Unemployment Duration and sport participation: evidence from Germany*

Charlotte Cabane†

Chapter 2

Abstract

In this study I use the German Socio-Economic Panel to evaluate the impact of leisure sport participation on the unemployment duration. The empirical literature on sport participation has focused on labour market outcomes and job quality while the impact of this activity on job search has not been studied. However, sports participation fosters socialization which, through the networking effect, accelerates the exit from unemployment to employment. Furthermore, there may be a selection effect of individuals with higher non-cognitive skills (which may optimize their job search). These hypotheses are tested using a duration model, taking into account unobservable heterogeneity. Because the timing of participation in sports activities is relevant, various measure of sport participation are tested as well as other activities.

JEL Codes: J24, J64, L83

Keywords: Unemployment duration, Non-cognitive skills, Sport

*We would like to thank Wladimir Andreff, Jean-Bernard Chatelain, Andrew Clark, Luke Haywood, Yannick l'Horty, Rob Simmons, and the participants of the SIMA seminar, of the WIP seminar, of the DESport seminar and of the PSI-PSE seminar for helpful comments and discussions.

†CES - Université Paris 1 Panthéon Sorbonne - Maison des Sciences Economiques, 106-112, boulevard de l'Hôpital, 75013 Paris. charlotte.cabane@univ-paris1.fr

1 Introduction

The usual returns of sport participation studied in the literature are : health, education, labour market outcomes (especially wages), and schooling. The effects highlighted are: signalling effect, reputation effect and increase in human capital (non-cognitive skills[1]). One specific situation on the labour market which is highly concerned by these channels has been left behind: unemployment.

For the same reasons that sporty people fare better on the labour market, they should experience shorter unemployment spells. The channels are the same. Sporty people who have more non-cognitive skills have a more efficient job search and thus find a job quicker (or even a better job).[2] There are many non-cognitive skills associated to sporty people and they can roughly be divided into two types: non-cognitive skills specific to certain types of sport and non-cognitive skills linked to sporty people in general. The first type of non-cognitive skills includes team spirit and competitive spirit, for example. These skills can be learned. Concentration and motivation belong to the second type and are often considered as personal traits and, thus, as constant individuals' characteristics. Here, I consider this second type of non-cognitive skills to be more important in the job search process. Therefore, from now on I use the term personal traits and refer to skills such as motivation and concentration.

There is another possible explanation which it does not involve a causal effect specific to sporting activities. One may argue that sporty people are different and signal to be so thanks to their sport participation. This mechanism refers to the signal theory developed by Spence (1973). It is assumed in the literature that sporty people send out *positive* signals to firms with respect to their unobservable level of productivity (Rooth, 2011). A priori, this signal is still effective for unemployed people with respect to their rivals on the labour market. Sporty people are chosen in priority and thus experience shorter unemployment spells. However, this effect is not necessarily caused by sports practice, it is just used as a signal and another leisure activity could

[1] Non-cognitive skills are personal preferences and personality traits which are valued in society but which do not involve technical or professional knowledge. Unlike cognitive skills, there is no school nor diploma which allows to learn and evaluate them.

[2] Question I already addressed in the previous chapter.

play the same role.

Besides these two channels[3], I decide to include a third one relying on networks. Sporty people have larger and more diversified social networks and thus, they have relatively more opportunities than non-sporty people. They have access to more information and benefit from more contacts and connections. This way, their unemployment spell can be reduced and/or they can find better jobs.

For all these reasons, I venture the hypothesis that sporty people should be unemployed for shorter periods than other individuals. The role played by selection effects is taken into account and discussed in more detail in the Data section.

The timing of the sporting activity is relevant and the analysis of unemployment spell duration involves three periods: the first one during which the individual is employed t_{w1}, the second one which begins once she lost her job t_u and the third one which begins when she finds another job t_{w2}.[4] Obviously, practising sports in t_{w2} will not have any impact on the length of t_u. Conversely, sport participation during t_u signals to firms that the individual is still physically active and thus healthy which is relevant with respect to her unobservable productivity at work. Moreover, it sends out a positive signal with respect to mental form too. A sporty unemployed individual is someone who does not give up -she is still socially active- once she is out of the labour market. And, for unemployed as well as for employed people, it sends out information linked to the specific sport practised. A rugby player is seen as someone which has a great team spirit whereas a dancer is considered as very rigorous and well-disciplined, for example. To sum up, sporting activity in t_u may reduce human capital depreciation and signal skills, and thus make the sporty unemployed individual more attractive. However, an over-investment in sport participation reduces the time available for job search and, moreover, may signal a loss of interest in the job market. Actually, Bougard *et al.* (2011) demonstrate -in a recent experimental study performed on the French labour market- that job applicants who are involved in associations are sig-

[3]I intentionaly do not develop on the reputation effect since it is specific to the US and I am working on German data.

[4]Actually, in the models used in this article, the unemployment spell are allowed to be right censored in order to be considered.

nificantly less often called than those who are not. Plus, sporting practice requires usually some financial investment that unemployed people are less able to afford. Even if they receive benefits and enjoy discounted access to sports facilities, their options in terms of sports can be limited (i.e. some sports are more expensive).

The efficiency of the networking effect depends on the reliability of the social network. Employed people are socially more attractive than unemployed ones and they are in interaction with many more people, therefore, sporting activity as a socialization process is more effective for employed people. Moreover, employed people can afford a broader range and more expensive sports than unemployed ones. This leads to more opportunities, or even to more interesting contacts (in terms of labour market opportunities) or even to meeting people in a privileged environment (such as private clubs or facilities). Once the network is built, people have to maintain it by going on with sport participation even once they become unemployed.

For the reasons exposed above, I refine my first hypothesis: sporting activities have a greater impact in reducing the unemployment spell if they are practised *before* and *during* the unemployment spell. I expect to observe a positive but smaller impact if the practice is only during pre-unemployment spell (in t_{w1}). The impact of a participation which only occurs during the unemployment spell (in t_u) has still to be determined.

The article is organized as follows: in the next section I review the literature before describing the model. The fourth section is dedicated to the presentation of the data and the fifth section contains the results. I conclude in the last section.

2 Literature review

Most of the literature on the relationship between sport participation and labour market outcomes use a US database which consists of information on sporting practice during college and labour market outcomes several years (10 to 13) later. Therefore, the methodology is often the same: the authors study the impact of being sporty in college on labour market outcomes.
In this article I am interested in the contemporary impact of sporting prac-

tice on individuals' unemployment duration. Thereby, I put the emphasis here on articles which study the immediate returns of sport participation. Studies using duration models and survival analysis are discussed in the third section.

Are sporty people more motivated?
Cornelißen & Pfeifer (2010) use the German Socio-Economic Panel (GSOEP) and demonstrate that sporty students are more successful than others. They explain the positive relationship they found by an increase in students' productivity at school. This increase is a result of an improvement of individuals' health and / or the acquisition of soft skills which are either rewarded at school or useful in the learning process. The soft skills credited to sporty people are: self-esteem, competitive spirit, tenacity, motivation, discipline and responsibility. Traditionally, it is assumed that boys and girls do not have the same character, that is, non-cognitive skills endowment. As a matter of fact, the authors find a larger effect of sport participation on girls. They are supposed to be less competitive and to have less self-esteem initially. Therefore, they have relatively more to learn from sports practice (i.e. sporty schoolgirls catch up with their male peers).

Lechner (2009) also uses the GSOEP and points out the positive relationship between sports participation and labour market outcomes. According to his study, being sporty is equal to an additional year of schooling in terms of labour market long-term outcomes. He clearly identifies three channels: health, "mental health" and individuals' unobservable characteristics. Sporty people are mentally and physically healthier thanks to sport participation.[5] Therefore they are more productive. Furthermore, they have unobservable specific characteristics which match with unobservable characteristics held by people who earn more.

Networking effect
I add to the effects previously cited the networking effect. Networks are essential in job-search because they ease the circulation of the information and multiply the sources of information.

[5]Labour market outcomes depend on individuals' productivity at work and a part of this productivity depends on health status. A healthy individual is less absent, more dynamic and more concentrated. By practising sports as an extracurricular activity, people maintain or increase their health status.

Rees (1966) highlights the relevance of networks and more precisely of informal networks in the hiring process. He argues that useful networks are reliable rather than large and diversified. The reason being that employers can more rely on recommendations of people they personally know[6] and who directly and personally commit themselves (like employees). Furthermore, it is costless for the firm and the future employee is more able to sell herself because she personally knows people who are working in the same area / firm.

Various models have been built in order to explain why and how networks influence labour market integration and outcomes but I focus on the one developed by Bramoullé & Saint-Paul (2010). Their model integrates the dynamism of the process which characterise labour market transition and social networks. Job status *and* social ties are interdependent and evolve through time. Bramoullé & Saint-Paul (2010) argue that the lower is the labour market turnover, the higher is the social segregation between employed and unemployed people. Therefore, individuals who have suffer from unemployment for a long time before being unemployed again experience lower exit rates from unemployment. And, people who stayed a long time employed before being unemployed experience higher exit rates from unemployment. These results are extremely interesting since they point out the fact that any activity which could connect unemployed people to employed people –such as sport practice– would partly break this time dependance / vicious circle.

Capperalli & Tatsiramos (2010) study the impact of friends networks on job finding rate, wages and employment stability using data from the British Household Panel Survey (BHPS). Focusing on friends position on the labour market, they find out that having best friends employed increases the probability to find a job.

By taking part in a team sport or by practising sport in a club, people socialize and get to know people who are not necessarily part of their original environment. Also, when people practise sports with their colleagues, it changes their usual professional relationship into a more personal relationship, which is stronger because it does not necessarily reproduce the hierarchy of the relationship set by the company's organization. Sport participation contributes to create other types of relationships between people who know each other and even to develop new relationships (with people who are not

[6]In the sense that people recommend people of their kind, friends.

already part of the individual's social environment).

Signalling effect
The presence of a signalling effect with respect to sporty people is clearly demonstrated by Rooth (2011). A part of his analysis is based on an experimentation on the Swedish labour market. This kind of study is called testing or correspondence study and allows to measure the impact of individuals' specific characteristics during the hiring process. Rooth (2011) found that people who declare practising sport as a leisure in their curriculum vitae have a higher probability of getting an interview. And, being sporty is equal to 1.5 additional years of work experience. He also estimates the impact of a variation of the physical fitness on earnings and finds a positive effect (4%). For this last impact it is harder to precise which effect is at work.
Unlike the previous cited studies, Rooth (2011) is able to differentiate types of sports (football, fitness etc.). This is very important in order to pinpoint impacts.

Evidence from Germany, previous studies
Most of the studies which demonstrate a positive relationship between sport participation and labour market outcomes (or education returns) are done using American data. Knowing the role sports has in the USA (social promoter, integrator, etc.), one can fairly question the relevance of such type of analysis with respect to European countries such as Germany. As a matter of fact, Germany (mainly because of the availability of the data) has already been at least twice investigated on the subject. The two studies - Lechner (2009) and Cornelißen & Pfeifer (2010)- have been realized on the GSOEP (German Socio-Economic Panel) and both outlined a positive impact of sport participation on labour market outcomes and school returns.

3 Model

The hypothesis tested here is that sporty people experience shorter unemployment spells than non-sporty people. Sport participation fosters socialization which, through the networking effect, accelerates the exit from unemployment to employment. Furthermore, sporty people may have non-cognitive skills -learned through sport participation- which optimize their job search. In order to value the effetc as precisely as possible, two specifi-

cations of duration model are used successively. They assume time as being continuous.
The selection within sporty people into unemployment is treated in the fourth section.

3.1 Weibull model

In a first step, I use a parametric model because it is more efficient -if well specified- and easier to interpret. Among parametric models, I choose the Weibull model. As the exponential model, it assumes a proportional relationship between the baseline hazard and the influence of individuals'characteristics. For the proportional hazard model, the hazard rate at time t for the subject i is written as follows:

$$h(t|x_i) = h_0(t)\varphi(X_i, \beta_x) \tag{1}$$

with

$$\varphi(X_i, \beta_x) = \exp(X_i\beta_x)$$

The probability of transiting from unemployment to employment in t for the individual i (1) is the product of the baseline hazard $h_0(t)$ and her individual characteristics X_i. The baseline hazard is the probability for everyone in the sample to exit at the time t, knowing they survived (they stayed unemployed) until time $t-1$. The sample is assumed to be homogeneous with respect to this baseline hazard. The proportionality of the hazard means that, for every individuals, the impact of x years of schooling is $\beta_{yearsofschooling} * h_0(t)$, for example. Individuals who have an x twice bigger automatically have a probability of exit twice bigger (ceteris paribus). In other words, the shape of the survival function remains the same and only its level changes.

Furthermore, the baseline hazard $h_0(t)$ for all individuals is parameterised as follows:

$$h_0(t) = pt_i^{p-1}\exp(\beta_0)$$

p being an ancillary shape parameter estimated from the data, and the scale

parameter being $exp(\beta_0)$.
The systemic part of the hazard rate for the subject i is still:

$$\varphi(X_i, \beta_x) = \exp(x_i \beta_x)$$

Hence, the equation of the hazard ratio in a Weibull model is the following:

$$h(t|x_i) = p t_i^{p-1} \exp(\beta_0 + x_i \beta_x) \tag{2}$$

This model allows to have different baseline hazards for the different sub-samples which is useful since the sample is rarely homogeneous. This means that each sub-sample has its own shape of baseline hazard and thus its own time dependence.[7]

More than half of the sample experiences more than one unemployment spell. Unobservable characteristics which influence the risk of getting unemployed must be taken into account. In fact, forgetting to consider it leads to overestimate (underestimate) the degree of negative (positive) duration dependance. Individuals (observations) with a high level of frailty -which means unobservable characteristics which increase their chances to find a job- get out faster from unemployment. Therefore, there are within the survivors more individuals with a low level of frailty and this proportion increases with time. Because the level of frailty is unobserved, the impact of this selection is directly imputed to time. In other words, the influence of t being over-estimated, the impact of the covariates is ***automatically*** under-estimated.
A way to adress this problem is to introduce individual frailty modeled as a parameter α which is Gamma distributed.[8] The unobserved characteristics are assumed to be independent from the covariates which comes to add individuals' random effect in the model. The hazard ratio is thus written as follows:

$$h(t|x_{ij}) = p t_i^{p-1} \alpha_j \exp(\beta_0 + x_{ij} \beta_x) \tag{3}$$

α_j being the group-level frailty (here a group j is an individual and i is an observation),
$\alpha_j > 0$ and $\alpha_j \sim \Gamma(1, \theta)$ (θ being estimated from the data).

[7] The systemic part does not change.
[8] The belonging of a duration to a group is estimated and not specified ex-ante.

For $\nu_j = \log \alpha_j$, the hazard can be written as follows:

$$h(t|x_{ij}) = h_0(t) \exp(X_i \beta_x + \nu_j)$$

The use of parametrical forms is interesting but less flexible than semi-parametric model. Therefore it is relevant to check that with less restrictions -on the baseline hazard form and on the frailty- the effects measured are still existent. This is why in the next sub-section I present another model. Using both models allows me to compare the results and to present more precise and reliable conclusions.

3.2 Heckman and Singer

The Heckman and Singer semi-parametric frailty (Heckman & Singer (1984)) allows to take into account heterogeneity without giving any functional form to the distribution of unobservables. This means that no assumption is done on the form of the frailty, thus the model is very flexible and *directly* fits the data. The procedure uses a semi-parametric maximum likelihood estimator which supposes a discrete multinomial distribution including mass points (which are estimated using the data).
The hazard can still be written as follows:

$$h(t|x_{ij}) = h_0(t) \exp(X_i \beta_x) \alpha_j$$

with α_j the frailty, which is now non-parametric and estimated using 2 mass points (one η is normalized to 1).[9]
This procedure allows for a piece-wise linear specification for duration dependence. However, I prefer to use a logarithmic form.

3.3 Deepening the relationship

Using these two different models, I test various sport participation timing as well as the impact of other activities. Indeed, a lot of activities can signal non-cognitive skills and enlarge social networks. In order to check the specificity of *sporting* activities, I include the following activities in the estimation: taking part in voluntary work, going to the cinema and being a churchgoer.

[9]The choice of the number of mass points results from the estimates I run.

The use of other activities allows to define more precisely the kind of relationship observed. Being a churchgoer and going to the cinema (both weekly) are related to socialization and networks. If being sporty is about being well-integrated, socially active and having friends, thus, its impact should be lower or even disappear once these two activities are included in the estimates. "Going to the cinema" also includes going to pop concerts, to dance or sporting events but it excludes going to museums, to the theatre or to classical concerts. Therefore, besides the social aspect of being with friends, they also relates to a specific kind of habits and budget. And, since both activities do not appear in one's CV, introducing these activities in my estimation cannot help me to detect the signalling effect. However, the weekly activity: taking part in voluntary work is almost always specified in a CV. This activity is used as a signal and thus, enables me to test the hypothesis that being sporty is a signal. They do not send out the same type of signal but both give information on the job applicant. I am not sure that the signalling effect impacts the exit rate from unemployment. Therefore, I infer that if being a volunteer has an influence it means that the signalling effect matters and thus, if being sporty is a signal too, it should also have an impact. If taking part in voluntary work has no impact, then, the influence of being sporty may not result from the signalling effect.

The use of shared frailty allows to control for unobservable heterogeneity which is not related to covariates. Therefore, personal traits –which highly influence individuals' development– can still be reflected by covariates. Being sporty may be a proxy for such unobservables, and, if so, one way to detect it is to observe whether its impact is stable. Indeed, these personal traits are considered as constant, therefore, their influence –in terms of size, sign and significance- is the same (by gender) whatever the sample. Analysing the correlation of being sporty and the exit probability from unemployment in t enables me to find out if such selection on unobservables exists.

Since the signalling effect does not allow to give credit to the *sport* practice in particular, I develop a strategy to value its relevance. This effect is effective only if the population considered is heterogeneous with respect to the signal. Here, sporty people can signal themselves only among non-sporty people. Therefore, by testing the impact of a measure of the *relative* sport participation within a group, I test the relevance of this explanation.
Moreover, signals are used when firms cannot directly measure individuals'

productivity. This means that the relevance of the signal depends on the availability of other information such as work experience. Therefore, the importance of sport participation impact on the exit rate should be greater for young people (or people with little work experience).[10] And, since I introduce a measure of the health status, I control for this part of individuals' productivity which also vary with age in order to catch this effect.

4 Data

4.1 Overview

Data used in this article are the German Socio Economic Panel (GSOEP). The panel runs from 1984 until nowadays and contains around 20 000 individuals by wave. There are various yearly questionnaires which enable to have a great definition of the individuals' current and past situation. Labour market information is recorded monthly which allows me to use survival models. However, individuals' characteristics are recorded yearly and information about sports is actually asked even less frequently. Individuals are questioned about their sport practice frequency every two years except between 1994 and 1999. In this interval the information is available each year. The question is the following:

How frequently do you do sports?
– *once per week,*
– *once per month,*
– *less than once per month,*
– *never.*

There is no formal definition of what the interviewer understands by *doing sports*[11] thus, there is a risk of measurement error. Actually, 17% of the population sampled declares to practice a physical activity at least once a week. This figure is below national statistics about sport participation in Germany but it is thus coherent knowing that it concerns only people

[10] They have spent less time on the labour market thus, the amount of information available with respect to their level of productivity is relatively low.

[11] Some years, the expression is formulated as written in the previous chapter or even using the expression *physical activity* instead of *sports*.

who experience unemployment (sample in which sporty people are under-represented). Besides, since it is self-declared and being sporty is positively looked upon, people have incitations to lie about their sporting activity. But the figures size being reasonable, it leads to be more confident with respect to this information.

I consider a person to be sporty ($sporty = 1$) as long as she declares having a physical activity at least once a week (definition coherent with the European standard). In order to define as precisely as possible the type of influence sport participation have, various dummy variables are built with respect to the timing of the practice: being sporty before being unemployed $sporty(t_{w1})$, being sporty while being unemployed $sporty(t_u)$, being sporty at the beginning of the unemployment spell *and* at the end of it $sporty(t_{w1,u})$.

The comparison of the characteristics of sporty and non-sporty people (see in Table 2) is interesting since it confirms the literature prediction and thus supports my hypothesis. There is a significantly smaller proportion of unemployed people within sporty people. Furthermore, sporty people stay employed for 10 months longer in average than non-sporty people and their unemployment spell lasts almost 7 months less. They are unemployed for significantly shorter periods. Sporty people are significantly more likely to be men, they are almost 3.5 years younger in average and they are significantly less living in former East-Germany. They have in average a higher level of education and practice more often activities such as visiting friends, taking part in voluntary work and helping in the neighbourhood (but they are less often churchgoers). And, as expected, they are healthier.[12]

Within the sample, people are between 17 and 60 years old (two thirds of the population is less than 40 years old), 39% are women, 85% are German and they have studied -in average- 11.4 years.
The analysed transitions are the ones between being unemployed and having a part-time or a full-time job. 39% of the individuals experience only one unemployment spell which lasts in average 7.4 months. 28.8% are subjected twice to unemployment and 17.1% 3 times[13] and the spells last in average re-

[12]Each difference in means commented here are statistically significant, details on the level are available in the Table 2.

[13]The maximum being 7 spells between 1994 and 2000.

spectively 7.4 and 6.4 months. In average, the first unemployment spell lasts 7.2 months, the second one lasts half a month less (6.7) and this keeps decreasing (the third spell lasts less than 5.5 months). The sample being bound within time, the more one experiences unemployment spells, the shorter the spells are.
Sporty individuals experience significantly less unemployment than the others: in average in total 2 unemployment spells versus 2.23. And their unemployment spells are shorter: they last in average 6 months for the first and unique spell versus 7.7 months for non-sporty people. When non-sporty people are twice unemployed within the sample, the first spell lasts in average 5.5 months versus 7.8 (and the second, respectively 6.5 months versus 7.4 months).

The covariates used in the model are defined with respect to time as follows. Gender, nationality, level of education[14] and total of unemployment spells are assumed to be constant over the sample. Characteristics which can change over the unemployment spell and influence the exit are: place of residence (by region: former East/West-Germany), marital status (dummy), health status[15] and daily hours dedicated to childcare are measured in the last month of unemployment. Values of variables which proportionally increase with time like age, work experience and unemployment experience are measured in the first month of unemployment.

4.2 Processus of selection into unemployment

A direct consequence of the hypothesis I make and test here -*sporty people have a higher level of productivity and/or easiness to integrate the labour market*- is that sporty people are under-represented among unemployed individuals. Therefore, the selection into unemployment is heterogeneous with respect to the sporting practice. In order to study this selection process, tests of equality of the means are ran on the characteristics of employed and unemployed people in different samples (sporty and non-sporty).
As expected, unemployed people with respect to employed people are likely

[14]They are detailed in 5 completed levels: failed (nothing completed), elementary school (Hauptschule, Volksschule), secondary school (Realschule), high school (Gymnasium - Abitur), university and the likes (Fachhochschule, Hochschule).

[15]The variable *health status* is a discrete variable equal to 1 if the individual is extremely satisfied by her health and equal to 5 if she is extremely dissatisfied.

to be older, to live in former East-Germany and to be foreigners (see in Table 2). Among the unemployed sub-population, there is an over-representation of women and people with a lower level of education. Moreover, people from this sub-sample appear to be less healthy and less often married. Finally, they less often practise activities[16] such as meeting friends, going to cultural events, going to church, taking part in voluntary work and practising sports.

This under-representation of sporty people within unemployed people and the fact that they have different socio-economic characteristics than non-sporty people (see Table 2) lead to analyze the selection into unemployment of sporty people in comparison to the selection within non-sporty people. Since the models used include frailty, they presuppose the observation of at least two unemployment spells by individual. Actually, frailty can be introduced for the whole sample, however, in presence of within cluster correlation, the standard errors are incorrect. Therefore, in order to avoid this risk I run the estimates on a restricted sub-sample (but I also present results obtained for the whole sample in order to get an idea of the selection and more information on the mechanisms). Since more rigorous results are expected from estimates on the restricted sub-sample, I study the difference of selection into being at least twice unemployed and having suffered less from unemployment over the period. Figures are reported in Table 3. For the sake of clarity, I called people who experienced at least two unemployment spells: type $i_{U>1}$ and the others type $i_{U=1}$ (which have suffered only once from unemployment over the period).

First, there is an over-representation of people with the two lowest education levels within people type $i_{U>1}$ in the sporty sample. Then, by selecting people type $i_{U>1}$, I select more non-German people in the sporty sample and conversely more German people in the non-sporty sample. The difference in age between the two sub-populations ($i_{U>1}$ and $i_{U=1}$) is greater within the sporty sample than within the non-sporty one (i.e. people type $i_{U>1}$ are much younger than people type $i_{U=1}$ in the sporty sample). And, people type $i_{U>1}$ in the sporty sample are significantly less satisfied with their health when there is no difference in the non-sporty sample. A last interesting difference to point out is the one between activities people are doing weekly. Sporty people are more active and the selection on unemployment experience criteria

[16]Except helping neighbours.

does not redefine two very different groups with respect to weekly activities. On the opposite, in the non-sporty sample, people type $i_{U>1}$ are significantly much more active. Here again one has to remember that non-sporty people tend to be more often unemployed than sporty people thus, they more often have more time to spend in weekly activities.

The question was: *Do I focus on specifically deprived people by focusing on sporty people who -against all expectations- experienced unemployment at least twice (people type $i_{U>1}$)?* And if the selection is different, how would it bias my results? In the sporty sample, the selection results in choosing young people, people with a very low level of education, people who are less satisfied by their health and foreigners significantly more often. This specific selection favours the hypothesis that "yes", by focusing on *unemployed people more than once*, I focus on less privileged people (with respect to usual individuals' productivity characteristics). They have less valuable characteristics with respect to the labour market. Therefore, being sporty can be one specificity among just as few others that can help them get out of unemployment.[17] Consequently, my results might be over-estimated.

5 Results

5.1 Baseline

Sport participation timing
In a first step, I test the impact of different sport participation timing. Results presented in Tables 4 show the correlation between being sporty before being unemployed $sporty(t_{w1})$ and being sporty at the end of the unemployment spell $sporty(t_u)$. Actually, for men as for women, there is no evidence of the existence of a correlation between being sporty in t_{w1} or in t_u and exiting unemployment. Punctual sporting activities in these circumstances are not rewarded. However, Table 5 shows that there is a strong evidence of a correlation between being persistently sporty $sporty(t_{w1,u})$ for men, and also for women. People have to be persistent or in other words, they have to be sporty. Since this characteristic is constant for each unemployment spell, it can easily be a proxy for personal traits –which are often considered as time invariant– for individuals who experience only one unemployment spell

[17] This specificity can be seen as a signal but it can also lead people to behave differently.

over the period. Analysis by sub-samples built with respect to the number of unemployment spells is then relevant. Differences due to sport participation timing are not surprising since I observe changes in sport participation contemporaneous to changes in labour-market status.[18] At least 15% of people who are sporty in t_{w1} stop practising sport weekly in t_u. And the same proportion starts to be sporty in t_u while they practised sport less than weekly in t_{w1}.

I answer my first interrogation about the impact of being sporty only while being unemployed: weekly sporting activity matters only when it is persistent. Therefore, from now on, I estimate the impact of the variable $sporty(t_{w1,u})$.

In Tables 4 and 5, the first three columns are ran without introducing shared frailty. It allows me to roughly compare the impact of sport participation with respect to the sample used. Figures are the coefficients, therefore the interpretation is as follows: for a woman, being sporty before and during the unemployment spell is correlated with an increase in the exit from unemployment to employment rate in t by 55% (Table 5, first column). In the women sub-sample, it already appears that whatever the profile of the individuals (i.e. the number of unemployment spells), sport practice is correlated with the probability of leaving unemployment in t. It is interesting to outline that, the correlation is much stronger in terms of size in this sub-sample than in the male sub-sample. Furthermore, and still for women, coefficients are always highly significant. The impact is slightly higher for women who experienced more than one unemployment spell in the period. Conversely, there is no significant impact for men who experienced more than one unemployment spell. It is important to recall here that people type $i_{U=1}$ are significantly different from people type $i_{U>1}$. Therefore, two reasons can explain these differences. First, the practised sport differs by gender and thus confers different advantages. Second, the effect at play is not the same for both gender. Indeed, if being sporty is a proxy for a personal trait its impact should not vary with the number of unemployment spells. Following this interpretation, being sporty would be a proxy for personality for women. And, it would be a signal or a proxy for social network for men, which is consistent with the fact that different sub-samples do not benefit in the same way from being

[18] These changes are undoubtedly linked and thus sport participation is endogeneous. But, since I choose to focus on persistent sport participation, I remove part of this impact.

sporty.

Including shared frailty
The introduction of the frailty is necessary for the reasons presented before and since heterogeneity is observed significantly (the likelihood-ratio test indicates that θ -the frailty variance- is significantly different from 0). Also, the results have to be interpreted knowing that each individual has a fixed ν_i (i.e. individuals' level of frailty).[19] As outlined before, the most rigorous results are obtained using the sub-sample restricted to people which experience at least two unemployment spells over the period. However, in order to have a rough idea of the unobservable heterogeneity in the whole sample, I also present results on it. And actually, results are fairly similar, except for the size of the coefficients. The level of significance and the sign do not change but the size really increases when controlling for one source of potential endogeneity. This is relevant since it means that being sporty is not a proxy for unobservable characteristics (which would be linked to the covariates).

As announced, I compare the results given by including two different types of frailty (parametric and semi-parametric). The results obtained are summed up in Tables 7 and 8. The level of significance and the sign do not change. But the introduction of shared frailty dramatically increases the size of the correlation between being sporty and the probability of leaving uenmployment in t, whatever the sub-sample. On the female sample, the two models support the hypothesis tested here: practising sports at least once a week during the whole unemployment spell is correlated with a higher level of hazard rate. This increase is equal to 96% in the Weibull model and exceeds 100% in the Heckman and Singer procedure for women who experienced at least two unemployment spells over the period. I remind here that, due to the sample restriction, I may over-estimate the impact. For men, there is no impact on this sub-sample. In other words, being sporty is rewarded only for men who have been unemployed only once over the period. And in this case, the increase in the exit probability from unemployment in t rate does not exceed 34%.

Interpretation of baseline general results
It is useful to have coefficients of covariates in order to compare impacts. I

[19]Results are conditional on this level.

also report results obtained by running the Weibull model with shared frailty excluding the variable of interest *sporty* in Table 6. In the Weibull model, the parameter p represents the shape of the curve of hazard rate with respect to time. For women who have experienced more than one unemployment spell (fifth column, Table 8) the equation of the hazard rate can be written as follows:

$$h(t\backslash x_{ij}) = 1.514t^{0.514} \exp(\beta_0 + X_i\beta_x + \nu_j)$$

p being less than 2 ($p = 1.514$), the baseline hazard is increasing with t but less than proportionally. Furthermore, this positive time dependance decreases with t. The probability of exit in t increases with t but to an extent which decreases every month.

In each estimation, covariates have the expected correlation: being married, being German, having more work experience and having higher levels of education than the elementary school level is positively correlated to the probability of exiting unemployment in t. Conversely, feeling unhealthy, spending more hours taking care of children, having more experiences of unemployment, being older and living in former East-Germany are linked to a decrease in the probability of leaving unemployment in t. In both subsamples, comparing coefficients of the covariates when the variable of interest is included in the estimation (Tables 7 and 8) with coefficients of the same covariates when it is not (Table 6) gives information on the channels through which the effect of sporting activities on the exit rate from unemployment passes. Part of the impact of being sporty is captured by variables on education and social-demographic characteristics. Indeed, coefficients associated to education are lower once included the dummy variable *sporty*. And the same is true for covariates indicating age, work experience and nationality. This is not surprising since sporty people are younger -and thus less experienced- and mostly Germans.

In the female sample, for the sub-sample restricted to people type $i_{U>1}$ (fifth column of the Table 8), being sporty has a huge impact on the probability of exiting unemployment in t but to a lower extent than having been to high school (instead of having quit after the elementary school). Indeed, practising sports weekly is correlated to an increase by 96% in this probability when having a high school level increases this probability by more than

175%. Finishing the secondary school would increase the exit rate by 29% and having a university degree by 19%. Figures seems surprising but, remember here that the majority of the sub-sample did not study further than the elementary school (reference category in the Tables). And being sporty is rewarded less than having achieved high school which is consistent with ones' expectations. The results are similar when looking at the Heckman and Singer specification. Being sporty increases the exit rate from unemployment in t by more than 100% but the impact of achieving high school is greater, increasing this same rate by almost 147%. As expected, living in former East-Germany has a negative influence on the probability of leaving unemployment in t. This probability decreases by 68% in the Weibull model and by 38% in the Heckman and Singer model.

In the male sample, as already outlined, there is no correlation between being sporty and the exit rate for people who experienced unemployment more than once over the period. Therefore, in order to build an understanding of the mechanisms, I focus on results on the whole sample, without forgetting that they may be much less reliable. The correlation between my variable of interest and the probability of exiting unemployment in t is sizeable but less important than correlations with levels of education or even the nationality or the marital status. Results estimated with the Heckman and Singer are similar and draw the same pattern. Therefore, in addition to the fact that being sporty is an advantage only for men who experience only one unemployment spell, it is relatively little rewarded. This leads me to infer that the influence of being sporty has to be associated to an impact which is neither very influent (relatively to the traditional ones), nor constant. The characteristics arguably suit the signalling effect.

The use of the Heckman and Singer model gives results which are almost similar to those obtained using a Weibull model with a parametric shared frailty. But, since the use of a semi-parametric framework for the frailty is heavier –the design of the sample is different and the estimation often fails to converge–, from now on I use the Weibull model (including shared frailty). Furthermore, I comment in priority on the results on the restricted sub-sample since, including frailty, these are the most rigorous results.

5.2 Deepening the relationship

I test the existence of the signalling effect in two different ways. First, since one can signal oneself only relatively to the others, I construct a variable of the relative sport participation. However, the introduction of such variable into the estimates does not give any interesting results. There is no relationship between sport practice frequency and the length of the unemployment spell. In other words, people do not benefit from being relatively sportier than the others. Firms are not interested in individuals' relative sport participation. This result considerably weakens the likelihood of the signalling.

Another specificity of the signalling effect is that it is more effective for people who do not have much work experience. Since the models used are not linear, the use and then the interpretation of interaction terms is complicated. Therefore, I test the relationship on various sub-samples built with respect to a specific number of years of work experience. I use the Weibull model including shared frailty. Since I need to have reasonable sample size, I chose values of work experience which enables to divide the sample in equivalent parts (in terms of number of observations).

One third of men have less than 5.8 years of experience, one third has between 5.8 and 17 years of work experience and the last third has more than 17 years of experience. Being sporty has a no impact on the first and last sub-sample, but the correlation with the exit rate is positive and significant for the men who have between 5.8 and 17 years of work experience. And this result is still significant on the restricted sub-sample of people type $i_{U>1}$. Being sporty increases the probability of leaving unemployment by 48%.[20] When the male sample is divided (by the the median which corresponds to 10.7 years of work experience) in two parts, then men who have more than 10.7 years of work experience benefit from being sporty when the others do not. These results confirm the hypothesis according to which being sporty plays a role as a signal for men. The interpretation of this signal is not easy since it plays a role for people who are in the middle of their career.

One third of the women have less than 4.4 years of experience, one third have between 4.4 and 14 years of work experience and the last one third has more than 14 years of experience. For them, being sporty has almost always the same influence in terms of sign, size en significance. This is a strong result which supports the idea that being sporty is not a signal for women.

[20] Results are not displayed here but they are available upon request.

5.3 Sporting practice versus other activities

Because other activities could have similar advantages to sporting activities, it is relevant to look at the impact of our variable of interest once controlling for the participation in other activities. A first idea of the impact is given by the four graphics in the Appendix (Figures 1 to 4). The graphic 1 represents the hazard function by sub-samples (sporty versus non-sporty) of men, estimated by using a Weibull model and without including shared frailty. The highest curve draws the probability for sporty men to get out from unemployment in t, knowing that they stayed unemployed until $t-1$. The lowest curve draws the probability for non-sporty men to get out from unemployment in t, knowing that they stayed unemployed until $t-1$. The graphic 2 represents the same functions but controlling for other activities, i.e. including other activities in the estimation. The highest curve still corresponds to the sporty sub-sample. By comparing the two graphics, one can see that including other activities tends to lower the gap between the two curves.[21] The phenomenon is the same within the female sample (see figures 3 and 4), therefore, it is relevant to estimate the equation including these controls.

First, I run a "beauty contest" in order to find out which activity is the most efficient (the reference being the sporting activity) with respect to unemployment exit (see Table 9). Using the Weibull model, I compare the log likelihood among each estimation. I find out that weekly sporting activities is –for both gender– the variable which gives the greater explanation power (i.e. the log likelihood the nearest to zero). But it is still interesting to look at the activities which are significantly correlated to the exit rate.

In the male sample, it appears that outside of sporting activities, all the activities still have a significant impact once reducing the sample to people who experienced at least two unemployment spells over the period. Going to the cinema weekly (i.e. the proxy for the social network social habits) significantly increases the exit rate in t by 76%. Taking part in weekly voluntary work significantly increases the exit rate in t by 68%, meaning that the signalling effect matters. A less significant impact is associated to "being a churchgoer" (another proxy for the social network) but it is linked to an

[21] Without reconsidering the fact that sporty people have a higher hazard rate, i.e. that they get out faster from unemployment

increase in the exit rate in t by 41%. In the women sample, taking part in weekly voluntary work or being a churchgoer do not significantly impact the probability of leaving unemployment in t. But going to the cinema weekly is linked with an increase in this probability by more than 100% which overcomes the impact of being sporty. Therefore, if being sporty is the activity which -among others- explains a bigger part of the exit rate in t (i.e. which is the more efficient in terms of explanation power), it is not the activity which has the greater impact in term of size.

The results of the addition of other activities in the baseline estimation are presented in Table 10. For men who experienced at least two unemployment spells over the period, taking part in activities other than sport significantly increases the exit rate from unemployment in t by 55% for volunteering, by 38% for being churchgoers and by 68% for weekly cinemagoers. Since the last activity has the most significant and greater impact (in terms of size), I infer that it is the most relevant.[22] Another interesting point is that in the sub-sample restricted to men who experienced exactly one unemployment spell over the period, the unique beneficial activity in terms of exit probability from unemployment is the sporting activity. Therefore, I infer that for men, being sporty is neither a proxy for social network[23], nor for personal traits. Indeed, it is a signal which is effective only for people who do not experience more than one unemployment spell.

For women which have been at least twice unemployed over the period, adding other activities lowers the impact of being sporty but does not cancel it. Actually, it remains the most rewarded activity in terms of increase in the exit rate from unemployment in t. It increases this rate by 90% when going weekly to the cinema increases this rate by 80%. The fact that the coefficients of both activities outlast means that their impact passes trough different channels. As already inferred, it can be deduced that being sporty for women is a proxy for personal traits when going weekly to the cinema is a proxy for the magnitude of the social network. The channel through which "going to

[22] Actually, it is also the only one which remains significant when I use the Heckman and Singer procedure on this specific sub-sample.

[23] Unless I consider that the social network developed while being sporty is the only useful one for people who experience only one unemployment spell over the period. But it seems too precise to be interpreted in this way since I have no information on the type of sport practised.

the cinema" passes in order to favour unemployment exit is the networking effect. But, going to the cinema (or to concerts) weekly requires time and money. Unemployed people benefit from discounts for such activities which could increase -ceteris paribus- their demand for cultural events. However, behaviours do not change so fast. Therefore, this activity is also a proxy for the socio-economic background (I did not control for the initial position people held on the labour market before becoming unemployed).

6 Conclusion

The two models presented and the various estimates and hypothesis testing outlined in this paper lead to the conclusion that practising sport weekly during unemployment is significantly correlated to the probability of leaving unemployment. This result concerns German people, aged between 17 and 60 between 1994 and 2000, it has to be interpreted differently for men and women. The covariates include measures of the health status as well as the level of education, therefore, the channel of education and health can be moved apart.

For women, the impact is huge – but does not overcome the effect of education– and almost equivalent whatever the number of unemployment spells. Other activities which involve socialization do also positively impact the exit rate. However, it appears that being sporty is not just a proxy for being emotionally intelligent or well-integrated. The other types of activitis found to play a role reflect social network, cultural habits and the standard of living, indicating that sport has an impact across socio-economic groups. Since I control for the level of cognitive-skills (via the variables on education and work experience), these observations lead me to conclude that the impact of sporting practice is due to non-cognitive skills or act as a signal of non-cognitive skills. Since the effect of being sporty is huge and constant, I infer that, more than a signal, it conveys an effect of selection. Indeed, since I cannot deduce that by practising sport, women develop non-cognitive skills and personal traits such as motivation and concentration, I have to admit the potential effect of selection. Sporty women have personal traits which increase the productivity of their job-search.

For men, practising sport weekly is an advantage only for specific sub-

sample: for those who experienced exactly one unemployment spell over the period and for all those who have between 5.8 and 17 years of work experience. Furthermore, the impact is sizeable but not truly big. Sport participation is a way among others to accelerate the exit from unemployment to employment when participation in other activities is not useful. These partial and very sub-samples specific correlations leads me to infer that sporting activities are used as a signal.

There is a selection bias and it has been argued here that this will lead to an over-estimation of the effect of being sporty. However, even if the results cannot be generalized to the whole sample of unemployed people, it is relevant and positive to have such findings.

References

Bougard, Jonathan, Brodaty, Thomas, Emond, Céline, L'Horty, Yannick, du Parquet, Loïc, & Petit, Pascale. 2011. Tech. rept.

Bramoullé, Yann, & Saint-Paul, Gilles. 2010. Social networks and labor market transitions. *Labour Economics*, **17(1)**, 188–195.

Capperalli, Lorenzo, & Tatsiramos, Konstantinos. 2010. *Friends' Networks and Job Finding Rates*. IZA DP. IZA.

Cornelißen, Thomas, & Pfeifer, Christian. 2010. The Impact of Participation in Sports on Educational Attainment: New Evidence from Germany. *Economics of Education Review*, **29(1)**, 94–103.

Heckman, James, & Singer, Burton. 1984. A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica*, **52(2)**, 271–320.

Lechner, Michael. 2009. Long-run labour market and health effects of individual sports activities. *Journal of Health Economics*, **28(4)**, 839–854.

Rees, Albert. 1966. Information Networks in Labor Markets. *The American Economic Review*, **56(1)**, 556–566.

Rooth, Dan Olof. 2011. Work Out or Out of Work: The Labor Market Return to Physical Fitness and Leisure Sport Activities. *Labour Economics*, **18(3)**, 399–409.

Spence, James. 1973. Job Market Signaling. *Quarterly Journal of Economics*, **87**, 355–374.

7 Appendix

Figure 1: Weibull model: hazard rate by *sporty*, male sample.

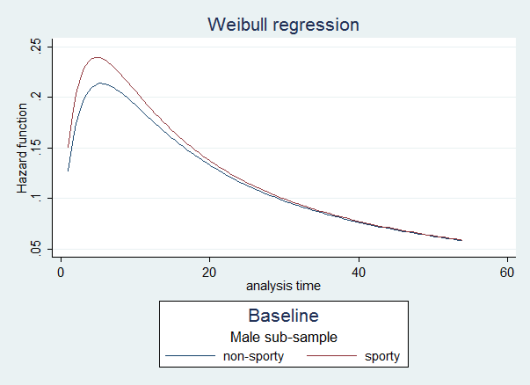

Figure 2: Weibull model: hazard rate by *sporty* controling for other activities, male sample

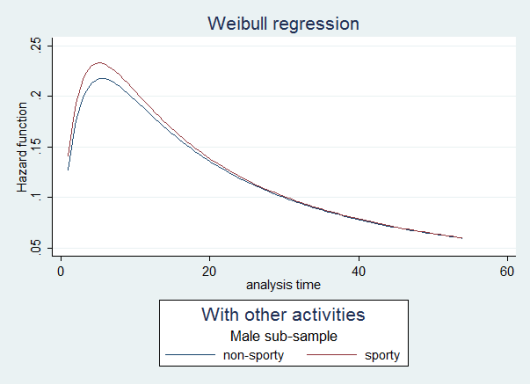

Figure 3: Weibull model: hazard rate by *sporty*, female sample.

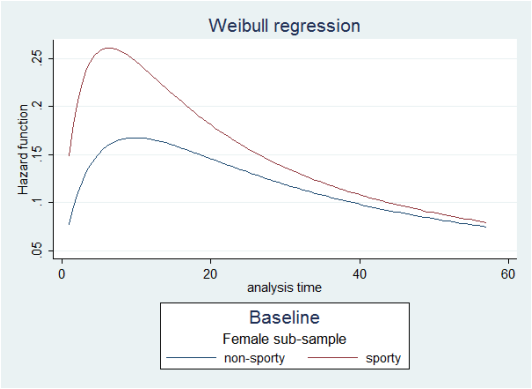

Figure 4: Weibull model: hazard rate by *sporty* controling for other activities, female sample

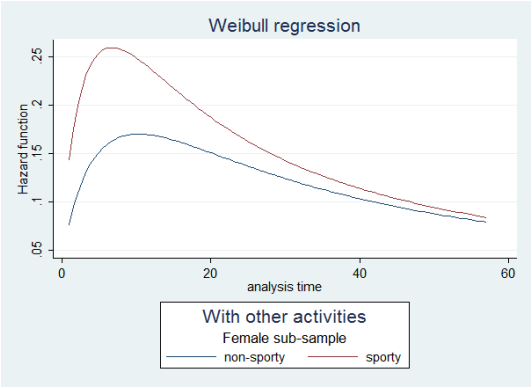

Table 1: Variables description

Variable	Obs	Mean	Std. Dev.	Min	Max
Individuals' activities					
sporty($t_{w1,u}$)	3318	.14	.35	0	1
being volunteer($t_{w1,u}$)	3316	.03	.18	0	1
being churchgoer($t_{w1,u}$)	3323	.04	.20	0	1
going out to cinema($t_{w1,u}$)	3321	.07	.26	0	1
daily hours dedicated to child care	3063	1.45	3.01	0	24
Individuals socio-economic characteristics					
woman	3331	.39	.49	0	1
German	3331	.85	.35	0	1
age	3331	35.09	10.84	17	60
married	3331	.55	.50	0	1
health status	3326	2.42	.86	1	5
work experience	3314	12.44	10.61	0	42.8
experience of unemployment	3314	1.13	1.62	0	20
total of unemployment spells	3331	2.19	1.34	1	7
no school	3279	.04	.20	0	1
elementary school	3140	.43	.49	0	1
secondary school	3282	.37	.48	0	1
high school	3282	.05	.22	0	1
university	3279	.12	.33	0	1
Living place					
Years					
1994	3331	.09	.28	0	1
1995	3331	.13	.34	0	1
1996	3331	.14	.35	0	1
1997	3331	.15	.36	0	1
1998	3331	.17	.38	0	1
1999	3331	.14	.35	0	1
2000	3331	.16	.37	0	1

B: at the beginning of the unemployment spell
E: at the end of the unemployment spell
BE: during the whole unemployment spell

Table 2: Selection

	Selection into unemployment				Selection into sport participation			
	empl.	**unempl.**	**sign**	**% var if unempl.**	**no sporty**	**sporty**	**sign**	**% var if sporty**
spell duration	98.81	21.51	***	-78.2%				
german	0.86	0.81	***	-5.8%	0.84	0.89	***	6.0%
no school	0.03	0.06	***	100.0%	0.04	0.02	***	-50.0%
elementary school	0.39	0.49	***	25.6%	0.43	0.31	***	-27.9%
secondary school	0.32	0.32			0.32	0.32	***	0.0%
high school	0.08	0.04	***	-50.0%	0.07	0.12		
university	0.18	0.11	***	-38.9%	0.15	0.23	***	53.3%
vocational education	0.63	0.61	***	-3.2%	0.63	0.60	***	-4.8%
general education	0.16	0.22	***	37.5%	0.17	0.14	***	-17.6%
woman	0.45	0.49	***	8.9%	0.47	0.43	***	-8.5%
age	38.71	39.14	***	1.1%	39.47	36	***	-8.8%
east	0.27	0.44	***	63.0%	0.31	0.21	***	-32.3%
religion	3.65	3.78	***	3.6%	3.66	3.70	***	1.1%
cinema	3.24	3.38	***	4.3%	3.39	2.87	-	-15.3%
volunteer	3.85	4.03	***	4.7%	3.94	3.67	***	-6.9%
sport	3.11	3.43	***	10.3%	3.79	1.32	***	-65.2%
volunteer week	0.06	0.03	***	-50.0%	0.03	0.12	***	300.0%
religion week	0.06	0.06	-	0%	0.05	0.07	***	40.0%
cinema week	0.06	0.07	***	16.7%	0.03	0.16	***	433.3%
sport week	0.21	0.15	***	-28.6%	-			
unemployed					0.15	0.11	***	-26.7%
married	0.67	0.6	***	-10.4%	0.68	0.56	***	-17.6%
health	2.41	2.64	***	9.5%	2.49	2.25	***	-9.6%

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3: Selection into unemployment at least twice by sample.

	SPORTY					NON SPORTY				
	$U \leqslant 1$	$U > 1$	**sign**	**dif if** $U > 1$	**% var**	$U \leqslant 1$	$U > 1$	**sign**	**dif if** $U > 1$	**% var**
spell duration	6,85	6,50		-0,35	-5,1%	8,44	7,22	***	-1,22	-14,5%
german	0,90	0,85	***	-0,05	-5,6%	0,84	0,88	***	0,04	4,8%
no school	0,02	0,04	***	0,02	100,0%	0,04	0,05	***	0,01	25,0%
elementary school	0,30	0,38	***	0,08	26,7%	0,43	0,43		0,00	0,0%
secondary school	0,31	0,35	***	0,04	12,9%	0,31	0,39	***	0,08	25,8%
high school	0,13	0,07	***	-0,06	-46,2%	0,07	0,04	***	-0,03	-42,9%
university	0,24	0,17	***	-0,07	-29,2%	0,16	0,11	***	-0,05	-31,3%
vocational education	0,59	0,65	***	0,06	10,2%	0,62	0,69	***	0,07	11,3%
general education	0,14	0,14		0,00	0,0%	0,17	0,15	***	-0,02	-11,8%
woman	0,44	0,40	***	-0,04	-9,1%	0,47	0,40	***	-0,07	-14,9%
age	36,40	32,53	***	-3,87	-10,6%	39,83	37,06	***	-2,77	-7,0%
east	0,18	0,41	***	0,23	127,8%	0,28	0,56	***	0,28	100,0%
religion	3,67	3,93	***	0,26	7,1%	3,62	3,88	***	0,26	7,2%
cinema	2,88	2,81	***	-0,07	-2,4%	3,40	3,33	***	-0,07	-2,1%
volunteer	3,66	3,76	***	0,10	2,7%	3,92	4,06	***	0,14	3,6%
married	0,57	0,45	***	-0,12	-21,1%	0,7	0,6	***	-0,10	-14,3%
health	2,24	2,35	***	0,11	4,9%	2,49	2,49		0,00	0,0%
volunteer weekly	0,12	0,11	*	-0,01	-8,3%	0,03	0,03	***	0,00	0,0%
religion weekly	0,07	0,05	***	-0,02	-28,6%	0,05	0,04	***	-0,01	-20,0%
cinema weekly	0,15	0,19	***	0,04	26,7%	0,03	0,05	***	0,02	66,7%

* significant at 10%; ** significant at 5%; *** significant at 1%
U is the number of unemployment spells experienced over the period.

Table 4: Results for different sport timings: Weibull model with and without shared frailty.

	No frailty			Shared frailty		
MEN	**All**	**U=1**	**U>1**	**All**	**U=1**	**U>1**
$sporty(t_{w1})$	0.045	0.189	-0.046	-0.035	-0.035	-0.040
	(0.116)	(0.166)	(0.148)	(0.139)	(0.246)	(0.176)
$sporty(t_u)$	-0.057	-0.016	-0.075	0.027	0.206	-0.107
	(0.122)	(0.191)	(0.149)	(0.141)	(0.253)	(0.179)
Constant	-0.238	-0.561	0.004	-0.068	-0.044	0.034
	(0.295)	(0.395)	(0.404)	(0.324)	(0.506)	(0.466)
$ln(p)$	0.185***	0.202***	0.194***	0.431***	0.503***	0.432***
	(0.017)	(0.026)	(0.021)	(0.028)	(0.067)	(0.032)
$ln(\theta)$				-0.743***	-0.528**	-0.756***
				(0.129)	(0.254)	(0.156)
Observations	1,783	652	1,131	1,783	652	1,131
Number of groups				1,185	652	533
Log likelihood	-2375	-859	-1496	-2316	-844	-1452

	No frailty			Shared frailty		
WOMEN	**All**	**U=1**	**U>1**	**All**	**U=1**	**U>1**
$sporty(t_{w1})$	0.172	-0.040	0.235	0.235*	0.072	0.261
	(0.123)	(0.206)	(0.147)	(0.136)	(0.250)	(0.166)
$sporty(t_u)$	0.008	0.263	-0.095	0.084	0.258	-0.012
	(0.117)	(0.179)	(0.148)	(0.134)	(0.240)	(0.166)
Constant	-0.809***	-1.087***	-1.117***	-0.723**	-1.486***	-0.768*
	(0.259)	(0.358)	(0.371)	(0.301)	(0.445)	(0.460)
$ln(p)$	0.208***	0.210***	0.240***	0.387***	0.404***	0.398***
	(0.021)	(0.031)	(0.027)	(0.035)	(0.070)	(0.042)
$ln(\theta)$				-1.109***	-0.979***	-1.203***
				(0.198)	(0.370)	(0.262)
Observations	1,165	492	673	1,165	492	673
Number of groups				883	491	392
Log likelihood	-1531	-648	-864	-1509	-643	-852

* significant at 10%; ** significant at 5%; *** significant at 1%
Robust standard errors in parentheses.
U is the number of unemployment spells experienced over the period.
Results include covariates: years, region, age, work experience, experience of unemployment, nationality, hours dedicated to children care, health status, marital status, total of unemployment spells over the period, and level of education.

Table 5: Results for persistent sporting activities: Weibull model with and without shared frailty.

	No frailty			Shared frailty		
MEN	**All**	**U=1**	**U>1**	**All**	**U=1**	**U>1**
$sporty(t_{w1,u})$	0.211***	0.365***	0.124	0.281***	0.417**	0.176
	(0.081)	(0.116)	(0.109)	(0.099)	(0.167)	(0.129)
Constant	-0.087	-0.680*	-0.109	0.095	-0.209	-0.084
	(0.279)	(0.389)	(0.402)	(0.319)	(0.499)	(0.462)
$ln(p)$	0.187***	0.207***	0.194***	0.431***	0.497***	0.431***
	(0.017)	(0.027)	(0.021)	(0.028)	(0.067)	(0.032)
$ln(\theta)$				-0.763***	-0.572**	-0.767***
				(0.129)	(0.262)	(0.156)
Observations	1,783	652	1,131	1,783	652	1,131
Number of groups				1,185	652	533
Log likelihood	-2371	-856	-1496	-2312	-842	-1452

	No frailty			Shared frailty		
WOMEN	**All**	**U=1**	**U>1**	**All**	**U=1**	**U>1**
$sporty(t_{w1,u})$	0.439***	0.432***	0.446***	0.670***	0.614***	0.673***
	(0.114)	(0.145)	(0.165)	(0.122)	(0.200)	(0.161)
Constant	-0.917***	-1.086***	-2.128***	-0.869***	-1.535***	-0.974**
	(0.267)	(0.363)	(0.420)	(0.305)	(0.450)	(0.463)
$ln(p)$	0.212***	0.214***	0.243***	0.402***	0.421***	0.415***
	(0.021)	(0.031)	(0.028)	(0.034)	(0.069)	(0.041)
$ln(\theta)$				-1.067***	-0.913***	-1.160***
				(0.188)	(0.345)	(0.242)
Observations	1,165	492	673	1,165	492	673
Number of groups				883	491	392
Log Likelihood	-1523	-646	-860	-1498	-640	-846

* significant at 10%; ** significant at 5%; *** significant at 1%
Robust standard errors in parentheses.
U is the number of unemployment spells experienced over the period.
Results include covariates: years, region, age, work experience, experience of unemployment, nationality, hours dedicated to children care, health status, marital status, total of unemployment spells over the period, and level of education.

Table 6: Baseline, results excluding the sport variable, using a Weibull model with shared frailty.

	MEN		WOMEN	
VARIABLES	**All**	**U>1**	**All**	**U>1**
german	0.671***	0.327**	0.323**	0.164
	(0.113)	(0.154)	(0.144)	(0.227)
exp work	0.099***	0.110***	0.039***	0.058***
	(0.013)	(0.018)	(0.010)	(0.016)
no school	-0.309*	-0.350*	-0.0352	-0.573
	(0.174)	(0.208)	(0.256)	(0.376)
secondary school	0.270***	0.282**	0.113	0.260*
	(0.094)	(0.121)	(0.103)	(0.141)
high school	0.668***	0.617**	0.813***	1.004***
	(0.173)	(0.251)	(0.198)	(0.311)
university	0.527***	0.337*	0.529***	0.663***
	(0.138)	(0.191)	(0.133)	(0.178)
total u spell	0.141***	0.126***	0.180***	0.245***
	(0.031)	(0.046)	(0.039)	(0.056)
exp of u	0.029	0.0575	0.020	0.092**
	(0.025)	(0.038)	(0.029)	(0.045)
daily hrs child	-0.062***	-0.072***	-0.063***	-0.057***
	(0.015)	(0.020)	(0.012)	(0.017)
married	0.348***	0.382***	0.082	0.179
	(0.088)	(0.115)	(0.094)	(0.126)
health status	-0.048	-0.020	-0.109**	-0.096
	(0.042)	(0.054)	(0.047)	(0.060)
age	-0.130***	-0.135***	-0.060***	-0.075***
	(0.014)	(0.019)	(0.010)	(0.015)
east	-0.222**	-0.327***	-0.471***	-0.544***
	(0.088)	(0.115)	(0.0995)	(0.142)
Constant	0.298	0.631	-1.202***	-1.638***
	(0.314)	(0.453)	(0.295)	(0.480)
$ln(p)$	0.435***	0.436***	0.378***	0.396***
	(0.028)	(0.032)	(0.035	(0.042)
$ln(\theta)$	-0.723***	-0.734***	-1.131***	-1.188***
	(0.127)	(0.153)	(0.206)	(0.266)
Observations	1,792	1,137	1,172	676
Number of groups	1,189	534	889	394
Log likelihood	-2328	-1459	-1525	-859

* significant at 10%; ** significant at 5%; *** significant at 1%
Robust standard errors in parentheses.
U is the number of unemployment spells experienced over the period.
Results include year dummies.

Table 7: Baseline results using a Weibull model, a Weibull model with shared frailty and an Heckman and Singer model including semi-parametric frailty; male sample.

	All			U>1		
	Weibull	**Weibull**	**Heckman & Singer**	**Weibull**	**Weibull**	**Heckman & Singer**
$sporty(t_{w1,u})$	0.211***	0.281***	0.295**	0.124	0.176	0.200
	(0.081)	(0.099)	(0.101)	(0.109)	(0.129)	(0.160)
german	0.511***	0.658***	0.602***	0.320**	0.332**	0.244
	(0.098)	(0.112)	(0.109)	(0.139)	(0.153)	(0.164)
exp work	0.083***	0.095***	0.069***	0.082***	0.106***	0.052**
	(0.013)	(0.013)	(0.013)	(0.018)	(0.018)	(0.020)
no school	-0.281**	-0.297*	-0.478**	-0.296**	-0.329	-0.678**
	(0.129)	(0.172)	(0.186)	(0.147)	(0.205)	(0.257)
secondary school	0.149**	0.240***	0.170*	0.147	0.255**	0.156
	(0.073)	(0.093)	(0.089)	(0.093)	(0.121)	(0.132)
high school	0.518***	0.638***	0.535**	0.418**	0.586**	0.421*
	(0.134)	(0.171)	(0.156)	(0.201)	(0.250)	(0.252)
university	0.379***	0.453***	0.421**	0.283**	0.308	0.294
	(0.108)	(0.138)	(0.132)	(0.135)	(0.190)	(0.213)
tot u spell	0.132***	0.145***	-0.396***	0.141***	0.130***	-0.285***
	(0.023)	(0.030)	(0.044)	(0.032)	(0.046)	(0.064)
exp of u	-0.001	0.025	-0.037	-0.014	0.052	-0.135*
	(0.023)	(0.025)	(0.028)	(0.030)	(0.037)	(0.046)
daily hrs child	-0.036***	-0.061***	-0.027*	-0.041***	-0.072***	-0.010
	(0.011)	(0.015)	(0.015)	(0.014)	(0.019)	(0.021)
married	0.244***	0.374***	0.233**	0.261***	0.389***	0.173
	(0.066)	(0.088)	(0.086)	(0.084)	(0.114)	(0.128)
health status	-0.023	-0.036	-0.049	-0.035	-0.010	-0.086
	(0.036)	(0.042)	(0.044)	(0.045)	(0.054)	(0.068)
age	-0.102***	-0.125***	-0.096***	-0.096***	-0.131***	-0.071***
	(0.013)	(0.014)	(0.014)	(0.018)	(0.019)	(0.020)
east	-0.126*	-0.206**	-0.077	-0.183**	-0.315***	-0.138
	(0.068)	(0.087)	(0.084)	(0.084)	(0.114)	(0.129)
Constant	-0.087	0.095	-0.430	-0.109	-0.084	-1.221**
	(0.279)	(0.319)	(0.389)	(0.402)	(0.462)	(0.604)
$ln(p)$	0.187***	0.431***		0.194***	0.431***	
	(0.017)	(0.028)		(0.021)	(0.032)	
$ln(\theta)$		-0.763***			-0.767***	
		(0.129)			(0.156)	
log t			0.313***			0.538***
			(0.070)			(0.114)
M2 const			0.951***			0.943**
			(0.193)			(0.319)
logitp2			0.982			0.960
			(0.703)			(1.024)
Shared frailty	No	Yes, Gamma	Yes, semi-parametric	No	Yes, Gamma	Yes, semi-parametric
Observations	1,783	1,783	11428	1,131	1,131	6750
Number of groups		1,185			533	
Log likelihood	-2371	-2312	-3463	-1496	-1452	-1698

* significant at 10%; ** significant at 5%; *** significant at 1%

Robust standard errors in parentheses.

U is the number of unemployment spells experienced over the period.

Results include covariates: years

Table 8: Baseline results using a Weibull model, a Weibull model with shared frailty and an Heckman and Singer model including semi-parametric frailty; female sample.

	All			U>1		
	Weibull	**Weibull**	**Heckman & Singer**	**Weibull**	**Weibull**	**Heckman & Singer**
$sporty(t_{w1,u})$	0.439***	0.670***	0.677***	0.446***	0.673***	0.773***
	(0.114)	(0.122)	(0.121)	(0.165)	(0.161)	(0.186)
german	0.213*	0.285*	0.290**	0.142	0.178	0.256
	(0.120)	(0.146)	(0.136)	(0.190)	(0.228)	(0.243)
exp work	0.033***	0.037***	0.021**	0.044***	0.057***	0.035**
	(0.010)	(0.011)	(0.010)	(0.013)	(0.016)	(0.016)
no school	-0.075	-0.004	0.171	-0.324	-0.461	-0.035
	(0.256)	(0.257)	(0.241)	(0.314)	(0.378)	(0.399)
secondary school	0.050	0.122	0.049	0.148	0.255*	0.135
	(0.089)	(0.105)	(0.101)	(0.116)	(0.142)	(0.151)
high school	0.598***	0.781***	0.640***	0.752**	1.014***	0.903**
	(0.194)	(0.202)	(0.183)	(0.337)	(0.310)	(0.372)
university	0.401***	0.495***	0.374**	0.454***	0.597***	0.428**
	(0.112)	(0.134)	(0.126)	(0.148)	(0.178)	(0.191)
tot u spell	0.165***	0.181***	-0.298***	0.237***	0.244***	-0.148**
	(0.033)	(0.040)	(0.048)	(0.036)	(0.056)	(0.073)
exp of u	0.011	0.022	-0.037	0.049	0.092**	-0.041
	(0.025)	(0.030)	(0.030)	(0.035)	(0.045)	(0.050)
daily hrs child	-0.042***	-0.060***	-0.052***	-0.034***	-0.050***	-0.032*
	(0.009)	(0.012)	(0.011)	(0.012)	(0.017)	(0.018)
married	0.072	0.061	0.112	0.147	0.133	0.269*
	(0.078)	(0.096)	(0.092)	(0.106)	(0.127)	(0.139)
health status	-0.074*	-0.102**	-0.102**	-0.059	-0.089	-0.090
	(0.039)	(0.048)	(0.047)	(0.051)	(0.061)	(0.068)
age	-0.049***	-0.057***	-0.042***	-0.055***	-0.071***	-0.050**
	(0.010)	(0.010)	(0.010)	(0.014)	(0.016)	(0.017)
east	-0.245***	-0.429***	-0.329**	-0.286**	-0.519***	-0.324**
	(0.085)	(0.101)	(0.097)	(0.126)	(0.142)	(0.156)
Constant	-0.917***	-0.869***	-1.691**	-2.128***	-0.974**	-1.853**
	(0.267)	(0.305)	(0.494)	(0.420)	(0.463)	(0.783)
$ln(p)$	0.212***	0.402***		0.243***	0.415***	
	(0.021)	(0.034)		(0.028)	(0.041)	
$ln(\theta)$		-1.067***			-1.160***	
		(0.188)			(0.242)	
log t			0.273***			0.379***
			(0.494)			(0.104)
M2 const			1.458***			0.860
			(0.345)			(0.581)
logitp2			2.724***			1.857
			(0.613)			(2.028)
Shared frailty	No	Yes, Gamma	Yes, semi-parametric	No	Yes, Gamma	Yes, semi-parametric
Observations	1,165	1,165	9268	673	673	5196
Number of groups		883			392	
Log likelihood	-1523	-1498	-2662	-860	-846	-1269

* significant at 10%; ** significant at 5%; *** significant at 1%
Robust standard errors in parentheses.
U is the number of unemployment spells experienced over the period.
Results include year dummies.

Table 9: Beauty contest between 4 extra-professional activities, Weibull model including shared frailty.

MEN	**All**				**T>1**			
$sporty(t_{w1,u})$	0.281***				0.176			
	(0.099)				(0.129)			
$cinema(t_{w1,u})$			0.470***				0.564***	
			(0.133)				(0.166)	
$volunteer(t_{w1,u})$				0.525***				0.520**
				(0.170)				(0.238)
$churchgoer(t_{w1,u})$		0.415***				0.342*		
		(0.161)				(0.191)		
Constant	0.095	-0.062	0.079	0.609*	-0.084	0.044	0.387	0.185
	(0.319)	(0.314)	(0.320)	(0.318)	(0.462)	(0.458)	(0.455)	(0.460)
$ln(p)$	0.431***	0.437***	0.438***	0.439***	0.431***	0.435***	0.438***	0.439***
	(0.028)	(0.028)	(0.028)	(0.028)	(0.032)	(0.031)	(0.031)	(0.032)
$ln(\theta)$	-0.763***	-0.723***	-0.724***	-0.650***	-0.767***	-0.750***	-0.749***	-0.710***
	(0.129)	(0.127)	(0.127)	(0.123)	(0.156)	(0.154)	(0.153)	(0.152)
Observations	1,783	1,790	1,790	1,790	1,131	1,136	1,135	1,135
Number of groups	1,185	1,188	1,189	1,188	533	534	534	533
Log likelihood	-2312	-2322	-2320	-2340	-1452	-1457	-1452	-1458

WOMEN	**All**				**T>1**			
$sporty(t_{w1,u})$	0.670***				0.673***			
	(0.122)				(0.161)			
$volunteer(t_{w1,u})$				0.098				-0.247
				(0.273)				(0.307)
$churchgoer(t_{w1,u})$		-0.286				-0.379		
		(0.269)				(0.405)		
$cinema(t_{w1,u})$			0.705***				0.727***	
			(0.170)				(0.243)	
Constant	-0.869***	-0.587**	-1.483***	-1.037***	-0.974**	-1.650***	-0.872*	-0.567
	(0.305)	(0.295)	(0.299)	(0.280)	(0.463)	(0.480)	(0.451)	(0.445)
$ln(p)$	0.402***	0.377***	0.379***	0.379***	0.415***	0.396***	0.389***	0.400***
	(0.034)	(0.035)	(0.035)	(0.035)	(0.041)	(0.042)	(0.042)	(0.042)
$ln(\theta)$	-1.067***	-1.151***	-1.184***	-1.125***	-1.160***	-1.191***	-1.309***	-1.157***
	(0.188)	(0.209)	(0.215)	(0.207)	(0.242)	(0.264)	(0.293)	(0.262)
Observations	1,165	1,171	1,168	1,168	673	676	674	675
Number of groups	883	888	887	886	392	394	394	394
Log likelihood	-1498	-1522	-1511	-1520	-846	-859	-853	-857

* significant at 10%; ** significant at 5%; *** significant at 1%
Robust standard errors in parentheses.
U is the number of unemployment spells experienced over the period.
Results include covariates: years, region, age, work experience, experience of unemployment, nationality, hours dedicated to children care, health status, marital status, total of unemployment, spells over the period, and level of education.

Table 10: Sporting activity versus other activities, Weibull model including shared frailty.

	MEN			WOMEN		
	all	**U=1**	**U>1**	**all**	**U=1**	**U>1**
$sporty(t_{w1,u})$	0.203**	0.349**	0.106	0.626***	0.497**	0.642***
	(0.102)	(0.175)	(0.131)	(0.123)	(0.204)	(0.164)
$volunteer(t_{w1,u})$	0.341**	0.230	0.440*	-0.062	1.734***	-0.368
	(0.170)	(0.269)	(0.234)	(0.276)	(0.592)	(0.314)
$churchgoer(t_{w1,u})$	0.392**	0.485	0.321*	-0.462*	-0.184	-0.601
	(0.160)	(0.314)	(0.191)	(0.279)	(0.385)	(0.433)
$cinema(t_{w1,u})$	0.397***	0.212	0.520***	0.542***	0.516**	0.590**
	(0.134)	(0.237)	(0.166)	(0.174)	(0.259)	(0.251)
Constant	-0.094	-0.794	-0.248	-1.712***	-1.771***	-2.030***
	(0.325)	(0.522)	(0.463)	(0.306)	(0.464)	(0.481)
$ln(p)$	0.439***	0.508***	0.436***	0.398***	0.435***	0.416***
	(0.028)	(0.067)	(0.031)	(0.034)	(0.069)	(0.041)
$ln(\theta)$	-0.751***	-0.546**	-0.782***	-1.149***	-0.926***	-1.221***
	(0.128)	(0.257)	(0.157)	(0.200)	(0.344)	(0.254)
Observations	1,781	651	1,130	1,161	491	670
Number of groups	1,183	651	532	882	490	392
log likelihood	-2299	-838	-1443	-1486	-632	-837

* significant at 10%; ** significant at 5%; *** significant at 1%

Robust standard errors in parentheses.

U is the number of unemployment spells experienced over the period.

Results include covariates: years, region, age, work experience, experience of unemployment, nationality, hours dedicated to children care, health status, marital status, total of unemployment spells over the period, and level of education.