



**HAL**  
open science

# Unemployment Duration and Sport participation : evidence from Germany

Charlotte Cabane

► **To cite this version:**

Charlotte Cabane. Unemployment Duration and Sport participation : evidence from Germany. 2011.  
halshs-00639307v1

**HAL Id: halshs-00639307**

**<https://shs.hal.science/halshs-00639307v1>**

Submitted on 8 Nov 2011 (v1), last revised 7 Apr 2014 (v2)

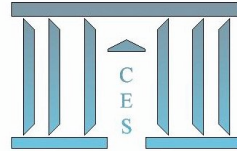
**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Documents de Travail du Centre d'Économie de la Sorbonne

C  
E  
S  
W  
o  
r  
k  
i  
n  
g  
P  
a  
p  
e  
r  
s



## Unemployment Duration and sport participation : evidence from Germany

Charlotte CABANE

2011.49



# Unemployment Duration and sport participation: evidence from Germany\*

Charlotte Cabane<sup>†</sup>

## 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. 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.

<sup>†</sup>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<sup>1</sup>). One specific situation on the labour market which is highly concerned by these channels has been left behind: unemployment.

For the same reasons why sporty people have a better situation 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).<sup>2</sup> There are many non-cognitive skills associated to sporty people and it can roughly be divided in two types: non-cognitive skills specific to certain types of sport and non-cognitive skills linked to sporty persons. The first type of non-cognitive skills includes team spirit and competitive spirit, for example. Skills which can be learned. Concentration and motivation form part to the second type and are often considered as personal traits and, thus, as constant individuals' characteristics. Here, I consider this type of non-cognitive skills are more pregnant into 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 explanation but it does not involves a causal effect peculiar 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, it is not necessarily caused by sports practice, it is just used as a signal and another activity could play the same role.

---

<sup>1</sup>Non-cognitive skills are personal preferences and personality traits which are valued in the society but which do not involve technical or professional knowledge. Unlike for cognitive skills, there is no school neither diploma which allows to learn and evaluate its.

<sup>2</sup>Question I already addressed in the previous chapter.

Besides these two channels<sup>3</sup>, 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 of 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  $t_{w2}$ .<sup>4</sup> 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 the 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 makes 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 applicant being involved in association are significantly less often called than people who are

---

<sup>3</sup>I intentionally do not develop on the reputation effect since it is specific to the US and I am working on German data.

<sup>4</sup>Actually, in the models used in this article, the unemployment spell are allowed to be right censored in order to be considered.

not. Plus, sporting practice requires usually some financial investment that unemployed people can less afford. Even if they receive helps and beneficiate from special access to sports facilities, they may be limited with respect to the type of sports they can practise.

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 much more people, therefore, sporting activity as a socialization process is more effective for employed people. Moreover, employed people can afford more types of sports than unemployed ones and also more expensive sports. This leads to more opportunities, or even more interesting contacts (in terms of labour market opportunities) or even to meet 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 sports 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 occur *before* and *during* the unemployment spell. I expect to observe a positive but smaller impact if the practice is only 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 sixth section contains the results. I conclude in the last section.

## 2 Literature review

Most of the literature on the relationship between sports 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 practice on individuals' unemployment duration. Thereby, I here put the emphasis

on articles which study the immediate returns of sports 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), they 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 rewarded at school or useful for the learning process. The soft skills attributed 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 for girls. They are supposed to be less competitive and to have less self-esteem. Therefore, they have relatively more to learn from sports practice so sporty schoolgirls catch up.

Lechner (2009) also uses the GSOEP and points out the positive relationship which exists 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.<sup>5</sup> Therefore they are more productive. Furthermore, they have unobservable specific characteristics which match with unobservable characteristics hold 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. By doing team sport, by practising sport in a club, people socialize and get to know people who are not

---

<sup>5</sup>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.

necessarily part of their original environment. And when people practice sports with their colleagues, it changes their usual professional relationship to personal relationship which are stronger because they do not necessarily reproduce the hierarchy of the relationship set by firm organization. Sport participation contribute to create other types of relationship between people and even with others (which do not already form part to individuals' 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 is the effect at work. Unlike the previous cited studies, Rooth (2011) is able to differentiate by different 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. Furthermore, sporty people may have non-cognitive skills -learned through sport participation- which optimize their job-search. In order to value as precisely as possible the effect, two specifications of duration model are successively used. They assume the 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, only the level changes.

Furthermore, the baseline hazard  $h_0(t)$  for all individuals is parameterized 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:

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

This model allows to have different baseline hazards for the different sub-samples with 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.<sup>6</sup>

More than the half of the sample experiences more than one unemployment spell. Unobservable characteristics which influence the risk of falling unemployed must be taken into account. In fact, forgetting to consider it leads to overestimate (underestimate) the degree of negative (positive) duration dependence. 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 the time. Because the level of frailty is unobserved, the impact of this selection is directly imputed to the 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  $\alpha$  which is Gamma distributed.<sup>7</sup> 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

---

<sup>6</sup>The systemic part does not change.

<sup>7</sup>The belonging of a duration to a group is estimated and not specified ex-ante.

follows:

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

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

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 suppose 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  $\alpha_j$  the frailty, which is now non-parametric and estimated using 2 mass points (one  $\eta$  is normalized to 1).<sup>8</sup>

This procedure allows for a piece-wise linear specification for duration dependence. However, I prefer to use a logarithmic form.

---

<sup>8</sup>The choice of the number of mass points results from the estimations I run.

### 3.3 Deepening the relationship

Using this 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: being volunteer, going to the cinema and being churchgoer.

The use of other activities allows to define more precisely the kind of relationship observed. Being 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 estimations. “Going to the cinema” also includes going to pop concerts, to dance or sporting events but it excludes going to museums, to theatre or classical concert. Therefore, besides the social aspect of being with friends, it also traduces a specific kinds of habits and budget. And, since both activities do not appear in curriculum, it cannot help to detect the signalling effect. However, the weekly activity: being a volunteer is almost always specified in a cv. This activity is used as a signal and thus, enables 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 being a volunteer 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 dedect 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 probability to exit from unemployment in  $t$  enables to find out if such selection on unobservables exists.

Since the signalling effect does not allow to give credit to the *sporting* 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 signaled 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 working experience. Therefore, the importance of sport participation impact on the exit rate should be greater for young people (or people with few working experience).<sup>9</sup> And, since I introduce a measure of the health status, I control for this part of individuals' productivity which also vary with the age. in order to catch this effect.

## 4 Data

### 4.1 Overview

The 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 questionnaires by year which enable to have a great definition of individuals' current and past situation. Labour market information is monthly recorded which allows to use survival models. However, individuals' characteristics are yearly recorded and information about sports is actually even less frequently asked. Individuals are questioned about their sports 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.*

---

<sup>9</sup>They pass less time on the labour market thus, the amount of information available with respect to their level of productivity is relatively low.

There is no formal definition of what the interviewer understands by *doing sports*<sup>10</sup> 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 who experience unemployment (sample in which sporty people are under-represented). Besides, since it is self-declared and well-judged to be sporty, 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 ( $sparty = 1$ ) 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 sports participation have, various dummy variables are built with respect to the timing of the practice: being sporty before being unemployed  $sparty(t_{w1})$ , being sporty while being unemployed  $sparty(t_u)$ , being sporty at the beginning of the unemployment spell *and* at the end of it  $sparty(t_{w1,u})$ .

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 goes in my sense. There is a significantly smaller proportion of unemployed people within sporty people. Furthermore, sporty people stay employed for a duration in average 10 months longer than non-sporty people and their unemployment spell last almost 7 months less. They are significantly unemployed for 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 the former East-Germany. They have in average a higher level of education and practice more often activities such as visiting friends, being volunteer, helping in the neighbourhood (but they are less churchgoer). And, as expected, they are healthier.<sup>11</sup>

Within the sample, people are between 17 and 60 years old (the two third of the population is less than 40 years old), 39% are women, 85% are German and they have studied -in mean- 11.4 years.

---

<sup>10</sup>Some years, the expression is formulated as written in the previous chapter or even using the expression *physical activity* instead of *sports*.

<sup>11</sup>Each differences in means commented here above are statistically significant, details on the level are available in the Table 2.

Transitions analyzed 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 mean 7.4 months. 28.8% are subjected twice to unemployment and 17.1% 3 times<sup>12</sup> and they last in mean respectively 7.4 and 6.4 months. In the mean, the first unemployment spell lasts 7.2 months, the second one lasts a half month less (6.7) and it keeps decreasing (the third spell lasts already less than 5.5 months). The sample being bound within time, the more one experiences unemployment spells, the shorter are the spells.

Sporty individuals experience significantly less unemployment than the others: in mean in total 2 unemployment spells versus 2.23. And, their unemployment spells are shorter: they last in mean 6 months for the first and unique spell versus 7.7 months for non-sporty. When they are twice unemployed within the sample, the first spell lasts in mean 5.5 months and 7.8 for non-sporty (and the second, respectively 6.5 month versus 7.4 months).

The covariates used in the model are defined with respect to the time as follows. Gender, nationality, level of education<sup>13</sup> 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 such as: place of residence (by region: former East/West-Germany), marital status (dummy), health status<sup>14</sup> and daily hours dedicated to child care are taken the last month of unemployment.

Values of variables which proportionally increase with time like age, working experience and unemployment experience are taken 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 in-

---

<sup>12</sup>The maximum being 7 spells between 1994 and 2000.

<sup>13</sup>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).

<sup>14</sup>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.

dividuals. 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 to be older, to live in the 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. And, at last, they practice less often activities<sup>15</sup> such as meeting friends, going to cultural events, going to the church, being volunteer 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, it presuppose the observation of at least two unemployment spells by individuals. 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 estimations on a restricted sub-sample (but I also present results obtained for the whole sample in order to have an idea of the selection and more information on the mechanisms). Since more rigorous results are expected from estimations on the restricted sub-sample, I study the difference of selection into being at least twice unemployed and having suffering less from unemployment over the period. Figures are reported in Table 3. For a 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 suffer 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

---

<sup>15</sup>Except helping neighbours.



much more younger than people type  $i_{U=1}$  in the sporty sample). And, people type  $i_{U>1}$  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 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 have to remember that non-sporty people tend to be more often unemployed than sporty people thus, they have more often 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 expectation- experienced at least twice unemployment (people type  $i_{U>1}$ )?* And if the selection is different, how would it bias my results. In the sporty sample, the selection comes to choose significantly more often young people, people with a very low level of education, people who are less satisfied by their health and foreigners. This specific selection favours the hypothesis that “yes”, by focusing on *more than once unemployed people*, I focus on less privileged people (with respect to usual individuals’ productivity characteristics). They have less characteristics valuable on the labour market. Therefore, being sporty can be one specificity among not many others that can help them for getting out from unemployment.<sup>16</sup> Consequently, my results risk to be over-estimated.

## 5 Results

### 5.1 Baseline

#### *Sport participation timing*

In a first step, I test the impact of different sports 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 exit from unemployment. Punctual sporting activities in this circumstances are not rewarded. However, Table 5 shows that there is a strong evidence

---

<sup>16</sup>This specificity can be seen as a signal but it can also lead people to behave differently.

of a correlation between being persistently sporty  $sporty(t_{w1,u})$  for men, but also for women. People have to be persistent or in other words, they have to be sporty. Since this characteristic is constant by 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 over the period. Analysis by sub-samples built with respect to the number of unemployment spells is then relevant. Differences due to timing sport participation is not surprising since I observe changes in sport participation contemporaneous to changes in labour-market status.<sup>17</sup> At least 15% of people who are sporty in  $t_{w1}$  stop to practise 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 to 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 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 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), sporting practice is correlated with the probability to exit from unemployment in  $t$ . It is interesting to outline that, the correlation is much more 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 on 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 explained these differences. First, the sport practised differs by gender and thus confers / gives different advantages. Second, the playing effect is not the same for

---

<sup>17</sup>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.

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 of social network for men, which is coherent with the fact that different sub-samples do not benefit the same from being sporty.

#### *Including shared frailty*

The introduction of the frailty is necessary -for the reasons presented before and- since the heterogeneity is significantly observed (the likelihood-ratio test indicates that  $\theta$  -the frailty variance- is significantly different from 0). Also, the results have to be interpreted knowing that each individual has a fixed  $\nu_i$  (i.e. individuals' level of frailty).<sup>18</sup> As outlined before, the most rigorous results are obtained using the sub-sample restricted to people which experience at least twice unemployment spell over the period. However, in order to have a roughly idea of the unobservable heterogeneity in the whole sample, I also present results on its. 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 increase 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 exit from unemployment 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 exceed 100% in the Heckman and Singer procedure for women who experienced at least two unemployment spell over the period. I recall 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

---

<sup>18</sup>Results are conditional on this level.

who have been only once unemployed over the period. And in this case, the increase in the exit 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. And, I 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 the time. For women which have experienced more than one unemployment spell (fifth column, Table 8) the equation of the hazard rate can be written as follows:

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

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

In each estimations, covariates have the expected correlation: being married, being German, having more working experience and having higher levels of education than the elementary school level is positively correlated to the probability of exit from unemployment in  $t$ . Conversely, feeling unhealthy, spending more hours to take care of children, having more experiences of unemployment, being older and living in the former East-Germany are linked to a decrease in the probability to exit from unemployment in  $t$ . In both sub-samples, 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 passes the effect of sporting activities on the exit rate from unemployment. 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, working 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 proba-

bility to exit from 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 seem surprising but, remember here that the majority of the sub-sample did not go further than the elementary school (reference category in the Tables). And being sporty is less rewarded than having achieved high school which is coherent with one's 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, it increases the same rate by almost 147%. As expected, living in the former East-Germany has a negative influence on the probability to exit from 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 have some intuitions on 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 to exit from 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 low / little rewarded. This leads me to infer that the influence of being sporty has to be associated to an impact which is neither very influential (relatively to the traditional ones), nor constant. Characteristics which arguably suits the signalling effect.

The use of the Heckman and Singer model gives results which are slightly 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 results on the restricted sub-sample since,

including frailty, there 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 estimations does not give any interesting results. There is no relationship between sporting practice frequency and the length of the unemployment spell. In other words, people do not profit by 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 working experience. Since the estimated regression 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 working experience. I use the Weibull model including shared frailty. Since I need to have reasonable sample size, I chose values of working experience which enables to divide the sample in equivalent parts (in terms of number of observations).

One third of the men have less than 5.8 years of experience, one third has between 5.8 and 17 years of working experience and the last one 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 working experience. And this result is still significant on the restricted sub-sample of people type  $i_{U>1}$ . Being sporty increases the probability to exit from unemployment by 48%. When the male sample is divided in two almost equal parts (the median corresponds to 10.7 years of working experience), then, people who have more than 10.7 benefit from being sporty when the others do not. I find similar results when I run the model on sub-samples built with respect to the age. This results confirms 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 working 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 and significance. This is a strong result which supports the idea that being sporty is not a signal for women. Here again results are equivalent when I run the estimations on sub-samples built by age.

### 5.3 Sporting practice versus other activities

Because other activities could have similar advantages as 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.<sup>19</sup> 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 sort of 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 estimations. 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.

---

<sup>19</sup>Without reconsidering the fact that sporty people have a higher hazard rate, i.e. that they get out faster from unemployment

In the male sample, it appears that except 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%. Being volunteer weekly significantly increases the exit rate in  $t$  by 68%, meaning that the signalling effect matters. A less significant impact is associated to “being churchgoer” (another proxy for the social network) but it is linked to an increase of the exit rate in  $t$  by 41%. In the women sample, being a volunteer or being churchgoer do not significantly impacts the probability to exit from 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 terms 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 which are not sport significantly increases the exit rate from unemployment in  $t$  by 55% if it is volunteering, by 38% if it is being churchgoer and by 68% if it is going weekly to the cinema. Since the last activity has the most significant and greater impact (in terms of size), I infer that it is the most relevant.<sup>20</sup> 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 from unemployment is the sporting activity. Therefore, I infer that for men, being sporty is neither a proxy for social network<sup>21</sup>, 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,

---

<sup>20</sup>Actually, it is also the only one which remains significant when I use the Heckman and Singer procedure on this specific sub-sample.

<sup>21</sup>Unless I consider that social network developed while being sporty is the only one usefull 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.



adding other activities lower 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 pass through 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 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 reductions for participating in 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 hold on the labour market before falling unemployed).

## 6 Conclusion

The two models presented and the various estimations and hypothesis testing outlined in this paper lead to the conclusion that practising sport weekly during unemployment is significantly correlated to the probability of exit from unemployment. This result concerns German people, aged between 17 and 60 between 1994 and 2000, it have to be interpreted separately 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 move 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 activity 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 working 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 constant, I infer that, more than a signal, it traduces 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 working experience. Furthermore, the impact is sizeable but not surprisingly big. Sport participation is a way among other to accelerate the exit from unemployment and when participation in other activities are 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.
- 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.
- 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.

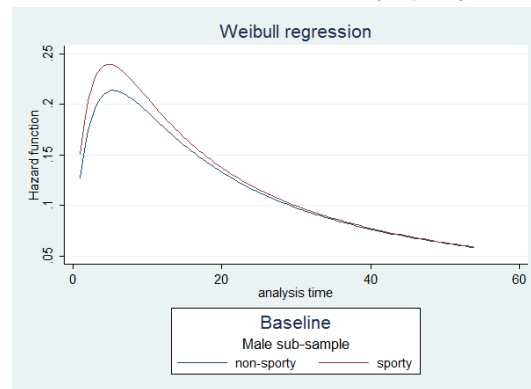


Figure 2: Weibull model: hazard rate by *sporty* controlling for other activities, male sample

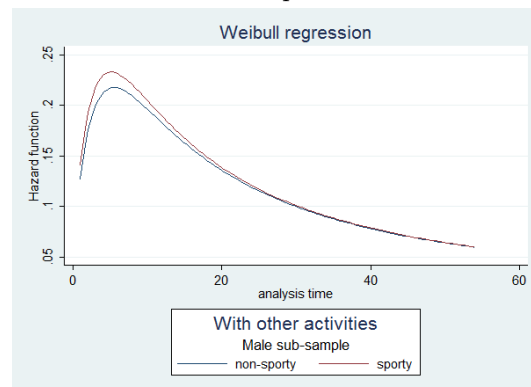


Figure 3: Weibull model: hazard rate by *sporty*, female sample.

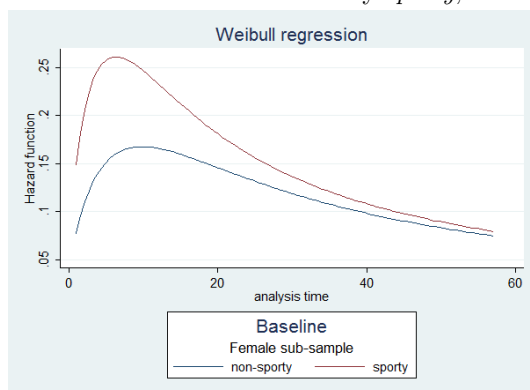


Figure 4: Weibull model: hazard rate by *sporty* controlling for other activities, female sample

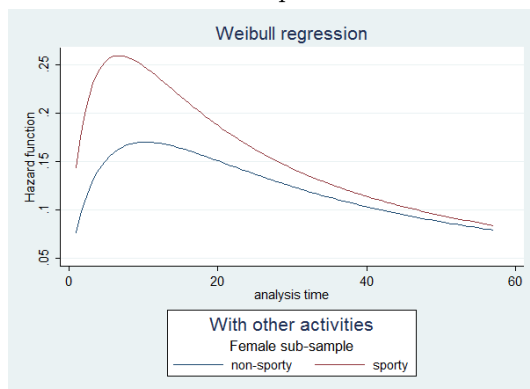


Table 1: Variables description

Variable	Obs	Mean	Std. Dev.	Min	Max
<b>Individuals' activities</b>					
sporty( $t_{w1,u}$ )	3318	.141953	.3490545	0	1
being volunteer( $t_{w1,u}$ )	3316	.0325694	.1775334	0	1
being churchgoer( $t_{w1,u}$ )	3323	.0412278	.1988466	0	1
going out to cinema( $t_{w1,u}$ )	3321	.0707618	.2564652	0	1
daily hours dedicated to housework E	3209	2.374572	2.28923	0	15
daily hours dedicated to child-care E	3063	1.45413	3.010937	0	24
daily hours dedicated to hobbies E	3220	1.704658	1.616563	0	10
<b>Individuals socio-economic characteristics</b>					
woman	3331	.3944761	.4888113	0	1
German	3331	.8537977	.353362	0	1
age B	3331	35.09152	10.84181	17	60
married E	3331	.5502852	.4975397	0	1
health status E	3326	2.422129	.8561729	1	5
working experience B	3314	12.44037	10.60942	0	42.8
experience of unemployment B	3314	1.134581	1.621302	0	20
total of unemployment spells	3331	2.191834	1.337877	1	7
no school	3279	.042391	.2015102	0	1
elementary school	3140	.4350318	.4958401	0	1
secondary school	3282	.3674589	.4821864	0	1
high school	3282	.0508836	.2197935	0	1
university	3279	.1222934	.3276743	0	1
<b>Living place</b>					
<b>Years</b>					
1994	3331	.0897628	.2858846	0	1
1995	3331	.1341939	.3409117	0	1
1996	3331	.1453017	.3524577	0	1
1997	3331	.1516061	.3586925	0	1
1998	3331	.1723206	.3777156	0	1
1999	3331	.1422996	.3494096	0	1
2000	3331	.1639147	.3702538	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				
	$u \leq 1$	$u > 1$	sign	dif if $u > 1$	% var	$u \leq 1$	$u > 1$	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%



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

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

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

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

Robust standard errors in parentheses.

Results include covariates: years, region, age, working 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.

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

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

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

Robust standard errors in parentheses.

Results include covariates: years, region, age, working 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.

VARIABLES	MEN		WOMEN	
	All	T>1	All	T>1
<b>german</b>	0.671*** (0.113)	0.327** (0.154)	0.323** (0.144)	0.164 (0.227)
<b>exp work</b>	0.0990*** (0.0134)	0.110*** (0.0183)	0.0393*** (0.0104)	0.0581*** (0.0156)
<b>no school</b>	-0.309* (0.174)	-0.350* (0.208)	-0.0352 (0.256)	-0.573 (0.376)
<b>secondary school</b>	0.270*** (0.0936)	0.282** (0.121)	0.113 (0.103)	0.260* (0.141)
<b>high school</b>	0.668*** (0.173)	0.617** (0.251)	0.813*** (0.198)	1.004*** (0.311)
<b>university</b>	0.527*** (0.138)	0.337* (0.191)	0.529*** (0.133)	0.663*** (0.178)
<b>total u spell</b>	0.141*** (0.0307)	0.126*** (0.0463)	0.180*** (0.0388)	0.245*** (0.0558)
<b>exp of u</b>	0.0290 (0.0254)	0.0575 (0.0376)	0.0198 (0.0292)	0.0919** (0.0446)
<b>daily hrs child</b>	-0.0621*** (0.0155)	-0.0720*** (0.0196)	-0.0627*** (0.0120)	-0.0575*** (0.0167)
<b>married</b>	0.348*** (0.0884)	0.382*** (0.115)	0.0819 (0.0944)	0.179 (0.126)
<b>health status</b>	-0.0477 (0.0422)	-0.0205 (0.0538)	-0.109** (0.0470)	-0.0958 (0.0603)
<b>age</b>	-0.130*** (0.0137)	-0.135*** (0.0189)	-0.0605*** (0.0101)	-0.0752*** (0.0155)
<b>east</b>	-0.222** (0.0878)	-0.327*** (0.115)	-0.471*** (0.0995)	-0.544*** (0.142)
<b>Constant</b>	0.298 (0.314)	0.631 (0.453)	-1.202*** (0.295)	-1.638*** (0.480)
<i>ln(p)</i>	0.435*** (0.0279)	0.436*** (0.0316)	0.378*** (0.0350)	0.396*** (0.0421)
<i>ln(θ)</i>	-0.723*** (0.127)	-0.734*** (0.153)	-1.131*** (0.206)	-1.188*** (0.266)
<b>Observations</b>	1,792	1,137	1,172	676
<b>Number of groups</b>	1,189	534	889	394
<b>Log likelihood</b>	-2328	-1459	-1525	-859

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

Robust standard errors in parentheses.

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.

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

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.

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

Results include year dummies.

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

MEN	All				T>1			
<i>sporty</i> ( $t_{w1,u}$ )	0.281*** (0.099)				0.176 (0.129)			
<i>cinema</i> ( $t_{w1,u}$ )			0.470*** (0.133)				0.564*** (0.166)	
<i>volunteer</i> ( $t_{w1,u}$ )				0.525*** (0.170)				0.520** (0.238)
<i>churchgoer</i> ( $t_{w1,u}$ )		0.415*** (0.161)				0.342* (0.191)		
<b>Constant</b>	0.095 (0.319)	-0.062 (0.314)	0.079 (0.320)	0.609* (0.318)	-0.084 (0.462)	0.044 (0.458)	0.387 (0.455)	0.185 (0.460)
<i>ln</i> ( $p$ )	0.431*** (0.028)	0.437*** (0.028)	0.438*** (0.028)	0.439*** (0.028)	0.431*** (0.032)	0.435*** (0.031)	0.438*** (0.031)	0.439*** (0.032)
<i>ln</i> ( $\theta$ )	-0.763*** (0.129)	-0.723*** (0.127)	-0.724*** (0.127)	-0.650*** (0.123)	-0.767*** (0.156)	-0.750*** (0.154)	-0.749*** (0.153)	-0.710*** (0.152)
<b>Observations</b>	1,783	1,790	1,790	1,790	1,131	1,136	1,135	1,135
<b>Number of groups</b>	1,185	1,188	1,189	1,188	533	534	534	533
<b>Log likelihood</b>	-2312	-2322	-2320	-2340	-1452	-1457	-1452	-1458
WOMEN	All				T>1			
<i>sporty</i> ( $t_{w1,u}$ )	0.670*** (0.122)				0.673*** (0.161)			
<i>volunteer</i> ( $t_{w1,u}$ )				0.098 (0.273)				-0.247 (0.307)
<i>churchgoer</i> ( $t_{w1,u}$ )		-0.286 (0.269)				-0.379 (0.405)		
<i>cinema</i> ( $t_{w1,u}$ )			0.705*** (0.170)				0.727*** (0.243)	
<b>Constant</b>	-0.869*** (0.305)	-0.587** (0.295)	-1.483*** (0.299)	-1.037*** (0.280)	-0.974** (0.463)	-1.650*** (0.480)	-0.872* (0.451)	-0.567 (0.445)
<i>ln</i> ( $p$ )	0.402*** (0.034)	0.377*** (0.035)	0.379*** (0.035)	0.379*** (0.035)	0.415*** (0.041)	0.396*** (0.042)	0.389*** (0.042)	0.400*** (0.042)
<i>ln</i> ( $\theta$ )	-1.067*** (0.188)	-1.151*** (0.209)	-1.184*** (0.215)	-1.125*** (0.207)	-1.160*** (0.242)	-1.191*** (0.264)	-1.309*** (0.293)	-1.157*** (0.262)
<b>Observations</b>	1,165	1,171	1,168	1,168	673	676	674	675
<b>Number of groups</b>	883	888	887	886	392	394	394	394
<b>Log likelihood</b>	-1498	-1522	-1511	-1520	-846	-859	-853	-857

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

Robust standard errors in parentheses.

Results include covariates: years, region, age, working 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.

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

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

Robust standard errors in parentheses.

Results include covariates: years, region, age, working 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.