1}|W_{t-1}, W_t)l(W_{t-1}|W_t)dZ_{t-1}dW_{t-1} \end{aligned}$$

Then, by Lemma 5 in Appendix E, the absence of Granger causality in Assumption 2 implies that $l(Z_{t-1}|W_t, W_{t-1}) = l(Z_{t-1}|W_{t-1})$. Then:

$$l(Z_t|W_t) = \int\int l(Z_t|Z_{t-1}, W_{t-1}, W_t)l(Z_{t-1}|W_{t-1})l(W_{t-1}|W_t)dZ_{t-1}dW_{t-1}.$$

Now, by replacing the definition of the conditional density in (2.13), and in (2.14), we get:

$$l(Z_t|W_t) = \int\int g(Z_t;\theta(W_t))c[G(Z_t;\theta(W_t)), G(Z_{t-1};\theta(W_{t-1})); \rho(\cdot, W_t)]\cdot \\ g(Z_{t-1};\theta(W_{t-1}))\cdot l(W_{t-1}|W_t)dZ_{t-1}dW_{t-1} \tag{2.15}$$

We can further simplify the r.h.s. of (2.15) by using

$$\int c(u,v)du = 1 \quad \forall v. \tag{2.16}$$

Hence, after a change of variable we obtain:

$$\begin{aligned} l((Z_t|W_t) &= \int\int g(Z_t;\theta(W_t))c[G(Z_t;\theta(W_t)), v]l(W_{t-1}|W_t)dvdW_{t-1} \\ &= \int g(Z_t;\theta(W_t))l(W_{t-1}|W_t)dW_{t-1} \\ &= g(Z_t;\theta(W_t)) \end{aligned}$$

This concludes the proof.

$\square$

We come to the main result of this Section.

Proposition 1. *Assume that the initial distribution is such that:*

$$Z_0|W_0 \sim g(\cdot;\theta(W_0)) \tag{2.17}$$

Then, under Assumptions 1-4:

$$l(Z_t) = \phi(Z_t)$$

for all t.

Note that we do not specify the distribution of W_t, since we aim at deriving a result that holds for any W_t. The only requirement that we impose here is that W_t is stationary.

Proof. Let us write the marginal density of Z_t as follows:

$$\begin{aligned} l(Z_t) &= \int\int\int l(Z_t|Z_{t-1}, W_t, W_{t-1}) l(Z_{t-1}, W_t, W_{t-1}) dZ_{t-1} dW_t dW_{t-1} \\ &= \int\int\int l(Z_t|Z_{t-1}, W_t, W_{t-1}) l(Z_{t-1}|W_t W_{t-1}) l(W_t, W_{t-1}) dZ_{t-1} dW_t dW_{t-1} \\ &= \int\int\int l(Z_t|Z_{t-1}, W_t, W_{t-1}) l(Z_{t-1}|W_{t-1}) l(W_t, W_{t-1}) dZ_{t-1} dW_t dW_{t-1}, \end{aligned}$$

where we use the absence of Granger causality in Assumption 2. We want to verify that the following condition is satisfied:

$$l(Z_t) = \int\int\int l(Z_t|Z_{t-1}, W_t, W_{t-1}) l(W_t, W_{t-1}) l(Z_{t-1}|W_{t-1}) dZ_{t-1} dW_t dW_{t-1}$$

Now we insert our expression for the conditional density of the present rank from (2.13). By (2.13) and recursive application of Lemma 1 we have $l(Z_{t-1}|W_{t-1}) = g(Z_{t-1}; \theta(W_{t-1}))$ for all t. Note that from equation (2.13), and $l(Z_{t-1}|W_t, W_{t-1}) = g(Z_{t-1}; \theta(W_{t-1}))$ from Granger non-causality and initial condition (2.17) and Lemma 1, the joint density of (Z_t, Z_{t-1}) given (W_t, W_{t-1}) is

$$l(Z_t, Z_{t-1}|W_t, W_{t-1}) = g(Z_t; \theta(W_t)) g(Z_{t-1}; \theta(W_{t-1})) \cdot c[G(Z_t; \theta(W_t)), G(Z_{t-1}; \theta(W_{t-1})); \rho(\cdot, W_t)],$$

we deduce that $c(\cdot, \cdot, \rho(\cdot, W_t))$ is the copula of Z_t, Z_{t-1} conditional on W_t, W_{t-1}.
Hence, we get:

$$\begin{aligned} l(Z_t) = \int\int\int g(Z_t; \theta(W_t)) \cdot c[G(Z_t; \theta(W_t)), G(Z_{t-1}; \theta(W_{t-1})); \rho(\cdot, W_t)] \cdot \\ g(Z_{t-1}; \theta(W_{t-1})) l(W_t, W_{t-1}) dZ_{t-1} dW_t dW_{t-1}. \end{aligned}$$

We now apply a change of variable:

$$G(Z_{t-1}; \theta(W_{t-1})) = v$$

and we get:

$$l(Z_t) = \int \int g(Z_t;\theta(W_t)) \left(\int c[G(Z_t;\theta(W_t)), v; \rho(\cdot, W_t)] dv \right) l(W_t, W_{t-1}) dW_t dW_{t-1}$$

The expression in the round brackets in the formula above is equal to one for the copula property reported in Equation (2.16), and thus:

$$l(Z_t) = \int g(Z_t;\theta(W_t)) \left(\int l(W_t, W_{t-1}) dW_{t-1} \right) dW_t = \int g(Z_t;\theta(W_t)) l(W_t) dW_t = \phi(Z_t)$$

By (2.12), which is exactly what we wanted to show.

□

Proposition 1 shows that, if the Gaussian rank process is initialized for a conditional distribution $g(\cdot|\theta(W_0))$ satisfying property (2.12), and the transition density is as in (2.13), then the condition of standard Gaussian marginal distribution for the rank process is met. We introduced a model which is compatible with the condition of standard gaussianity of the ranks in a very general framework; it applies to any copula family, possibly with a functional parameter. This functional parameter, in turn, is allowed to depend on W_t.

2.3.3 The autoregressive family of copulas

In the previous section we have shown how to specify a joint dynamics of rank Z_t and score W_t, by means of a generic copula function that can be indexed by a functional parameter. In this section we introduce a flexible nonparametric family of copula functions to be used in this setting. These copula functions are inspired by nonlinear first-order autoregressive processes. Let us consider the nonlinear autoregressive dynamics[13]:

$$Z_t = \Lambda(\rho(Z_{t-1}) + \varepsilon_t) \tag{2.18}$$

[13]The model presented in this Section differs from that presented in Section 2. Indeed, here we specify a constrained model, such that the stationary distribution of the ranks is standard normal. In Section 2, instead, we just run a preliminary unconstrained model as a form of explorative data analysis, without making any assumption on the laws of the ranks or of the score.

where by hypothesis $\varepsilon_t \sim IIN(0,1)$, Λ is a strictly monotonic function and ρ is a function that expresses the dependence between the past and the present individual ranks, where Z_t is rank, the larger is the partial derivative of the function $\rho(\cdot)$ with respect to the past rank, the higher the degree of positional persistence. Equation (2.18) defines a time-homogeneous Markov process. We assume that ρ is such that Z_t is a strictly stationary process with unique invariant distribution. Let us now derive the constraints to impose on functions Λ and ρ such that the invariant distribution of Markov process (Z_t) is $N(0,1)$.

Proposition 2. *In the process defined by equation* (2.18)*, the invariant distribution of Z_t is standard normal if, and only if, the function $\Lambda(\cdot)$ is given by the following expression:*

$$\Lambda(k) = \Phi^{-1}\left[\int_{-\infty}^{\infty} \Phi(k-\rho(Z_{t-1}))\phi(Z_{t-1})dZ_{t-1}\right], \tag{2.19}$$

for all k.

Proof. By the law of iterated expectations we have:

$$\begin{aligned}
P(Z_t \le z) &= E[P(Z_t \le z|Z_{t-1})] \\
&= E\left[P\{\Lambda(\rho(Z_{t-1})+\varepsilon_t) \le z|Z_{t-1}\}\right] \\
&= E\left[P\{\varepsilon_t \le \Lambda^{-1}(z)-\rho(Z_{t-1})|Z_{t-1}\}\right] \\
&= \int_{-\infty}^{\infty} \Phi[\Lambda^{-1}(z)-\rho(Z_{t-1})]\phi(Z_{t-1})dZ_{t-1},
\end{aligned}$$

where Φ is the cdf of the standard normal distribution. Therefore, the following condition is imposed:

$$\int_{-\infty}^{\infty} \Phi(\Lambda^{-1}(z)-\rho(Z_{t-1}))\phi(Z_{t-1})dZ_{t-1} = \Phi(z).$$

By applying now a change of variable: $\Lambda^{-1}(z) = k \Leftrightarrow z = \Lambda(k)$, we get the following restriction, i.e. that $\Lambda(k)$ is such that

$$\Phi(\Lambda(k)) = \int_{-\infty}^{\infty} \Phi(k-\rho(Z_{t-1}))\phi(Z_{t-1})dZ_{t-1},$$

i.e. equation (2.19).

□

Thus, function Λ is a functional of function ρ. The integral in the RHS of Equation (2.19) has to be computed numerically by simulation or quadrature. We get the conditional c.d.f. and pdf:

$$P(Z_t \leq z | Z_{t-1} = \zeta) = \Phi[\Lambda^{-1}(z) - \rho(\zeta)],$$

$$f_{Z_t|Z_{t-1}}(z|\zeta) = \phi[\Lambda^{-1}(z) - \rho(\zeta)] \cdot \frac{1}{\lambda(\Lambda^{-1}(z))},$$

where λ is the first-order derivative of the function Λ.

Therefore, we obtain the following joint pdf:

$$f_{Z_t,Z_{t-1}}(z,\zeta) = f_{Z_t|Z_{t-1}}(z|\zeta) \cdot f_{Z_{t-1}}(\zeta) = \phi[\Lambda^{-1}(z) - \rho(\zeta)] \cdot \frac{1}{\lambda(\Lambda^{-1}(z))} \cdot \phi(\zeta).$$

We can now get the explicit copula density of Z_t and Z_{t-1}. The copula pdf is:

$$\begin{aligned} c(u, v; \rho(\cdot)) &= \frac{f_{Z_t,Z_{t-1}}(\Phi^{-1}(u), \Phi^{-1}(v))}{f_{Z_{t-1}}(\Phi^{-1}(v)) \cdot f_{Z_t}(\Phi^{-1}(u))} \\ &= \frac{\phi[\Lambda^{-1}(\Phi^{-1}(u)) - \rho(\Phi^{-1}(v))]}{\phi(\Phi^{-1}(u))\lambda(\Lambda^{-1}(\Phi^{-1}(u)))}, \end{aligned} \tag{2.20}$$

for the arguments $u, v \in [0, 1]$. This copula family is parametrized by the autoregressive function $\rho(\cdot)$. Here, we combined the results from Section 3.2 and 3.3 to obtain our model. The family of copulas represented by equation (2.20) is the one used in equation (2.13). We can now make the functional parameter which indexes the copula depend on W_t.

Note that our results hold for any function $\rho(\cdot)$. Hence, we can easily introduce the score W_t: $\rho(\cdot) = \rho(\cdot, W_t)$ and $c(u, v; \rho(\cdot)) \equiv c(u, v, \rho(\cdot, W_t))$. Finally note that, in order to keep the notation plain, we do not make explicit the dependence of Λ on ρ. However, this dependence is immediate from Proposition 2. It is worth noting that the constrained model thus obtained is not a nested version of the unconstrained model presented in equation (2.3). In the constrained model, indeed, an additional transformation, $\Lambda(\cdot)$, is applied, in order to ensure that the stationary distribution of the ranks is standard normal.

2.3.4 A functional mobility measure

The aim of this section is to introduce a measure of positional mobility. Summarizing the previous section, the model specification is as follows. The distribution of Z_{it} given $Z_{i,t-1}$, W_{it}, $W_{i,t-1}$ is:

$$l(Z_{it}|Z_{i,t-1},W_{it},W_{i,t-1}) = g(Z_{it};\theta(W_{it}))c[G(Z_{it};\theta(W_{it})),G(Z_{i,t-1},\theta(W_{i,t-1}));\rho(\cdot,W_{it})]$$

where $c(\cdot,\cdot;\rho(\cdot))$ is the copula function defined in (2.20), $g(\cdot;\theta(W_{it}))$ is a p.d.f. that satisfies the condition in equation (2.12). Let us define the variable:

$$\xi_{it} \equiv \Phi^{-1}[G(Z_{it};\theta(W_{it}))] \tag{2.21}$$

We write the stochastic representation of our model as follows:

$$\xi_{it} = \Lambda[\rho(\xi_{i,t-1};W_{it}) + \varepsilon_{it}] \tag{2.22}$$

where $\varepsilon_{it} \sim IIN(0,1)$, and $\Lambda(\cdot) = \Lambda(\cdot;W_{it})$ is given by:

$$\Lambda(k;W_{it}) = \Phi^{-1}\left[\int_{-\infty}^{\infty} \Phi[k-\rho(Z,W_{it})]\phi(Z)dZ\right]. \tag{2.23}$$

We now want to derive an adequate measure of positional mobility. We start from the conditional expectation:

$$\begin{aligned} E(Z_{it}|Z_{i,t-1},W_{it},W_{i,t-1}) &= E[G^{-1}(U_{it};\theta(W_{it}))|Z_{i,t-1},W_{it},W_{i,t-1}] \\ &= E[G^{-1}(\tilde{\Lambda}[\tilde{\rho}(U_{i,t-1};W_{it})+\varepsilon_{it}]|Z_{i,t-1},W_{it},W_{i,t-1}] \\ &= \int_{-\infty}^{\infty} G^{-1}\left(\tilde{\Lambda}[\tilde{\rho}(G(Z_{i,t-1};\theta(W_{i,t-1}));W_{it})+\varepsilon]\right)\phi(\varepsilon)d\varepsilon, \end{aligned}$$

where $Z_{it} = G^{-1}(U_{it};\theta(W_{it})) \iff U_{it} = G(Z_{it};\theta(W_{it}))$, and this variable U_{it} follows the stochastic representation $U_{it} = \tilde{\Lambda}[\tilde{\rho}(U_{i,t-1};W_{it}) + \varepsilon_{it}]$ from (2.22), where $\tilde{\rho}(u;W_{it}) = \rho(\Phi^{-1}(u), W_{it})$ and $\tilde{\Lambda}(k;W_{it}) = \int_0^1 \Phi(k - \tilde{\rho}(v;W_{it}))dv$.

So that we can compute the partial derivative with respect to the past rank, which stands for positional mobility:

$$\frac{\partial E(Z_{it}|Z_{i,t-1},W_{it},W_{i,t-1})}{\partial Z_{i,t-1}} = \int_{-\infty}^{\infty} \frac{1}{g\left[G^{-1}\left[\tilde{\Lambda}[\tilde{\rho}(G(Z_{i,t-1};\theta(W_{i,t-1}));W_{it}) + \varepsilon]\right]\right]} \cdot$$
$$\tilde{\Lambda}'[\tilde{\rho}(G(Z_{i,t-1};\theta(W_{i,t-1}));W_{it}) + \varepsilon] \cdot \tilde{\rho}'(G(Z_{i,t-1};\theta(W_{i,t-1}));W_{it}) \cdot g(Z_{i,t-1};\theta(W_{i,t-1})) \cdot \phi(\varepsilon)d\varepsilon. \tag{2.24}$$

This measure corresponds to the partial derivative of the autoregressive function with respect to the past rank (equation (2.4)) in the context of the unconstrained model presented in the previous Section.

2.4 Estimation

We estimate the copula model of Section 2.3, in which mobility is represented by a function of the past rank and of some explanatory variables. As explained in Section 2.3, the distribution of the covariates has been modelled in order to ensure that present and past ranks are admissible copula arguments. The model is estimated in a semi-nonparametric way, with the method of Sieves (see e.g. Chen (2007) for a survey). The main idea of this method, which has been first developed by Grenander (1981) and Geman and Hwang (1982), is to estimate a function by means of the approximating space generated by a set of basis functions.

Similarly to what has been done for the unconstrained estimates presented in Section 2, we use as explanatory variables age, age squared at the beginning of the sample period[14] and a qualitative variable representing the highest education level achieved by the individual (Bonhomme and Robin (2009)). We argue that these dummy variables are exogenous, i.e. that they are not

[14]This is necessary, since all the explanatory variables included in the score have to be stationary over time.

influenced by the individual position in the wage distribution. Indeed, we only consider education that takes place before labor market entry. We do not include among the explanatory variables any form of on-the-job training, due to its potential endogeneity. This means that we are in the particular case in which all of our explanatory variables are constant over time. For the sake of generality, we maintain the focus on a general setting with regressors X_{it} dependent on individual and time in the presentation of the estimation method.

2.4.1 Estimation strategy

Let us define $W_{it} = X'_{it}\beta_1^0$ and:

$$l(Z_{i,t}|Z_{i,t-1}, W_{i,t}, W_{i,t-1}) = g(Z_{i,t}; X'_{i,t}\beta_1^0) \cdot c[G(Z_{i,t}; X'_{i,t}\beta_1^0), G(Z_{i,t-1}; X'_{i,t-1}\beta_1^0), \rho(\cdot, X'_{i,t}\beta_2^0)]$$

where $G(Z_{i,t}; X'_{i,t}\beta_1^0)$ is the c.d.f. of the distribution of the rank, conditional on the individual variables, $g(Z_{i,t}; X'_{i,t}\beta_1^0)$ is the corresponding p.d.f., $\beta_1^0 \in \mathbb{R}^{p_1}$ and $\beta_2^0 \in \mathbb{R}^{p_2}$ are parameter vectors, and $c[\cdot, \cdot, \rho(\cdot)]$ is the copula density in (2.20).

We need to estimate both the univariate distributions of the ranks and the joint one. Let us first consider the estimation of the univariate conditional distribution $G(Z_{i,t}; X'_{i,t}\beta_1^0)$. The use of a linear combination $X'_{it}\beta_1^0$ of the explanatory variables corresponds to a semiparametric single-index model. We estimate the coefficient vector β_1^0 with a kernel single-index Maximum Likelihood approach:

$$\hat{\beta}_1 = \arg\max_{\beta_1 \in B_1} \sum_{i=1}^{N} \sum_{t=1}^{T} k_{it} \log\ \hat{g}_{-(i,t)}(Z_{i,t}|\beta_1' X_{it}; \beta_1), \tag{2.25}$$

where $B_1 \subset \mathbb{R}^{p_1}$ is the parameter set, $\hat{g}_{-(i,t)}(z|w; \beta_1^0)$ is the leave-one-out conditional kernel density of Z_{it} given $W_{it}(\beta_1^0) = w$, where $W_{it}(\beta_1^0) = X'_{it}\beta_1^0$ and

$$\hat{g}_{-(i,t)}(z|w;\beta_1^0) = \frac{\frac{1}{h}\sum_{s=1,s\neq t}^{T}\sum_{j=1,j\neq i}^{N} K\left(\frac{Z_{jt}-z}{h}\right) K\left(\frac{W_{jt}(\beta_1^0)-w}{h}\right)}{\sum_{s=1,s\neq t}^{T}\sum_{j=1,j\neq i}^{N} K\left(\frac{W_{jt}(\beta_1^0)-w}{h}\right)} \tag{2.26}$$

where K is a kernel density that satisfies Assumption 5 in Appendix C and h is the bandwidth or smoothing parameter that satisfies Assumption 6 in Appendix C. Moreover, k_{it} is the trimming term introduced by Rosemarin (2012) [15]. The maximization in (2.25) is performed by a Newton-Raphson algorithm, starting from an initial guess for parameter β_1^0. This initial guess has been obtained via an iterated maximum likelihood procedure, which is described in detail in Appendix B. With this maximization, we obtain an estimator of β_1^0 that is consistent for $N \to \infty$ and T fixed (see Section 2.4.2). By plugging in this estimate of β_1^0, we are now able to compute the desired estimated cdfs:

$$\tilde{G}(Y_{it},\hat{f}) = \int_{-\infty}^{Z_{it}} \hat{f}(Z|X_{it})dZ, \tag{2.27}$$

where $Y_{it} = (Z_{it}, X'_{it})'$, $\hat{f}(Z|X) = \hat{g}(Z|\hat{\beta}'_1 X)$ and:

$$\hat{g}(z|\hat{\beta}'_1 x) = \frac{\frac{1}{h}\sum_{t=1}^{T}\sum_{i=1}^{N} K\left(\frac{Z_{it}-z}{h}\right) K\left(\frac{\hat{\beta}'_1(X_{it}-x)}{h}\right)}{\sum_{t=1}^{T}\sum_{i=1}^{N} K\left(\frac{\hat{\beta}'_1(X_{it}-x)}{h}\right)}. \tag{2.28}$$

We now estimate the copula function, $c(\cdot,\cdot,\rho(\cdot,X'_{i,t}\beta_2^0))$. We perform a simultaneous M-

---

[15] This trimming term is defined as follows: $k_{it} = \mathbb{I}_{it} \times \left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{I}_{it}\right)^{-1}$, where $\mathbb{I}_{it} = 1$ if $min\left\{\frac{1}{h}\sum_{s=1,s\neq t}^{T}\sum_{j=1,j\neq i}^{N} K\left(\frac{Z_{jt}-z}{h}\right) K\left(\frac{W_{jt}(\beta_1^0)-w}{h}\right), \sum_{s=1,s\neq t}^{T}\sum_{j=1,j\neq i}^{N} K\left(\frac{W_{jt}(\beta_1^0)-w}{h}\right)\right\} >$ $a_0 n^{-c}$ and zero otherwise, for some small constants $a_0, c > 0$. Thus, the indicator $\mathbb{I}_{it}$ is normalized to account for the actual number of observations considered in the computation of the log-likelihood.

estimation of the parameters β_2^0 and of the autoregressive function, in the spirit of e.g. Wong and Severini (1991) and Chen and Shen (1998). This estimation is performed via the following Sieve Maximum Likelihood procedure:

$$\hat{\theta} = \arg\max_{\theta \in \Theta_N} \sum_{i=1}^{N} l(Y_i, \theta, \hat{f}) \tag{2.29}$$

where:

$$l(Y_i, \theta, f) \equiv \sum_{t=1}^{T} \log c[\tilde{G}(Y_{it}, f), \tilde{G}(Y_{i,t-1}, f); \rho(\cdot|X_{it}'\beta_2)], \tag{2.30}$$

$Y_i = (Z_{i,1}, X_{i,1}, ..., Z_{i,T}, X_{i,T})$, and $\tilde{G}(Y_{it}, f) = \int_{-\infty}^{Z_{it}} f(Z|X_{it})dZ$ is the conditional c.d.f. of Z_{it} given $X_{it}'\beta_1^0$, and f is the parameter governing the univariate distribution, and $\hat{f}(z|x) = \hat{g}(z|\hat{\beta}_1'x)$ is the estimator defined by equations (2.25), (2.27) and (2.28), $\theta = (\beta_2, \rho(\cdot|\cdot)) \in B_2 \times \mathbb{H} \equiv \Theta$, $\Theta_N = B_2 \times \mathbb{H}_N$ where $\beta_2 \in B_2 \subset \mathbb{R}^{p_2}$, is the parameter governing the copula. Note that θ includes both a finite-dimensional and an infinite-dimensional component. $\mathbb{H}$ is an infinite-dimensional space of bivariate functions and $\mathbb{H}_N$ is a bivariate sieve space (made up by Hermite polynomials in our implementation) whose dimension depends on sample size N. More specifically, the approximation $\rho \in \mathbb{H}_N$ is

$$\rho(Z_{i,t-1}, W_{it}) = \sum_{k,l=0}^{m} \lambda_{k,l} H_k(W_{it}) H_l(Z_{i,t-1}) \tag{2.31}$$

where λ is the vector of coefficients of the polynomial basis used to approximate $\rho(\cdot)$. The number of polynomials used to approximate the autoregressive function depends on the dimension of the sample: $m = m(N)$ (Chen (2007)).

2.4.2 Consistency of the estimators

We adopt the standard "micro" panel asymptotics with T fixed and N going to infinity. As for the convergence of our estimated univariate cdfs to the true ones, we suppose that Assumptions

1-4 in Section 3.1 and 5-7 in Appendix C hold. By building on the results of Rosemarin (2012), who extended the work by Delecroix et al. (2003) on semi-nonparametric estimation of index models, the following theorem holds:

Theorem 1. *Under Assumptions 1-7: (i) The estimator $\hat{\beta}_1$ defined in equation* (2.25) *is consistent:*

$$\hat{\beta}_1 \xrightarrow[N\to\infty]{p} \beta_1^0.$$

(ii) The estimator $\hat{g}$ defined in (2.28) *satisfies:*

$$\sup_{(z,x)\in\mathbb{S}} |\hat{g}(z|\hat{\beta}_1' x) - g(z|\beta_1^{0\prime} x)| = O_p\left(\left(\frac{\log N}{Nh^2}\right)^{1/2}\right),$$

where $\mathbb{S} \subset \mathbb{R}^{p_1+1}$ is a compact set introduced to control boundary effects.
(iii) $\sqrt{N}(\hat{\beta}_1 - \beta_1^0) \xrightarrow[N\to\infty]{d} N(0, V(\beta_1^0))$, where $V(\beta_1^0) = \Omega(\beta_1^0)^{-}\Psi(\beta_1^0)\Omega(\beta_1^0)^{-}$, $\Psi(\beta_1^0) = E_{\mathbb{S}_\delta}[\nabla_{\beta_1} \log g(Z_{it}|\beta_1^{0\prime} X_{it})\nabla_{\beta_1} \log g(Z_{it}|\beta_1^{0\prime} X_{it})']$, $\Omega(\beta_1^0)^{-}$ is the generalized inverse of $\Omega(\beta_1^0)$, and $\Omega(\beta_1^0) = E_{\mathbb{S}_\delta}[-\nabla^2_{\beta_1\beta_1} \log g(Z_{it}|\beta_1^{0\prime} X_{it})]$.

Let us now consider the estimation of the copula parameter. We suppose that Assumptions 8-10, which are reported in Appendix C, hold. In particular, we consider a norm $||\cdot||_\Theta$ on parameter set Θ which satisfies the requirements of Assumption 8. Hence, the following Theorem, which directly follows from Theorem 2.1 in Hahn et al. (2016), is valid:

Theorem 2. *If Assumptions 1-10 hold, then the second-step Sieve M-estimator $\hat{\theta}_N$ defined in* (2.29) *is consistent:*

$$||\hat{\theta}_N - \theta_0||_\Theta \xrightarrow[N\to\infty]{p} 0.$$

2.4.3 Empirical results

We apply the above estimation procedure to the sample of US data described in Section 2. Analogously to what has been done for the unconstrained estimates, the polynomial basis used to approximate the functional parameter $\rho(\cdot|\cdot)$ in (2.31) is a Hermite basis of degree $m = 2$. We also used another polynomial basis, e.g. Hermite polynomials of degree $m = 3$, and the estimation results did not significantly change. In particular, the estimated coefficients of the terms of order higher than 2 were very close to zero. For our implementation, the kernel K is the standard Gaussian pdf. In the following, we focus on the estimated copula parameters, i.e. vector β_2 and function $\rho(\cdot|\cdot)$. From the new estimated coefficients for the variable score, we deduce that the latter is increasing in age and also increasing in the education level (with the only exception of graduate school, $Edu5_{it}$). This finding is completely consistent with what we found in Section 2 by performing the unconstrained estimation procedure:

$$\begin{aligned}\hat{W}_{it} = -0.0004Age_{it} + 0.0048Age_{it}^2 + Edu1_{it} + 1.0155Edu2_{it} + 1.4691Edu3_{it}\\ +1.6757Edu4_{it} + 1.3786Edu5_{it}.\end{aligned}$$

The constrained estimate of the autoregressive function is the following[16]:

$$\begin{aligned}\rho(Z_{i,t-1}, \hat{W}_{i,t}) = 0.3441 + 0.1296\hat{W}_{i,t} + 1.7348Z_{i,t-1} - 0.1548\hat{W}_{i,t}Z_{i,t-1}+\\ -0.0496(\hat{W}_{i,t}^{\,2} - 1) - 0.1211(Z_{i,t-1}^2 - 1) + 0.1076(Z_{i,t-1}^2 - 1)\hat{W}_{i,t}\\ +0.0702(\hat{W}_{i,t}^{\,2} - 1)Z_{i,t-1} - 0.0345(\hat{W}_{i,t}^{\,2} - 1)(Z_{i,t-1}^2 - 1)\end{aligned} \tag{2.32}$$

In Figure 6, we display the autoregressive function (2.32). It appears that, for any value of the score, this function is monotonically increasing in the past rank and it takes the highest values for people with high levels of the score and high past ranks. On the other side, we notice that the estimated autoregressive function is also (weakly) increasing in the score for

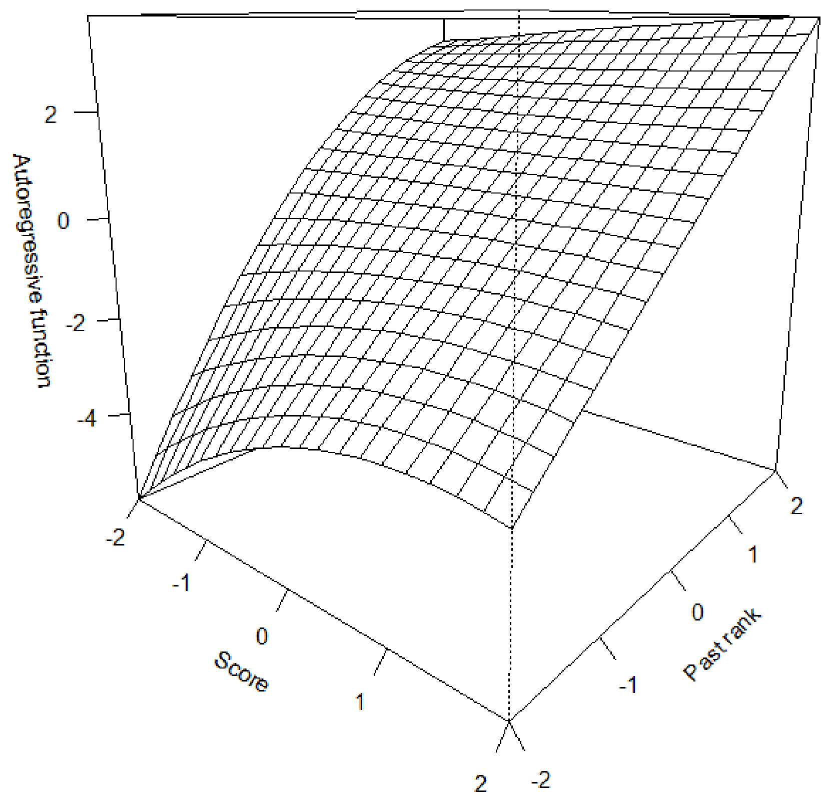
[16]Note that the score $\hat{W}_{it}$ is centered and standardized so that it is approximately distributed as a standard normal variable.

all the past ranks. The shape of the autoregressive function is similar to that found with the unconstrained estimation procedure. However, with respect to the unconstrained estimates, we notice that in the constrained case the score has a far larger role in determining the present rank, in particular for low values of the past rank. Figures 7 and 8 show the sections of the constrained autoregressive function, for constant score and constant past rank, respectively.

In Figure 7, the slope of the estimated function $\rho(Z_{it}, W_{it})$ is in some cases greater than 1. However, it is more informative to consider the composed autoregressive function $\rho \circ \Lambda(\tilde{U}_{i,t-1}, W_{i,t})$, whose sections are reported in Figure 9, which refers to the additive autoregressive representation of our model (see Section 2.3.4). From Figure 9, we see that the slope of this composed function is always less than 1[17].

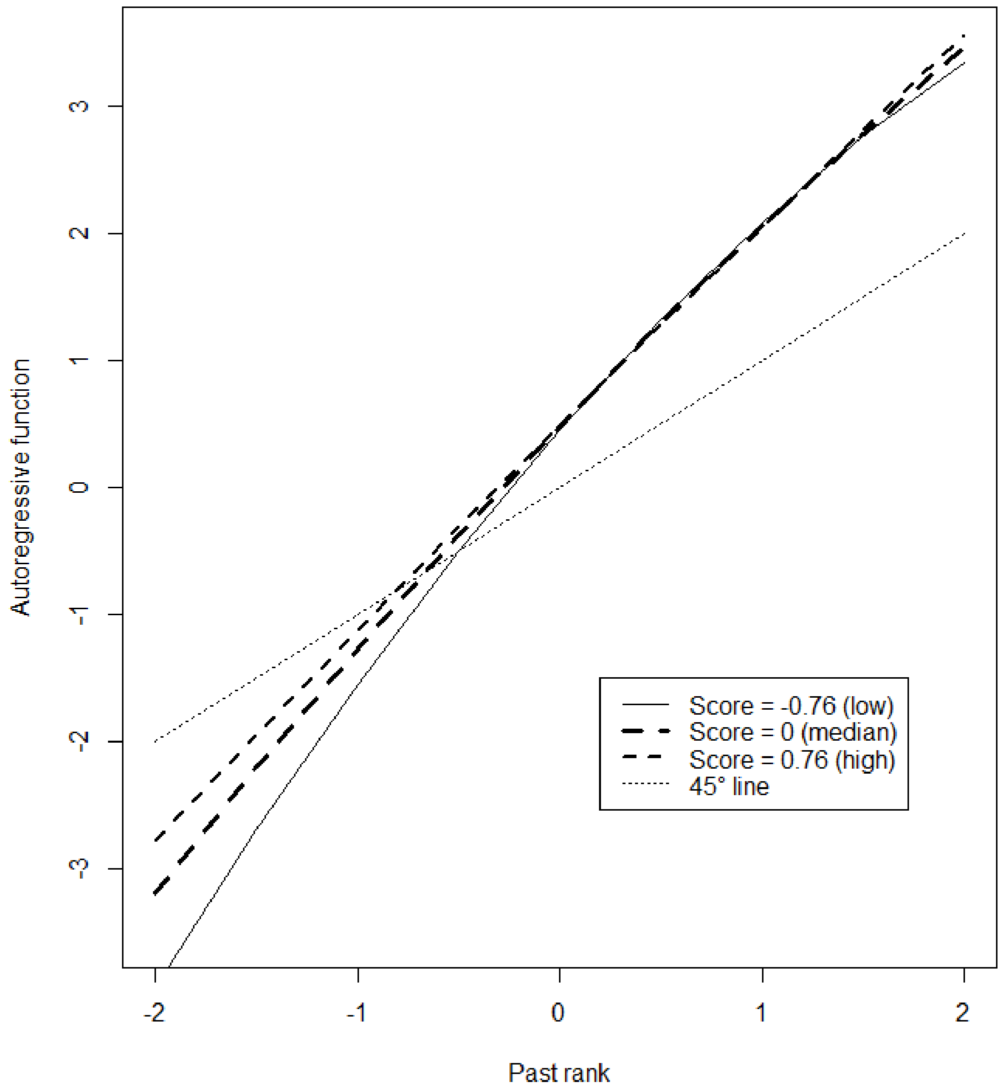
[17]The study of the stochastic properties of our model, included stationarity, is beyond the scope of the present paper.

Figure 6: Constrained estimate of the function $\rho(Z_{i,t-1}, W_{i,t})$



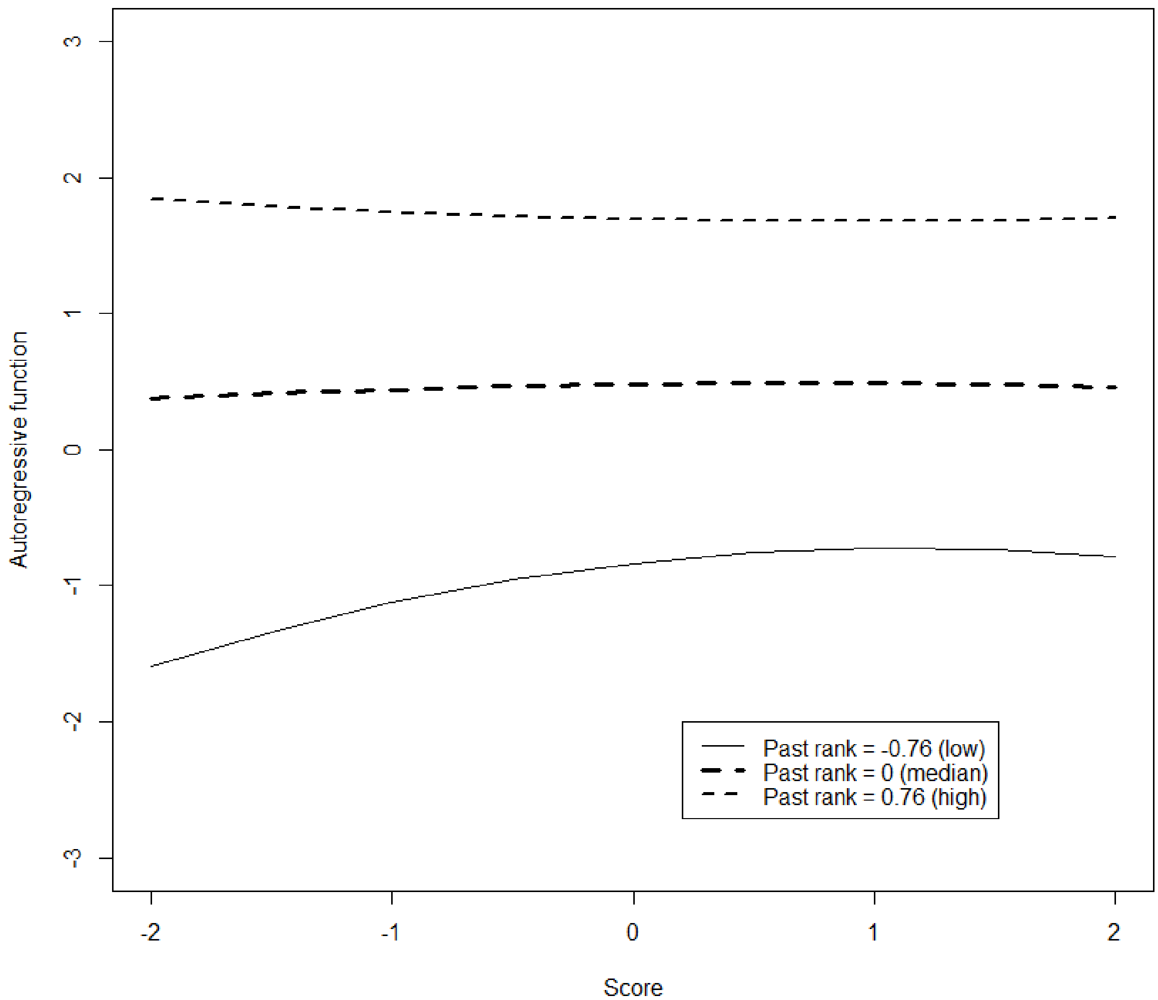
This figure shows the three-dimensional representation of the autoregressive function. This function has been approximated with a Hermite basis of degree 2.

Figure 7: Estimated constrained function $\rho(Z_{it}, W_{it})$ for different values of the score



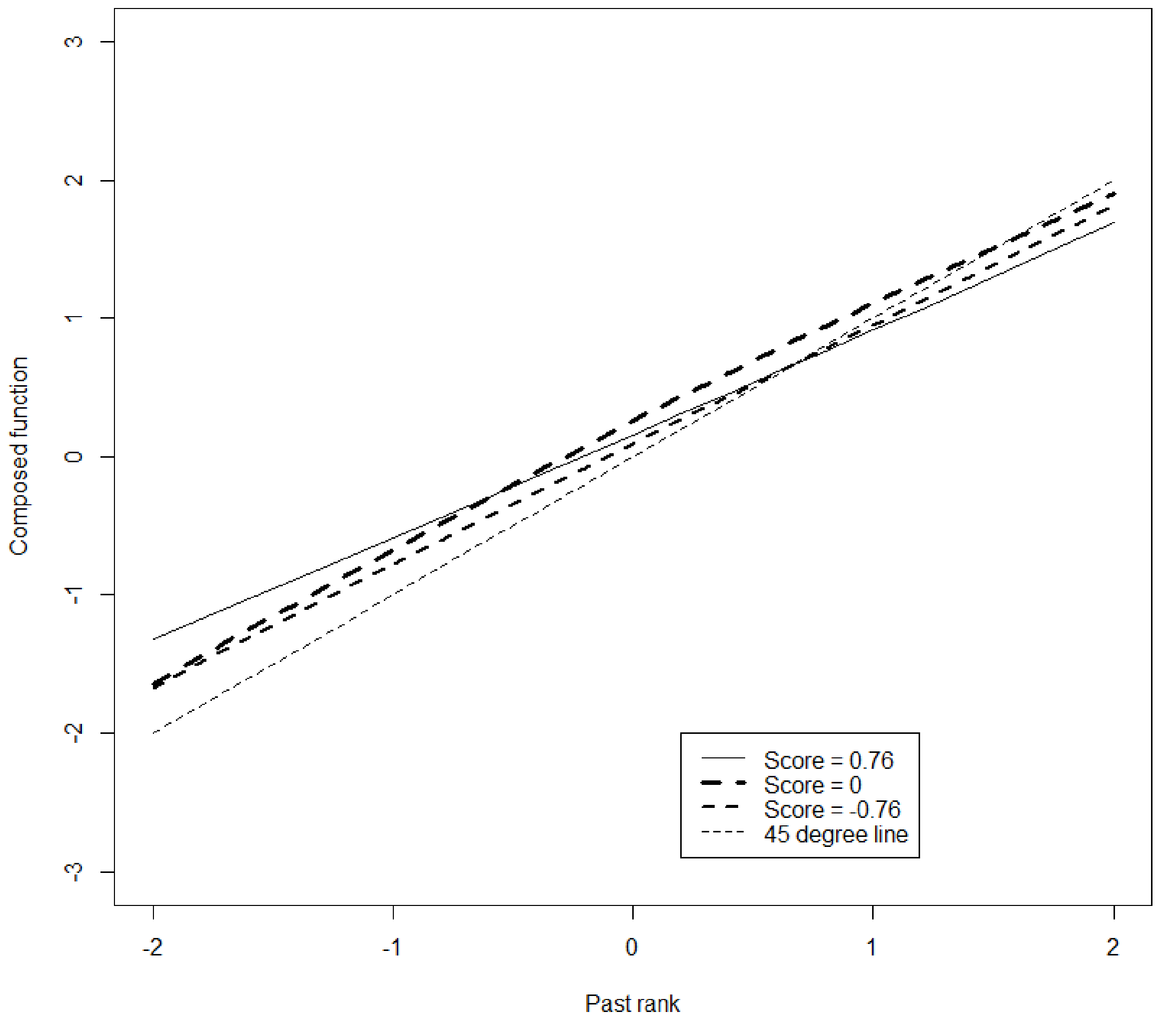
In this figure we display the value of the estimated function $\rho(Z_{i,t-1}, W_{it})$, as a function of past rank $Z_{i,t-1}$,for different values of the score W_{it}, which correspond to quantile 25%, 50% and 75% of the score distribution.

Figure 8: Estimated constrained function $\rho(Z_{it}, W_{it})$ for different values of the initial rank



In this figure we display the value of the estimated function $\rho(Z_{i,t-1}, W_{it})$ as a function of the score W_{it}, for different values of the past rank, which correspond to percentile 25%, 50% and 75%.

Figure 9: Estimated constrained autoregressive function for different values of the score



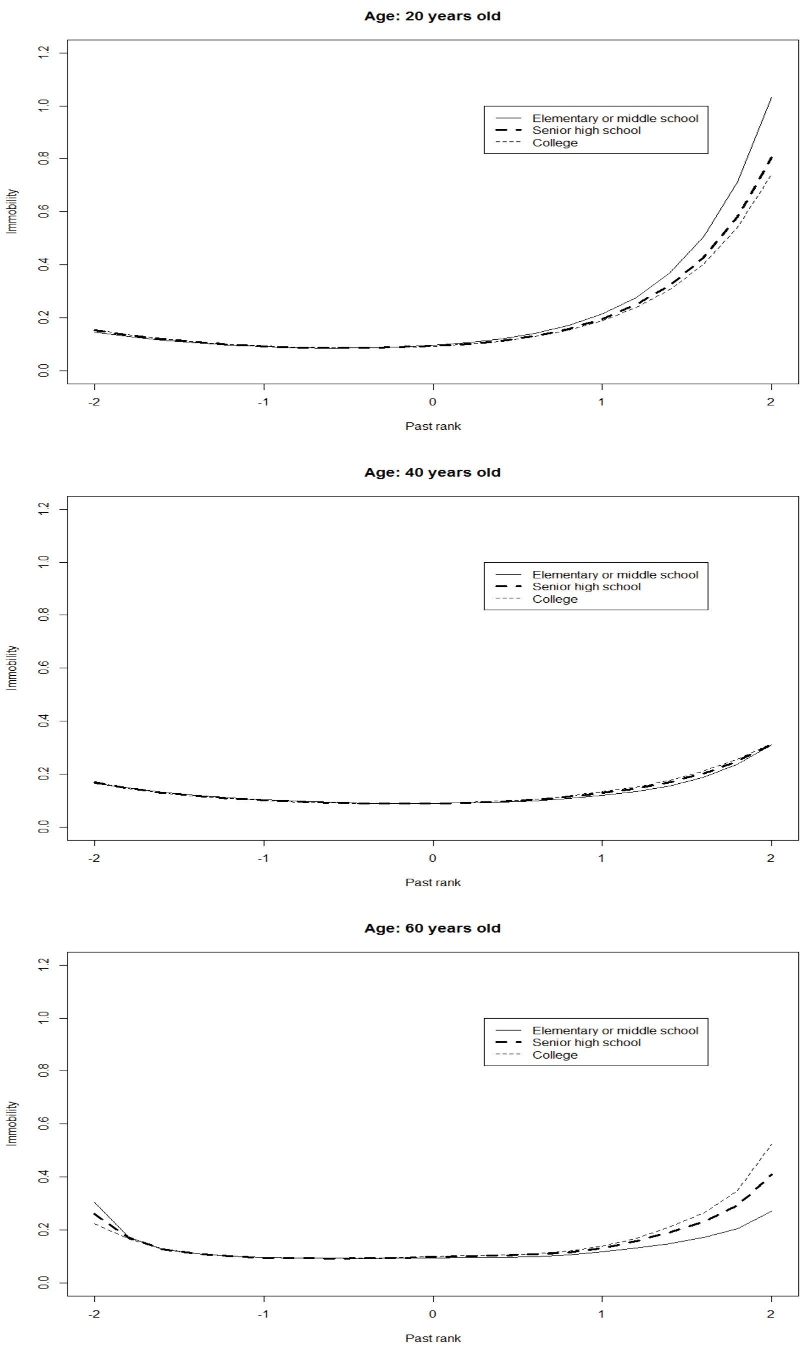
In this figure we display the value of the estimated autoregressive function $\rho \circ \Lambda(\tilde{U}_{i,t-1}, W_{i,t})$, for different values of the score W_{it}, which correspond to quantile 25%, 50% and 75% of the score distribution.

As explained before, the score is influenced by both individual age and the highest education level achieved. We now want to analyze the estimated patterns of our mobility measure, as defined by equation (3.10), which are represented in Figure 10. From Figure 10, we deduce that positional mobility patterns do not seem to significantly change according to the highest education level achieved by the individual. These findings suggest that, in the US labor market for the period considered, experience or other individual factors play a more relevant role than education in determining the degree of individual positional mobility. For each education level, the mobility patterns of workers in their 40s and in their 60s almost fully overlap. We notice that workers in their 60s are less mobile in the upper part of the distribution if they have completed either senior high school or college. This hints at the fact that downward mobility for older workers is likely to be reduced by higher levels of formal education achieved.

Note that, in Figure 10, the y-axis records the value of our mobility measure, as defined by equation (3.10). The higher this quantity is, the higher is the association between the present and the past rank and hence the lower the degree of positional mobility is. This is the reason why the y-axis is labelled as ”immobility”. A remarkable difference with respect to what the unconstrained estimation results suggested (Figure 5) consists in the fact that young workers exhibit significantly higher mobility levels in the bottom and middle part of the distribution, thus suggesting that they are not subject to the low-pay trap; this is true regardless of their education level.

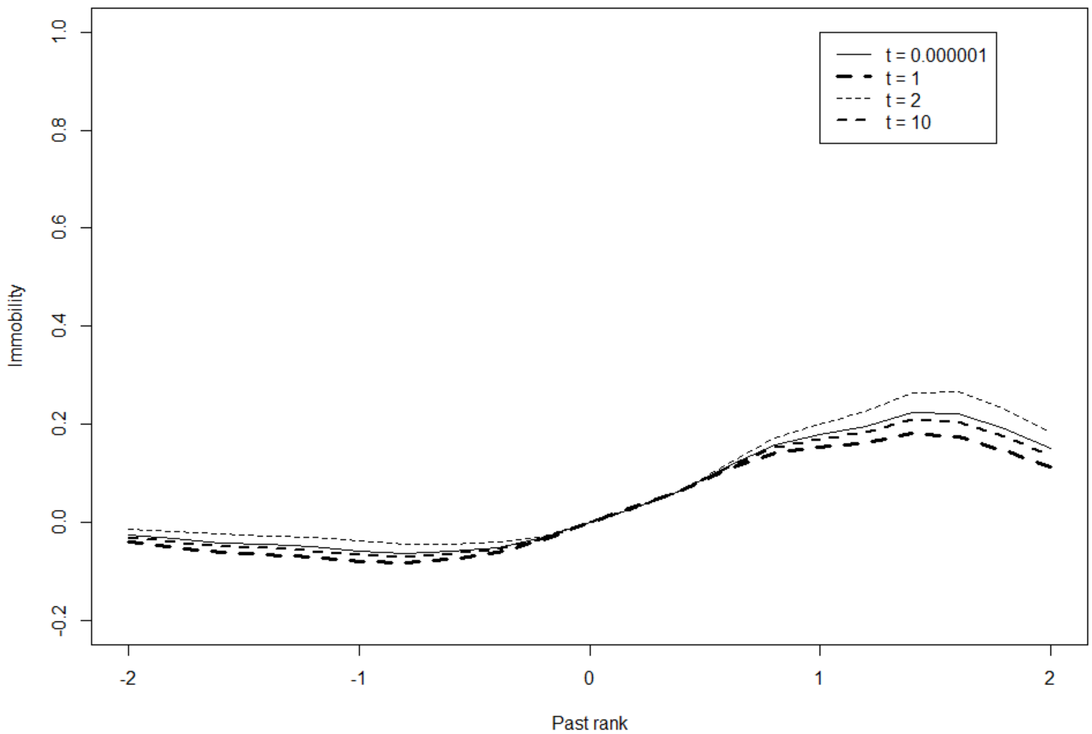
To summarize, we find almost no indication of the existence of a low-wage trap for any age group of workers in the US labor market. Regardless of their education, indeed, people in their 20s and older show a rather high degree of mobility in the lowest half of the distribution. On the contrary, they exhibit a high degree of positional persistence at the top of the distribution. This finding suggests that they are not exposed to the risk of downward positional mobility. For people in their 40s and in their 60s, too, mobility is generally decreasing with the past position. We found that positional mobility is not heavily influenced by the highest education level completed by the individual. Also in Bonhomme and Robin (2009), indeed, the estimated coefficients of the impact of the different education dummies on the mobility parameter were rather close to each other. To conclude, Figure 11 shows our mobility measure in the case of a fully parametric Plackett copula, the one used by Bonhomme and Robin (2009), which can be compared to our constrained mobility measure. Both with our model and with a fully parametric copula, it emerges that positional mobility is higher in the middle of the distribution and lower at the extremes of it, and in particular that there is a higher degree of immobility at the top of the wage distribution. We notice that, by changing this unique copula parameter, the concavity/convexity of this function is only slightly altered. Therefore, we conclude that a unique parameter is not enough to control how mobility changes across the whole wage distribution. Indeed, by comparing Figures 11 and 12, we deduce that our functional model allows for more flexibility and, hence, for a more realistic and precise assessment of positional mobility in different parts of the wage distribution. We also perform the same computations with different parametric copulas and the results are similar, i.e. by changing the unique copula parameter it is not possible to precisely determine the mobility pattern in the different parts of the rank distribution (see Figure 18 in Appendix A for another example).

Figure 10: Mobility patterns for different education levels



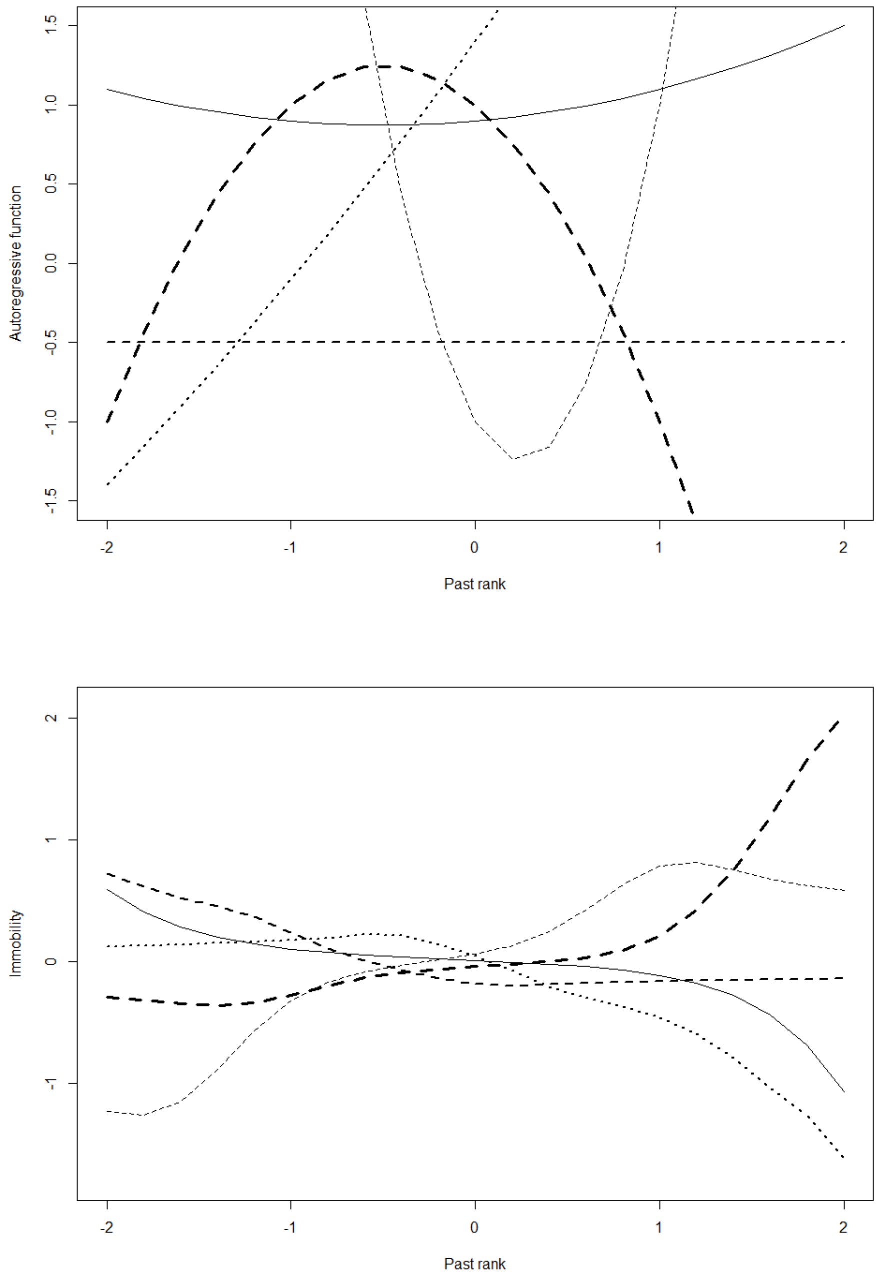
In this figure, each panel represents the pattern of our mobility measure, as defined by equation (3.10), for workers of the same age and different education level.

Figure 11: Partial derivative of the Plackett copula for different values of the copula parameter t



This figure shows how positional mobility, as it can be derived by the Plackett copula used by Bonhomme and Robin (2009), changes across the distribution. The measure of mobility represented is: $\frac{\partial}{\partial Z_{i,t-1}} E(Z_{i,t}|Z_{i,t-1}, W_{i,t}, W_{i,t-1}) = \int_0^1 G^{-1}(u;\theta(W_{it}))\frac{\partial c}{\partial v}(u, G(Z_{i,t-1};\theta(W_{i,t-1})) \cdot g(Z_{i,t-1};\theta(W_{i,t-1}))du$, where $c(u,v)$ is the Plackett copula.

Figure 12: Autoregressive function $\rho(\cdot;\cdot)$ (upper panel) and the associated mobility measure (bottom panel) for different configurations of the coefficients of the polynomial basis



The upper panel of this figure shows the pattern of the autoregressive function $\rho(\cdot;\cdot)$ for a fixed value of the score W_{it} and different configurations of the coefficients of the polinomial basis used for the approximation (Hermite polynomials of degree 2). The bottom panel shows the pattern of the mobility measure defined by our equation (3.10) for each parameter configuration represented in the upper panel.

2.5 Concluding remarks

As explained by Schumpeter (1955), the wage scale can be compared to a hotel, in which there are both luxury rooms and cheaper rooms. At any point in time, all the rooms are occupied by some guest; however, they do not always occupy the same room. The study of wage mobility aims at understanding how frequently do guests "switch rooms" and which are the factors underlying this process. In this paper we presented a new model for the wage rank dynamics.
With a new autoregressive copula model, which allows greater flexibility than a fully parametric one, we estimate the autoregressive function which links the present and the past Gaussian ranks, conditional on some individual characteristics. We get evidence that, in the US labor market, there is a rather high degree of wage mobility for both young and old individuals, initially being at the bottom of the wage scale. We find no evidence of workers of any age or education level being stuck in the so called low-wage trap.
A limitation of the present work is that we rely on the missing-at-random assumption for our unbalanced panel data. Moreover, in the present paper we only considered age and the education level achieved by the individual as explanatory variables; however, it would be interesting to also control for other variables such as gender, or the presence of a migration background of the worker. This constitutes scope for future research. As a future outlook, this constrained model may be used to simulate individual wage trajectories. Another promising extension would be to re-write the model in order to allow also for transitions into and out of unemployment.

References

1. Altonji, J. G., Smith, A. A., & Vidangos, I. (2013). Modeling earnings dynamics. Econometrica, 81(4), 1395-1454.

2. Arellano, M., Blundell, R., & Bonhomme, S. (2017). Earnings and consumption dynamics: a nonlinear panel data framework. Econometrica, 85(3), 693-734.

3. Arellano, M., & Bonhomme, S. (2017). Quantile Selection Models with an Application to Understanding Changes in Wage Inequality. Econometrica, 85(1), 1-28.

4. Atkinson, A. B., Bourguignon, F., and Morrisson, C. (1988). Earnings mobility. European Economic Review, 32(2), 619-632.

5. Auten, G., Gee, G., & Turner, N. (2013). Income Inequality, Mobility, and Turnover at the Top in the US, 1987–2010. The American Economic Review, 103(3), 168-172.

6. Auten, G., and Gee, G. (2009). Income mobility in the United States: New evidence from income tax data. National Tax Journal, 301-328.

7. Autor, D. H., Katz, L. F., & Kearney, M. S. (2006). The Polarization of the US Labor Market. The American Economic Review, 96(2), 189-194.

8. Bachmann, R., Bechara, P., and Schaffner, S. (2014). Wage inequality and wage mobility in Europe. Review of Income and Wealth.

9. Behrman, J.,& Taubman, P. (1985). Intergenerational earnings mobility in the United States: some estimates and a test of Becker's intergenerational endowments model. The Review of Economics and Statistics, 144-151.

10. Bonhomme, S., & Robin, J. M. (2010). Generalized non-parametric deconvolution with an application to earnings dynamics. The Review of Economic Studies, 77(2), 491-533.

11. Bonhomme, S., and Robin, J. M. (2009). Assessing the equalizing force of mobility using short panels: France, 1990-2000. The Review of Economic Studies, 76(1), 63-92.

12. Bratsberg, B., Roed, K., Raaum, O., Naylor, R., & Eriksson, T. (2007). Nonlinearities in Intergenerational Earnings Mobility: Consequences for Cross-Country Comparisons. The Economic Journal, 117(519), C72-C92.

13. Broyden, C. G. (1970). The convergence of single-rank quasi-Newton methods. Mathematics of Computation, 24(110), 365-382.

14. Buchinsky, M., and Hunt, J. (1999). Wage mobility in the United States. Review of Economics and Statistics, 81(3), 351-368.

15. Burkhauser, R. V., Holtz-Eakin, D., and Rhody, S. E. (1997). Labor Earnings Mobility and Inequality in the United States and Germany during the Growth Years of the 1980s. International Economic Review, 775-794.

16. Cardoso, A. R. (2006). Wage mobility: do institutions make a difference?. Labour economics, 13(3), 387-404.

17. Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. Handbook of econometrics, 6, 5549-5632.

18. Chen, X., Liao, Z., & Sun, Y. (2014). Sieve inference on possibly misspecified semi-nonparametric time series models. Journal of Econometrics, 178, 639-658.

19. Chen, X., & Pouzo, D. (2012). Estimation of nonparametric conditional moment models with possibly nonsmooth generalized residuals. Econometrica, 80(1), 277-321.

20. Chen, X., Koenker, R., & Xiao, Z. (2009). Copula-based nonlinear quantile autoregression. The Econometrics Journal, 12(s1), S50-S67.

21. Chen, X., & Fan, Y. (2006). Estimation of copula-based semiparametric time series models. Journal of Econometrics, 130(2), 307-335.

22. Chen, X., & Shen, X. (1998). Sieve extremum estimates for weakly dependent data. Econometrica, 289-314.

23. Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. The Quarterly Journal of Economics, 129(4), 1553-1623.

24. Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. The Journal of Economic Perspectives, 27(3), 79-102.

25. David, H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. The American Economic Review, 103(5), 1553-1597.

26. De Coulon, A., and Zuercher, B. A. (2004). Low-Pay Mobility in the Swiss Labour Market. In ”Minimum Wages, Low Pay and Unemployment”, part of the series Applied Econometrics Association Series pp 57-86. Eds. Meulders D., Plasman R., and Rycx F., Palgrave Macmillan UK.

27. Delecroix, M., Hardle, W., & Hristache, M. (2003). Efficient estimation in conditional single-index regression. Journal of Multivariate Analysis, 86(2), 213-226.

28. Dickens, R. (2000). Caught in a trap? Wage mobility in Great Britain: 1975-1994. Economica, 67, 477-197.

29. Dickens, W. T., & Lang, K. (1988). The reemergence of segmented labor market theory. The American Economic Review, 78(2), 129-134.

30. Fachinger, U., and Himmelreicher, R. K. (2012). Income Mobility - Curse or Blessing? Mobility in Social Security Earnings: Data on West-German Men since 1950. Schmollers Jahrbuch: Journal of Applied Social Science Studies, 132(2), 175-203.

31. Ferrie, J. P. (2005). History lessons: The end of American exceptionalism? Mobility in the United States since 1850. The Journal of Economic Perspectives, 19(3), 199-215.

32. Fields, G. S., and Ok, E. A. (1999). The measurement of income mobility: an introduction to the literature (pp. 557-598). Springer Netherlands.

33. Fletcher, R., & Reeves, C. M. (1964). Function minimization by conjugate gradients. The computer journal, 7(2), 149-154.

34. Foote, C. L., & Ryan, R. W. (2014). Labor-Market Polarization over the Business Cycle. In NBER Macroeconomics Annual 2014, Volume 29 (pp. 371-413). University of Chicago Press.

35. Formby, J. P., Smith, W. J., and Zheng, B. (2004). Mobility measurement, transition matrices and statistical inference. Journal of Econometrics, 120(1), 181-205.

36. Friedman, M. (1962). Capitalism and Freedom (Chicago, 1962).

37. Gabaix, X., Lasry, J. M., Lions, P. L., & Moll, B. (2016). The dynamics of inequality. Econometrica, 84(6), 2071-2111.

38. Gagliardini, P., & Gourieroux, C. (2007). An efficient nonparametric estimator for models with nonlinear dependence. Journal of Econometrics, 137(1), 189-229.

39. Geman, S., and Hwang, C. R. (1982). Nonparametric maximum likelihood estimation by the method of sieves. The Annals of Statistics, 401-414.

40. Gittleman, M., and Joyce, M. (1996). Earnings Mobility and Long-Run Inequality: An Analysis Using Matched CPS Data. Industrial Relations: A Journal of Economy and Society, 35(2), 180-196.

41. Goos, M., Manning, A., & Salomons, A. (2009). Job polarization in Europe. The American Economic Review, 99(2), 58-63.

42. Gourieroux, C., and Monfort, A. (1995). Statistics and Econometric Models (Vol. 1). Cambridge University Press.

43. Gottschalk, P. (1997). Inequality, income growth, and mobility: The basic facts. The Journal of Economic Perspectives, 11(2), 21-40.

44. Gottschalk, P. (1982). Earnings mobility: Permanent change or transitory fluctuations?. The Review of Economics and Statistics, 450-456.

45. Gottschalk, P., and Spolaore, E. (2002). On the evaluation of economic mobility. The Review of Economic Studies, 69(1), 191-208.

46. Germandt, J. (2009). Decreasing Wage Mobility in Germany. Discussion Paper no. 09-044. Mannheim Centre for European Economic Research (ZEW).

47. Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica: Journal of the Econometric Society, 424-438.

48. Grenander, U. (1981). Abstract inference. New York: Wiley.

49. Hahn, J., Liao, Z., & Ridder, G. (2016). Nonparametric two-step sieve M estimation and inference.

50. Hall, P., & Yao, Q. (2005). Approximating conditional distribution functions using dimension reduction. Annals of Statistics, 1404-1421.

51. Hansen, B. E. (2008). Uniform convergence rates for kernel estimation with dependent data. Econometric Theory, 24(3), 726-748.

52. Hardle, W., Hall, P., & Ichimura, H. (1993). Optimal smoothing in single-index models. The annals of Statistics, 21(1), 157-178.

53. Hart, P. E. (1983). The size mobility of earnings. Oxford Bulletin of Economics and Statistics, 45(2), 181-193.

54. Hautsch, N., Okhrin, O.,& Ristig, A. (2014). Efficient iterative maximum likelihood estimation of high-parameterized time series models. Working paper

55. Hungerford, T. L. (1993). US income mobility in the seventies and eighties. Review of Income and Wealth, 39(4), 403-417.

56. Ichimura, H. (1987). Estimation of single index models (Doctoral dissertation, Massachusetts Institute of Technology).

57. Ichimura, H., & Lee, S. (2010). Characterization of the asymptotic distribution of semiparametric M-estimators. Journal of Econometrics, 159(2), 252-266.

58. Jäntti, M., & Jenkins, S. P. (2015). Income mobility. In Handbook of income distribution (Vol. 2, pp. 807-935). Elsevier.

59. Joe, H. (1997). Multivariate models and multivariate dependence concepts. CRC Press

60. Kalwij, A. S., and Alessie, R. (2007). Permanent and transitory wages of British men, 1975-2001: Year, age and cohort effects. Journal of Applied Econometrics, 22(6), 1063-1093.

61. Kim, S. (2013). Wage mobility of foreign-born workers in the united states. Journal of Human Resources, 48(3), 628-658.

62. Kopczuk, W., Saez, E., and Song, J. (2010). Earnings inequality and mobility in the United States: evidence from social security data since 1937. The Quarterly Journal of Economics, 125(1), 91-128.

63. Li, Q., & Racine, J. S. (2007). Nonparametric econometrics: theory and practice. Princeton University Press.

64. Moffitt, R. A., and Gottschalk, P. (2002). Trends in the transitory variance of earnings in the United States. The Economic Journal, 112(478), C68-C73.

65. Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. The computer journal, 7(4), 308-313.

66. Nelsen, R. B. (1999). An introduction to copulas (Vol. 139). Springer Science and Business Media.

67. Otto, W. (2005). Measuring income mobility. Przeglad Statystyczny, 52(1), 5-35.

68. Patton, A. J. (2012). A review of copula models for economic time series. Journal of Multivariate Analysis, 110, 4-18.

69. Pavlopoulos, D. et al. 2007. Job Mobility and Wage Mobility of High- and Low-Paid Workers. Schmollers Jahrbuch, 127, 47-58, Berlin.

70. Piketty, T., & Saez, E. (2003). Income Inequality in the United States, 1913-1998. The Quarterly Journal of Economics, 1-39.

71. Piore, M. J. (1975). Notes for a theory of labour market segmentation. R. Edwards et al.

72. Powell, J. L., Stock, J. H., & Stoker, T. M. (1989). Semiparametric estimation of index coefficients. Econometrica: Journal of the Econometric Society, 1403-1430.

73. Reich, M., Gordon, D. M., & Edwards, R. C. (1973). A theory of labor market segmentation. The American Economic Review, 63(2), 359-365.

74. Rosemarin, R. (2012). Semiparametric estimation of single-index conditional densities for dependent data. Unpublished doctoral dissertation.

75. Schiller, B. R. (1977). Relative earnings mobility in the United States. The American Economic Review, 67(5), 926-941.

76. Schluter, C. (1997). On the non-stationarity of German income mobility (and some observations on poverty dynamics). LSE STICERD Research Paper, (30).

77. Schumpeter, J. A. (1955). Imperialism and Social Classes: Two Essays (Vol. 4). Ludwig von Mises Institute.

78. Shorrocks, A. F. (1978). The measurement of mobility. Econometrica: Journal of the Econometric Society, 1013-1024.

79. Sims, C. A. (1972). Money, income, and causality. The American economic review, 540-552.

80. Sklar, A. (1959). Fonctions de repartition a n dimensions et leurs marges. Inst. Statist. Univ. Paris Publ. 8 229-231

81. Solon, G. (1992). Intergenerational income mobility in the United States. The American Economic Review, 393-408.

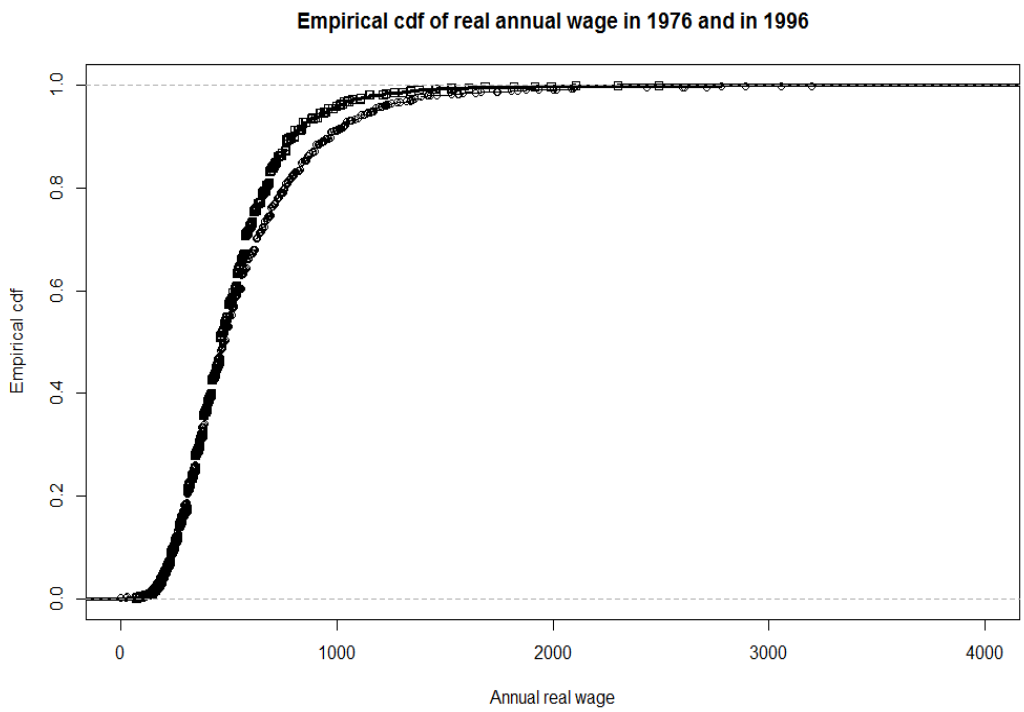
82. Sousa-Poza, A. (2004). Is the Swiss labor market segmented? An analysis using alternative approaches. Labour, 18(1), 131-161.

83. Tansel, A., Dalgic, B., & Guven, A. (2014). "Wage Inequality and Wage Mobility in Turkey." IZA working paper.

84. Wong, W. H., & Severini, T. A. (1991). On maximum likelihood estimation in infinite dimensional parameter spaces. The Annals of Statistics, 603-632.

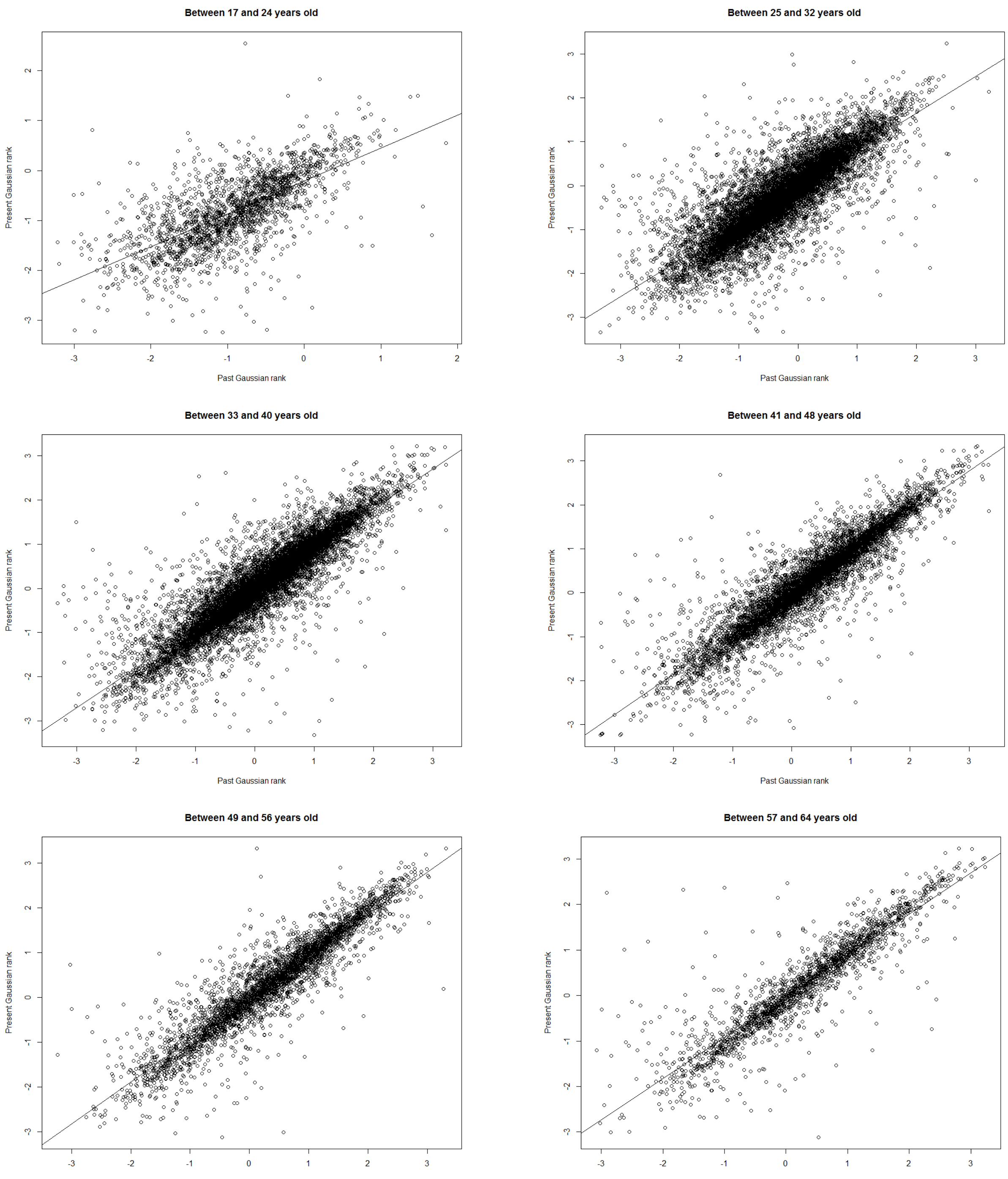
2.6 Appendix A Robustness checks

Figure 13: Evolution of the wage distribution in our sample, from 1976 to 1996



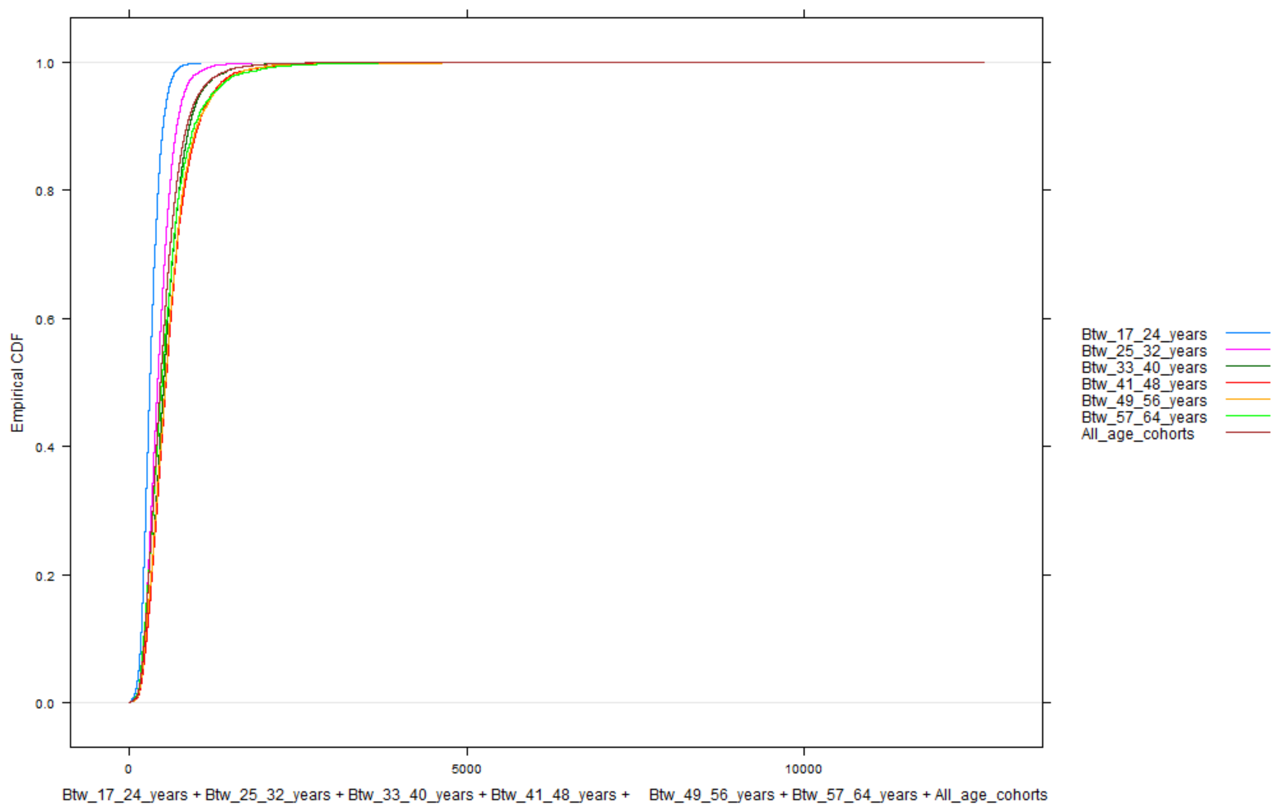
Note: real wages have been computed by dividing nominal wages by the value of the Consumer Price Index in the United States in each of the two years considered.

Figure 14: Relationship between the present and the past Gaussian ranks, for different age cohorts, PSID data, 1975-1996



In this figure, each panel represents the association (linear fit) between the present and the past Gaussian ranks, for each of the six age cohorts in which we divided our sample.

Figure 15: Empirical cdf of real annual wage, by age cohort



Note: real wages have been computed by dividing nominal wages by the value of the Consumer Price Index in the United States in each of the years included in the sample.

Table 2: Descriptive statistics for education and log wage for pooled data and for age cohorts, PSID data, 1975-1996

		17-24	25-32	33-40	41-48	49-56	57-64	Pooled data
Education	average	12.087	12.874	13.001	12.714	11.734	11.002	12.721
(years)	st.deviation	1.607	2.176	2.485	3.033	3.543	3.907	2.681
	min	1	0	0	0	0	0	0
	max	17	17	17	17	17	17	17
	median	12	12	12	12	12	12	12
	25% quantile	12	12	12	12	10	9	12
	75% quantile	12	14	16	16	14	14	15
Log wage	average	9.384	9.823	10.131	10.204	10.068	9.978	10.076
	st.deviation	0.578	0.600	0.622	0.703	0.691	0.756	0.629
	min	2.996	0	0	0	0	0	0
	max	13.665	12.724	13.176	12.995	13.181	13.459	12.969
	median	9.393	9.852	10.167	10.258	10.086	10.021	10.094
	25% quantile	9.210	9.582	9.852	9.888	9.741	9.649	9.680
	75% quantile	9.826	10.258	10.597	10.714	10.592	10.485	10.477

This table reports some descriptive statistics for education and wage, for pooled data and divided by age cohorts, in our PSID sample for the period 1975-1996.

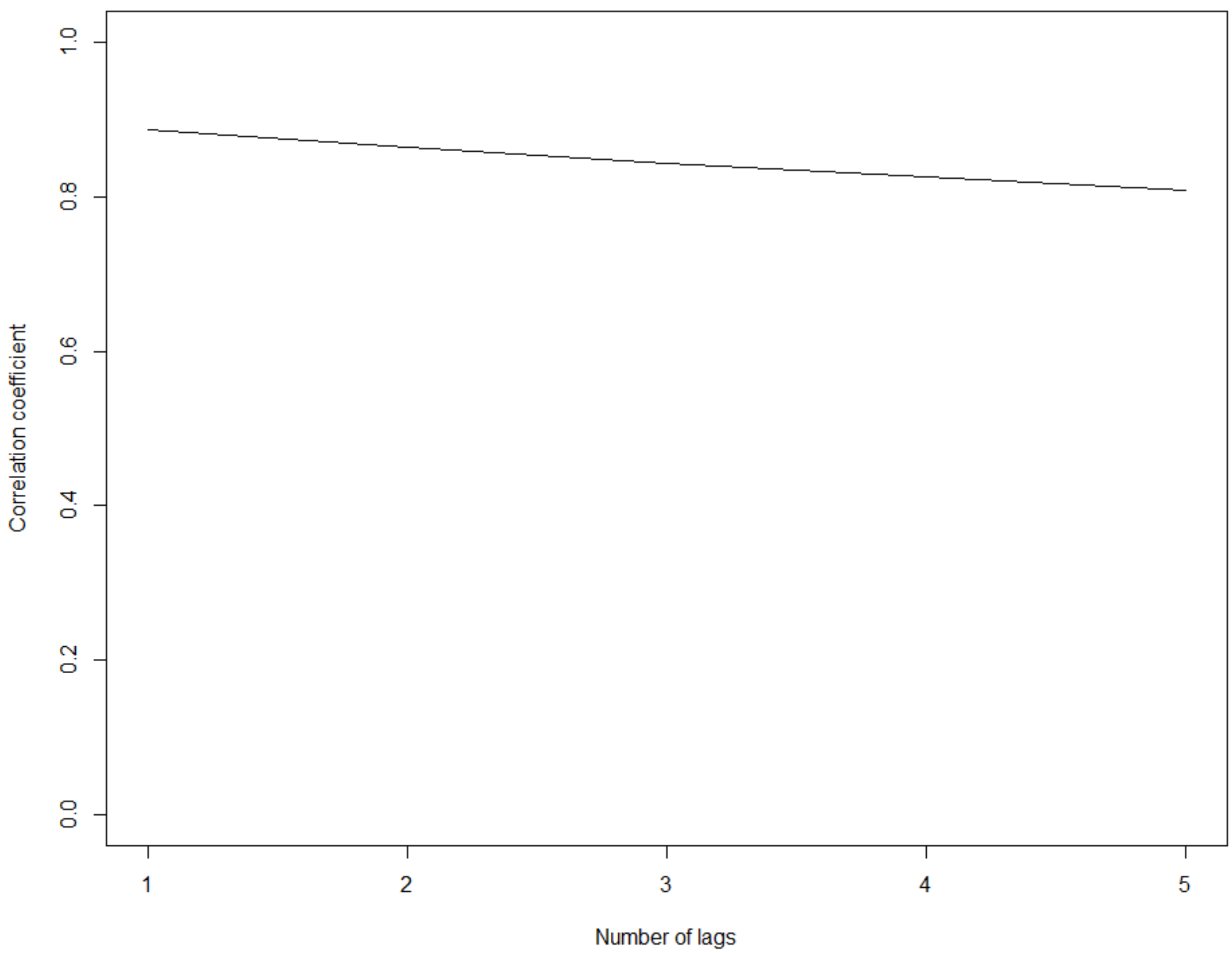
Table 3: P-values of the unit-root tests for the Gaussian ranks Z_t

H_0: Panels contain unit roots
H_A: Panels are stationary

Test	p-value
Levin-Lin-Chu unit-root test	0.0000
Harris-Tzavalis unit-root test	0.0000
Breitung unit-root test	0.0013

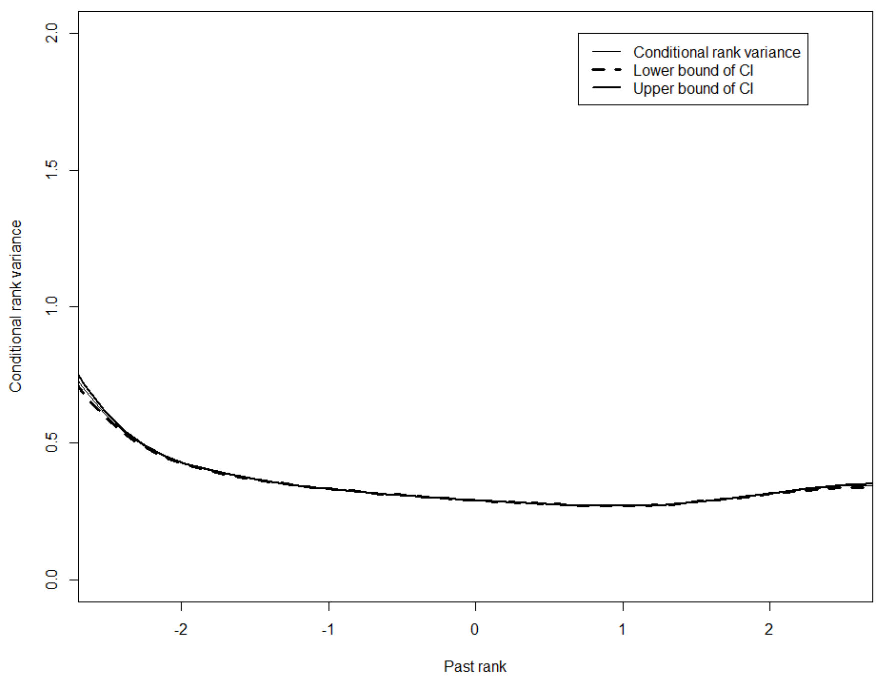
The p-values reported in this table confirm that our rank processes are stationary.

Figure 16: Correlation coefficient between the present individual Gaussian rank and the past individual Gaussian rank at the specified lag



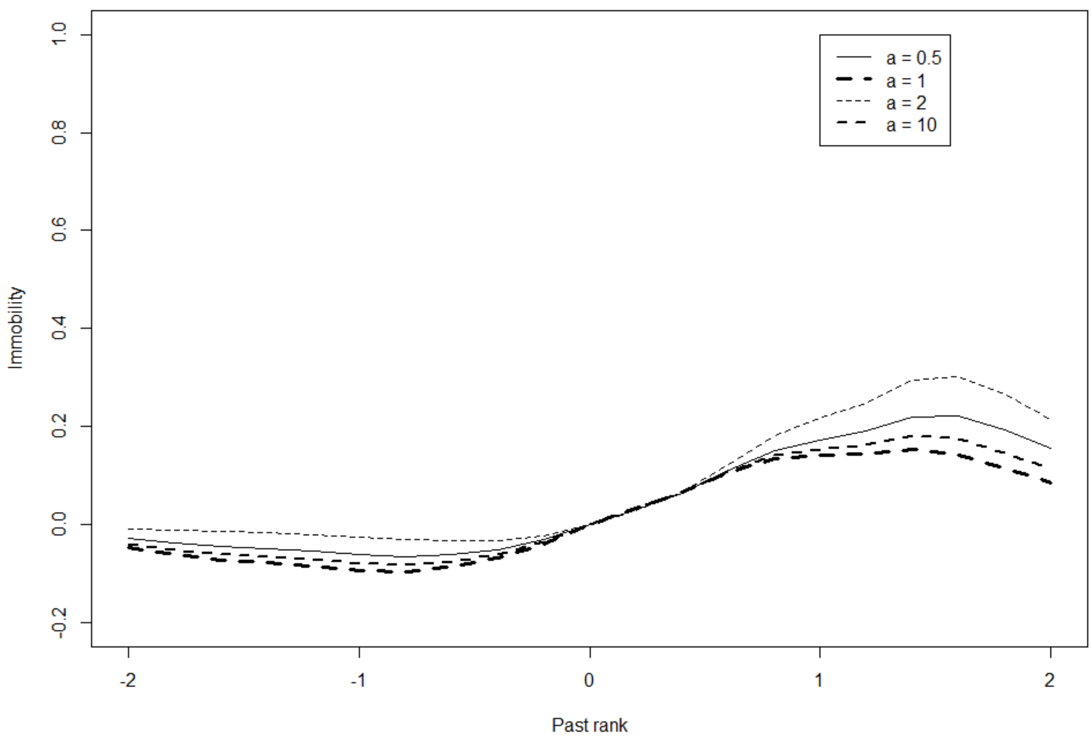
As expected, the correlation between present and past rank declines with time lags and it is lower and lower as time passes.

Figure 17: Conditional rank variance, US data, nT=35'376



This figure shows the estimated conditional variance of the Gaussian ranks in the US, together with the confidence interval for such variance. From this Figure, we deduce that it is reasonable to assume that the rank variance is constant.

Figure 18: Partial derivative of the log copula for different values of the copula parameter a



This figure shows how positional mobility, as it can be derived by the log copula, changes across the distribution. The measure of mobility represented is: $\frac{\partial}{\partial Z_{i,t-1}} E(Z_{i,t}|Z_{i,t-1}, W_{i,t}, W_{i,t-1}) = \int_0^1 G^{-1}(u; \theta(W_{it})) \frac{\partial c}{\partial v}(u, G(Z_{i,t-1}; \theta(W_{i,t-1})) \cdot g(Z_{i,t-1}; \theta(W_{i,t-1}))du$, where $c(u, v)$ is the log copula.

2.7 Appendix B Unconstrained estimation

As a first step, we obtain an initial guess for the β parameters in the variable score, $W_{i,t} = \beta' X_{i,t}$, via a nonlinear least square estimation. The model is the following:

$$Z_{i,t} = \Psi(W_{i,t}) Z_{i,t-1} + \varepsilon_{i,t} \tag{2.33}$$

where:

$$\Psi(s) = \frac{e^{2s} - 1}{e^{2s} + 1}$$

In Table 4, the results of this preliminary nonlinear least squares estimate (equation (2.33)) are reported. We find that the variable score is increasing in age and decreasing in age squared. Moreover, the score is increasing in all education levels, the highest values being associated with senior elementary and middle school (Edu1) and graduate education (Edu5).

Table 4: Nonlinear Least Square estimate (NLS)

variable	estimated coefficient
Age	.0188*** (.0033)
Age squared$\cdot 10^{-2}$	-.0173*** (.0060)
Edu1	1.3542*** (.0877)
Edu2	1.0431*** (.0485)
Edu3	.9255*** (.0415)
Edu4	1.1438*** (.0552)
Edu5	1.5022*** (.0892)
n. of obs.	35'346

This table reports the coefficients estimated via NLS on pooled PSID data for the period 1975-1996.

We now perform a Sieve (unconstrained) estimation by using a Hermite basis of second degree, $H_j(X)$. Therefore, once we have got a first guess for β, say, $\hat{\beta}$, we construct the following quantity: $\widetilde{W_{i,t}} = \hat{\beta}' X_{i,t}$. We then standardize this quantity and get the following:

$$\hat{W_{i,t}} = (\widetilde{W_{i,t}} - average(\widetilde{W_{i,t}}))/st.dev(\widetilde{W_{i,t}})$$

Now, we consider:

$$min_\theta \sum_{t=1}^{T} \sum_{i=1}^{n} [Z_{i,t} - \sum_{j=1}^{J} \sum_{k=1}^{K} \theta_{jk} H_j(Z_{i,t-1}) H_k(\hat{W}_{i,t})]^2$$

from this minimization program, we get an estimate for θ, say $\hat{\theta}$, by Ordinary Least Squares:

$$\hat{\theta} = (H'H)^{-1} H'Y$$

where H is the Sieve basis: $H = [H_j(Z_{i,t}) H_k(\hat{W}_{i,t})]$. We can now perform the third step of the estimation procedure:

$$min_\beta \sum_{t=1}^{T} \sum_{i=1}^{n} [Z_{i,t} - \sum_{j=1}^{J} \sum_{k=1}^{K} \hat{\theta}_{jk} H_j(Z_{i,t-1}) H_k(\beta' X_{1,t})]^2$$

Note that we are now minimizing with respect to β and not to θ. With this improved estimate of β we are now able to repeat the subsequent estimation step. Note that, each time that we obtain a different estimate of our beta parameters, we use this new result to estimate the (cross-sectional) mean and the standard deviation of the score and then we use these estimates to center and standardize the score itself again, so that it is approximately distributed as a standard normal variable. We proceed in this way in an iterative fashion, until the estimation results are substantially unchanged by another step of the procedure (we iterate until convergence). Note that $J = K = 3$, since we use a basis of degree 2. This iterative procedure yields consistent estimates of both the thetas and the betas. Indeed, in a time series context, Hautsch et al. (2014) show that, when the number of parameters to be estimated is high, it is possible to divide the full vector of parameters into some sub-vectors of arbitrary length, to be iteratively estimated in an alternate manner by MLE. The only requisite for these procedure to yield consistent parameter estimates is to start from a first guess which has been obtained by a consistent estimator. This is indeed our case, since our first guess has been obtained by a consistent estimation technique, namely NLS.

We stop our iterative procedure when the percentage change from the previous estimate to the following one is less than 2% for each of the estimated coefficients. When we use a Hermite series of degree 2 to approximate the correlation function, convergence is achieved after 26 iterations are performed. In order to ensure the robustness of the results found, we also tried to do another iteration step after convergence was achieved, in order to compare the results obtained with different optimization methods, such as a standard Newton-type algorithm, the method developed by Nelder and Mead (1965), the quasi-Newton method developed in 1970 by Broyden, and the conjugate gradients method based on the work by Fletcher and Reeves (1964). In each of those cases, the difference between the estimated coefficient obtained with different methods was less than 2% of the value of each coefficient. This findings confirms the validity of the results found.

We also tried to estimate the thetas and the betas (i.e, respectively, the coefficients determining the shape of the autoregressive function and those defining the influence of the individual explanatory variables on the score) simultaneously, via a maximum likelihood procedure. In this case, again, the difference in the estimation results between the two methods was equal at most to 2.2% of the value of each coefficient and the difference in the estimated coefficients was equal, on average, to 0.0016% of the value of the coefficient itself. This findings confirm the validity of our results.

2.8 Appendix C Consistency of the estimators

G.1. Assumptions

Assumption 5. *(i) Function $K(\cdot)$ is a symmetric, compactly supported kernel in $\mathbb{R}$ and it is three times differentiable with bounded derivative.*

(ii) Further $K(\cdot)$ is either a second-order or a fourth-order kernel function, i.e. $\int u^j K(u)du = 0 \;\; forj = 1, ..., p-1$, and $\int u^p K(u)du \neq 0$, for either $p = 2$, or $p = 4$.

Assumption 6. *(i) The bandwidth $h = h_N > 0$ satisfies $h = o(1)$ and $n^{2-\delta}h^6 \to \infty$ for some arbitrarily small $\delta > 0$.*

(ii) For the trimming operator we require that a_0,$c > 0$ and $n^c h^2 = o(1)$ and $n^{1-2c-\delta}h^2 \to \infty$ for some arbitrarily small $\delta > 0$.

(iii) $h < n^{-0.5p}$ and $h^4 > n^{\delta-1}$, where $p = 4$ is the order of the kernel function.

Assumption 7. *(i) For all $\beta_1 \in B_1$, $(Z_{it}, \beta_1' X_{it})$ has probability density $g_{Z,\beta_1'X}(z, w)$ with respect to Lebesgue measure on $\mathbb{S}_\delta$, this density is bounded, and such that $\inf\limits_{(z,w)\in\mathbb{S}_\delta} g_{Z,\beta_1'X}(z, w) > 0$, where $\mathbb{S}_\delta = \{\delta \in \mathbb{R}^{P_1} : \exists y \in \mathbb{R} \;\; s.t.(y, \delta) \in \mathbb{S}\}$. Moreover, $g_{Z,\beta_1'X}(z, w)$ and $E(X_{it}|Z_{it} = z, \beta_1' X_{it} = w)$ and $E(X_{it}X_{it}'|Z_{it} = z, \beta_1' X_{it} = w)$ are $(2 + p)$-times continously differentiable with respect to $(z, w) \in \mathbb{S}_\delta$.*

(ii) For all $\beta_1 \in B_1$, let $E_{\mathbb{S}_\delta}$ be the conditional expectation given $(Z_{it}, X_{it}) \in \mathbb{S}_\delta$, $E_{\mathbb{S}_\delta}[\log g(Z_{it}|\beta_1' X_{it})]$ is finite and has a unique global maximum β_1^0 that lies in the interior of B_1.

(iii) $q'\Omega(\beta_1^0)q > 0$ for any vector $q \in \mathbb{R}^{p_1}$ s.t. $q \perp \beta_1^0$, where $\Omega(\beta_1^0) = E_{\mathbb{S}_\delta}[-\nabla^2_{\beta_1\beta_1} \log g(Z_{it}|\beta_1^{0'} X_{it})]$ and $\nabla^2_{\beta_1\beta_1}$ denotes the matrix of second-order partial derivatives with respect to β_1.

Assumption 8. *(i)$E[\tilde{l}(\tilde{Y}_{it}, \theta_0, f_0)]$ exists in $\mathbb{R}$, where $\tilde{l}(\tilde{Y}_{it}, \theta_0, f_0) = \log c[\tilde{G}(Y_{it}, f_0), \tilde{G}(Y_{i,t-1}, f_0); \rho_0(\cdot|X_{it}'\beta_2^0)]$ and $\tilde{Y_{it}} = (Y_{it}, Y_{i,t-1})$.*

(ii) For all $\epsilon > 0$, there exists some non-increasing positive sequence $c_N(\epsilon)$ such that, for all $N \geq 1$,

$$E[\tilde{l}(Y_{it},\theta_0,f_0)] - \sup_{\{\theta\in\Theta_N:||\theta-\theta_0||_\Theta\geq\epsilon\}} E[\tilde{l}(Y_{it},\theta,f_0)] \geq c_N(\epsilon) \tag{2.34}$$

where $\liminf_N c_N(\epsilon) > 0$, $\forall\epsilon > 0$, *and* $||\cdot||_\Theta$ *is a metric defined on* Θ *or some metric space containing* Θ.

Assumption 9. $\theta_0 \in \Theta$, *and* $\Theta_N \subset \Theta_{N+1} \subset \Theta$ *for all* $N \geq 1$ *and there exists some* $\theta_N \in \Theta_N$ *such that:*

$$|E[l(Y_i,\theta_N,f_0) - l(Y_i,\theta_0,f_0)]| = O(\eta_{2,N})$$

where $\eta_{2,N}$ *is some positive non-increasing sequence.*

Assumption 10. *(i)* $\sup_{\theta\in\Theta_N,f\in\mathcal{N}_{f,N}} |\mu_N[l(Y_i,\theta,f)]| = O_p(\eta_{0,N})$ *where* $\eta_{0,N}$ *is some finite positive non-increasing sequence going to zero,* $\mu_N[l(Y_i,\theta,f)] = \frac{1}{N}\times\sum_{i=1}^{N}\{l(Y_i,\theta,f) - E[l(Y_i,\theta,f)]\}$ *and* $\mathcal{N}_{f,N} = \{f \in \mathbb{F}_N : ||f-f_0||_\mathbb{F} \leq \delta_{f,N}\}$ *with* $\delta_{f,N} = o(1)$, *where* $||f-f_0||_\mathbb{F} = \sup_{(z,x)\in\mathbb{S}} |f(z|x)-f_0(z|x)|$.

(ii) There is a finite positive non-increasing sequence $\{\eta_{1,N}\}$ *going to zero such that*

$$\sup_{\theta\in\Theta_N,f\in\mathcal{N}_{f,N}} |E[l(Y_i,\theta,f) - l(Y_i,\theta,f_0)]| = O(\eta_{1,N}). \tag{2.35}$$

Delecroix et al. (2003) present the assumptions under which a semi-parametric single-index model yields a consistent estimator in a cross-sectional framework. Their work has been extended by Rosemarin (2012). Assumptions 5-7 are needed to ensure consistency and asymptotic normality of our semi-parametric single-index model in a panel data setting. These assumptions build on the conditions used in Rosemarin (2012). Assumption 5 imposes conditions on the kernel function. Assumption 6 imposes conditions on the bandwidth h in order to obtain uniform convergence of the kernel density estimators. In particular, Assumption 6 (iii) is needed to show $\sqrt{N}$-convergence and asymptotic normality of the $\hat{\beta}_1$. Note that Assumption 6 (iii) only holds if the kernel is of fourth order. Assumption 7 includes some requirements for the continuity and differentiability of expectations and density functions. Assumption 7 (ii) is a global identifiability condition for β_1^0 and $E_{\mathbb{S}_\delta}[\log g(Z_{it}|\beta_1' X_{it})]$ is the limit value of the criterion which is maximized in the sample. Matrix Ω in Assumption 7 (iii) is a sort of Hessian matrix and Assumption 7 (iii) itself is a common requirement in the dimension reduction literature (see Hall and Yao (2005)).

Assumptions 8-10 are adapted to our framework from Hahn et al. (2016) and are needed to ensure consistency of the second-step Sieve Maximum Likelihood estimator. As underlined by Hahn et al. (2016), Assumption 8 essentially requires that the second-step Sieve M estimation problem is well-posed. Intuitively, our second-step estimation is a well-posed problem, since, as it appears from equations (2.20) and (2.29)-(2.30), the estimation criterion does not only depend on the approximating function $\rho(\cdot)$ via an integral of $\rho(\cdot)$, but also directly, i.e. via the values of ρ at sample points. Verification of Assumption 8 from more primitive conditions is not simple since the estimation problem is nonlinear. Assumption 9 is a condition on the accuracy of the approximation of Θ by the sequence of Sieve spaces, it imposes a bound on the approximation error and requires that the Sieve space is expanding. According to Hahn et al. (2016), it is essentially condition b of Lemma A.2 in Chen and Pouzo (2012). Hahn et al. (2016) also point out that Assumption 9 is implied by conditions 3.2 and 3.3 in Chen (2007). Assumption 10 is similar to condition 3.5 of Theorem 3.1 in Chen (2007) and the first part of Condition d of Lemma A.2 in Chen and Pouzo (2012); it requires uniform convergence of the empirical process $\mu_N[l(Y_i, \theta, f)]$.

G2. Proof of Theorem 1

Part (i)

To show consistency of the estimator $\hat{\beta}_1$ in our single-index model, we apply Theorem 1 in Rosemarin (2012). In our model, under Assumptions 1-7, the hypotheses of Rosemarin (2012) hold. In particular, Assumption A1 in Rosemarin (2012), which requires the processes to be strong mixing, is implied by our Assumption 1, i.e. that the processes (Z_{it}, X_{it}) are iid across individuals. Indeed, given that T is fixed, we can re-name the observations as follows, so that our data are a particular case of time-series with finite-order dependence. To obtain this result we rely on our Assumption 1. We have T observations for each of the N individuals included in our sample. We re-arrange these observations in a unique vector, such that the first T observations correspond to the first individual, the observations from $T+1$ to $2T$ correspond to the second individual, and so on, until the $NT-th$ observation, corresponding to period T for individual N. More specifically, observations are re-indexed as follows: $Z_j = Z_{i,t} \Leftrightarrow j = T(i-1)+t, \quad j = 1, ..., NT$. And similarily for X.

In this way, by Assumption 1 we ensure that there is dependence at most for T lags in our observations, hence we are in a particular case of strong mixing time series and hence the Assumption A1 in Rosemarin (2012) applies to our case, with constants $\eta_1 = \eta_2 = \infty$. Note that Assumption A1 in Rosemarin (2012) also requires data X to be strictly stationary. This is implied by our Assumption 3. The fact of re-indexing the series does not invalidates the result. Indeed, by our Assumptions 1 and 3, we can write $l(Z_j, X_j), \ j = 1, ..., NT$, which denotes the invariant density of processes (Z_{it}, X_{it}). Having rewritten the model in this way, our Assumption 6 (i) implies Assumption A3 in Rosemarin (2012). Similarly, our Assumption 5 (i) implies Assumptions A2 in Rosemarin (2012). Our Assumption 6 (ii) implies Assumption A5 in Rosemarin (2012), and our Assumption 7 (ii) implies Assumption A6 in Rosemarin (2012). Further, our Assumption 7 (i) implies Assumption A4 in Rosemarin (2012) because, by our

Assumption 1 (iid individuals), the joint density can be written as the product of the marginal densities and boundedness of this product follows from boundedness of the marginal densities.

Part (ii)

The second part of our Theorem 1 in Section 4.2 follows directly from Theorem 3 in Rosemarin (2012). With respect to the previous part of our Theorem 1, here some additional assumptions are required. In particular, we also need Assumptions A7-A10 in Rosemarin (2012). Our Assumption 5 (i) implies Assumption A7 in Rosemarin (2012). Assumption 6 (i) implies Assumption A8 in Rosemarin (2012). Assumption 7 (i) implies Assumption A9 and Assumption 7 (iii) implies Assumption A10 in Rosemarin (2012).

Part (iii)

To show $\sqrt{N}$-convergence and asymptotic normality of the $\hat{\beta}_1$, we need again to assume that our Assumptions 1-10 hold. The only additional requirement with respect to Parts (i) and (ii) is given by Condition 2.8 in Rosemarin (2012), which is implied by our Assumption 6 (iii).

G3. Proof of Theorem 2

To show consistency of our second-step Sieve M-estimator, we apply Theorem 2.1 by Hahn et al. (2016). In particular, Assumption 2.1 in Hahn et al. (2016) is verified if our Assumption 8 holds. Note that to get our Assumption 8 we re-wrote Assumption 2.1 in Hahn et al. (2016) in a slightly different way to accommodate the panel structure of our data. Starting from their expression (adapted to our case): $E[\Psi(Y_i, \theta, f)] = E[l(Y_i, \theta, f)] = E[\sum_{t=1}^{T} l(Y_{it}, \theta, f)] = \sum_{t=1}^{T} E[l(Y_{it}, \theta, f)]$, then by our Assumptions 1 (iid individuals) and 3 (stationarity) we can finally write $E[\Psi(Y_i, \theta, f)] = T \times E[l(Y_{it}, \theta, f)]$, independent of i and t. Then the fixed factor T in the r.h.s. in equation (2.34) simplifies by redefinition of the constant $c_N(\epsilon)$. Assumption 2.2 in Hahn et al. (2016) holds if we adopt our Assumption 9 and Assumption 2.3 in Hahn et al. (2016) is valid if our Assumption 10 holds. Note that Theorem 2.1 by Hahn et al. (2016) requires that the first-step estimate $\hat{f}$ is consistent in norm $||\cdot||_{\mathbb{F}}$, which is implied by our Theorem 1 part (ii).

Supplementary material

2.9 Appendix D Copula models

Joe (1997) and Nelsen (1999) provide extensive surveys on copula theory. A copula function couples marginal distributions to get a joint distribution. By definition, a joint cumulative function C on $[0;1]^2$ with uniform marginal distributions on $[0,1]$ is a copula. Thus, a function $C : [0;1]^2 \to [0,1]$ is a copula if the following conditions hold:

$$C(0,v) = C(u,0) = 0, \forall u, v \in [0,1]$$

$$C(1,v) = v, C(u,1) = u, \forall u, v \in [0,1]$$

For any rectangle $R = [u_1, u_2] \times [v_1, v_2] \subset [0;1]^2$ the following holds:

$$\int\int C(du, dv) = C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0.$$

When the distribution C is continuous, the associated density is called the copula density:

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v},\ u, v \in [0,1].$$

Sklar's theorem (1959)

Let $F(x,y)$ be the bivariate c.d.f. of random variables X and Y, with marginal c.d.f. equal to F_x and F_y. Let then $U = F_x(X)$ and $V = F_y(Y)$. The joint c.d.f. is given by:

$$P(U \leq u, V \leq v) = P(X \leq F_x^{-1}(U), Y \leq F_y^{-1}(V)) = F(F_x^{-1}(u), F_y^{-1}(v)),\ \forall u, v \in [0,1]$$

Then there exists a copula such that:

$$C(u,v) = F(F_x^{-1}(u), F_y^{-1}(v)), \forall u, v \in [0,1]$$

i.e.$F(x,y) = C[F_x(x), F_Y(y)]\ \forall x, y.$

This copula is unique if F is a continuous distribution. Assuming sufficient regularity, it is possible to obtain the following conditional c.d.f.:

$$P(U \leq u \mid V = v) = \partial C(u,v)/\partial v$$

The conditional density function is obtained by deriving once more with respect to u. The copula is limited by the so-called Frechet-Hoeffding bounds; the upper bound is: $C_U(u,v) = max(u+v-1;0)$, while the lower bound is: $C_L(u,v) = min(u,v)$. In the case of independence we have $C_I(u,v) = u \cdot v$. There are many different families of copulas, for example the Gaussian, the Frank or the Gumbel.

The Gaussian copula

A Gaussian copula is associated with a bivariate Gaussian distribution with zero expected values, unitary variances and correlation equal to ρ. For $0 \leq \rho \leq 1$, the Gaussian copula is defined as:

$$C(u,v,\rho) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v), \rho)$$

where Φ is the$N(0,1)$ c.d.f., Φ^{-1} is the quantile function (the inverse of Φ) and Φ_ρ is the bivariate standard normal c.d.f. with correlation parameter equal to ρ. The Gaussian copula density is defined as:

$$c(u,v,\rho) = \frac{\varphi_2(\Phi^{-1}(u), \Phi^{-1}(v), \rho)}{\varphi(\Phi^{-1}(u))\varphi(\Phi^{-1}(v))}$$

Where Φ is the standard normal distribution, φ is its density and $\varphi_2(x,y,\rho)$ is the density of the bivariate normal distribution with correlation coefficient ρ, expected values equal to zero and unitary variances. More explicitly,

$$c(u,v,\rho) = \frac{1}{\sqrt{1-\rho^2}} \cdot exp\left(\frac{x^2+y_2^2}{2} + \frac{2\rho xy - x^2 - y^2}{2(1-\rho^2)}\right)$$

with $x = \Phi^{-1}(u)$ and $y = \Phi^{-1}(v)$. It is known that C_U for $\rho = 1$, C_L for $\rho = -1$ and C_I for$\rho = 0$ (Joe 1997 and Nelsen 1999). The lower bound, the minimum copula denoted by C_L, models perfect negative dependence, while the upper bound, the maximum copula denoted by C_U, models perfect positive dependence. The product copula C_I stands for independence between the two variables. The bivariate Gaussian copula generates zero tail dependence.

2.10 Appendix E Non-causality definitions

Definition 1.1 Granger non-causality (1969)

Let us consider two stochastic processes, (Z_t) and (W_t). If the linear predictor of the current value of W, given its own past and the past of Z, does not depend on the latter, then there is absence of Granger causality from Z to W.

This definition, which has been originally formulated for the linear predictor case, can be generalized to density functions. Z is said not to cause W in the Granger sense if the following holds:

$$l(W_t|\underline{Z_{t-1}}, \underline{W_{t-1}}) = l(W_t|\underline{W_{t-1}}) \tag{2.36}$$

where $\underline{Z_{t-1}} = (Z_{t-1}, Z_{t-2}, ...)$ and $\underline{W_{t-1}} = (W_{t-1}, W_{t-2}, ...)$, and $l(\cdot)$ represents the conditional density.

Definition 1.2 Sims non-causality (1972)

Let us consider two stochastic processes, (Z_t) and (W_t). Z does not cause W in the Sims definition if the linear predictor of Z_t, based on $..., W_{t-1}, W_t, W_{t+1}, ...$ is identical to the linear predictor of Z_t based on $W_t, W_{t-1}, ...$ alone. In this case, too, a generalization of the above-presented definition is possible. We say that W does not Sims-cause Z if the following holds:

$$l(Z_t|\underline{Z_{t-1}}, \underline{W_T}) = l(Z_t|\underline{Z_{t-1}}, \underline{W_t}) \tag{2.37}$$

where $\underline{W_T} = (W_1, ...W_t, ...W_T)$ and $\underline{W_t} = (W_1, ...W_t)$.

With reference to our model, we consider two stochastic processes, (Z_t) and (W_t), where Z_t stands for the individual wage rank and W_t represents the score, i.e. the weighted sum of the explanatory variables.

Lemma 2

If there is absence of Granger causality from Z to W, then the following holds:

$$l(\overline{W^{t+1}}|\underline{Z_t}, \underline{W_t}) = l(\overline{W^{t+1}}|\underline{W_t})$$

Proof.

$$\begin{aligned} l(\overline{W^{t+1}}|\underline{Z_t}, \underline{W_t}) &= l(W_{t+1}, W_{t+2}, ..., W_T|\underline{Z_t}, \underline{W_t}) \\ &= l(W_T|W_{t+1}, ..., W_{T-1}, \underline{Z_t}, \underline{W_t})l(W_{t+1}, ..., W_{T-1}|\underline{Z_t}, \underline{W_t}) \\ &= l(W_T|\underline{Z_t}, \underline{W_{T-1}})l(W_{t+1}, ..., W_{T-1}|\underline{Z_t}, \underline{W_t}) \end{aligned} \tag{2.38}$$

The first term in the right hand side of equation (2.38) is equal to $l(W_T|\underline{W_{T-1}})$ due to the assumption of Granger non-causality. The same computations can be performed in turn for every W in the interval $[W_{t+1}, ..., W_{T-1}]$. Therefore, we have proved that:

$$l(\overline{W^{t+1}}|\underline{Z_t}, \underline{W_t}) = l(\overline{W^{t+1}}|\underline{W_t})$$

$\square$

Lemma 3

Granger non-causality implies Sims non-causality (in the sense of equations (2.36) and (2.37)). It is known that the converse can be proved similarly (Gourieroux and Monfort (1995)); the two concepts are equivalent. For completeness, we present here a proof of Lemma 3, which is similar to that presented in Gourieroux and Monfort (1995).

Proof. We assume that Z_t does not cause W_t in the definition of Granger and we want to show that this implies that Z_t does not cause W_t in the Sims sense. Let us consider the conditional density of Z_t given its own past history, $\underline{Z_{t-1}}$ and the whole history of W, denoted by $\underline{W_T}$. The following decomposition is always valid:

$$l(Z_t|\underline{Z_{t-1}}, \underline{W_T}) = \frac{l(\underline{Z_t}, \underline{W_T})}{l(\underline{Z_{t-1}}, \underline{W_T})} = \frac{l(\underline{Z_t}, \underline{W_t})l(\overline{W^{t+1}}|\underline{Z_t}, \underline{W_t})}{l(\underline{Z_{t-1}}, \underline{W_t})l(\overline{W^{t+1}}|\underline{Z_{t-1}}, \underline{W_t})}, \tag{2.39}$$

where $\overline{W^{t+1}} = (W_{t+1}, ..., W_T)$. We need now to introduce the following Lemma:
Using Lemma 3, equation (2.39) becomes:

$$l(Z_t|\underline{Z_{t-1}}, \underline{W_T}) = \frac{l(\underline{Z_t}, \underline{W_t})l(\overline{W^{t+1}}|\underline{W_t})}{l(\underline{Z_{t-1}}, \underline{W_t})l(\overline{W^{t+1}}|\underline{W_t})} = \frac{l(\underline{Z_t}, \underline{W_t})}{l(\underline{Z_{t-1}}, \underline{W_t})} = l(Z_t|\underline{Z_{t-1}}, \underline{W_t}) \tag{2.40}$$

We have shown that:

$$l(Z_t \mid \underline{Z_{t-1}}, \underline{W_T}) = l(Z_t \mid \underline{Z_{t-1}}, \underline{W}_t)$$

i.e. that Z does not cause W in the Sims sense. Therefore, Granger non-causality implies Sims non-causality. $\square$

Lemma 4

If (Z_t) is a first-order markovian process conditional on (W_t) and if there is absence of Sims-causality from Z to W, then the following holds:

$$l(Z_t|\underline{Z_{t-1}}, \underline{W_t}) = l(Z_t|Z_{t-1}, \underline{W_t})$$

Proof. From the assumption of Sims non-causality from Z to W we can write the following:

$$l(Z_t \mid \underline{Z_{t-1}}, \underline{W_t}) = l(Z_t \mid \underline{Z_{t-1}}, \underline{W_T})$$

From the assumption of first-order markovianity of Z_t conditional on W_t we know that:

$$l(Z_t|\underline{Z_{t-1}}, \underline{W_T}) = l(Z_t|Z_{t-1}, \underline{W_T})$$

This last conditional density can be rewritten as it follows:

$$l(Z_t|Z_{t-1}, \underline{W_T}) = \frac{l(Z_t, Z_{t-1}, \underline{W_T})}{l(Z_{t-1}, \underline{W_T})} = \frac{l(Z_t, Z_{t-1}, \underline{W_t})l(\overline{W^{t+1}}|Z_t, Z_{t-1}, \underline{W_t})}{l(Z_{t-1}, \underline{W_t})l(\overline{W^{t+1}}|Z_{t-1}, \underline{W_t})}.$$

From Lemma 3 we know that Sims non-causality implies Granger non-causality (the two concepts are equivalent). Therefore, from the assumption of Granger non-causality from Z to W and Lemma 2 we obtain:

$$l(Z_t|Z_{t-1}, \underline{W_T}) = \frac{l(Z_t, Z_{t-1}, \underline{W_t})l(\overline{W^{t+1}}|\underline{W_t})}{l(Z_{t-1}, \underline{W_t})l(\overline{W^{t+1}}|\underline{W_t})} = \frac{l(Z_t, Z_{t-1}, \underline{W_t})}{l(Z_{t-1}, \underline{W_t})} = l(Z_t|Z_{t-1}, \underline{W_t})$$

We have shown that:

$$l(Z_t|\underline{Z_{t-1}}, \underline{W_t}) = l(Z_t|Z_{t-1}, \underline{W_T}) = l(Z_t|Z_{t-1}, \underline{W_t})$$

$\square$

Lemma 5

If there is absence of Granger causality from Z to W, then the following holds:

$$Z_t|W_{t-p}, ..., W_{t+q} \sim Z_t|W_{t-p}, ..., W_t$$

Proof.

$$l(Z_t|W_{t-p}, ..., W_{t+q}) = \frac{l(Z_t, W_{t-p}, ..., W_{t+q})}{l(W_{t-p}, ..., W_{t+q})} = \frac{l(W_{t+1}, ..., W_{t+q}|Z_t, W_t, ...W_{t-p})l(Z_t, W_t, ...W_{t-p})}{l(W_{t+1}, ..., W_{t+q}|, W_t, ...W_{t-p})l(W_t, ..., W_{t-p})}$$

from the assumption of Granger non-causality and Lemma 2, we obtain the following result:

$$l(Z_t|W_{t-p}, ..., W_{t+q}) = \frac{l(Z_t, W_t, ...W_{t-p})}{l(W_t, ..., W_{t-p})} = l(Z_t|W_t, ...W_{t-p})$$

This proofs that:

$$Z_t \mid \underline{W_T} \sim Z_t \mid \underline{W_t}$$

$\square$

Lemma 6

If there is absence of Granger causality from Z to W, then:

$$Z_{t-1}|W_{t-1}, W_t \sim Z_{t-1}|W_{t-1}$$

Proof.

$$l(Z_{t-1}|W_{t-1}, W_t) = \frac{l(Z_{t-1}, W_{t-1}, W_t)}{l(W_{t-1}, W_t)} = \frac{l(W_t|W_{t-1}, Z_{t-1})l(Z_{t-1}, W_{t-1})}{l(W_t|W_{t-1})l(W_{t-1})}$$

where $l(W_t|W_{t-1}, Z_{t-1}) = l(W_t|W_{t-1})$

from the assumption of Granger non-causality. Therefore we get:

$$l(Z_{t-1}|W_{t-1}, W_t) = \frac{l(Z_{t-1}, W_{t-1})}{l(W_{t-1})} = l(Z_{t-1}|W_{t-1})$$

$\square$

2.11 Appendix F Examples of simple autoregressive models

1.

This example is based on the assumption of joint normality of the ranks and of the univariate score W_t: $(Z_{it}, W_{it}) \sim N(0, I_d)$. Let us consider the following laws for the dynamics of the ranks and for the evolution of the explanatory variables:

$$Z_t = \rho Z_{t-1} + W_t + \varepsilon_t$$

and

$$W_t = rW_{t-1} + \eta_t$$

Note that in this example Z_t is univariate. We then make the following assumptions:

1. $\varepsilon_t \sim IIN(0, \omega^2)$
2. $\eta_t \sim IIN(0, \sigma^2)$
3. $\varepsilon_t \perp\!\!\!\perp \eta_t$

From Assumption 1, it follows that:

$$Z_t | \underline{W_t}, \underline{Z_{t-1}} \sim N(\rho Z_{t-1} + W_t, \omega^2)$$

and, from Assumption 2:

$$W_t | \underline{W_{t-1}}, \underline{Z_{t-1}} \sim N(rW_{t-1}, \sigma^2)$$

The two conditional laws that characterize the joint process of the ranks and the explanatory variables are normal. Let us now examine more closely the distribution of the present rank Z_t, which is recursively defined as:

$$Z_t = \frac{1}{1-\rho L}(W_t + \varepsilon_t) = \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + ... + W_t + \rho W_{t-1} + \rho^2 W_{t-2} + ...$$

where L stands for the lag operator.

The distribution of $Z_t|\underline{W_t}$, which for the hypotheses made in the previous Section is equal to the distribution of $Z_t|\underline{W_T}$, is the following:

$$Z_t|\underline{W_t} \sim N(W_t + \rho W_{t-1} + \rho^2 W_{t-2} + ..., \frac{\omega^2}{1-\rho^2})$$

We can now compute the variance of Z_t as it follows:

$$V(Z_t) = V(E(Z_t|\underline{W_t})) + E(V(Z_t|\underline{W_t}))$$

We get that the following condition must hold:

$$V(Z_t) = V(W_t + \rho W_{t-1} + \rho^2 W_{t-2} + ...) + \frac{\omega^2}{1-\rho^2} = 1$$

which is equivalent to:

$$(1-\rho^2)V(W_t+\rho W_{t-1}+\rho^2 W_{t-2}+...)+\omega^2+\rho^2=1 \tag{2.41}$$

With this example, we have shown that, under the hypothesis of joint normality of the ranks and the score (weighted sum of explanatory variables), Z_t and Z_{t-1} conditional on $\underline{W_t}$ have Gaussian copula with correlation ρ. We assumed that W_t is normally distributed and that Z_t, conditional on W_t, is also normally distributed. In this setting, we were able to show that Z_t is also distributed as a standard normal. Furthermore, the present rank Z_t has for sure zero expected value (this result can be shown by a recursive procedure). In order for Z_t to be an acceptable copula argument, Equation (2.41) must hold. In this example we preserved standard gaussianity of the present rank Z_t, but at the cost of assuming a normal distribution for the score W_t.

2.

In order to understand the characteristics that the non-linear autoregressive function should have, let us consider the following example, where z represents the past rank and w stands for the score:

$$\rho(z,w)=\rho_0(z+\alpha)^{-w}$$

Where ρ_0 and α are constants, with $\rho_0>0$ and $0\leq\alpha<1$. We consider the partial derivative $\frac{\partial\rho}{\partial z}(z,w)$ as the relevant measure of local positional persistence, which is function of both z and w.

We have:

$$\frac{\partial\rho}{\partial z}=-w\rho_0(z+\alpha)^{-(w+1)}$$

Now we can compute the mixed derivative:

$$\frac{\partial^2\rho}{\partial z\partial w}=\rho_0(z+\alpha)^{-(w+1)}\cdot[wlog(z+\alpha)-1]$$

the first element is surely non-negative, while for the second factor there exists a turning point:

$$wlog(z+\alpha)-1\geq 0 \Leftrightarrow z\geq e^{1/w}-\alpha$$

If the past wage rank is higher than a certain threshold (which inversely depends on the score), then the mixed derivative is positive. This means that, if the past rank is high enough, a higher score is associated with a higher degree of positional persistence, given that there is little room for further improvements in the relative position.

When w goes to infinity, the turning point is very close to $1-\alpha$, while when w goes to zero, the turning point is close to infinity (there is no turning point in the sign of the derivative). On the other side, if we look more closely at the role of the score, we get that an alternative condition for the mixed derivative to be positive:

$$w\geq \frac{1}{log(z+\alpha)}$$

If the value of the score is greater than a certain threshold level (which inversely depends on the past rank), then a higher initial rank is associated with a higher degree of positional persistence; this means that a high score protects individuals at the top of the distribution from downward mobility and, on the contrary, enhances upward mobility for workers at the bottom of the wage scale.

2.12 Appendix G An example of function $g(Z_t;\theta(W_t))$

Let us suppose the following:

$$W_t\sim N(0,\Sigma)$$

for the conditional density of the present rank we assume the following simple form:

$$Z_t|W_t\sim N(W_t,\omega^2)$$

Z_t is for sure Gaussian. However, in order for Z_t to be an acceptable copula argument, we also have to ensure that it has zero expected value and unitary variance:

$$E(Z_t) = E(E(Z_t|W_t)) = E(W_t) = 0,$$

by the law of iterated expectations. For the variance we have:

$$V(Z_t) = V(E(Z_t|W_t)) + E(V(Z_t|W_t)) = V(W_t) + E(\omega^2) = \Sigma + \omega^2$$

Hence, the condition for unitary variance is:

$$\Sigma + \omega^2 = 1$$

3 Wage mobility before and after the financial crisis: an empirical analysis for Germany and the UK

3.1 Introduction

[18] The aim of this paper is to analyze and compare the degree of relative wage mobility before and after the latest financial crisis (2007-2009) in Germany and in the United Kingdom. Due to the long-lasting effects of this crisis on the economic structure of many industrialized countries, several scholars call it the ”Great Recession” (e.g. Caliendo and Hogenacker 2012, Rinne and Zimmermann 2012, Blanchard et al. 2014). It is likely that, besides its effects on employment and wage levels, the financial turmoil also influenced relative wage moblity. Earnings mobility is a distinctive feature of an economy (Friedman 1962, Shorrocks 1978, Gregg et al. 2015), which gives more insights than point-in-time inequality measures. The study of wage dynamics allows to assess whether a ”low-wage trap” (Dickens 2000) exists or not, i.e. whether low-paid workers have the opportunity of improving their rank over time. The contributions of this paper are manifold. First, it is the first study devoted to the analysis of how wage mobility changed after the recent financial crisis. Second, it tackles the issue of measuring differences in wage mobility between immigrants and natives by means of a new semi-nonparametric method, which allows to accommodate for complex wage dynamics. In the previous literature, indeed, evidence on differences in wage mobility due to nationality is scarce and the few existing studies only performed mean estimations with fully parametric models. Up to now, there is a large body of literature on the effect of migration on employment probabilities and wage level of natives (see e.g. Card 2001, Borjas 2003, Aydemir and Borjas 2011, Glitz 2012, Peri 2012, Smith 2012, Cadena 2013, Patel and Vella 2013, Abramitzky et al. 2014, Olney 2015, Peri 2016, Dustmann

[18] I would like to thank Patrick Gagliardini, Christian Gourieroux, Bo Honore, Thierry Kamionka, Rico Maggi, Alain Monfort, Jean-Marc Robin, Sebastien Roux, Mark Watson, the participants to the IRE seminar at USI, Lugano, 13-14 October 2016 and the participants to the junior seminar at CREST, Paris, on 11^{th} April 2017, for their insightful advice.

et al. 2016, Dustmann et al. 2017). However, there is still room for investigation on whether immigrant workers have a higher or lower degree of wage mobility than their native colleagues. Lastly, in the present paper we analyze the influence of different institutional frameworks on the degree of relative wage mobility with the flexible model mentioned above. To the best of our knowledge, the few existing cross-country studies on wage mobility only rely on summary mobility indices, without analyzing the degree of (micro) wage mobility in different parts of the earnings distribution (see e.g. Ayala and Sastre 2007). From a policy viewpoint, we are mostly interested in the evolution of wage mobility at the bottom of the distribution; hence, a flexible enough modellization of wage dynamics is fundamental, since we cannot assume that the degree of positional persistence is the same across the whole wage distribution.

3.1.1 Literature review

The institutional framework in Germany and in the UK

In the literature, a distinction is usually made between three main models of labor market institutions (Blanchard et al. 2014). The first one, called ”Anglo-Saxon”, is based on a low level of Employment Protection Legislation (EPL) and a low unemployment insurance. The UK belongs to this first case. The second model, called ”Nordic” or ”Scandinavian” (the flexicurity model) is based on a medium-to-high degree of EPL, a generous but rigidly-defined unemployment insurance and strong active labor market policies. Lastly, there is the so-called ”Continental” model, based on high EPL, generous unemployment insurance and limited active labor market policies. EPL includes laws about hiring and firing and on the use of fixed-term contracts. Germany is the typical ”Continental” European welfare state: in 2013 Germany had the most stringent restriction for the layoff of permanent workers, whereas the United Kingdom had the least stringent regulations in Europe (Holmlund 2014). The level of employment protection is significantly higher in Germany than in the UK (Bruecker et al. 2014). The UK is also characterized by low unemployment benefits. A legal minimum wage exists in the UK as well as in Germany. However, minimum wage has been introduced in Germany only in 2015, whereas it exists since 1999 in the UK (Metcalf 1999).

In Table 5 we summarize the main features of the labor markets in the two countries under scrutiny. As mentioned before, Germany has a stricter Employment Protection Legislation[19] and a higher tax wedge. However, minimum wage expressed as a percentage of the median full-time wage is now almost identical in the two countries. Data on unemployment rate and on labor force participation are also quite similar across the two countries. Incidence of temporary employment is higher in Germany than in the UK, both before and after the crisis; this finding has often been mentioned by the scholars as one of the reasons why Germany was able to suffer a relatively limited job loss during the Great Recession. Note that, of course, the economic variables reported in Table 5 are subject to oscillations linked to the business cycle. Hence, variations from 2006 to 2013 do not necessarily reflect a change in the underlying economic structure caused by the crisis.

[19] This is an aggregate index ranging from 1 to 3.5 and including different aspects of EPL.

Table 5: Overview of labor market institutions in 2006 and in 2015

Before the crisis		
	Germany	United Kingdom
Employment Protection Legislation index (2006)	2.68	1.26
Public unemployment speding (as % of GDP)	1.6	0.2
Minimum wage (% of median full-time wage)	-	45
Tax wedge (as % of labor cost)	52.3	34
Trade union density (2006)	20.7	28.2
Unemployment rate	10.3	5.3
Labor force participation rate	75	76.8
Incidence of temporary employment (%)	14.5	5.8
Share of part-time workers on the working population (%)	21.8	23.1
After the crisis		
	Germany	United Kingdom
Employment Protection Legislation index (2013)	2.68	1.10
Public unemployment speding (as % of GDP)	1	0.3
Minimum wage (% of median full-time wage)	48	49
Tax wedge (as % of labor cost)	49.44	30.82
Trade union density (2013)	18.1	25.8
Unemployment rate	4.4	5.1
Labor force participation rate	77.6	77.6
Incidence of temporary employment (%)	13.1	6.2
Share of part-time workers on the working population (%)	22.4	24

This table provides a summary of the main indicators of labor market institutions and labor market performance in Germany and in the United Kingdom before and after the Great Recession (in 2006 and 2015 or latest available year). The source of the data is the OECD Employment outlook 2007 and 2016.

It is worth noting that the degree of strictness of EPL did not dramatically change in any of the two countries considered between 2006 and 2015. On the other hand, in 2015 both labor force participation and unemployment had recovered and bounced back to their pre-crisis levels in both countries (in the UK the unemployment rate was almost identical in 2006 and in 2015, yet it was slightly lower in 2015 than in 2006). However, in 2015 in the UK both the incidence of temporary employment and that of part-time work is higher than it was in 2006. In Germany, the incidence of temporary employment is lower in 2015 than it was in 2006. The Gini coefficient computed on wages is equal to 0.3 in the UK, whereas it equals 0.29 in Germany (OECD 2015). Eichhorst et al. (2010) find evidence that countries with a higher degree of labor maket flexibility fared better in terms of employment during and after the financial crisis. However, they suggest that such good results may have been obtained at the cost of a higher degree of polarization of the labor market, i.e. a stronger divide between well-paid and low-pais jobs. In this paper, we are interested in assessing what happened to the degree of individual positional mobility between 2002 and 2014.

Wage mobility in Germany and in the UK

In the previous literature, only a few studies tried to perform a cross-country comparison of wage mobility (exceptions are, for example, Gottschalk and Spolaore 2002 or Chen 2009). Moreover, those papers exclusively focus on the computation of some summary measures of wage inequality and wage mobility, without allowing the degree of wage mobility to change across the distribution. Atkinson et al. (1992 and 1998) evidenced the existence of national-level differences in earnings mobility patterns. At a first glance, one may argue that the lower the degree of Government intervention in the labor market, the higher the degree of wage mobility. However, Salverda et al. (2001) find only partial support for the hypothesis that the Anglo-Saxon labor market is characterized by higher wage mobility than the continental one. Similarily, Maasoumi and Trede (2001) and Aaberge et al. (2002) find that earnings mobility is higher in Germany and in the Scandinavian countries than in the US. On the contrary, Burkhauser et al. (1997) find remarkable similarity in the degree of wage mobility in Germany and in the US. Cardoso (2006) as well as Clark and Kanellopoulos (2013) find no evidence of a systematic relationship between the type of labor market institutions and the degree of low pay persistence. Hence, the question of the influence of the instutional framework on wage mobility is still open. It is reasonable to assume that different features of the labor market have distinct and perhaps countervailing effects on wage mobility. In the present work, we analyze wage mobility in Germany and in the UK; these countries are selected because they are among the largest economies in Europe and because they represent different labor market models. As mentioned above, the British labor market is characterized by high flexibility (Cardoso 2006, Dustmann and Pereira 2008, Chen 2009, Aristei and Perugini 2015), whereas Germany has been historically characterized by a high degree of labor market regulation (Abraham and Housman 1993, Siebert 1997, Burkhauser et al. 1997). Nevertheless, when relying on aggregate mobility indices, Germany appears to have a somewhat larger degree of income mobility than the UK (Ayala and Sastre 2007). The analysis of the impact of each labor market feature on wage mobility lies beyond the scope of the present paper. Our aim, instead, is to assess which groups are more or less mobile in each of the two countries and whether wage mobility changed after the recent financial crisis. We propose a methodological innovation, by applying a functional copula model[20] in which positional persistence is represented by an autoregressive function which depends on both the past wage rank and on some individual characteristics. Until now, wage mobility has mostly been studied by means of transition matrices; however, such matrices rely on some unrealistic assumptions, such as individual homogenity or time-invariant transition probabilities from one quantile to another. This is the reason why, in the present work, we will present such transition matrices only as exploratory data analysis. We build on the work by Bonhomme and Robin (2009), who use a Plackett copula to analyze the joint dynamics of the present and past wage ranks. In their model, however, changes in mobility across the wage distribution are determined by the parametric copula chosen, rather than by underlying economic dynamics. Previous studies often found evidence that the degree of wage mobility is not constant across the wage distribution (e.g. (Hungerford 1993, Jarvis and Jenkins 1995, Gottschalk 1997, Cardoso 2006, Raferzeder and Winter-Ebmer 2004, Pavlopoulos et Al. 2007, Germandt 2009). The impor-

[20]This model has been extensively described in Naguib and Gagliardini (2016).

tance of nonlinear modelling of wage dynamics has also been recently stressed by Altonji et al. (2013), Arellano and Bonhomme (2017) and Arellano et al. (2017). This further supports our choice of a semi-nonparametric model for the estimation.

Wage mobility before and after the financial crisis

To the best of our knowledge, the analysis of how wage mobility patterns changed during the Great Recession (2007-2009) has never been attempted in the previous literature and in general empirical evidence on the impact of a financial crisis on wage mobility is rather scarce. Aristei and Perugini (2015) and Fields et al. (2015), for example, only investigate the degree of earnings mobility in the pre-crisis period, whereas Gregg et al. (2015) do not distinguish between pre- and post-crisis period. Inequality was on the rise in the US well before the financial crisis broke out (Wisman 2013) and the UK (as well as many other advanced economies) witnessed a similar trend starting from the '70s (Bell and Van Reenen 2013). Moreover, Bell and Van Reenen (2013) find evidence that, in the UK, the Great Recession left the earnings of those at the very top of the distribution almost unchanged. On the other hand, Cockx and Ghirelli (2016) find evidence that financial crisis (2007-2009) seriously worsened the career prospects of young graduates in a rigid labor market (Belgium). According to Rinne and Zimmermann (2012 and 2013) and to Caliendo and Hogenacker (2012), the German model showed remarkable resilience during the Great Recession also thanks to the major labor market reforms that had been introduced in 2003-2005. Implemented in four waves, the Hartz reforms led to the restructuring and downsizing of unemployment benefits and social assistance schemes, whereas fixed-term contracts and agency work were massively deregulated. Consequently, German employment remained almost unaffected during the Great Recession, despite a sharp decline in GDP in 2008-2009. (Dustmann et al. 2014). On the contrary, employment dropped by about 2% in 2009-2010 in the UK (Rinne and Zimmermann 2013). Indeed, according to Eichhorst et al. (2010), the impact of the crisis in terms of employment rate varied much across different countries, irrespectively to GDP losses, due to differences in labor market institutions. Until now, emphasis has been given to the analysis of the impact of the Great Recession on the absolute wage level or on the employment rate (e.g. Chodorow-Reich 2013). To the best of our knowledge, a comprehensive analysis on the influence of the crisis on the degree of relative wage mobility for different groups of workers (defined by age, gender, education and nationality) has never been performed in the previous literature. The remaining of the present paper is structured as follows. Section 2 will present the model used for estimation, Section 3 will be devoted to a preliminary descriptive data analysis, whereas in Section 4 there will be the the description and comment of the estimation results and Section 5 will conclude.

3.2 The model

In this Section, we aim at summarizing the main theoretical results of Naguib and Gagliardini (2016). We are interested in the analysis of the dynamics of the joint process $(Z_{i,t}, W_{i,t})$, where $Z_{i,t}$ stands for the cross-sectional Gaussian individual rank in the wage distribution and $W_{i,t} = \beta' X_{i,t}$ is the (univariate) score, i.e. a weighted sum of the individual explanatory variables. In the following, N stands for the total number of individuals in the sample. The processes $\{(Z_{it}, W_{it}), t \in \mathbb{N}, i = 1, ..., N\}$ are i.i.d. across individuals. Let us define $Z_t = (Z_{1,t}, ..., Z_{N,t})$. We want to study the joint distribution of the Gaussian rank Z_{it} and of the (standardized) score W_{it}. The sample density is:

$$l(\underline{Z_T}, \underline{W_T}) = \prod_{i=1}^{N} l(\underline{Z_{i,T}}, \underline{W_{i,T}}) = \prod_{i=1}^{N} \prod_{t=1}^{T} l(Z_{i,t}, W_{i,t} | \underline{Z_{i,t-1}}, \underline{W_{i,t-1}}) \tag{3.1}$$

where $\underline{Z_{i,t-1}} = (Z_{i,t-1}, Z_{i,t-2}, Z_{i,t-3}, ...)$ and $\underline{Z_T} = (Z_T, Z_{T-1}, Z_{T-2}, ...)$. In the remaining of this section, for explanatory purpose, we omit the subscript i. We then assume Granger non-causality from Z to W, that W_t is Markovian of first-order and strictly stationary and that

$l(Z_t|\underline{W_t}, \underline{Z_{t-1}}) = l(Z_t|W_t, W_{t-1}, Z_{t-1})$. Note that, to ensure stationarity of the score, only strictly stationary explanatory variables X_t are included in it. These assumptions allow us to focus on the two conditional densities, which are the transition density of the rank given the score:

$$l(Z_t|Z_{t-1}, W_t, W_{t-1}) \tag{3.2}$$

and the transition density of the score:

$$l(W_t|W_{t-1}). \tag{3.3}$$

The second of these two transition laws is exogenously given, hence we will exclusively focus on the first one (equation (2)). The details of the derivation are presented in Naguib and Gagliardini (2016). These two laws define a stationary process for (Z_{it}, W_{it}), which will be characterized by a certain joint stationary law (implied by the model): $l(Z_t, W_t)$. The conclusion to be ensured is that the stationary density of the Gaussian ranks is standard normal:

$$Z_t \sim N(0, 1) \tag{3.4}$$

Let us suppose that $g(\cdot;\theta), \theta \in \Theta$, is a family of probability density functions (p.d.f.), and that $G(\cdot;\theta), \theta \in \Theta$ is the corresponding family of cumulative distribution functions (c.d.f.). Let $l(W_t)$ be the stationary distribution of W_t. Under the following two additional hypothesis:

$$\int g(Z;\theta(W))l(W)dW = \phi(Z) \tag{3.5}$$

and:

$$l(Z_t|Z_{t-1}, W_t, W_{t-1}) = g(Z_t;\theta(W_t))c[G(Z_t;\theta(W_t)), G(Z_{t-1};\theta(W_{t-1})); \rho(\cdot, W_t)], \tag{3.6}$$

where $c(\cdot;\cdot;\rho)$ is a generic copula density, ϕ is the standard normal density and $\theta(W)$ is a function of W, it is possible to show that equation (3.4) holds. The proof of this result is reported in Naguib and Gagliardini (2016). We conclude that, conditional on W_t, the copula of (Z_t, Z_{t-1}) is $c(\cdot;\cdot;\rho(W_t))$. Our results hold for any copula.

Until now we have shown how to specify a joint dynamics of rank Z_t and score W_t, the functions $\theta(\cdot)$ and $l(\cdot)$ which satisfy standard Gaussian distribution of the ranks by using a generic copula function. As shown by Naguib and Gagliardini (2016), a flexible nonparametric family of copula functions to be used in this setting can be written as follows. Let us consider the nonlinear autoregressive dynamics:

$$Z_t = \Lambda(\rho(Z_{t-1}) + \varepsilon_t) \tag{3.7}$$

where by hypothesis $\varepsilon_t \sim IIN(0,1)$, Λ is a strictly monotonic function and ρ is a function that expresses the dependence between the past and the present individual ranks. The larger the value of the partial derivative of the function $\rho(\cdot)$ with respect to the past rank, the higher the degree of positional persistence. W_t is the score, as defined above. Under the condition that $\Lambda(k)$ is such that

$$\Phi(\mathit{\Lambda}(k)) = \int_{-\infty}^{\infty} \Phi(k - \rho(Z_{t-1}))\phi(Z_{t-1})dZ_{t-1}, \tag{3.8}$$

the invariant distribution of Markov process (Z_t) is $N(0,1)$ (the proof is reported in Naguib and Gagliardini 2016). Equation (3.8) implies that Λ is a function of $\rho(\cdot)$. We can now derive the explicit copula density of Z_t and Z_{t-1}. The copula p.d.f. is:

$$c(u,v;\rho(\cdot)) = \frac{\phi[\mathit{\Lambda}^{-1}(\Phi^{-1}(u)) - \rho(\Phi^{-1}(v))]}{\phi(\Phi^{-1}(u))\lambda(\mathit{\Lambda}^{-1}(\Phi^{-1}(u)))}, \tag{3.9}$$

for the arguments $u, v \in [0,1]$. This copula family is parametrized by the autoregressive function $\rho(\cdot)$. Note that our results hold for any function $\rho(\cdot)$. Hence, we can easily introduce the score W_t: $\rho(\cdot) = \rho(\cdot, W_t)$ and $c(u,v;\rho(\cdot)) \equiv c(u,v,\rho(\cdot,W_t))$.

We now consider an adequate measure of positional mobility. We start from the conditional expectation:

$$E(Z_{it}|Z_{i,t-1}, W_{it}, W_{i,t-1}) \;=\; \int_{-\infty}^{\infty} G^{-1}\left(\tilde{\Lambda}[\tilde{\rho}(G(Z_{i,t-1};\theta(W_{i,t-1}));W_{it})+\varepsilon]\right)\phi(\varepsilon)d\varepsilon,$$

where $Z_{it} = G^{-1}(U_{it};\theta(W_{it})) \iff U_{it} = G(Z_{it};\theta(W_{it}))$, and this variable U_{it} follows the stochastic representation $U_{it} = \tilde{\Lambda}[\tilde{\rho}(U_{i,t-1};W_{it}) + \varepsilon_{it}]$ (see Naguib and Gagliardini (2016)), where $\tilde{\rho}(u;W_{it}) = \rho(\Phi^{-1}(u), W_{it})$ and $\tilde{\Lambda}(k;W_{it}) = \int_0^1 \Phi(k - \tilde{\rho}(v;W_{it}))dv$.

So that we can compute the partial derivative with respect to the past rank, which stands for positional mobility:

$$\frac{\partial E(Z_{it}|Z_{i,t-1}, W_{it}, W_{i,t-1})}{\partial Z_{i,t-1}} = \int_{-\infty}^{\infty} \frac{1}{g\left[G^{-1}\left[\tilde{\Lambda}[\tilde{\rho}(G(Z_{i,t-1};\theta(W_{i,t-1}));W_{it})+\varepsilon]\right]\right]}\cdot$$
$$\tilde{\Lambda}'[\tilde{\rho}(G(Z_{i,t-1};\theta(W_{i,t-1}));W_{it})+\varepsilon]\cdot\tilde{\rho}'(G(Z_{i,t-1};\theta(W_{i,t-1}));W_{it})\cdot g(Z_{i,t-1};\theta(W_{i,t-1}))\cdot\phi(\varepsilon)d\varepsilon. \tag{3.10}$$

3.3 Data and exploratory analysis

Our dependent variable is gross annual wage. Similarly to what has been done by Bonhomme and Robin (2009), we define individual annual wage as the sum of two components, a permanent one and a transitory one. In order to get the individual Gaussian ranks, we follow the same procedure that has been applied in Naguib and Gagliardini (2016), i.e. we first regress log earnings on age and the ratio of experience to age minus 15 (the threshold for active population) in order to purify earnings from the permanent component, and then we study the evolution over time of the transitory component alone. This preliminary estimation is perfomed via the panel data fixed-effect technique, in order to adequately take into account the potential presence of unobserved heterogeneity across workers. We also include in this regression a time fixed-effect (λ_t) in order to take into account all the time-fixed effects, among them also the impact of inflation on wages.

$$Wage_{i,t} = \beta_1 Age_{it} + \beta_2 \frac{Experience_{it}}{Age_{it} - 15} + \eta_i + \lambda_t + \epsilon_{i,t} \tag{3.11}$$

where $Wage_{i.t}$ stands for log earnings, $\frac{Experience_{it}}{Age_{it}-15}$ is the proportion of years of activity in labor market to age in excess of 15, η_i represents the individual fixed effect and $\epsilon_{i,t}$ is the transitory wage component. From the residuals of this preliminary regression, we obtain the Gaussian wage ranks via the following formula:

$$Z_{i,t} = \Phi^{-1}(\hat{F}_t(\hat{\epsilon}_{i,t})) \tag{3.12}$$

We then compute earnings ranks using the empirical cumulative distribution function (c.d.f.) of earnings residuals, $\hat{F}_t$, and we apply the quantile function of the standard normal distribution, Φ^{-1}, to impose standard gaussianity (Gottschalk 1997, Moffitt and Gottschalk 2002, Kalwij and Alessie 2007, Bonhomme and Robin 2009). Following Bonhomme and Robin (2009), we drop observations for students, retirees and self-employed worker. In addition we only include observations relative to full-time employees, in order to limit the role of variations in the intensive margin of labor supply on wage dynamics (Bonhomme and Robin 2009, Bachmann et al. 2016). We consider individuals aged between 15 and 64, since these are the usual thresholds for the definition of the active population.

Table 6: Summary statistics

	British data											
	2002	2003	2004	2005	2006	2007	2009	2010	2011	2012	2013	2014
Mean age	40	41	42	42	43	43	42	43	44	44	45	44
Mean log wage	9.821	9.859	9.913	9.943	9.979	10.015	10.083	10.100	10.129	10.138	10.199	10.206
Variance log wage	0.259	0.282	0.271	0.261	0.269	0.283	0.291	0.275	0.291	0.290	0.294	0.266
n. of obs.	11'018	11'018	9902	9336	6397	7979	6362	6362	4560	2199	2961	4164
	German data											
	2002	2003	2004	2005	2006	2007	2009	2010	2011	2012	2013	2014
Mean age	42	43	43	44	44	44	44	45	45	45	45	46
Mean log wage	10.407	10.394	10.422	10.424	10.433	10.446	10.443	10.456	10.497	10.513	10.533	10.555
Variance log wage	0.238	0.246	0.242	0.250	0.244	0.246	0.259	0.261	0.252	0.261	0.247	0.246
n. of obs.	5588	5588	5121	5067	4598	4671	4521	4521	4838	5709	5886	5504

This table reports some summary statistics for age and log wage for each of the two countries considered, both before and after the crisis.

We perform this analysis on two different data sources. For the United Kingdom, we consider data from the British Labor Force Survey (BLFS), whereas for Germany we take data from the German Socio-Economic Panel (GSOEP). Then, for both countries we construct two sub-samples, one for the pre-crisis period (2002-2007) and one for the post-crisis period (2009-2014). For the UK, the sample size ($n * T$) is equal to 43'448 observations for the pre-crisis period and 24'197 observations for the post-crisis period. In the German case, we have a sample size ($n * T$) equal to 25'037 for the first period and 26'449 for the second one. This division into sub-samples is aimed at understanding whether wage dynamics in any of the two countries considered changed after the recent financial crisis. Table 6 contains some descriptive summary statistics on the samples considered. As explorative data analysis, in the remaining of this Section we perform some preliminary estimates with methods traditionally used in mobility studies, such as transition matrices and multinomial logit models. However, as explained in the previous Section, we will then resort to a more flexible specification, in order to more precisely assess the degree of individual positional persistence, conditional on some explanatory variables. In Tables 7 and 8 we report the transition matrices obtained on German data for the years 2005-6 and 2010-11. We notice that the German labor market is characterized by a substantial degree of decile immobility, since most of the observations lie on the main diagonal of the matrix in both cases. Before the crisis (2005-6) it seems that it was somewhat easier to exit from the poorest decile, since 30% of people being there in 2002 were no more in the same decile in the following year. On the contrary, for the years 2010-11 we find that only about 20% of workers managed to escape from the lowest-paid decile of the distribution.

With reference to British data (Table 9 and 10) we find similar results. In both countries, indeed, the percentage of stayers, i.e. workers remaining in the same decile after a year, are far higher than those of people experiencing upward or downward mobility. Moreover, in both countries the percentage of stayers is highest in the bottom and in the upper deciles, whereas it is lower in the middle of the distribution. In the UK, we notice that the percentage of stayers in the bottom decile remained almost unchanged before and after the crisis, whereas the percentage of stayers in the highest decile slightly increased. However, as explained before, this type of analysis only provides a first insight on the dynamics of the labor market, but could hide intra-decile wage mobility. Indeed, with transition matrices it is not possible to obtain information on the individual rank dynamics.

Table 7: Decile transition matrix: Germany pre-crisis

2003 / 2002	1	2	3	4	5	6	7	8	9	10
1	**0.699**	0.200	0.042	0.025	0.019	0.006	0.000	0.006	0.000	0.002
2	0.091	**0.569**	0.196	0.085	0.028	0.016	0.014	0.000	0.000	0.000
3	0.045	0.103	**0.547**	0.196	0.067	0.028	0.006	0.003	0.006	0.000
4	0.038	0.046	0.172	**0.519**	0.133	0.040	0.022	0.020	0.006	0.002
5	0.015	0.031	0.070	0.227	**0.452**	0.128	0.046	0.015	0.015	0.000
6	0.009	0.016	0.037	0.096	0.163	**0.443**	0.145	0.072	0.012	0.007
7	0.002	0.011	0.009	0.028	0.082	0.194	**0.496**	0.144	0.024	0.011
8	0.004	0.000	0.004	0.011	0.033	0.059	0.206	**0.518**	0.140	0.024
9	0.004	0.000	0.000	0.009	0.004	0.017	0.039	0.141	**0.683**	0.102
10	0.000	0.000	0.000	0.002	0.000	0.009	0.007	0.013	0.120	**0.850**

This table reports the empirical transition probabilities from one decile to another between 2005 and 2006 in Germany. The sample size is n=4598.

Table 8: Decile transition matrix, Germany post-crisis

2011 / 2010	1	2	3	4	5	6	7	8	9	10
1	**0.808**	0.138	0.027	0.012	0.004	0.004	0.004	0.000	0.002	0.000
2	0.165	**0.531**	0.190	0.062	0.021	0.019	0.007	0.004	0.000	0.002
3	0.039	0.129	**0.449**	0.265	0.052	0.041	0.013	0.011	0.000	0.002
4	0.025	0.049	0.186	**0.482**	0.145	0.073	0.027	0.008	0.004	0.000
5	0.011	0.027	0.063	0.175	**0.389**	0.225	0.069	0.023	0.015	0.002
6	0.008	0.010	0.033	0.040	0.163	**0.443**	0.228	0.060	0.018	0.000
7	0.008	0.010	0.014	0.029	0.070	0.170	**0.482**	0.164	0.045	0.008
8	0.008	0.010	0.002	0.008	0.016	0.043	0.186	**0.482**	0.227	0.018
9	0.002	0.009	0.000	0.000	0.005	0.014	0.035	0.164	**0.647**	0.124
10	0.002	0.002	0.004	0.000	0.002	0.004	0.006	0.027	0.135	**0.818**

This table reports the empirical transition probabilities from one decile to another between 2010 and 2011 in Germany. The sample size is n=4100.

Table 9: Decile transition matrix, UK pre-crisis

2005 / 2006	1	2	3	4	5	6	7	8	9	10
1	**0.687**	0.176	0.039	0.031	0.019	0.006	0.006	0.011	0.006	0.019
2	0.219	**0.446**	0.205	0.078	0.027	0.009	0.007	0.001	0.003	0.004
3	0.067	0.217	**0.406**	0.198	0.067	0.024	0.013	0.006	0.002	0.002
4	0.033	0.094	0.181	**0.407**	0.176	0.067	0.020	0.015	0.005	0.003
5	0.019	0.034	0.079	0.165	**0.432**	0.177	0.065	0.018	0.010	0.002
6	0.008	0.017	0.033	0.075	0.166	**0.455**	0.160	0.056	0.022	0.008
7	0.006	0.004	0.015	0.031	0.060	0.189	**0.440**	0.189	0.051	0.016
8	0.004	0.006	0.004	0.004	0.022	0.048	0.157	**0.524**	0.199	0.031
9	0.007	0.002	0.005	0.009	0.015	0.022	0.042	0.168	**0.571**	0.159
10	0.002	0.002	0.002	0.002	0.008	0.006	0.017	0.041	0.144	**0.777**

This table reports the empirical transition probabilities from one decile to another between 2005 and 2006 in the UK (BLFS data). The sample size is n=6397.

Table 10: Decile transition matrix, UK post crisis

2011 / 2010	1	2	3	4	5	6	7	8	9	10
1	**0.661**	0.233	0.051	0.016	0.012	0.010	0.008	0.002	0.002	0.004
2	0.189	**0.511**	0.205	0.057	0.018	0.014	0.002	0.002	0.000	0.000
3	0.057	0.155	**0.514**	0.204	0.042	0.013	0.013	0.000	0.002	0.000
4	0.022	0.060	0.172	**0.507**	0.161	0.049	0.020	0.007	0.002	0.000
5	0.018	0.022	0.062	0.222	**0.422**	0.191	0.047	0.004	0.011	0.000
6	0.004	0.009	0.011	0.057	0.196	**0.510**	0.148	0.046	0.013	0.004
7	0.007	0.007	0.004	0.024	0.065	0.254	**0.457**	0.117	0.059	0.007
8	0.002	0.000	0.002	0.004	0.020	0.055	0.243	**0.517**	0.137	0.020
9	0.006	0.000	0.002	0.002	0.013	0.023	0.051	0.148	**0.582**	0.173
10	0.011	0.002	0.000	0.000	0.000	0.014	0.002	0.021	0.144	**0.805**

This table reports the empirical transition probabilities from one decile to another between 2010 and 2011 in the UK (BLFS data). The sample size is n=4560.

As a further exploratory analysis, we also run some multinomial logit estimates on both samples. We construct a dependent variable which takes value 1 in case of relative upward mobility, it is equal to -1 in case of relative downward mobility and zero otherwise. We define relative upward mobility as a situation in which the individual Gaussian rank in year t is greater than the Gaussian rank of the same individual in year $t-1$ and vice versa for relative downward mobility. The estimation results (reported in Appendix B) show that education and the past rank are relevant drivers of wage mobility. However, in these estimates we do not take into acount the size of the rank change (if we were to include all the possibile values of the rank change, then the algorithm would not converge). This is the reason why in the following we resort to the full model presented in the previous Section.

3.4 Estimation results

3.4.1 Estimation strategy

As mentioned above, we estimate a copula model in which mobility is represented by a function of the past rank and of some explanatory variables, in order to take explicitly into account the fact that the degree of mobility may differ across the wage distribution. The model is estimated in a semi-nonparametric way, with the method of sieves (Chen 2007). We assume that the data is independent across the i index (i.e. that different individuals are i.i.d.). As explanatory variables, we consider age, gender and a qualitative variable representing the highest education level achieved by the individual (Bonhomme and Robin 2009). We also include a dummy which takes value 1 if the individual is foreign-born. All these variables are strictly exogenous and stationary (age has been de-trended). Recall from Section 3.2 that $W_{it}^0 = X_{i,t}'\beta$, i.e. the score or index is a weghted sum of the individual explanatory variables. Let $G(Z_{i,t}; X_{i,t}'\beta_1)$ be the distribution of the rank, conditional on the individual variables, and $c[\cdot,\cdot,\rho(\cdot)]$ the copula density. The univariate conditional distributions are estimated via a kernel single-index model. We estimate the coefficients β_1 with a Maximum Likelihood approach:

$$argmax_{\beta_1} \sum_t \sum_i log\ \hat{g}_{-(it)}(Z_{i,t}|\beta_1' X_{it})\mathbf{I}_{\{Z_{it},X_{it}\}\in S}. \tag{3.13}$$

where $\hat{g}_{-(it)}(z|w)$ is the estimated conditional kernel density, $w_{jt}(\beta_1) = \beta_1' X_{jt}$, Z_{jt} represents the Gaussian rank and K is a standard Gaussian p.d.f.:

$$\hat{g}_{-(it)}(z|w;\beta_1) = \frac{\frac{1}{h}\sum_t \sum_{j,j\neq i} K\left(\frac{Z_{jt}-z}{h}\right) K\left(\frac{\hat{w}_{jt}(\beta_1)-w}{h}\right)}{\sum_t \sum_{j,j\neq i} K\left(\frac{\hat{w}_{jt}(\beta_1)-w}{h}\right)}. \tag{3.14}$$

With this procedure, we obtain an estimator of the betas that is consistent under some regularity conditions (see Chapter 2). By plugging in this estimate of the betas, the desired estimated c.d.f. can be easily recovered by integration. Let us now resort to the estimation of the copula function, $c(\cdot,\cdot,\rho(\cdot, X_{i,t}'\beta_2))$, which is performed via the following Sieve Maximum Likelihood procedure:

$$\max_{\theta\in\Theta_N} \sum_{i=1}^{N} l(Y_i,\theta) \tag{3.15}$$

$$l(Y_i,\theta) \equiv \sum_{t=1}^{T} logc[\hat{U}_{i,t}, \hat{U}_{i,t-1}, \rho(Z_{i,t-1}, X_{i,t}'\beta_2)] \tag{3.16}$$

where $Y_i = (Z_{it}, X_{it}, t = 1, ..., T)$. We have $\theta = (\beta_2, \rho) \in B_2 \times H \equiv \Theta$, $\Theta_N = B_2 \times H_N$ where H_N is a bivariate sieve space whose dimension depends on N. Moreover, $\hat{U}_{i,t} = \hat{G}(Z_{i,t}; X_{i,t}'\hat{\beta}_1)$, $\hat{U}_{i,t-1} = \hat{G}(Z_{i,t-1}; X_{i,t-1}'\hat{\beta}_1)$ are the estimated (empirical) c.d.f. of the present and of the past Gaussian ranks, conditional on the individual explanatory variables, and $c(u, v; \rho(\cdot, W_t))$ is the copula density presented in equation (3.9) of Section 3.2. We estimate the function $\rho(Z_{i,t-1}, X_{i,t}'\beta_2)$ non-parametrically via the method of sieves and λ is the vector of coefficients of the polynomial basis used to approximate $\rho(\cdot)$:

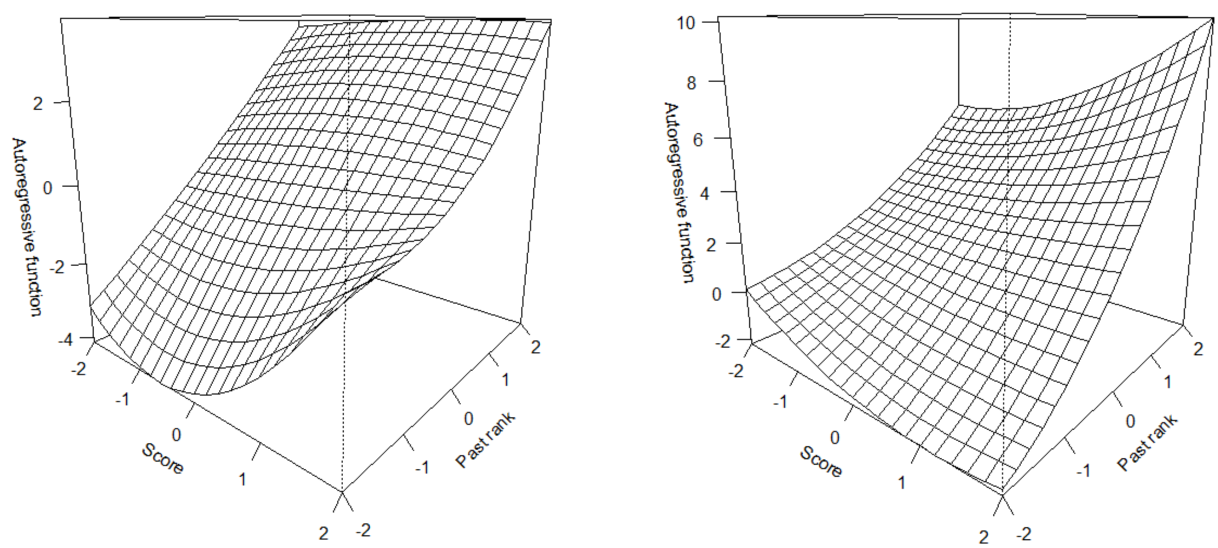
$$\rho(Z_{t-1}, W_t) \approx \sum_{k,l=0}^{m} \lambda_{k,l} H_k(\hat{W}_{it}) H_l(Z_{i,t-1}) \tag{3.17}$$

The number of Hermite polynomials used to approximate the autoregressive function depends on the dimension of the sample: $m = m(N)$ (Chen 2007). Convergence of this estimator has been proved in Gagliardini and Naguib (2016) for $N \to \infty$ and T fixed.

3.4.2 Estimation results on the German sample

As explained above, we apply the model presented in Section 3.2 to the German data (GSOEP) for the pre-crisis (2002-2007) and for the post-crisis period (2009-2014). We obtain the following estimates of the autoregressive function.

Figure 19: Constrained estimate of the function $\rho(Z_{i,t-1}, W_{i,t})$, Germany pre-crisis (left) and post-crisis (right)



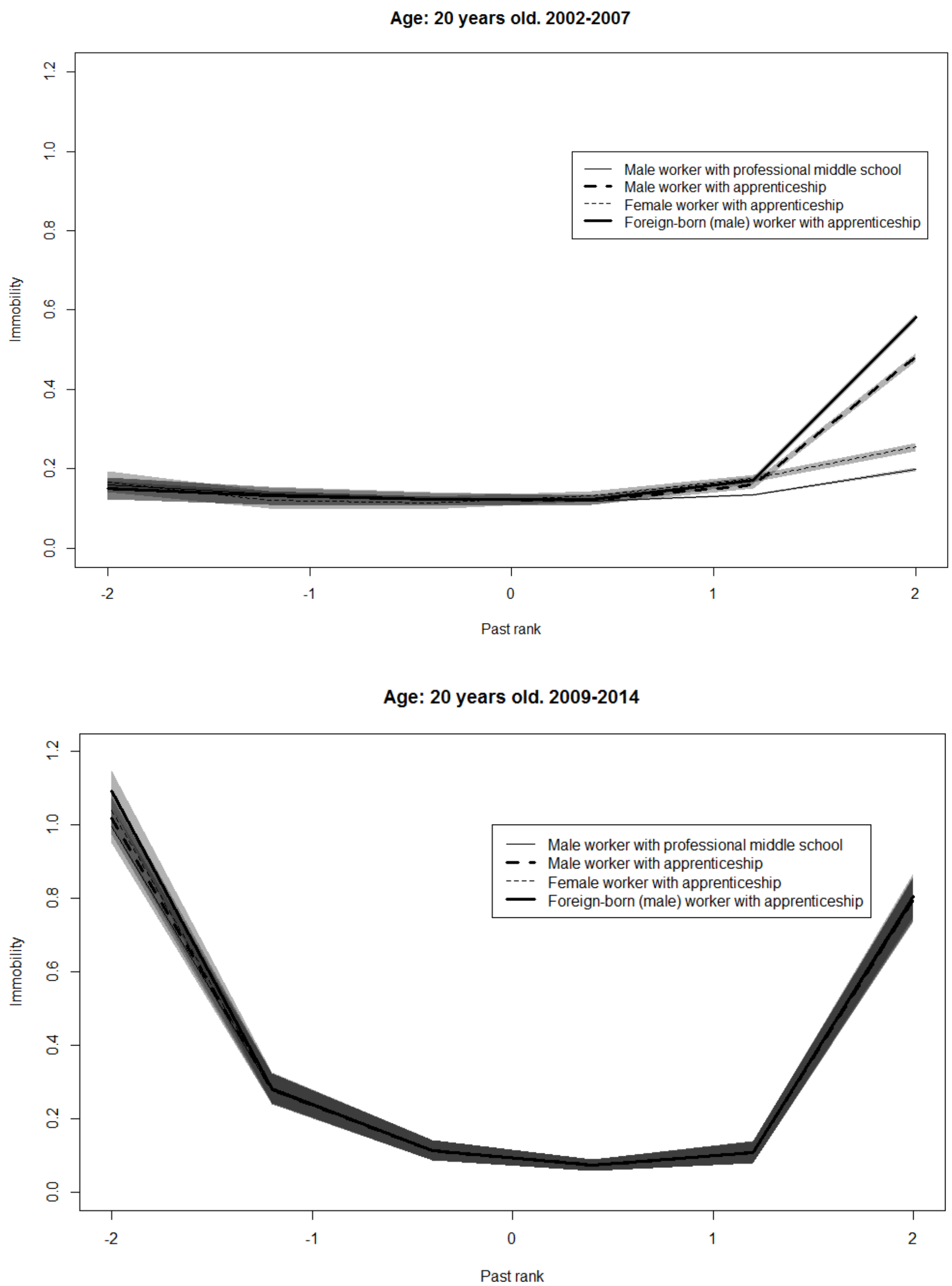
This figure shows the three-dimensional representation of the autoregressive function. This function has been approximated with a Hermite basis of degree 2.

The function $\rho(\cdot)$ is always increasing in the past rank in both periods, as expected; this means that, the higher the past rank, the higher is the expected present rank, conditional on the set of explanatory variables considered. From Figure 19, we deduce that the score plays a more relevant role in determining individual positions in the distribution after the crisis than before it. In the left side panel of Figure 19, indeed, we notice that the present Gaussian rank is only slightly increasing in the score in the lower part of the wage distribution. This means, for example, that low-paid younger workers are more likely to end up in a high rank than their older colleagues with the same past rank. Since different individual explanatory variables are likely to have a countervailing impact on the autoregressive function $\rho(\cdot)$ and on the degree of individual rank mobility, in the remaining of this Section we compute the mobility measure defined by equation (3.10) for different groups of workers, in order to be able to analyze the impact of each variable on the degree of positional mobilty. Each panel in Figures 20-21 represents the mobility measure presented in the previous Section in equation (3.10), for workers of the same age with different individual characteristics (education level, gender or the fact of being foregin-born), before and after the crisis. In Figures 20-21 and in Figures 23-24, the y-axis records indeed this mobility measure. The larger this quantity is, the higher is the association between the present and the past rank and hence the lower the degree of positional mobility is. this is the reason why the y-axis is labelled as 'immobility". Note that, since the mobility pattern of 40-year-old workers almost fully coincide, for all education levels, with those of 60-year-old workers, we omit the mobility pattern of the older group, in order to improve the readibility of the graphs. Moreover, only some education levels have been represented in the graphs for reasons of brevity and readibility. The summary of our findings on the mobility patterns for all the age and education groups both in Germany and in the UK is reported in Table 11 and in Table 12. For the pre-crisis period, we notice that positional mobility in Germany is decreasing in the past rank for individuals both in their 20s and in their 40s, almost regardless their education level. This finding suggests that young and middle-aged workers are not subject to the so-called low-wage trap. Even if they are currently occupying a rather low rank in the wage distribution, indeed, their degree of rank persistence is rather low and hence they have rather high probabilities of ending up in a higher rank in the following year; this is consistent with Clark and Kanellopoulos (2013), who find evidence of a low degree of low pay persistence in Germany for the period 1994-2001.

The higher degree of positional persistence at the upper end of the wage distribution suggests that, once a worker has reached a high position, the risk of experiencing downward mobility is moderate. Education does not seem to have a significant role in shielding workers from the low-wage trap, since the degree of positional mobility at the bottom of the distribution for workers who achieved different educational qualification is almost identical. We further notice that, among young workers, females and individuals who only completed professional middle school are more mobile at the top of the distribution. This suggests that they are more exposed to the risk of downward mobility, with respect to their male colleagues who achieved apprenticeship. However, this finding does not hold any more among workers in their 40s; on the contrary, middle-aged female workers are less mobile than their male colleagues at the top of the distribution. Lastly, we note that the fact of being foreign-born does not seem to significantly influence the individual patterns of positional mobility.

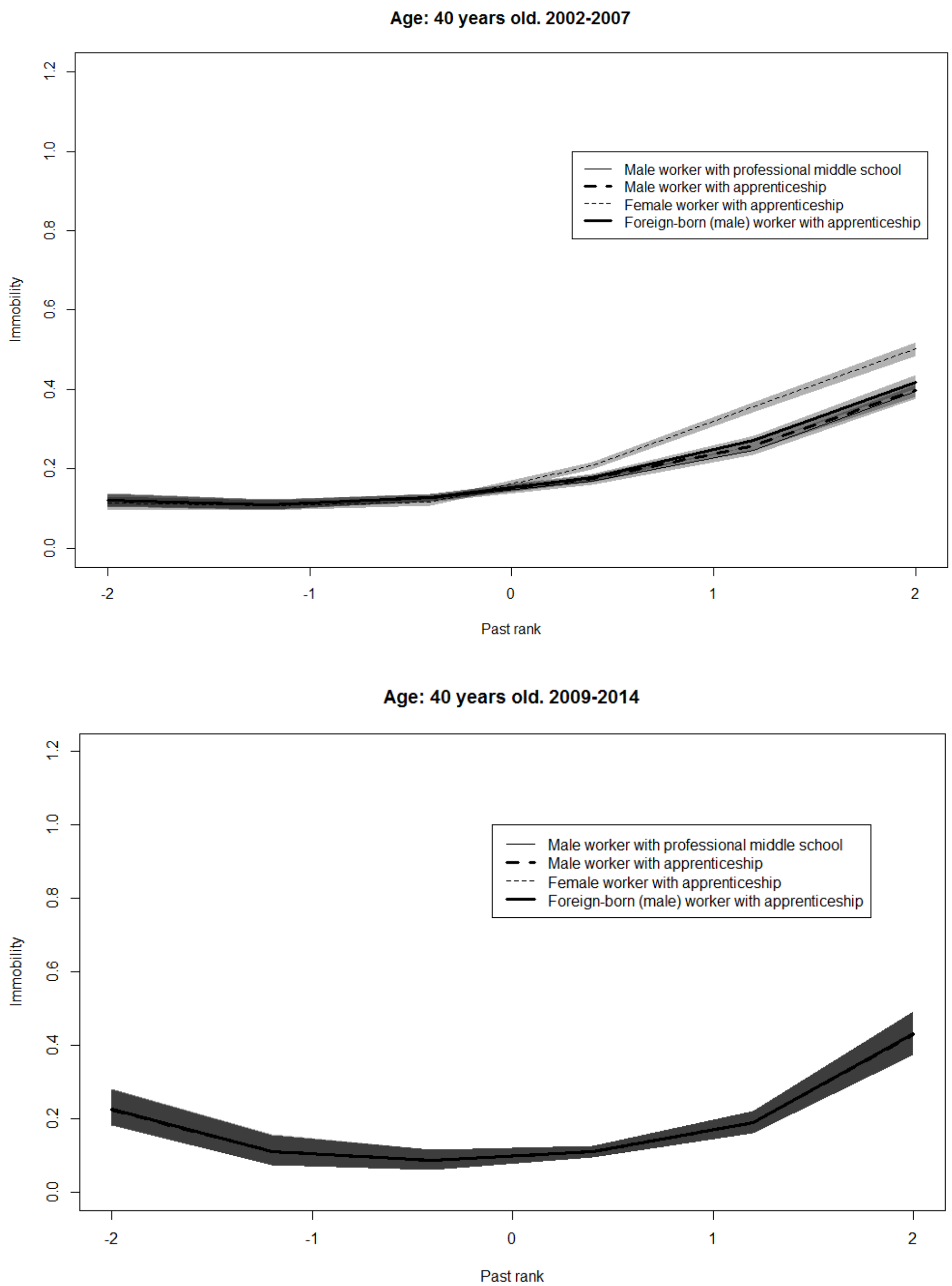
If we look at the estimation results after the crisis in Germany, we notice that education seems to have lost its role as a driver of positional mobility. For the post-crisis period in Germany, the mobility patterns of the different age groups for different education levels are rather close to each other and almost fully overlap. We now find that young workers (in their 20s) are far less mobile than their older colleagues at the bottom of the distribution, this suggests that this age group is exposed to the risk of being stuck in a low-wage trap after the crisis. Gender differences in the mobility patterns have disappeared after the crisis. Moreover, the fact of being foreign-born has lost its significant role in determining the degree of individual mobility. Considering the workers in their 40s and 60s, their degree of mobility appears as more stable across the distribution than it was before the crisis, i.e. the value of the partial derivative is almost constant for different values of the past rank. This suggests that, for these age groups, the importance of the past rank in determining individual mobility diminished after the financial crisis.

Figure 20: Mobility patterns, Germany pre-crisis (top panel) and post-crisis (bottom panel), 20-years old workers



In this figure we display the immobility measure defined in equation (3.10) for different groups of workers. On the x-axis the past rank is reported, whereas the y-axis measures rank (positional) immobility. If this partial derivative is low, this means that the association between present and past ranks is low and hence the degree of rank mobility is high. On the contrary, if the value of the partial derivative is high, this means that the association between the present and the past ranks is close, and hence the degree of rank mobility is low. The shaded areas represent the 95% pointwise confidence intervals for the partial derivatives and have been obtained by bootstrapping.

Figure 21: Mobility patterns, Germany pre-crisis (top panel) and post-crisis (bottom panel), 40-years old workers

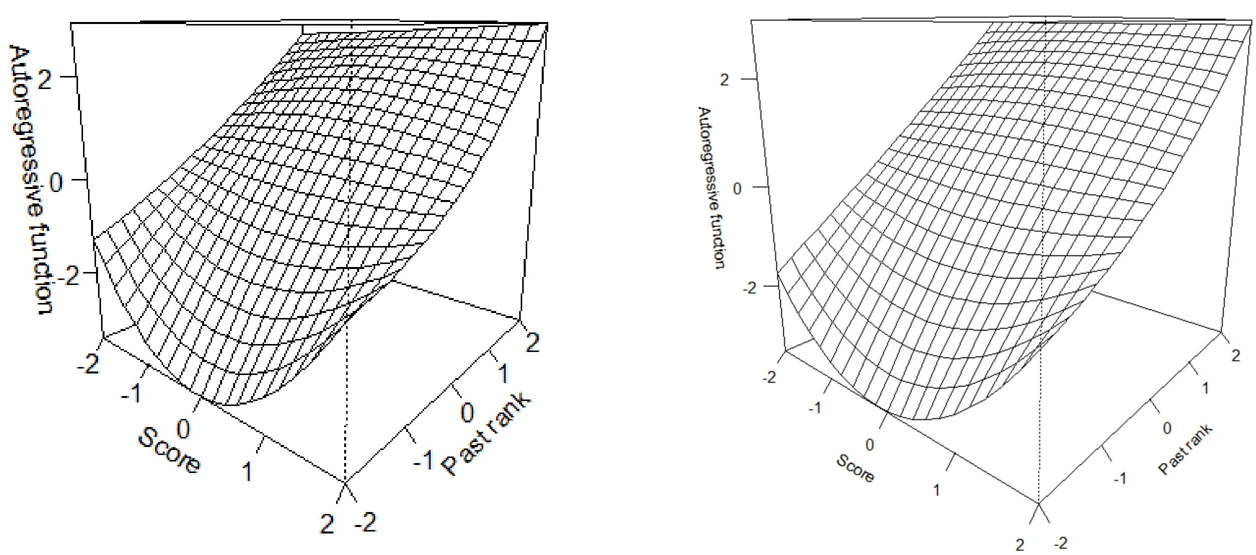


In this figure we display the immobility measure defined in equation (3.10) for different groups of workers. On the x-axis the past rank is reported, whereas the y-axis measures rank (positional) immobility. If this partial derivative is low, this means that the association between present and past ranks is low and hence the degree of rank mobility is high. On the contrary, if the value of the partial derivative is high, this means that the association between the present and the past ranks is close, and hence the degree of rank mobility is low. The shaded areas represent the 95% pointwise confidence intervals for the partial derivatives and have been obtained by bootstrapping.

3.4.3 Estimation results on the British sample

Similarly to what we did above for the case of Germany, in this Section we compare mobility patters before and after the Great Recession for different groups of workers in the UK. From Figure 22, we immediately notice that the shape of the autoregressive function differs from that obtained in the German sample. In the British case the estimated autoregressive function is closer to a flat surface, especially in the post-crisis period. In both countries, the autoregressive function is always increasing in the past rank. Moreover, in the UK this estimated function is U-shaped in the score for middle and low past ranks, whereas it is almost flat, regardless the value of the score, for the highest past ranks.

Figure 22: Constrained estimate of the function $\rho(Z_{i,t-1}, W_{i,t})$, United Kingdom pre-crisis (left) and post-crisis (right)



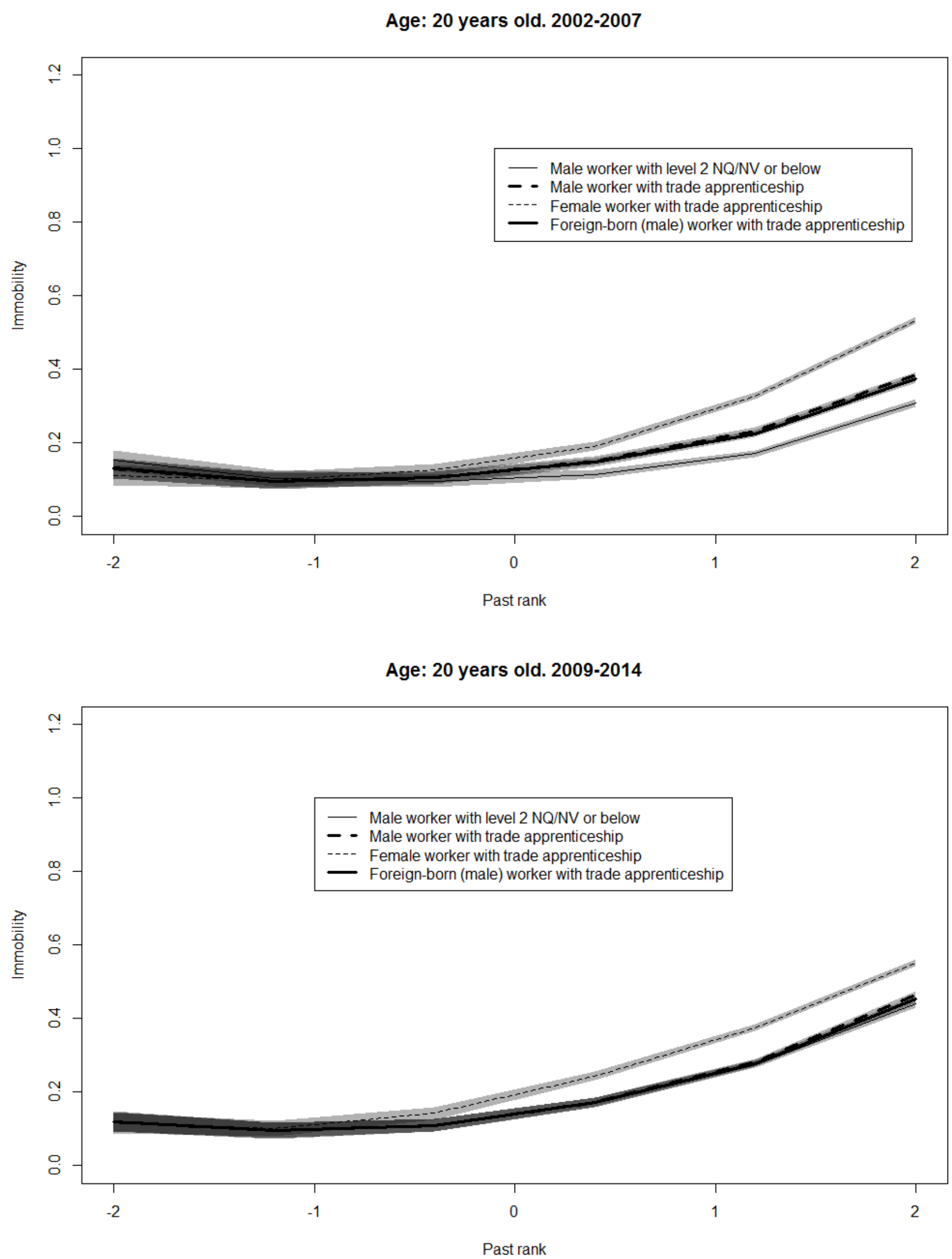
This figure shows the three-dimensional representation of the autoregressive function. This function has been approximated with a Hermite basis of degree 2.

We observe from Figure 22 that the shape of the function did not dramatically change after 2007. It seems that the impact of the crisis on positional mobility was smaller in the UK than in

Germany. In order to better understand the implications of this change in the shape of the auto-regressive function, let us consider, as it has been done above for Germany, the pattern of our mobility measure, as defined by equation (3.10). This measure, computed for different groups of individuals (defined by age, gender, education level and the fact of being foreign-born or not) before and after the crisis in the UK, is reported in Figures 23 - 24. Differently from the case of Germany, here we observe that age has almost no significant role in determining the degree of individual positional mobility. For all the education groups, the patterns of our mobility measure computed for different age groups often overlap. In the UK, mobility is always decreasing in the past rank and the degree of positional mobility is generally higher than what we found for Germany, in particular for workers in their 20s after the crisis. Neither education nor the fact of being foreign born play a relevant role in determining positional mobility in the UK and this holds both for the pre-crisis and the post-crisis period. We notice that female workers who completed apprenticeship are less mobile than their male colleagues in the middle-upper part of the distribution; this holds for all age groups, both before and after the crisis, and suggests that female employees face a lower risk of downward mobility than their male colleagues once they have reached a high position in the distribution. Before the crisis, young workers who only completed the lowest educational level were characterized by higher mobility at the top of the distribution and hence a higher risk of falling back in the wage distribution in the following years. However, this pattern disappeared after the crisis.

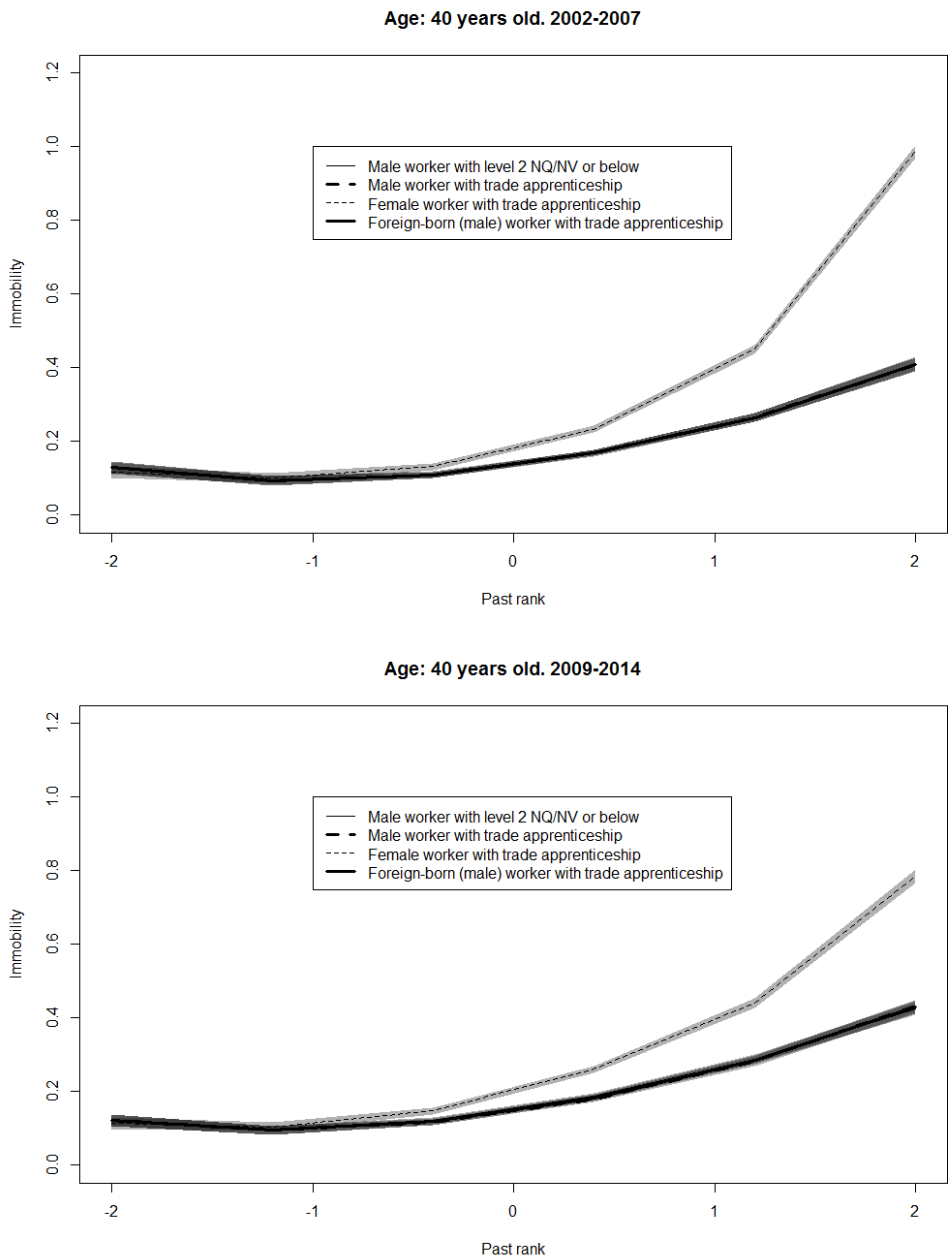
In Germany, the change in the mobility patterns following the crisis was more dramatic than in the UK. As mentioned in Section 3.1, indeed, the UK are characterized by a more flexible labor market than Germany, in particular with less strict hiring and firing rules. Hence, it is likely that, in the UK, adjustments in the aftermath of the financial crisis took place both in terms of wages and in terms of employment, with the result that the influence of the crisis on the degree of wage mobility was smaller in the UK than in Germany.

Figure 23: Mobility patterns, United Kingdom pre-crisis (top panel) and post-crisis (bottom panel), 20-years old workers



In this figure we display the immobility measure defined in (3.10) for different groups of workers. On the x-axis the past rank is reported, whereas the y-axis measures rank (positional) immobility. If this partial derivative is low, this means that the association between present and past ranks is low and hence the degree of rank mobility is high. On the contrary, if the value of the partial derivative is high, this means that the association between the present and the past ranks is close, and hence the degree of rank mobility is low. The shaded areas represent the 95% pointwise confidence intervals for the partial derivatives and have been obtained by bootstrapping.

Figure 24: Mobility patters, United Kingdom pre-crisis (top panel) and post-crisis (bottom panel), 40-years old workers



In this figure we display the immobility measure defined in equation (3.10) for different groups of workers. On the x-axis the past rank is reported, whereas the y-axis measures rank (positional) immobility. If this partial derivative is low, this means that the association between present and past ranks is low and hence the degree of rank mobility is high. On the contrary, if the value of the partial derivative is high, this means that the association between the present and the past ranks is close, and hence the degree of rank mobility is low. The shaded areas represent the 95% pointwise confidence intervals for the partial derivatives and have been obtained by bootstrapping.

On the other hand, in Germany wage was the main channel of adjustment to the new economic conditions, and this is the reason why we witness non-negligible changes in the wage mobility patterns before and after the crisis. In Germany, indeed, some groups of young workers witnessed a decrease in mobility in the low part of the distribution after the crisis (namely, workers in their 20), whereas the mobility degree of all the age groups remained almost unchanged after the crisis in the UK. Our findings support the hypothesis that the institutional context plays a fundamental role in determining the degree of relative wage mobility of different groups of workers in an economy and how the mobility patterns change after a shock. Our main findings are summarized in Table 11 and 12 below.

Table 11: Overview: the more mobile groups

	Before the crisis	After the crisis
Germany	20-year-old male workers with professional middle school	40-year-old workers with any education
UK	20-year-old male workers with level 2 NQ/NV or below	20-year-old male workers with any education

This table provides a summary of the groups that exhibit the highest positional mobility, i.e. the lowest values of our mobility measure defined in equation (3.10), in the two countries considered, before and after the crisis.

Table 12: Overview: the less mobile groups

	Before the crisis	After the crisis
Germany	20-year-old foreign-born workers with high past rank	20-year-old workers at the extremes of the distribution, any education
UK	40-year-old female workers with high past rank	40-year-old female workers with high past rank

This table provides a summary of the groups that exhibit the lowest positional mobility, i.e. the lowest values of our mobility measure defined in equation (3.10), in the two countries considered, before and after the crisis.

3.5 Concluding remarks

In this paper we analyzed wage mobility before and after the Great Recession in two European countries, Germany and the UK. We find evidence that the mobility patterns of different age and education groups of workers dramatically changed from 2002-7 to 2009-14 in Germany, but not in the UK. It emerges that, after the crisis, young workers being in a low rank were less mobile than their older colleagues in Germany, and that the crisis caused for them an increase in the risk of being stuck in the low ranks. On the other hand, the degree of mobility of all the age and education groups remained almost unchanged in the UK. The financial crisis had only a limited impact on positional mobility patterns in the UK, whereas it had stronger effects in Germany.

The finding of a higher degree of wage mobility in Germany than in the UK is consistent with the work of Shimer (2012). According to Shimer, indeed, if wages are rigid, the negative effect of a shock (here: the financial crisis) on employment will be more long-lasting. Germany is often taken as an example of a country whose employment rate fully bounced back after the crisis. A rather high degree of wage mobility may have contributed to make this recovery possible. With the innovative estimation method used, i.e. a functional copula estimated by means of semi-nonparametric methods, we are able to estimate the value of the autoregressive function which links the present and the past individual wage rank for each individual in every point of the wage distribution. As mentioned above, we notice that the shape of this autoregressive function, which is different in each country, notably changed in Germany after the crisis, whereas remained rather close to its previous shape in the UK. Further developments of this research may include a deeper analysis of some of the factors at the roots of cross-country differences in wage mobility, i.e. some specific features of the institutional framework, or the structure of the national pension system.

References

1. Aaberge, R., Bjorklund, A., Jantti, M., Palme, M., Pedersen, P. J., Smith, N., and Wennemo, T. (2002). Income inequality and income mobility in the Scandinavian countries compared to the United States. Review of Income and Wealth, 48(4), 443-469.

2. Abraham, K. G., & Houseman, S. N. (1993). Does employment protection inhibit labor market flexibility? Lessons from Germany, France, and Belgium (No. w4390). National Bureau of Economic Research.

3. Abramitzky, R., Boustan, L. P., & Eriksson, K. (2014). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. Journal of Political Economy, 122(3), 467-506.

4. Altonji, J. G., Smith, A. A., & Vidangos, I. (2013). Modeling earnings dynamics. Econometrica, 81(4), 1395-1454.

5. Arellano, M., Blundell, R., & Bonhomme, S. (2017). Earnings and consumption dynamics: a nonlinear panel data framework. Econometrica, 85(3), 693-734.

6. Arellano, M., & Bonhomme, S. (2017). Quantile Selection Models with an Application to Understanding Changes in Wage Inequality. Econometrica, 85(1), 1-28.

7. Aristei, D., & Perugini, C. (2015). The drivers of income mobility in Europe. Economic Systems, 39(2), 197-224.

8. Atkinson, A. B., Bourguignon, F., and Morrisson, C. (1992). Empirical studies of earnings mobility (Vol. 52). Taylor and Francis.

9. Atkinson, A. B., Bourguignon, F., & Morrisson, C. (1988). Earnings mobility. European Economic Review, 32(2-3), 619-632.

10. Ayala, L., & Sastre, M. (2008). The structure of income mobility: empirical evidence from five UE countries. Empirical Economics, 35(3), 451-473.

11. Aydemir, A., & Borjas, G. J. (2011). Attenuation bias in measuring the wage impact of immigration. Journal of Labor Economics, 29(1), 69-112.

12. Bachmann, R., Bechara, P., & Schaffner, S. (2016). Wage inequality and wage mobility in Europe. Review of Income and Wealth, 62(1), 181-197.

13. Bell, B. D., & Reenen, J. V. (2013). Extreme wage inequality: pay at the very top. The American Economic Review, 103(3), 153-157.

14. Bigard, A., Guillotin, Y., and Lucifora, C. (1998). Earnings mobility: an international comparison of Italy and France. Review of Income and Wealth, 44(4), 535-554.

15. Blanchard, O. J., Jaumotte, F., & Loungani, P. (2014). Labor market policies and IMF advice in advanced economies during the Great Recession. IZA Journal of Labor Policy, 3(1), 2.

16. Bonhomme, S., and Robin, J. M. (2009). Assessing the equalizing force of mobility using short panels: France, 1990 - 2000. The Review of Economic Studies, 76(1), 63-92.

17. Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. The quarterly journal of economics, 118(4), 1335-1374.

18. Bruecker, H., Hauptmann, A., Jahn, E. J., & Upward, R. (2014). Migration and imperfect labor markets: Theory and cross-country evidence from Denmark, Germany and the UK. European Economic Review, 66, 205-225.

19. Burkhauser, R. V., Holtz-Eakin, D., and Rhody, S. E. (1997). Labor Earnings Mobility and Inequality in the United States and Germany during the Growth Years of the 1980s. International Economic Review, 775-794.

20. Cadena, B. C. (2013). Native competition and low-skilled immigrant inflows. Journal of Human Resources, 48(4), 910-944.

21. Caliendo, M., & Hogenacker, J. (2012). The German labor market after the Great Recession: successful reforms and future challenges. IZA Journal of European Labor Studies, 1(1), 3.

22. Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. Journal of Labor Economics, 19(1), 22-64.

23. Cardoso, A. R. (2006). Wage mobility: do institutions make a difference?. Labour economics, 13(3), 387-404.

24. Chen, W. H. (2009). Cross-national differences in income mobility: Evidence from Canada, the United States, Great Britain and Germany. Review of Income and Wealth, 55(1), 75-100.

25. Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. Handbook of econometrics, 6, 5549-5632.

26. Chodorow-Reich, G. (2013). The employment effects of credit market disruptions: Firm-level evidence from the 2008-9 financial crisis. The Quarterly Journal of Economics, 129(1), 1-59.

27. Clark, K., & Kanellopoulos, N. C. (2013). Low pay persistence in europe. Labour Economics, 23, 122-134.

28. Cockx, B., & Ghirelli, C. (2016). Scars of recessions in a rigid labor market. Labour Economics, 41, 162-176.

29. Constant, A., & Massey, D. (2005). Labor Market Segmentation and the Earnings of German Guestworkers. Population Research & Policy Review, 24(5).

30. Dickens, R. 2000. Caught in a trap? Wage mobility in Great Britain: 1975-1994. Economica, 67, 477-197.

31. Dustmann, C., Schoenberg, U., & Stuhler, J. (2017). Labor supply shocks, native wages, and the adjustment of local employment. The Quarterly Journal of Economics, 132(1), 435-483.

32. Dustmann, C., Schoenberg, U., & Stuhler, J. (2016). The impact of immigration: Why do studies reach such different results?. The Journal of Economic Perspectives, 30(4), 31-56.

33. Dustmann, C., Fitzenberger, B., Schoenberg, U., & Spitz-Oener, A. (2014). From sick man of Europe to economic superstar: Germany's resurgent economy. The Journal of Economic Perspectives, 28(1), 167-188.

34. Dustmann, C., & Pereira, S. C. (2008). Wage growth and job mobility in the United Kingdom and Germany. Industrial & Labor Relations Review, 61(3), 374-393.

35. Eichhorst, W., Escudero, V., Marx, P., & Tobin, S. (2010). The impact of the crisis on employment and the role of labour market institutions. IZA working paper

36. Fields, G. S., Duval-Hernandez, R., Freije, S., & Puerta, M. L. S. (2015). Earnings mobility, inequality, and economic growth in Argentina, Mexico, and Venezuela. The Journal of Economic Inequality, 13(1), 103-128.

37. Friedman, M. (1962). Capitalism and Freedom (Chicago, 1962).

38. Gregg, P., Scutella, R., & Vittori, C. (2015). Individual Earnings Mobility and the Persistence of Earnings Inequalities in Australia. Economic Record, 91(292), 16-37.

39. Gottschalk, P. (1997). Inequality, income growth, and mobility: The basic facts. The Journal of Economic Perspectives, 11(2), 21-40.

40. Gottschalk, P., & Spolaore, E. (2002). On the evaluation of economic mobility. The Review of Economic Studies, 69(1), 191-208.

41. Germandt, J. (2009). Decreasing Wage Mobility in Germany. Discussion Paper no. 09-044. Mannheim Centre for European Economic Research (ZEW).

42. Glitz, A. (2012). The labor market impact of immigration: A quasi-experiment exploiting immigrant location rules in Germany. Journal of Labor Economics, 30(1), 175-213.

43. Holmlund, B. (2014). What do labor market institutions do?. Labour Economics, 30, 62-69.

44. Hungerford, T. L. (1993). US income mobility in the seventies and eighties. Review of Income and Wealth, 39(4), 403-417.

45. Jarvis, S., & Jenkins, S. (1998). How much income mobility is there in Britain?. The Economic Journal, 108(447), 428-443.

46. Jarvis, S. J., and Jenkins, S. P. (1995). Do the poor stay poor?: new evidence about income dynamics from the British Household Panel Survey. University of Essex.

47. Joe, H. (1997). Multivariate models and multivariate dependence concepts. CRC Press

48. Kalwij, A. S., & Alessie, R. (2007). Permanent and transitory wages of British men, 1975-2001: year, age and cohort effects. Journal of Applied Econometrics, 22(6), 1063-1093.

49. Kim, S. (2013). Wage mobility of foreign-born workers in the united states. Journal of Human Resources, 48(3), 628-658.

50. Maasoumi, E., & Trede, M. (2001). Comparing income mobility in Germany and the United States using generalized entropy mobility measures. Review of Economics and Statistics, 83(3), 551-559.

51. Mastromatteo, G., & Destefanis, S. (2009). Labor-market performance in the OECD. An assessment of recent evidence. The Levy Economics Institute of Bard College, 559(Aprile), 1-45.

52. Metcalf, D. (1999). The British national minimum wage. British Journal of Industrial Relations, 37(2), 171-201.

53. Naguib, C., Gagliardini, P. Wage Mobility: A Functional Copula Approach, unpublished manuscript, 2016.

54. Nelsen, R. B. (1999). An introduction to copulas (Vol. 139). Springer Science and Business Media.

55. OECD (2016), OECD Employment Outlook 2016, OECD Publishing, Paris.

56. OECD (2015), OECD Employment Outlook 2015, OECD Publishing, Paris.

57. OECD (2007), OECD Employment Outlook 2007, OECD Publishing, Paris.

58. Olney, W. W. (2015). Remittances and the wage impact of immigration. Journal of Human Resources, 50(3), 694-727.

59. Patel, K., & Vella, F. (2013). Immigrant networks and their implications for occupational choice and wages. Review of Economics and Statistics, 95(4), 1249-1277

60. Pavlopoulos, D. et al. 2007. Job Mobility and Wage Mobility of High- and Low-Paid Workers. Schmollers Jahrbuch, 127, 47-58, Berlin.

61. Peri, G. (2016). Immigrants, productivity, and labor markets. The Journal of Economic Perspectives, 30(4), 3-29.

62. Peri, G. (2012). The effect of immigration on productivity: Evidence from US states. Review of Economics and Statistics, 94(1), 348-358.

63. Raferzeder, T., and Winter-Ebmer, R. (2007). Who is on the rise in Austria: Wage mobility and mobility risk. The Journal of Economic Inequality, 5(1), 39-51.

64. Rinne, U., & Zimmermann, K. F. (2013). Is Germany the north star of labor market policy?. IMF Economic Review, 61(4), 702-729.

65. Rinne, U., & Zimmermann, K. F. (2012). Another economic miracle? The German labor market and the Great Recession. IZA Journal of Labor Policy, 1(1), 3.

66. Salverda, Wiemer, et al. "The European-American employment gap, wage inequality, earnings mobility and skill: A study for France, Germany, the Netherlands, the United Kingdom and the United States. Final report." (2001). European Commission, Employment and social affairs.

67. Shimer, R. (2012). Wage rigidities and jobless recoveries. Journal of Monetary Economics, 59, S65-S77.

68. Shorrocks, A. F. (1978). The measurement of mobility. Econometrica: Journal of the Econometric Society, 1013-1024.

69. Siebert, H. (1997). Labor market rigidities: at the root of unemployment in Europe. The Journal of Economic Perspectives, 11(3), 37-54. ISO 690

70. Sklar, A. (1959). Fonctions de repartition a n dimensions et leurs marges. Inst. Statist. Univ. Paris Publ. 8 229-231

71. Smith, C. L. (2012). The impact of low-skilled immigration on the youth labor market. Journal of Labor Economics, 30(1), 55-89.

72. Sologon, D. M., and O'Donoghue, C. (2011). Shaping Earnings Mobility: Policy and Institutional Factors. The European Journal of Comparative Economics, 8(2), 175.

73. Wisman, J. D. (2013). Wage stagnation, rising inequality and the financial crisis of 2008. Cambridge Journal of Economics, 37(4), 921-945.

Appendix A Copula models[21]

Joe (1997) and Nelsen (1999) provide extensive surveys on copula theory. A copula function couples marginal distributions to get a joint distribution. By definition, a joint cumulative function C on $[0;1]^2$ with uniform marginal distributions on $[0,1]$ is a copula. Thus, a function $C : [0;1]^2 \to [0,1]$ is a copula if the following conditions hold:

$$C(0,v) = C(u,0) = 0, \forall u, v \in [0,1] \tag{3.18}$$

$$C(1,v) = v, C(u,1) = u, \forall u, v \in [0,1] \tag{3.19}$$

[21] This Appendix heavily relies on Naguib and Gagliardini (2016), Appendix D.

For any rectangle $R = [u_1, u_2] \times [v_1, v_2] \subset [0; 1]^2$ the following holds:

$$\int\int C(du, dv) = C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0. \tag{3.20}$$

When the distribution C is continuous, the associated density is called the copula density:

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v},\ u, v \in [0, 1]. \tag{3.21}$$

Sklar's theorem (1959)

Let $F(x, y)$ be the bivariate c.d.f. of random variables X and Y, with marginal c.d.f. equal to F_x and F_y. Let then $U = F_x(X)$ and $V = F_y(Y)$. The joint c.d.f. is given by:

$$P(U \leq u, V \leq v) = P(X \leq F_x^{-1}(U), Y \leq F_y^{-1}(V)) = F(F_x^{-1}(u), F_y^{-1}(v)),\ \forall u, v \in [0, 1] \tag{3.22}$$

Then there exists a copula such that:

$$C(u, v) = F(F_x^{-1}(u), F_y^{-1}(v)), \forall u, v \in [0, 1] \tag{3.23}$$

i.e.$F(x, y) = C[F_x(x), F_Y(y)]\ \forall x, y$.

This copula is unique if F is a continuous distribution. Assuming sufficient regularity, it is possible to obtain the following conditional c.d.f.:

$$P(U \leq u \mid V = v) = \partial C(u, v)/\partial v \tag{3.24}$$

The conditional density function is obtained by deriving once more with respect to u. The copula is limited by the so-called Frechet-Hoeffding bounds; the upper bound is: $C_U(u, v) = max(u+v-1; 0)$, while the lower bound is: $C_L(u, v) = min(u, v)$. In the case of independence we have $C_I(u, v) = u \cdot v$. There are many different families of copulas, for example the Gaussian, the Frank or the Gumbel.

Appendix B Additional estimates

Table 13: Multinomial logistic regression, Germany

Downward mobility	2002-2007	2009-2014
Age	.0320 (.1022)	-.0646 (.0662)
Age^2	-.0007 (.0011)	.0010 (.0012)
Female dummy	.2622 (.2646)	-.4376* (.2408)
High school diploma	-.9792*** (.3077)	-.3975 (.2866)
Professional high school diploma	-.4523 (.5112)	-.1496 (.4351)
Apprenticeship	-.3704 (.3037)	-.2107 (.3185)
Foreign-born	.2533 (.7293)	-.3050 (.7278)
Past Gaussian rank	.3688*** (.1302)	-.1061 (.1392)
Constant	5.0891** (2.3064)	7.1340*** (2.0710)
Upward mobility	2002-2007	2009-2014
Age	.0712 (.1022)	-.0417 (.0663)
Age^2	-.0012 (.0011)	.0003 (.0012)
Female dummy	.0902 (.2646)	-.5668** (.2410)
High school diploma	-.5782* (.3077)	-.0759 (.2869)
Professional high school diploma	-.1811 (.5114)	.0880 (.4355)
Apprenticeship	-.5027* (.3038)	-.1766 (.3188)
Foreign-born	.1073 (.7298)	-.1470 (.7277)
Past Gaussian rank	-.0604 (.1302)	-.4534*** (.1394)
Constant	4.2404* (2.3073)	6.4557 (2.0736)
LR chi2(16)	584.44	338.63
Log likelihood	-12063.129	-11875.137
Prob > chi2	0.0000	0.0000
N. obs.	25'037	26'449

This table reports the results of a multinomial logistic regression in which the dependent variable is the change in the Gaussian wage rank. This variable takes value 1 if the present rank is higher than the past rank, value -1 if the present rank is lower than the past rank and value zero if the two ranks are identical. This estimate has been performed on GSOEP data, seprately for the pre-crisis period (left column) and for the post-crisis period (right column).

Table 14: Multinomial logistic regression, United Kingdom

Downward mobility	2002-2007	2009-2014
Age	-.0252 (.1399)	.0080 (.2050)
Age^2	-.0002 (.0017)	-.1693 (.3715)
Female dummy	-.0806 (.3335)	-.1693 (.3715)
Level 3 NQ/NV	.2462 (.5324)	.6431 (.6562)
Trade apprenticeship	.3916 (.7694)	.4064 (1.0575)
Level 4 NQ/NV and above	-.1841 (.3854)	.0210 (.4344)
Foreign-born	-.5466 (.5283)	-.0459 (.4268)
Past Gaussian rank	-.3081* (.1787)	-.0374 (.2050)
Constant	7.8275*** (2.8484)	6.8064** (3.3097)
Upward mobility	2002-2007	2009-2014
Age	-.3839*** (.1398)	-.3285** (.1574)
Age^2	.0050*** (.0017)	.0043** (.0018)
Female dummy	-.1817 (.3336)	-.2676 (.3716)
Level 3 NQ/NV	.3362 (.5326)	.6932 (.6563)
Trade apprenticeship	.3974 (.7695)	.3973 (1.0574)
Level 4 NQ/NV and above	.1964 (.3856)	.3115 (.4346)
Foreign-born	-.5908 (.5288)	-.0520 (.4269)
Past Gaussian rank	-.7720*** (.1789)	-.4408** (.2051)
Constant	13.1095*** (2.8481)	11.4767*** (3.3091)
LR chi2(16)	8350.95	5193.79
Log likelihood	-26230.765	-14357.139
Prob > chi2	0.0000	0.0000
N. obs.	43'448	24'197

This table reports the results of a multinomial logistic regression in which the dependent variable is the change in the Gaussian wage rank. This variable takes value 1 if the present rank is higher than the past rank, value -1 if the present rank is lower than the past rank and value zero if the two ranks are identical. This estimate has been performed on BLFS data, seprately for the pre-crisis period (left column) and for the post-crisis period (right column).

4 Rank matters: an analysis of relative wage mobility as a neglected determinant of job satisfaction

4.1 Introduction

[22] Job satisfaction is an important variable in social sciences, since it is a good predictor of individual well-being (Cabral Vieira 2005) and it is positively related to worker's productivity (Freeman 1977). The analysis of the factors determining the employee's level of job satisfaction have been widely studied by the past literature (e.g. Millan et al. 2013). However, some specific aspects still need investigation.

The aim of the present paper is to assess the influence of rank or positional wage mobility on individual job satisfaction. Whereas in the past literature the role of the absolute wage level as a driver of individual well-being or job satisfaction has often been put under scrutiny (see e.g. Easterlin 1995, Clark 1999, Clark and Oswald 1996, Judge et al. 2010), on the contrary economists have almost never included income ranks among the arguments of the utility function. Nevertheless, human beings are deeply interested in rankings (Brown et al. 2008). In this paper we want to put under test the hypothesis that workers care about their degree of relative (rank) wage mobility, i.e. that their job satisfaction is directly influenced by their prospects of going up or down in the wage distribution over time.

Some researchers have tried to investigate the impact of changes in the absolute wage level on subjective well-being or on individual job satisfaction (for example see Diener et al. (1993), Clark and Oswald (1996), Clark (1999), Leontaridi and Sloane (2004), Di Tella et al. (2010)) and have found sometimes contrasting results. For example, Diener et al. (1993) finds no sound econometric evidence of an influence of social comparison based on absolute wage changes in the determination of individual well-being. On the contrary, Clark and Oswald (1996) claim that such comparison effects exist and are highly relevant. However, in these studies wage

[22] We would like to thank Thierry Kamionka and Sebastien Roux (CREST, Paris) for their insightful advice.

mobility has always been defined uniquely as the change in the absolute level of pay, i.e. they only considered absolute wage mobility.

In the literature, there are two main approaches to wage mobility; the first one considers the growth rate of wages and salaries from one year to the following one (absolute mobility). The second one, instead, refers to the concept of relative or positional mobility. Suppose that all workers in a given country and in a given year are ranked according to their gross annual wage; the study of their relative mobility is the analysis of how their individual position (rank) in the distribution, i.e. with respect to the other workers, changes across time (Jantti and Jenkins 2013). Just to make an example, it is possible for a worker to experience absolute wage mobility, without experiencing, at the same time, relative wage mobility. This happens when all wages rise or decline in the same proportion. Rank is usually regarded as a fundamental variable in fields such as sport economics (see for example Macmillan and Smith, 2007) or education economics (e.g. for the university rankings, see Marginson and Van der Wende 2007). The importance of the income rank for individual well-being, despite being recognized since the seminal work of Parducci (1965) in the field of cognitive psychology (Brown et al. 2008, Boyce et al. 2010), has been until now rather neglected in the economic literature on job satisfaction. According to the range frequency theory, developed by Parducci in 1965, satisfaction will be predicted partly by the ordinal position of a wage within a comparison set, i.e. by the individual rank. Therefore, we expect that there is a significant association between relative wage mobility and job satisfaction. On the other side, we expect absolute wage mobility to have only a limited role in determining the degree of individual satisfaction.

To the best of our knowledge, rank (or relative) wage mobility has never been included among the determinants of job satisfaction in the past literature. Indeed, the analysis of the individual rank as a determinant of individual (life) satisfaction has remained until now almost in the exclusive domain of psychology. On the contrary, it has been rather neglected in the fields of labor economics. A first preliminary test of the range frequency theory has been performed by Brown et al. (2008). We extend their work in the following ways; first, the authors only consider the static, point-in-time individual rank in the wage distribution among the explanatory variables of individual satisfaction. We build on their approach by including different measures of relative wage mobility among the determinants of job satisfaction. Moreover, we extend the work by Brown et al. (2008) by using national survey data instead of experimental data, and thus having the opportunity of investigating the influence of relative wage mobility on job satisfaction on a larger scale, by considering a comprehensive measure of job satisfaction as our dependent variable and by employing panel data instead of a cross-section. Note that, when talking about relative wage mobility, we are not referring to the theory of relative or reference income. By relative or positional wage mobility, indeed, we mean the change in the individual ranks in the wage distribution from one year to the following one, as it will be explained in greater detail in Section 4.2. On the other hand, the relative or reference income theory claims that individual well-being not only depends on her absolute wage level, but also on her wage compared to some reference income measure (i.e. the average wage in a group of friends or colleagues, see e.g. Clark and Oswald (1996), Stutzer (2004), Ferrer-i-Carbonell (2005)). An empirical test of the reference income theory lies beyond the scope of the present paper. Instead, our focus consists in testing the range frequency theory, i.e. the importance of rank and in particular of rank changes (which we consider as a measure of relative wage mobility), as an alternative explanation of the link between wage and job satisfaction. Wage mobility may affect job satisfaction in two ways: first, since people are generally risk-averse, they prefer to earn a stable income over time. Therefore, the higher wage mobility is, the lower reported job satisfaction will be; this is the so called wage insecurity aspect (Fachinger and Himmelreicher 2010, Gottschalk and Spolaore 2002). On the other hand, wage mobility is often linked to the concept of equality of opportunity. If wage mobility in the lowest ranks of the distribution is high, low-paid workers are able to improve their positions from one year to the following one (Friedman 1962, Clark 2003, Clark et al. 2009). The degree of individual tolerance for current levels of income inequality will be higher, since workers will expect a future improvement in their relative income position in the distribution (Bjornskov et al. 2013). Hence, in principle, the sign of the impact of wage mobility on job satisfaction may be either positive or negative. This is the reason why it makes sense to include both absolute and relative wage mobility among the determinants of job satisfaction. The remainder of the present paper is organized as follows: Section 2 will present the data and the estimation methods used. Section 3 will be devoted to the presentation and discussion of the estimation results and Section 4 will conclude.

4.2 Methods

4.2.1 Model

Our dependent variable is indivudual self-assessed job satisfaction. In the previous literature, much attention has been devoted to the analysis of the influence of income on subjective well-being or on so-called general life satisfaction (e.g. Diener 1993, Clark and Oswald 1996, McBride 2001, Ferrer-i-Carbonell 2005). However, individual well-being is a multifaceted concept, which includes several aspects such as satisfaction with own health, job satisfaction or satisfaction with free time. General life satisfaction is a variable that has more informative value for the psychologists than for the economists. On the contrary, since the seminal work of Freeman (1978), the importance of job satisfaction in the economic analysis has been established. Indeed, the previous literature has found a clear negative relationship between the level of employee job satisfaction and the turnover rate within a firm (e.g. among the most recent, Brunetto et al. 2010 and Yucel 2012). Moreover, the existence of a negative and statistically significant relationship between the turnover rate and the firm profitability has also been proved (Staw 1980, Mobley 1982, Mowday et al. 1982, Darmon 1990, Hom and Griffeth 1995, Ton and

Huckman 2008). Therefore, firms have a strong interest in the study of the influence of wage mobility on job satisfaction, since the latter variable is directly associated with their profitability. In addition, by considering a broad concept such as general life satisfaction as the dependent variable there is a rather high risk of incurring in omitted-variable bias or endogeneity (for example, health status can both influence and be influenced by subjective well-being). On the other hand, the main drivers of job satisfaction, a narrower idea than life satisfaction, have been largely investigated and identified by the economic literature (e.g. Hamermesh (1977), Borjas (1979), Akerlof, Rose and Yellen (1988), Clark (1999)), hence the risk of omitted-variable bias should be lower. Moreover, psychologists claim that utility from work is one of the most important parts which constitute an individual's utility (Clark 1999). Therefore, in the present paper we will assume separability of the utility function in its work and non-work-related determinants and we will focus on the first group. Following the previous literature on the theme of job satisfaction, we will estimate a model of the form:

$$JS_{it} = \beta' X_{it} + \alpha WageMobility_{it} + \varepsilon_{it} \tag{4.1}$$

Where JS_{it} stands for reported job satisfaction, X_{it} is a vector of individual or job-related characteristics, such as age, gender, civil status, percentage of occupation or firm size and sector, whereas $WageMobility_{it}$ is one of our measures of wage mobility, which will be defined in the remainder of this Section, β is the vector of coefficients of the control variables (to be estimated), α is our main estimating coefficient of interest, and $\varepsilon_{it} \sim N(0,1)$ is the classical error term. As mentioned above, we aim at analyzing the impact of wage mobility on job satisfaction. We consider two kinds of earnings mobility measures; the first one refers to absolute mobility and is defined as the individual growth rate of wage from one year to the following one. Our second measure refers to the concept of relative or positional mobility (Jantti and Jerkins 2013); for each year, we rank individuals in the sample according to their wage. Their position in the wage distribution in a given year is called rank; by dividing this number by the sample size in that year, we find the uniform rank (U_{it}), i.e. an index that is uniformly distributed between zero and one. Finally, we apply to the uniform ranks the quantile function of the standard normal distribution, in order to obtain the Gaussian ranks, Z_{it}:

$$Z_{it} = \Phi^{-1}(U_{it}) \tag{4.2}$$

This last transformation is performed for consistency with the previous two chapters. The higher the value of the Gaussian rank, the higher the position occupied by the individual in the sample wage distribution in a given year. We define relative wage mobility as the change in the individual Gaussian rank from one year to the following one:

$$RelativeWageMobility_{it} = \Delta Z_{it} = Z_{it} - Z_{it-1} \tag{4.3}$$

Given the ordinal nature of our dependent variable and in the wake of the previous literature on job satisfaction, the estimation method employed is the random-effects ordered probit.

4.2.2 Data

We use data from the German Socio-Economic Panel (GSOEP) for the years 1985-2012; the total number of observations in this unbalanced panel is equal to 164'058. We choose to analyze the German labor market, since Germany is the largest economy in continental Europe. As usual in the literature, we eliminate all the observations with zero annual income (in addition to all those for which information on wages was missing); we trim data on the basis of age, so that only observations of people between 15 and 64 years of age are included (these are the usual thresholds to describe active population). We further exclude students, apprentices, self-employed individuals and people who are reported as "working only irregularly" or unemployed. Note that in the definition of individual income we do not include bonuses, nor other forms of windfall earnings. In the wake of the past literature, we consider as determinants of job

satisfaction the following explanatory variables: the logarithm of annual individual labor earnings ("wage, salaries from main job"), annual work hours, work experience (as measured by the number of years spent working full time), gender, age, age squared, education as measured by the number of years spent in school, marital status, the number of children in the household, the State of residence (Bundesland), change of job[23], employment level of the individual (part-time dummy), size of the firm, sector (NACE) and type of occupation (with more or less decisional autonomy). This last variable takes values from 1 ("No decisional autonomy at all") to 5, which stands for the largest possible degree of decisional autonomy at the workplace.
Table 15 below reports some descriptive statistics of the main variables included in the model. All the variables related to individual satisfaction are measured and recorded in the GSOEP dataset with a scale ranging from zero (not satisfied at all) to ten (fully satisfied), which makes the use of the ordered model a natural choice for the estimation.

[23]Of course, the decision of quitting a job is influenced by the degree of individual job satisfaction, hence the variable standing for job change may be in principle endogenous (if the change of job is voluntary). In order to solve this problem, we take a lag of this variable, i.e. our explanatory variable stands for the presence of a job change in the previous year, instead than in the present year.

Table 15: Descriptive statistics of selected variables, GSOEP data, 1985-2012 (164'058 obs.)

Variable	Mean	Std. Dev.	Min	Max
Age	41.622	10.553	17	64
Female dummy	0.423	0.494	0	1
Married dummy	0.681	0.466	0	1
Part time work dummy	0.22	0.414	0	1
Change of job dummy	0.1	0.3	0	1
Size of the firm (20-200 employees)	0.294	0.456	0	1
Size of the firm (201-2000 employees)	0.242	0.428	0	1
Size of the firm (more than 2000 employees)	0.261	0.439	0	1
Number of children in the household	0.728	0.98	0	10
Years of education	12.112	2.702	7	18
Total annual work hours	2065.769	586.235	52	7445
Total annual wage (in euros)	26633.05	16754.95	42	300000
Years of full time working experience	16.871	11.082	0	49.6
Decisional autonomy at work	2.646	1.098	1	5
Job satisfaction	6.663	2.348	0	10

4.3 Results and discussion

4.3.1 The effect of (absolute and relative) wage mobility on job satisfaction

Table 16 presents the estimation results of different model specifications. From Estimations 1-6 in Table 16, we find evidence that the estimated coefficient attached to relative wage mobility is positive and highly statistically significant (at a 99% confidence level) in all model specifications. This means that positional improvements (i.e. increases in the individual Gaussian ranks) are associated with a higher probability of reporting higher levels of job satisfaction. On the other hand, a worsening in the individual relative position (i.e. a decrease in the individual Gaussian rank over time) is associated with a lower probability of experiencing high levels of job satisfaction. This constitutes a first evidence that relative wage mobility is an important driver of satisfaction at the workplace. This effect is still relevant after we control for the absolute wage level. In Table 16, both equation (2) and (5) estimate the influence of relative wage mobility on job satisfaction, with the difference that in equation (5) the logarithm of total annual wage is added to the explanatory variables. From the estimation results, we deduce that the coefficient of relative wage mobility is almost identical in equation (2) and in equation (5). Hence, positional improvements have a positive value to the individual *per se*, regardless the absolute level of the wage earned in a year.

The relevance of our relative mobility measure in the job satisfaction equation is consistent with the previous literature. In particular, we find empirical support for the range frequency theory developed by Parducci (1965), according to which the position (rank) in the wage distribution is one of the main drivers of individual job satisfaction. In both equation (1) and (4) in Table 16, the only mobility measure included in the estimating equation is absolute wage mobility. The difference between these two estimations is that in equation (4) the logarithm of annual wage is included among the explanatory variables. From equation (1) and equation (4) in Table 16, we find evidence of a positive relationship between absolute wage mobility (i.e. percentage wage growth in a year) and the probability of reporting higher levels of job satisfaction. This finding is consistent with Clark (1999). However, when both mobility measures (relative and absolute wage mobility) are included in the model, i.e. in equation (6) and in equation (3) (respectively with and without log wage included among the explanatory variables), we find that both estimated coefficients are statistically significant. This provides further support to the hypothesis that these two measures correspond to different mobility concepts, each of which has a potentially different impact on individual job satisfaction. It is also worth noting that, when both absolute and relative wage mobility are included in the model, the estimated coefficient attached to absolute mobility is negative, whereas that attached to relative mobility is positive (this holds in different model specifications, i.e. both with and without the logarithm of total annual wage among the explanatory variables). According to Kahneman and Tversky (1979) people are loss averse, i.e. their decrease in utility due to a monetary loss is larger than the increase in utility following a monetary gain of the same size. This may explain why people dislike absolute wage mobility: the prospect of a potential monetary gain is more than compensated by the prospect of an equivalent potential monetary loss (Clark 2011). This is the reason why, once relative mobility has been included in the estimating equation, we find evidence of a negative and statistically significant relationship between absolute wage mobility and the probability of recording high levels of job satisfaction. On the other hand, relative wage mobility is positively associated to the probability of recording higher levels of individual job satisfaction in all the model specifications in Table 16 (i.e. with and without absolute wage mobility included among the explanatory variables), thus suggesting that the relationship between relative wage mobility and individual job satisfaction is robust.

To obtain further insights on the role of relative mobility as a driver of job satisfaction, in equation (7) reported in Table 16 we consider an alternative measure of relative wage mobility. In this last model specification, we take as a determinant of job satisfaction the absolute value of the change in the Gaussian ranks from one year to the following one. This means that we take the absolute value of the variable standing for relative wage mobility in equations (1)-(6) in Table 16. The estimation results show that there is a positive and statistically significant relationship (at a 99% confidence level) between this new variable (i.e. the absolute value of the rank change) and the probability of experiencing higher levels of job satisfaction. This means that positional mobility is perceived as a value *per se* by the workers. A higher degree of positional mobility is often associated with more flexible career patterns and greater opportunities of climbing the wage scale. Moreover, if the individual is currently employed in a low-pay job, higher relative (positional) mobility means that there are higher chances for him or her to escape the so-called low-wage trap. As mentioned in the introduction, a labor market in which there is a higher degree of relative wage mobility is commonly perceived by the workers as more fair, since it should be characterized by a greater equality of opportunity.

Table 16: Estimation results, random-effects ordered probit model

	1	2	3	4	5	6	7
Abs. mobility (% change)	0.0110*** (0.0021)	-	-0.0044* (0.0026)	0.0111*** (0.0021)	-	-0.0043* (0.0026)	-
Rel. mobility 1	-	0.0618*** (0.0054)	0.0682*** (0.0067)	-	0.0618*** (0.0054)	0.0681*** (0.0062)	
Rel. mobility 2 - (Alternative def.)	-	-	-	-	-	-	0.0254*** (0.0074)
Prob >Wald chi2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Number of obs	164’058	164’058	164’058	164’058	164’058	164’058	164’058

The dependent variable is job satisfaction in all the estimations reported in Table 2. (1) Only absolute mobility. (2) Only relative mobility. (3) Both absolute and relative mobility. (4) Only absolute mobility, with log of annual income among the control variables. (5) Only relative mobility, with log of annual income among the control variables. (6) Both absolute and relative mobility, with log of annual income among the control variables. (7) In this equation the absolute value of our measure of relative mobility (i.e. change in the Gaussian ranks) has been included in the model as the unique wage mobility measure.

Indeed, relative downward mobility represents a worsening of the individual position in the wage distribution, but does not necessarily imply a decrease in the absolute wage level. Relative wage mobility is perceived as a value by the workers and hence has a positive impact on the probability of reporting higher levels of job satisfaction. As a robustness check, we performed the estimates with robust standard errors, in order to take into account the possibility of heterogeneity, and no significant changes in the estimation results were found (results not reported for reasons of brevity). In order to be able to analyze in greater detail the impact of both absolute and relative mobility on job satisfaction, in Table 17 below we computed the marginal effects of our variables of interest on the probability of reporting each of the 11 job satisfaction levels for the main estimations of Table 16. From these results, we notice that in estimations (1)-(2) the marginal impact of relative wage mobility is about five times larger than that of absolute wage mobility. Moreover, from estimation (3) we deduce that the marginal impact of relative wage mobility is around fifteen times larger than that of absolute wage mobility. When only one wage mobility measure (either absolute or relative) is introduced in the estimated equation (n. (1)-(2),

(5) and (7) in Table 3 below), we find that wage mobility has a negative marginal effect on the probability of experiencing low to middle satisfaction levels (from zero to seven). On the other hand, it has a positive marginal effect on the probability of recording the three highest levels of satisfaction. As mentioned before, the effect of a marginal change in the relative wage mobility measure (i.e. difference in the Gaussian ranks from one year to the following one) is on average around ten times larger than the effect of a marginal change in the annual growth rate of wage (absolute mobility).

Table 17: Marginal effects, random-effects ordered probit model

	1	2	3	3	5	7
	Abs mob	Rel mob	Abs mob	Rel mob	Rel mob	Rel mob (2)
Pr(Satisf)=0	-0.0002*** (0.00004)	-0.0013*** (0.0001)	0.0001* (0.0001)	-0.0014*** (0.0001)	-0.0013*** (0.0001)	-0.0005*** (0.0002)
Pr(Satisf)=1	-0.0004*** (0.0001)	-0.0024*** (0.0002)	0.0002* (0.0001)	-0.0027*** (0.0003)	-0.0024*** (0.0002)	-0.0010*** (0.0003)
Pr(Satisf)=2	-0.0003*** (0.0001)	-0.0015*** (0.0001)	0.0001* (0.0001)	-0.0017*** (0.0002)	-0.0015*** (0.0001)	-0.0006*** (0.0002)
Pr(Satisf)=3	-0.0005*** (0.0001)	-0.0026*** (0.0002)	0.0002* (0.0001)	-0.0028*** (0.0003)	-0.0026*** (0.0002)	-0.0011*** (0.0003)
Pr(Satisf)=4	-0.0006*** (0.0001)	-0.0031*** (0.0003)	0.0002* (0.0001)	-0.0035*** (0.0003)	-0.0031*** (0.0003)	-0.0013*** (0.0004)
Pr(Satisf)=5	-0.0012*** (0.0002)	-0.0067*** (0.0006)	0.0005* (0.0003)	-0.0074*** (0.0007)	-0.0067*** (0.0006)	-0.0028*** (0.0008)
Pr(Satisf)=6	-0.0008*** (0.0002)	-0.0044*** (0.0004)	0.0003* (0.0002)	-0.0049*** (0.0005)	-0.0044*** (0.0004)	-0.0018*** (0.0005)
Pr(Satisf)=7	-0.0004*** (0.0001)	-0.0021*** (0.0002)	0.0002* (0.0001)	-0.0023*** (0.0002)	-0.0021*** (0.0002)	-0.0009*** (0.0003)
Pr(Satisf)=8	0.0017*** (0.0003)	0.0093*** (0.0008)	-0.0007* (0.0004)	0.0103*** (0.0010)	0.0093*** (0.0008)	0.0038*** (0.0011)
Pr(Satisf)=9	0.0016*** (0.0003)	0.0091*** (0.0008)	-0.0006* (0.0004)	0.0100*** (0.0010)	0.0091*** (0.0008)	0.0037*** (0.0011)
Pr(Satisf)=10	0.0010*** (0.0002)	0.0057*** (0.0005)	-0.0004* (0.0002)	0.0063*** (0.0006)	0.0058*** (0.0005)	0.0024*** (0.0007)

The dependent variable is job satisfaction in all the estimations. Numbering of the estimations is the same as in Table 2. (1) Only absolute mobility. (2) Only relative mobility. (3) Both absolute and relative mobility. (5) Only relative mobility, logarithm of annual income included among the control variables. (7) In this equation the absolute value of our measure of relative mobility (i.e. change in the Gaussian ranks) has been included in the model as the unique wage mobility measure.

Let us now consider the estimations in which both mobility measures are included, i.e. equations (3) and (6) in Table 16. In this case, we are able to disentangle the marginal effects of relative wage mobility from those of its absolute counterpart. Indeed, from Table 17 (estimation (3)), we deduce that absolute mobility has a positive marginal effect on the probability of experiencing low and middle levels of job satisfaction, whereas it has a negative marginal effect on the probability of reporting the three highest satisfaction levels (from 8 to 10). The reverse is true for relative wage mobility. From equation (3) we find evidence that relative wage mobility has a negative marginal impact on the probability of having a level of job satisfaction below 8, whereas it has a positive impact on the probability of recording a satisfaction level equal or greater than 8. The same result is confirmed if we consider an alternative definition of relative wage mobility, as it has been done in equation (7) in Table 16 and 17. As explained above, in this estimation we consider the absolute value of the change in Gaussian ranks from one year to another as a measure of relative wage mobility. In this last case, we still find that there is a positive association between relative wage mobility and the probability of having one of the three highest job satisfaction levels.

These findings provide further evidence that relative or positional wage mobility (also called rank mobility) is linked to the idea of equality of opportunity and hence it positively influences the degree of self-assessed job satisfaction of the worker. On the other side, absolute wage mobility, once it is disentangled from its relative counterpart, is negatively related to job satisfaction, because individuals are loss averse (Kahneman and Tversky 1979). The estimation results reported in Table 16 and 17 confirm the hypothesis that relative wage mobility is a relevant determinant of individual satisfaction at the workplace. Estimation results for the control variables are broadly consistent with the previous literature and are only reported in Appendix A for reasons of brevity. For example, we find evidence that job satisfaction is increasing in the number of children (Forgionne and Peeters 1982) and that is quadratic in age (Clark and Oswald 1996). Moreover, our estimates show that job satisfaction is lower for married individuals (Agho et al. 1993) and higher for those with higher decisional autonomy and who completed higher education levels (see e.g. Mottaz 1984, Lee and Wilbur 1985 and Rogers 1991). Land and sector dummies have also been included among the explanatory variables.

4.3.2 Robustness checks

As additional robustness checks, we performed some other estimations, in which alternative definitions of relative wage mobility were used. In particular, we tried to assess whether the reference group of the individual corresponds to the entire sample or it is smaller. In estimations (1)-(3) reported in Table 18 below, relative wage mobility is defined in the same way as it was in estimations (1)-(6) of Table 16 (i.e. the change in the Gaussian rank from one year to another). Here, however, the individual rank is not computed on the whole sample, but it represents how the individual position changed in a year whitin a restricted group of workers.

In estimation 1 in Table 18, for each worker in the sample this reference group consists of the 100 workers immediately below and the 100 immediately above the individual. In estimation 2 this group is extended to the 250 workers below and the 250 workers above and finally in estimation 3 we consider the 500 workers above and the 500 workers below each individual. The idea is that people could care more about mobility whitin a restricted group of people that are close to them in the wage distribution, rather than about their position in the whole wage distribution.

Table 18: Robustness checks, random-effects ordered probit model

	1	2	3
Rel. mobility	0.0345*** (0.0028)	0.0589*** (0.0051)	0.0340*** (0.0028)
Prob >Wald chi2	0.0000	0.0000	0.0000
Number of obs	163’858	163’558	163’058

The dependent variable is job satisfaction in all the estimations reported in Table 4. (1) Reference group of 200 people. (2) Reference group of 500 people. (3) Reference group of 1000 people.

However, from our estimation results reported in Tables 18 and 19, we find evidence that both the sign and the size of the influence of these alternative definitions of wage mobility on job satisfaction are rather close to those found in the previous Section (Table 16 and 17). It appears that relative wage mobility, either defined with respect to the full sample or with respect to a smaller group of workers, has a positve and highly statistically significant impact on individual job satisfaction. The marginal effect of relative wage mobility on the probability of recording the different values of job satisfaction is rather close in the different estimations (both those reported in Table 17 and those in Table 19).

The differences in the marginal effects of relative wage mobility on job satisfaction, depending on the size of the reference group used, are almost negligible. The estimated coefficients for the control variables reported in Table 18 are rather close, both in sign and in size, to those obtained in the estimations presented in Table 16, and are only reported in Appendix A for reasons of brevity. The estimation results reported in Table 18 and 19 provide further evidence of the robustness of our main estimation results, which have been presented in the previous Section.

Table 19: Marginal effects, robustness checks, random-effects ordered probit model

	1	2	3
Pr(Satisf)=0	-0.0007*** (0.0001)	-0.0012*** (0.0001)	-0.0007*** (0.0001)
Pr(Satisf)=1	-0.0014*** (0.0001)	-0.0023*** (0.0002)	-0.0013*** (0.0001)
Pr(Satisf)=2	-0.0009*** (0.0001)	-0.0015*** (0.0001)	-0.0009*** (0.0001)
Pr(Satisf)=3	-0.0014*** (0.0001)	-0.0024*** (0.0002)	-0.0014*** (0.0001)
Pr(Satisf)=4	-0.0018*** (0.0002)	-0.0030*** (0.0003)	-0.0017*** (0.0001)
Pr(Satisf)=5	-0.0037*** (0.0003)	-0.0064*** (0.0006)	-0.0037*** (0.0003)
Pr(Satisf)=6	-0.0025*** (0.0002)	-0.0042*** (0.0004)	-0.0024*** (0.0002)
Pr(Satisf)=7	-0.0012*** (0.0001)	-0.0020*** (0.0002)	-0.0011*** (0.0001)
Pr(Satisf)=8	0.0052*** (0.0004)	0.0089*** (0.0008)	0.0052*** (0.0004)
Pr(Satisf)=9	0.0051*** (0.0004)	0.0086*** (0.0008)	0.0050*** (0.0004)
Pr(Satisf)=10	0.0032*** (0.0003)	0.0055*** (0.0005)	0.0032*** (0.0003)

The dependent variable is job satisfaction in all the estimations. (1) Reference group of 200 people. (2) Reference group of 500 people. (3) Reference group of 1000 people.

4.4 Conclusions

In this paper, we aimed at assessing the relationship between relative wage mobility and individual job satisfaction. Whereas the past literature exclusively focused on the role of the absolute income level and on that of absolute wage mobility in determining job satisfaction, we claim that also relative or rank mobility matters, i.e. how the individual position in the wage distribution changes over time, relative to that of the other workers. In order to do this, we construct a measure of relative wage mobility, which is based on rank change. By the word rank, we mean the individual position in the wage distribution at a given point in time.

In all the different model specifications, we find that the coefficient attached to relative wage mobility is positive and highly statistically significant. Even after controlling for the logarithm of total annual income and for a measure of absolute mobility (i.e. the annual rate of wage growth), the positive impact of relative mobility on job satisfaction remains significant at a 99% confidence level. Our results are also robust to different definitions of the reference group considered for the computation of the ranks. The marginal effect of a change in relative wage mobility is around ten times bigger than the marginal effect of a change in absolute wage mobility; this provides a further indication of the relevance of our variable of interest as a driver of individual job satisfaction. Moreover, we defined an alternative variable standing for relative wage mobility. This new variable is equal to the absolute value of the Gaussian rank change and we find evidence that it is, too, positively related to the probability of recording higher levels of job satisfaction. This suggests that relative wage mobility is perceived as an advantage by the worker, in the sense of allowing the transition towards better-paid jobs. This is the reason why the presence of relative mobility itself is associated with higher job satisfaction in our estimation results, i.e. high relative mobility is associated with high levels of job satisfaction and vice versa. Our results are entirely consistent with the range frequency theory (Parducci 1965, Brown et al. 2008, Boyce et al. 2010), which suggests that people not only care about their income and its evolution over time, but also about their rank, i.e. their position within a group of workers and about the evolution of their rank from one year to the following one. Note that we also found evidence that the whole wage distribution constitutes a relevant reference group for the individual. Indeed, job satisfaction turns out to be directly and significantly influenced by positional mobility, i.e. rank change within the entire wage distribution.

The implications of these results are manifold. From the employer's viewpoint, in order to keep the workers motivated it is not enough to ensure that their wages grow over time, but it is also necessary that there is an adequate degree of rank or relative wage mobility within the organization. A low degree of relative wage mobility, with some workers permanently lagging behind in the wage distribution, indeed, is likely to cause more dissatisfied and hence less productive employees. On the other hand, from a policy-making viewpoint, it clearly emerges that people care about the evolution of their relative position in the wage distribution and not only about the evolution of their absolute pay level. Therefore, it is of fundamental importance to advance in the study of the determinants of rank or positional wage mobility and to predispose policy interventions in order to enhance this type of mobility, in particular at the bottom part of the wage distribution.

References

1. Agho, A. O., Mueller, C. W., & Price, J. L. (1993). Determinants of employee job satisfaction: An empirical test of a causal model. Human relations, 46(8), 1007-1027.

2. Akerlof, G. A., Rose, A. K., Yellen, J. L., Ball, L., & Hall, R. E. (1988). Job switching and job satisfaction in the US labor market. Brookings papers on economic activity, 1988(2), 495-594.

3. Bjornskov, C., Dreher, A., Fischer, J. A., Schnellenbach, J., & Gehring, K. (2013). In-

equality and happiness: When perceived social mobility and economic reality do not match. Journal of Economic Behavior & Organization, 91, 75-92.

4. Borjas, G. J. (1979). Job satisfaction, wages, and unions. Journal of Human Resources, 21-40.

5. Boyce, C. J., Brown, G. D., & Moore, S. C. (2010). Money and happiness: Rank of income, not income, affects life satisfaction. Psychological Science, 21(4), 471-475.

6. Brown, G. D., Gardner, J., Oswald, A. J., & Qian, J. (2008). Does Wage Rank Affect Employees' Well-being?. Industrial Relations: A Journal of Economy and Society, 47(3), 355-389.

7. Brunetto, Y., Teo, S. T., Shacklock, K., & Farr-Wharton, R. (2012). Emotional intelligence, job satisfaction, well-being and engagement: explaining organisational commitment and turnover intentions in policing. Human Resource Management Journal, 22(4), 428-441.

8. Clark, A. E. (2011). Income and happiness: Getting the debate straight. Applied Research in Quality of Life, 6(3), 253.

9. Clark, A. E. (2003) Inequality-aversion and income mobility: A direct test. Delta.

10. Clark, A. E. (1999). Are wages habit-forming? Evidence from micro data. Journal of Economic Behavior & Organization, 39(2), 179-200.

11. Clark, A. E., Kristensen, N., & Westergard-Nielsen, N. (2009). Job Satisfaction and Coworker Wages: Status or Signal?. The Economic Journal, 119(536), 430-447.

12. Clark, A. E., & Oswald, A. J. (1996). Satisfaction and comparison income. Journal of public economics, 61(3), 359-381.

13. Darmon, R. Y. (1990). Identifying sources of turnover costs: A segmental approach. The Journal of Marketing, 46-56.

14. Diener, E., Sandvik, E., Seidlitz, L., & Diener, M. (1993). The relationship between income and subjective well-being: Relative or absolute?. Social Indicators Research, 28(3), 195-223.

15. Di Tella, R., Haisken-De New, J., & MacCulloch, R. (2010). Happiness adaptation to income and to status in an individual panel. Journal of Economic Behavior & Organization, 76(3), 834-852.

16. Easterlin, R. A. (1995). Will raising the incomes of all increase the happiness of all?. Journal of Economic Behavior & Organization, 27(1), 35-47.

17. Fachinger, U., & Himmelreicher, R. K. (2012). Income Mobility-Curse or Blessing? Mobility in Social Security Earnings: Data on West-German Men since 1950. Schmollers Jahrbuch, 132(2), 175-203.

18. Ferrer-i-Carbonell, A. (2005). Income and well-being: an empirical analysis of the comparison income effect. Journal of Public Economics, 89(5), 997-1019.

19. Forgionne, G.A. and Peeters, V.E. (1982). Differences in job motivation and satisfaction among female and male managers. Human Relations, 35, 101-18

20. Frank, R. H. (1991). Positional externalities (pp. 25-47). Strategy and Choice. Cambridge, Mass.: The MIT Press.

21. Freeman, R. (1978). Job Satisfaction as an Economic Variable. American Economic Review, 68(2).

22. Friedman, M. (1962). Capitalism and Freedom. Chicago.

23. Gottschalk, P., & Spolaore, E. (2002). On the evaluation of economic mobility. The Review of Economic Studies, 69(1), 191-208.

24. Griffeth, R. W., & Hom, P. W. (1995). The employee turnover process. Research in personnel and human resources management, 13(3), 245-293.

25. Hamermesh, D. S. (1977). Economic aspects of job satisfaction. Essays in Labor Market Analysis. New York: John Wiley, 53-72.

26. Jantti, M., & Jenkins, S. P. (2013). Income mobility. IZA Working Paper.

27. Judge, T. A., Piccolo, R. F., Podsakoff, N. P., Shaw, J. C., & Rich, B. L. (2010). The relationship between pay and job satisfaction: A meta-analysis of the literature. Journal of Vocational Behavior, 77(2), 157-167.

28. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica: Journal of the econometric society, 263-291.

29. Lee, T. W., & Johnson, D. R. (1991). The effects of work schedule and employment status on the organizational commitment and job satisfaction of full versus part time employees. Journal of vocational Behavior, 38(2), 208-224.

30. Leontaridi, R. M., & Sloane, P. J. (2004). Low Pay, Higher Pay, Earnings Mobility and Job Satisfaction in Britain. In Minimum Wages, Low Pay and Unemployment (pp. 87-140). Palgrave Macmillan UK.

31. Macmillan, P., & Smith, I. (2007). Explaining international soccer rankings. Journal of Sports Economics, 8(2), 202-213.

32. Marginson, S., & Van der Wende, M. (2007). To rank or to be ranked: The impact of global rankings in higher education. Journal of studies in international education, 11(3-4), 306-329.

33. McBride, M. (2001). Relative-income effects on subjective well-being in the cross-section. Journal of Economic Behavior & Organization, 45(3), 251-278.

34. Millan, J. M., Hessels, J., Thurik, R., & Aguado, R. (2013). Determinants of job satisfaction: a European comparison of self-employed and paid employees. Small business economics, 40(3), 651-670.

35. Mobley, W. H. (1982). Employee turnover, causes, consequences, and control. Addison-Wesley.

36. Mottaz, C. (1984). Education and work satisfaction. Human Relations, 37(11), 985-1004.

37. Mowday, R. T., Porter, L. W., & Steers, R. (1982). Organizational linkage: the psychology of commitment, absenteeism and turnover. Organizational and Occupational Psychology, 10(3), 2008.

38. Parducci, A. (1965). Category judgment: a range-frequency model. Psychological review, 72(6), 407.

39. Rogers, R. (1991). The effects of educational level on correctional officer job satisfaction. Journal of Criminal Justice, 19(2), 123-137.

40. Staw, B. M. (1980). The consequences of turnover. Journal of occupational Behaviour, 253-273.

41. Stutzer, A. (2004). The role of income aspirations in individual happiness. Journal of Economic Behavior & Organization, 54(1), 89-109.

42. Ton, Z., & Huckman, R. S. (2008). Managing the impact of employee turnover on performance: The role of process conformance. Organization Science, 19(1), 56-68.

43. Vieira, J. A. C. (2005). Skill mismatches and job satisfaction. Economics letters, 89(1), 39-47.

44. Yucel, I. (2012). Examining the relationships among job satisfaction, organizational commitment, and turnover intention: An empirical study. International Journal of Business and Management, 7(20), 44.

Appendix A: Full estimation results

Table 20: Estimation results, random-effects ordered probit model

	1	2	3	4	5	6	7
Female	-0.0177 (0.0126)	-0.0180 (0.0126)	-0.0177 (0.0126)	-0.0178 (0.0127)	-0.0181 (0.0127)	-0.0178 (0.0127)	-0.0158 (0.0126)
Age	0.0251*** (0.0027)	0.0259*** (0.0027)	0.0257*** (0.0027)	0.0250*** (0.0027)	0.0258*** (0.0027)	0.0256*** (0.0027)	.0248*** (0.0027)
Age^2	-0.0003*** (0.00003)	-0.0003*** (0.00003)	-0.0003*** (0.00003)	-0.0003*** (0.00003)	-0.0003*** (0.00003)	-0.0003*** (0.00003)	-.0003*** (0.00003)
Married	-0.0349*** (0.0096)	-0.0337*** (0.0096)	-0.0337*** (0.0096)	-0.0345*** (0.0096)	-0.0333*** (0.0096)	-0.0333*** (0.0096)	-0.0351*** (0.0096)
N. children	0.0079* (0.0043)	0.0077* (0.0043)	0.0079* (0.0043)	0.0085** (0.0043)	0.0083* (0.0043)	0.0085** (0.0043)	0.0083* (0.0043)
Education	0.0134*** (0.0023)	0.0132*** (0.0023)	0.0132*** (0.0023)	0.0140*** (0.0023)	0.0137*** (0.0023)	0.0137*** (0.0023)	0.0134*** (0.0023)
Work hours	-3.69e-06 (8.11e-06)	-0.00001 (8.13e-06)	-0.00001 (8.14e-06)	-3.40e-06 (8.11e-06)	-0.00001 (8.14e-06)	-0.00001 (8.14e-06)	-1.87e-06 (8.11e-06)
Part time	0.0112 (0.0117)	0.0165 (0.0118)	0.0173 (0.0118)	0.0116 (0.0118)	0.0169 (0.0118)	0.0177 (0.0118)	0.0102 (0.0118)
Autonomy	0.0990*** (0.0048)	0.0985*** (0.0048)	0.0983*** (0.0048)	0.0988*** (0.0048)	0.0983*** (0.0048)	0.0981*** (0.0048)	0.0987*** (0.0048)
Size 20-200	0.0154 (0.0104)	0.0132 (0.0104)	0.0129 (0.0104)	0.0155 (0.0104)	0.0133 (0.0104)	0.0130 (0.0104)	0.0155 (0.0104)
Size 200-2000	-0.0011 (0.0116)	-0.0032 (0.0116)	-0.0036 (0.0116)	-0.0011 (0.0116)	-0.0033 (0.0116)	-0.0037 (0.0116)	-0.0009 (0.0116)
Size >2000	0.0006 (0.0119)	-0.0020 (0.0119)	-0.0023 (0.0119)	0.0002 (0.0119)	-0.0023 (0.0119)	-0.0027 (0.0119)	0.0013 (0.0119)
Experience	-0.0001 (0.0008)	0.0003 (0.0008)	0.0003 (0.0008)	-0.0001 (0.0008)	0.0003 (0.0008)	0.0003 (0.0008)	-0.0002 (0.0008)
Change of job	0.0842*** (0.0098)	0.0927*** (0.0098)	0.0940*** (0.0098)	0.0843*** (0.0098)	0.0928*** (0.0098)	0.0940*** (0.0098)	0.0813*** (0.0098)
Abs. mobility (% change)	0.0110*** (0.0021)	-	-0.0044* (0.0026)	0.0111*** (0.0021)	-	-0.0043* (0.0026)	-
Annual wage (log)	-	-	-	-0.0204*** (0.0062)	-0.0203*** (0.0062)	-0.0202*** (0.0062)	-
Rel. mobility 1	-	0.0618*** (0.0054)	0.0682*** (0.0067)	-	0.0618*** (0.0054)	0.0681*** (0.0067)	-
Rel. mobility 2 - (Alternative def.)	-	-	-	-	-	-	0.0254*** (0.0074)
Land dummies	yes	yes	yes	yes	yes	yes	yes
Sector dummies	yes	yes	yes	yes	yes	yes	yes
Prob >Wald chi2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Number of obs	164'058	164'058	164'058	164'058	164'058	164'058	164'058

The dependent variable is job satisfaction in all the estimations. (1) Only absolute mobility. (2) Only relative mobility. (3) Both absolute and relative mobility. (4) Only absolute mobility, with log of annual income in the control variables. (5) Only relative mobility, with log of annual income in the control variables. (6) Both absolute and relative mobility, with log of annual income in the control variables. (7) Alternative definition of the relative wage mobility variable.

Table 21: Robustness checks, random-effects ordered probit model

	1	2	3
Female	-0.0153 (0.0126)	-0.0144 (0.0127)	-0.0145 (0.0127)
Age	0.0248*** (0.0027)	0.0251*** (0.0027)	0.0250*** (0.0027)
Age^2	-0.0003*** (0.00003)	-0.0003*** (0.00003)	-0.0003*** (0.00003)
Married	-0.0349*** (0.0096)	-0.0352*** (0.0096)	-0.0352*** (0.0096)
N. children	0.0083* (0.0043)	0.0083* (0.0043)	0.0083* (0.0043)
Education	0.0127*** (0.0023)	0.0127*** (0.0023)	0.0126*** (0.0023)
Work hours	-0.00001 (8.14e-06)	-0.00001 (8.15e-06)	-0.00001* (8.18e-06)
Part time	0.0183 (0.0118)	0.0165 (0.0118)	0.0173 (0.0118)
Autonomy	0.0965*** (0.0048)	0.0961*** (0.0048)	0.0966*** (0.0048)
Size 20-200	0.0122 (0.0104)	0.0131 (0.0104)	0.0127 (0.0104)
Size 200-2000	-0.0053 (0.0116)	-0.0046 (0.0116)	-0.0058 (0.0116)
Size >2000	-0.0042 (0.0119)	-0.0030 (0.0119)	-0.0037 (0.0120)
Experience	0.0001 (0.0008)	0.0001 (0.0008)	0.0001 (0.0008)
Change of job	0.0952*** (0.0098)	0.0914*** (0.0098)	0.0955*** (0.0099)
Rel. mobility	0.0345*** (0.0028)	0.0589*** (0.0051)	0.0340*** (0.0028)
Land dummies	yes	yes	yes
Sector dummies	yes	yes	yes
Prob >Wald chi2	0.0000	0.0000	0.0000
Number of obs	163’858	163’558	163’058

The dependent variable is job satisfaction in all the estimations. (1) Reference group of 200 people. (2) Reference group of 500 people. (3) Reference group of 1000 people.

Note that the negative coefficient attached to log wage in estimations 4-6 reported in Table 20, despite being somewhat surprising, is consistent with the esitmations results of Clark and Oswald (1996).

5 Conclusion

In this thesis we presented a new model for the wage rank dynamics. By using the new functional copula model developed in Chapter 1, we estimated the autoregressive function which links the present and the past Gaussian ranks, conditional on the individual characteristics. This semi-nonparametric approach allowed us to have a rather high degree of flexibility in the modellization, without being prone to the curse of dimensionality. Our estimates show that in the US for the period 1975-1996 the degree of relative wage mobility was rather high, both for young and for old individuals whose past rank was in the bottom half of the wage distribution. This finding endorses the common wisdom of the US being the "land of opportunity" and suggests that workers currently being in low ranks are not likely to remain stuck in a low-wage trap also in the following year. The degree of positional immobility results rather low for all age and education groups in the US labor market in the period considered. Education does not seem to play a relevant role as a driver of the individual patterns of positional mobility. These are the results that we obtain from a first application of the new functional copula model developed in Chapter 1 to PSID data. In Chapter 2, we extended the analysis of relative wage mobility to the comparison between two countries characterized by rather different institutional backgrounds. Indeed, relative wage mobility has been recognized in the previous literature as a fundamental feature in an economy. However, to the best of our knowledge, no investigation has been performed up to now about how relative wage mobility changed after the financial crisis started in 2007. In Chapter 2 we tackled this question. The aim of the chapter was to assess whether different labor market institutions were at the root of different responses of wage mobility to the financial crisis. To this aim, we considered the two largest European economies: Germany and the United Kingdom, and we applied our semi-nonparametric copula model separately to the pre-crisis and to the post-crisis period in both countries. Our estimation results show that the mobility patterns of different age and education groups of workers dramatically changed between 2002-7 and 2009-14 in Germany. On the contrary, the shape of the estimated autoregressive function linking the present and the past rank did not significantly change due to the crisis in the UK. Our estimates show that, after the crisis, young people being in a low rank were less mobile than their older colleagues in Germany, and education had lost its role as a driver of relative wage mobility. On the contrary, the degree of mobility of all the age and education groups remained almost unchanged before and after the crisis in the UK. We notice from these results that the financial crisis had only a limited impact on individual patterns of positional mobility in the UK, whereas it had stronger effects in Germany. It is also worth noting that the fact of being female was associated with different mobility patterns in Germany before the crisis, but then this gender-linked difference almost completely disappeared after 2009.

The use of the new semi-nonparametric functional copula model developed in Chapter 1 allowed us to estimate the value of the autoregressive function linking present and past ranks virtually for every individual in every point of the wage distribution. This is much more informative than summary mobility measures that have usually been used in the previous literature, such as aggregate mobility indices or transition matrices. We are then able to assess that the autoregressive function, which is different in each country, changed in Germany due to the financial crisis, whereas it remained close to its previous shape in the UK. We can therefore suppose that the presence of a more flexible and less regulated labor market such as the British one allowed the effects of the crisis to be absorbed in a shorter time frame and/or by different variables (i.e. employment and not only wage), so that the impact on wage mobility patterns after some years had already become difficult to identify. On the contrary, in a highly regulated labor market such as the German one, the effects of the financial crisis on individual relative wage mobility patterns were more long-lasting, even if Germany showed a higher degree of resilience than the UK in terms of employment rate during the financial crisis. However, it is still unclear which aspects of the institutional framework are more closely linked to the degree of wage mobility. A deep analysis of the factors at the roots of the cross-countries differences in wage mobility that we found in Chapter 2 constitutes scope for further research.

However, why is it so important to analyze the drivers of relative wage mobility, instead of those of its absolute counterpart? Chapter 3 provides a sound answer to this question. In this last chapter, indeed, we study the relationship between relative wage mobility and individual job satisfaction. Whereas the past literature exclusively focused on the role of the absolute income level and on that of absolute wage mobility in determining job satisfaction, we claim that also relative mobility matters, i.e. how the individual position in the wage distribution changes over time, relative to that of the other workers. In order to do this, we construct a measure of relative wage mobility, which is based on rank change. By the word rank, we mean the individual position in the wage distribution at a given point in time. In all the different model specifications, we find that the coefficient attached to relative wage mobility is positive and highly statistically significant. Even after controlling for the logarithm of total annual income and for a measure of absolute mobility (i.e. annual rate of wage growth), the positive impact of relative mobility on job satisfaction remains significant at a 99% confidence level. Our results are also robust to different definitions of the reference group considered for the computation of the ranks. The marginal effect of a change in relative wage mobility results around ten times bigger than the marginal effect of a change in absolute wage mobility; this provides a further indication of the relevance of our variable of interest as a driver of individual job satisfaction. Our results are consistent with the economic theory, which suggests that people not only care about their income and its evolution over time, but also about their relative or comparison income, i.e. how much they earn as compared to the members of a reference group. The concept of relative wage mobility, i.e. the change of positions in the wage distribution as compared to the other workers, incorporates this idea of comparison among individuals.

The results presented in Chapter 3 provide a sound empirical support to the research questions of both Chapter 1 and Chapter 2, by showing the fundamental role of relative wage mobility as a key determinant of individual job satisfaction. As a future research development, it would be interesting to determine in greater detail which kind of wage mobility is more likely to determine changes in the individual levels of job satisfaction. Once assessed (as it has been found in Chapter 3 of the present work) that workers care about relative wage mobility, it could be investigated how broad their reference group is, i.e. whether they care about gaining positions in the wage distribution with respect, for example to their colleagues working in the same division, firm, region, or sector.