LONGITUDINAL WORK ABILITY INDEX (WAI): A CASE STUDY WITH WORKERS IN AN AGRICULTURAL RESEARCH COMPANY

Lilian Ribeiro Lacerda da Silva
University of Araraquara, Brazil
E-mail: lilianrlacerda@gmail.com

Jorge Alberto Achcar
University of São Paulo, Brazil
E-mail: achcar@fmrp.usp.br

Jose Luis Garcia Hermosilla
University of Araraquara, Brazil
E-mail: jlghermosilla@hotmail.com

ABSTRACT

The main goal of this study is to describe the variability in work ability over time and to verify the influences of professional and sociodemographic variables. It is a longitudinal study conducted with 148 employees of an agricultural research company in a city located in the countryside of São Paulo state, Brazil. The Work Ability Index (WAI) questionnaire was applied to each individual in the sample in two times: December 2018 and June 2019. Different statistical analyzes of the data were performed using t-Student hypothesis tests for paired data, multiple linear regression models with normal errors, logistic regression models and Bayesian beta regression models assuming dependent WAI measures (the novelty of this study). From the obtained results, it is concluded that the Work Ability Index average decreases over time. Some covariates as education, physical activity, position in the company, number of children and nature of task have some influence on the variability of the WAI. The results of the data analysis show that the most important factor affecting the workers WAI index in both times is physical activity.

Keywords: Ability to work, WAI questionnaire, longitudinal study, Bayesian analysis.

470

1. INTRODUCTION

The ability to work is measured by the physical, mental and social conditions of the worker (Moreira, 2013; Cordeiro & Araujo, 2016). It is also defined as the balance between work and worker (Ilmarinen, 2006, 2007, 2009; Ahlstrom et al., 2010; Seitsamo, Tuomi & Martikainen, 2007). In order to maintain and promote work ability, it is important to avoid occupational diseases, remoteness and early retirement, that is, actions that are beneficial not only to the individual, but also to society (Martinez & Latorre, 2009; Martinez, Latorre & Fischer, 2010).

The Work Ability Index (WAI) has been used by organizations as a relevant tool to determine how well the worker is performing his activities, measuring his physical, mental and social capacity, and promoting the improvement of the work environment and to identify employees who need support (Silva Junior et al., 2013; Renosto et al., 2009; Tuomi et al., 2001; Tuomi et al., 2010; Martelin et al., 2008).

WAI longitudinal studies are needed as a way to obtain more concrete results on the availability for work of workers by analyzing its evolution and assessing the real cause, by monitoring the population, as well as directing the company in its strategies for maintaining and promoting the capacity to develop the work (Martinez, Latorre & Fischer; 2010; Cordeiro & Araujo, 2016).

The present study is aimed to describe the variability of WAI over time and to analyze the influence of some socio-demographic and professional variables on the evolution of WAI based on the obtained information of the workers sample considered in the study.

2. METHODS

It is a longitudinal study, considering employees of an agricultural research company located in the city of São Carlos, São Paulo State, Brazil. The company develops research in the agricultural area, such as increased milk production, meat improvement, and working with forages. The company structure is concentrated on a farm. In 2018, the company had 149 active employees, but only 148 worked in the city of São Carlos, so the survey was conducted with these 148 workers of the company.

The research was authorized by the head of the organization and approved by the Research Ethics Committee of the enterprise, since the study involves human beings. The information on some covariates (a list of the covariates used in the study is presented in an Appendix at the end of this manuscript) associated to the WAI questionnaire were obtained

471

from the company's database, leading to the research input variables such as gender (female or male), nature of the task (research, administrative, research support and field), position (researcher, analyst, technician and assistant), employment (effective or retired), length of employment (up to 10 years, 11 to 20 years, 21 to 30 years and over 30 years), marital status (married or single), education (elementary school, high school, higher education, and postgraduate), categorized age (31-40, 41-50, 51-60, and over 60), number of children (if yes or no and how much if so) and another covariate related to the frequency of physical activity practices (do not practice, rarely, a few times a month, every week and every day). The importance to detect important factors affecting the workers WAI indexes is the goal of many studies (Pohjonen, 2001).

The main response of interest was the WAI questionnaire response obtained for each worker in the company. According to Tuomi, et al. (2010) the WAI reaches a value ranging from 7 to 49 points, with 7 being the lowest rating and 49 the best rating of the index. Through the obtained results, it is possible to classify the ability to work as low, moderate, good and optimal. The WAI was applied in two times: December, 2018 and June, 2019 which characterizes longitudinal data. Several statistical techniques were used to analyze the WAI data. Initially it was used a standard statistical analysis for paired data to compare the WAI means in the two times.

The dataset was also analyzed using independent standard multiple linear regression models (Montgomery & Runger, 2010; Draper & Smith, 1981), leading to an analysis of the joint effect of covariates with the response given by the WAI assuming a transformation of the WAI data in the two periods (December 2018 and June 2019 denoted as WAI_1 and WAI_2), given by,

- $Y_1 = [(WAI_1 - 7)/42]$ (Y_1 is defined in the interval [0,1])

- $Y_2 = [(WAI_2 - 7)/42]$ (Y_2 is defined in the interval [0,1])   (1)

In addition to the use of a multiple linear regression model, a statistical treatment of the data was also performed using a logistic regression model, where the WAI response was categorized into a binary form, another way to analyze the relationship between the covariates and the WAI response. This type of regression model is used when the response variable is categorized into two classes usually referred to as success and failure (Bernoulli independent trials), quantified with values 1 and 0 (Montgomery & Runger, 2010). In the study, the optimal

and good WAI indexes were coded by the value 1 (success) and the low and moderate WAI indexes were coded by the value 0 (failure).

Finally, as the novel part of this study, it was assumed a Bayesian beta regression model (Ferrari & Cribari-Neto, 2004; Jorgensen, 1997) assuming dependent WAI measures. Since the WAI scores range is defined in the interval [7, 49], it was assumed the transformation (1) of the data to a [0,1] scale, where beta regression models were used in the data analysis assuming independent and dependent structures under a Bayesian approach (Box & Tiao, 1973). To capture the possible dependence between the repeated WAI measures it was considered the introduction of a latent variable in the regression models and the use of hierarchical Bayesian models using MCMC (Markov Chain Monte Carlo) methods (Gelfand & Smith, 1990; Chib & Greenberg, 1995).

2.1.   A Beta regression model

Since Y_1 and Y_2 introduced by (1) are defined in the interval [0,1], it is assumed beta probability distributions for each random variable, given by the density probability function,

$$f(y/a,b) = c\ y^{a-1}(1-y)^{b-1}\ ,\ \ 0 < y < 1 \tag{2}$$

where c is a beta function defined by, $c = B(a,b) = \Gamma(a+b)/\Gamma(a)\Gamma(b)$ where $\Gamma(v) = \int_0^\infty x^{v-1}e^{-x}\,dx$ (gamma function). The Beta distribution is denoted by $Y \sim Beta(a,b)$. The mean and variance of the beta distribution are given respectively by,

$$\mu = E(Y) = a/(a+b),$$

and $\tag{3}$

$$\sigma^2 = var(Y) = ab/[(a+b)^2(a+b+1)]$$

In presence of a vector of covariates associated to each unit, it is assumed a regression model considering a reparametrized form for the beta distribution with density (2) given by, $\mu = a/(a+b)$ and $\Phi = a + b$ (Ferrari & Cribari-Neto, 2004; Jorgensen, 1997). In this way we have, $a = \Phi\mu$, $b = (1 - \mu)\Phi$, $E(Y) = \mu$ and $var(Y) = V(\mu)/(1 + \Phi)$ where $V(\mu) = \mu(1 - \mu)$, so that μ is the mean of the response variable and Φ can be interpreted as a precision parameter in the sense that, for fixed μ, the larger the value of Φ, the smaller the variance of Y. The density of the random variable Y can be written, in the new parameterization, as,

$$f(y/\ \mu,\ \Phi) = \frac{\Gamma(\Phi)}{\Gamma(\Phi\mu)\Gamma[(1-\mu)\Phi]}\ y^{\Phi\mu-1}\,(1-y)^{(1-\mu)\Phi-1} \tag{4}$$

where $0 < \mu < 1$ and $\Phi > 0$.

473

Assuming the presence of a covariate vector $\mathbf{x} = (x_1, x_2,..., x_p)'$ with p covariates associated to each WAI observation, it is assumed the following regression model for the mean (Cepeda-Cuervo, Achcar & Lopera, 2013),

$$\log[\mu/(1-\mu)] = \text{logit}(\mu) = \boldsymbol{\beta'}\mathbf{x} \qquad (5)$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2,..., \beta_p)'$ is a vector of regression parameters.

This modeling approach was proposed by Cepeda and Gamerman (2005) considering the joint modeling of the mean and variance or dispersion parameters in the biparametric exponential family, including joint modeling of the mean and dispersion parameters in the beta distribution, under a Bayesian approach. The use of Bayesian methods has been extensively used in data analysis given the great flexibility of fit by combining a likelihood function with a prior distribution based on the Bayes formula (Box & Tiao, 1973) which could be informative using prior information of experts or non-informative when there is little prior information on the parameters of the proposed model.

In this study, it is introduced a Bayesian analysis for the WAI data assuming bivariate regression models. The dependence between the observed proportions is considered using hierarchical Bayesian models (Gelman et al., 2004) to capture the possible dependence between the two WAI measures. Combining the joint prior distribution for the parameters of the model, $\boldsymbol{\beta} = (\beta_1, \beta_2,..., \beta_p)'$ and Φ, with the likelihood function based on the reparametrized form of the beta density given by (4), the joint posterior distribution for $\boldsymbol{\beta}$ and Φ is determined from the Bayes formula (Box & Tiao, 1973). The posterior summaries of interest are obtained using Markov Chain Monte Carlo (MCMC) simulation methods as the popular Gibbs sampling algorithm or the Metropolis-Hastings algorithm (Gelfand & Smith, 1990; Chib & Greenberg, 1995) using the free existing OpenBugs software (Lunn et al., 2000).

3. COMPANY CHARACTERIZATION AND RESULTS

The company subject to this study has a predominantly older working class, with approximately 67% of its employees aged 41-60 and over 83% of workers aged over 41 years. This is partly due to the human resources policy that the company, of a public nature, has been adopting as a result of the budgetary restrictions imposed by the federal government, which no longer opens job vacancies for the replacement and renewal of its staff.

The last recruitment that the company opened was in in the year of 2009. Another aspect that deserves attention in this regard is the accumulation of activities for workers who stay and assume other responsibilities due to retirement and dismissal of many employees. In the

organization, there is a predominance of male workers, composed of 68.24% of the employees, while 31.76% are female. This characteristic is attributed to activities performed in the agricultural sector, such as animal handling, operation and maintenance of agricultural tractors, and even in the field of research, where the concentration of men is usually higher than that of women.

Related to education, 52.03% of the workers have graduate education, a necessary level of education for a company whose main activity is research. However, approximately 10% of workers have only elementary school and perform field activities to maintain the farm. Of the company's workers, 73.65% are married and 76.35% have at least one child. A relevant fact is that some workers have family ties with other workers, that is, in the company there are couples, brothers, brothers-in-law, who work together.

The company also provides homes for some field workers and their families. In general, 29% of the workers have up to 10 years in the job company, 26% have from 11 to 20 years, 23% from 21 to 30 years and 22% over 30 years. Regarding the type of the worker employment, about 78% of respondents are permanent employees of the company and 22%, although still working, are retired workers. This is a relevant information for the organization's managers, as about 20% of workers can leave the job as soon as they want.

With reference to the positions of the participants, 35% are assistants, 28% are researchers, 25% are analysts and 12% are technicians and at all job levels there are a lack of employees. However, a major concern of the company is in relation to assistant workers, who work directly in the field, because even with the opening of new recruitments there is difficulty in hiring this type of worker in São Paulo state.

So one of the great challenges of the company is the mechanization of some activities, facing the containment of expenses. Regarding the nature of the task, there is 30% of people working in the administrative area of the company, 28% working in the research area, 28% working in the field and 14% working in the research support. Although the company has as its purpose the research, being a public company, the processes have been very bureaucratic and therefore there is concentration of workers in the administrative area. The company has been working to improve and unify some systems to reduce rework and make activities less bureaucratic.

The first WAI dataset here denoted as WAI_1 was collected in December, 2018 with the participation of the 148 employees of the company. The second WAI dataset here denoted as

WAI_2 was collected in June, 2019, six months after the first collection with the participation of the same 148 individuals, that is, there was no dropout. With the obtained dataset initially was performed a paired data analysis to verify if there is statistical difference between the WAI means in the two considered times. Table 1 shows the obtained inference results (use of a paired Student-t test and the Minitab® software).

Table 1: Paired data analysis for WAI_1 and WAI_2

WAI	Sample Average	St-Deviation	p-value
WAI_1	40.561	4.756	0.046
WAI_2	39.882	4.999	

Source: The authors (2020)

From the results presented in Table 1, it is observed that the means for WAI_1 and WAI_2 are different, since the p-value is $0.046 < 0.05$ (significance level set at 5%). In addition, it is observed that the sample average of WAI_1 is higher than the sample average of WAI_2, that is, it is possible to conclude that there was a significant variability in the work ability index with a decrease in the mean of WAI over time.

Another statistical analysis was considered using independent multiple linear regression models relating all covariates with the WAI transformed responses (1) for each period. Tables 2 and 3 (use of the Minitab® software) show the obtained inference results (F-statistics and p-values associated to each regression parameter). The required assumptions for the use of linear regression models with normal errors (normality and constant variance for the residuals) were verified from standard residual plots.

As observed in the results presented in Table 2, WAI_1 is statistically affected by physical activity, with p-value equal to $0.006 < 0.05$, and by the number of children, with p-value equal to 0.054 (close to 0.05), assuming the usual significance level equal to 5%. The other covariates considered in the study (age, gender, education, marital status, nature of the job, job title, employment relationship, and length of employment) do not show significant effects on the WAI_1 response (p-value > 0.05).

The least squares estimator for the regression parameter associated with the number of children (use of a dummy variable in the multiple linear regression model (Draper & Smith, 1981) to compare workers without children with those who have at least one child, where the reference is not having children) is given by 0.0580 which shows that WAI_1 increases for those who have at least one child compared to those who have no children.

476

Table 2: Results of multiple linear regression analysis with response Y_1 (WAI_1)

Variables		F-Statistics	p-value
Socio-demographics	Age	0.46	0.501
	Gender	0.04	0.836
	Education	0.43	0.731
	Marital status	0.03	0.856
	Number of Chidren	2.61	0.054
	Physical activity	3.84	0.006
Professionals	Nature of task	0.36	0.785
	Position	1.17	0.325
	Employment bond	0.15	0.695
	Company Time	1.30	0.277

Source: The authors (2020)

Regarding to physical activity, the least squares estimator of the corresponding regression parameter (use of dummy variables where the reference is the one who practices physical activity every day) is equal to - 0.0691, implying that WAI_1 index has a significative decreasing for those who do not practice physical activity or practice few physical activities when compared to those who practice physical activity every day.

Considering the WAI_2 response, the only covariate that shows a significative effect on the response is physical activity, with p-value equal to $0.003 < 0.05$, as observed in Table 3. The least squares estimator for the associated regression parameter for physical activity (use of a dummy variable) is given by - 0.1494 showing that WAI_2 significantly decreases among those who do not engage in physical activity or have few physical activity compared to those who practice physical activities every day.

A third statistical analysis of the data was performed considering a logistic regression model, where WAI values were categorized with values 1 (optimal and good) and 0 (low and moderate), where maximum likelihood estimates (use of iterative numerical methods) for the regression parameters were obtained using the Minitab® software. Table 4 shows the p-values associated to each regression parameter.

Table 3: Results of Multiple Linear Regression with response Y2 (WAI2)

Variables		F-Snedecor	p-value
Sociodemographics	Age	2.33	0.129
	Gender	3.03	0.084
	Schooling	1.11	0.346
	Marital status	0.81	0.370
	Number of Chidren	2.10	0.104
	Physical activity	4.28	0.003
Professionals	Nature of task	2.37	0.074
	Position	0.11	0.956

Employment bond	0.13	0.716
Company Time	0.48	0.696

Source: The authors (2020)

Table 4. Binary Logistic Regression Results for WAI$_1$ and WAI$_2$

Covariates	p-value (WAI$_1$)	p-value (WAI$_2$)
Age	0.183	0.257
Gender	0.905	0.435
Schooling	0.069	0.022
Marital status	0.364	0.396
Number of Chidren	0.148	0.612
Physical activity	0.039	0.005
Nature of task	0.148	0.073
Position	0.012	0.260
Employment bond	0.351	0.831
Company Time	0.739	0.566

Source: The authors (2020)

From the results of Table 4, it is observed that there is statistical dependence of the binary WAI$_1$ with physical activity and position (p-values < 0.05), and binary WAI$_2$ with education and physical activity (p-values < 0.05).

3.1. Use of a Beta regression model under a Bayesian approach for the dependent WAI indexes

Since the two WAI indexes are evaluated to the same individual, it is possible to have some dependence between WAI$_1$ and WAI$_2$ (WAI$_1$ measured in December, 2018 and WAI$_2$ measured in June, 2019). Figure 1 shows the scatter plot of WAI$_1$ versus WAI$_2$, from where it is observed an approximately linear relation between the WAI$_1$ and WAI$_2$ indexes indicating dependence. The sample Pearson correlation coefficient between WAI$_1$ and WAI$_2$ is given to 0.647 (p-value < 0.001) indicating a significe correlation between the two responses measured for each individual. This result shows the need of better regression modeling approach in the data analysis incorporating the dependence structure. In this way, it is assumed a hierarchical Bayesian model to jointly analyze the WAI$_1$ and WAI$_2$ indexes.
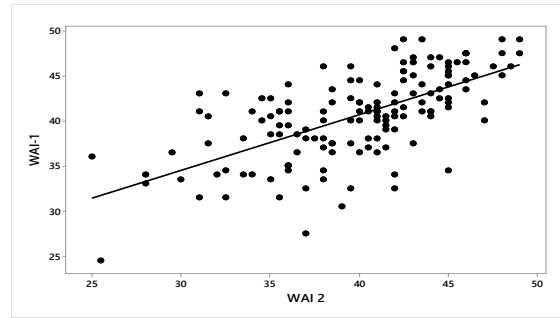
478

Figure 1: Scatterplot of WAI_1 versus WAI_2

Considering the regression models defined by (4) and (5), it is assumed dummy variables related to the unordered categorical variables given by nature of the task (reference is research) and position (reference is researcher). The physical activity was measured in the two times of WAI measures (December 2018 and June 2019) also in the transformed form Y_1 and Y_2 introduced in (1). For a Bayesian analysis of the data, it was considered the introduction of a latent variable W_i assuming a normal distribution with mean zero and variance σ_w^2 , that is, $N(0, \sigma_w^2)$ to capture the possible dependence between the two WAI measures, that is,

logit (μ_{ri}) = β_{r0} + β_{r1}gender$_i$+ β_{r2} education$_i$ + β_{r3}age$_i$ + β_{r4} marital.status$_i$ + β_{r5} number.children$_i$ + β_{r6} physical.activity$_i$ + β_{r7} employment.bond$_i$ + β_{r8} company.time$_i$+ β_{r9} nature.administrative$_i$+ β_{r10} nature.research.support$_i$ + β_{r11} nature.field$_i$ + β_{r12} analyst.position$_i$ + β_{r13} technical.position$_i$+ β_{r14} assistant.position$_i$ + w_i (6)

where r = 1 (WAI_1 measured in December 2018) and r = 2 (WAI_2 measured in June 2019) for i = 1,2,…, 148. The covariate age is considered in a standardized form (age - average)/SD.

For a hierarchical Bayesian analysis of the model it is assumed normal prior distributions with means equal to zero and fixed values for the variances for all regression parameters and an uniform U(0,10) prior distribution for the parameter $\zeta = 1/ \sigma_w^2$. Posterior summaries of interest were obtained using the OpenBugs software (LUNN et al., 2000) using standard MCMC procedures. It was generated 111,000 samples for each parameter of interest where the 11,000 first simulated samples were discarded as a burn-in period, which is usually used to minimize the effect of the initial values.

The posterior summaries of interest were based on 1,000 samples, taking every 100th sample to have approximately uncorrelated values. Convergence of the MCMC algorithm was monitored by usual time series plots for the simulated samples and also using some existing Bayesian convergence methods. Table 5 shows the posterior summaries of interest (posterior means, posterior standard deviations and 80% credible for each parameter).

479

Table 5: Posterior summaries (responses Y_1 and Y_2)

	mean	sd	Lower c.i.	Upper c.i.
β_{10}	1.328	0.4015	0.8523	1.866
β_{11}	0.01228	0.08055	-0.0846	0.1169
β_{110}	-0.009842	0.08272	-0.1171	0.09526
β_{111}	-0.01167	0.09183	-0.1314	0.1039
β_{112}	0.02969	0.08552	-0.07832	0.1369
β_{113}	-0.002856	0.09292	-0.1184	0.122
β_{114}	-0.03509	0.08989	-0.1487	0.07978
β_{12}	0.0653	0.07076	-0.02587	0.1575
β_{13}	-0.04396	0.06782	-0.1303	0.04054
β_{14}	-0.04536	0.08139	-0.1485	0.05345
β_{15}	0.08419	0.06043	0.005568	0.1604
β_{16}	-0.08535	0.04569	-0.1414	-0.02792
β_{17}	0.03183	0.08927	-0.08592	0.1475
β_{18}	0.06387	0.06131	-0.01083	0.142
β_{19}	0.02236	0.08075	-0.08632	0.1229
β_{20}	0.7473	0.3891	0.255	1.284
β_{21}	0.1026	0.07719	0.0277	0.1966
β_{210}	-0.01129	0.08847	-0.1247	0.1023
β_{211}	-0.03078	0.08808	-0.1396	0.08387
β_{212}	0.009325	0.08248	-0.1005	0.112
β_{213}	-0.02793	0.08693	-0.1396	0.08502
β_{214}	0.02925	0.08975	-0.08504	0.1407
β_{22}	0.1229	0.06164	0.04698	0.1992
β_{23}	0.06301	0.06873	-0.02559	0.1521
β_{24}	0.09413	0.08072	-0.00686	0.1953
β_{25}	0.02963	0.05819	-0.04703	0.1035
β_{26}	-0.06776	0.04216	-0.1222	-0.01324
β_{27}	0.01286	0.08837	-0.0989	0.1257
β_{28}	0.03506	0.06075	-0.04348	0.1102
β_{29}	0.05104	0.08285	-0.05611	0.1555
φ_1	22.31	4.296	17.1	27.68
φ_2	30.26	6.129	22.75	38.41
ζ	1.663	0.2591	1.34	2.013

Source: The authors (2020)

From the results of Table 5, we get the following interpretations:

- The covariates number of children and physical activity affect the mean of WAI_1 (in December 2018), since the zero value is not included in the 80% credible intervals for the regression parameter β_{15} and β_{16}.

480

- The covariates gender, education and physical activity affect the mean of WAI_2 (in June 2019), since the zero value is not included in the 80% credible intervals for the regression parameters β_{21}, β_{22} and β_{26}.

Table 6 shows a summary of the obtained inference results assuming the different statistical models used in this study.

Table 6: Significate covariates in the regression models (*5%) and (**10%)

Independent classical multiple regression model in the responses Y_1 and Y_2	
Y_1 (WAI_1)	Number of children*, physical activity*
Y_2 (WAI_2)	Gender**, physical activity*, nature of task**
Binary regression on binary WAI indexes	
WAI_1	Schooling degree**, physical activity*, position in company
WAI_2	Schooling degree*, physical activity*, nature of task**
Dependent Bayesian Beta regression model in the responses Y_1 and Y_2	
Y_1 (WAI_1)	Number of children, physical activity
Y_2 (WAI_2)	Gender, physical activity, schooling degree

Source: The authors (2020)

4. DISCUSSION OF THE OBTAINED RESULTS AND CONCLUDING REMARKS

From the obtained results using different statistical models, it is possible to see that the most important factor affecting WAI_1 and WAI_2 is physical activity (1: every day; 2: every week; 3: few times a month; 4: rarely; 5: do not practice physical activity). This effect is illustrated in Figure 2 from where it is possible to observe larger WAI values for workers that have physical activity every day and smaller WAI values for workers that do not practice physical activity in both times.

From Figure 3, it is possible to observe that WAI values are affected by the number of children in time 1 (WAI_1) but is not affected by the number of children in time 2 (WAI_2). This is in agreement with the results presented in Table 6. In Figures 4 and 5, it is possible to observe that WAI values are affected by gender (larger WAI value in time 2 for males) and education in times 1 (WAI_1) and 2 (WAI_2) where it is observed larger values of WAI for higher levels of education (see Appendix with the levels of the covariates). These results are also in agreement with the results presented in Table 6.
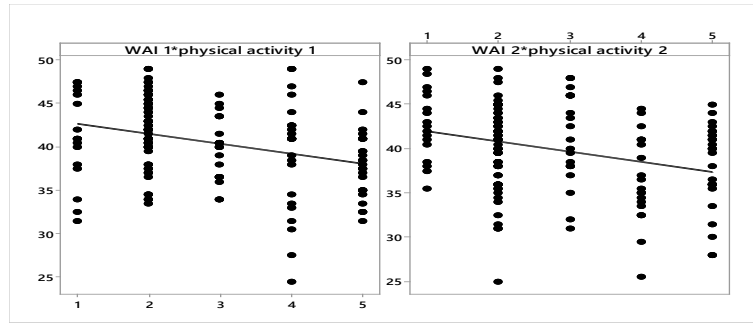
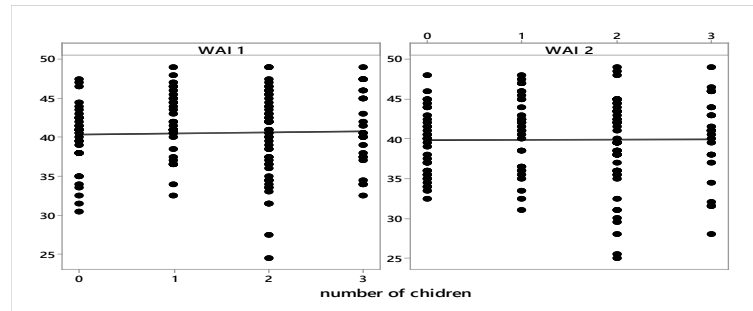Figure 2: Scatterplot of WAI_1 and WAI_2 versus physical activity

Figure 3: Scatterplot of WAI1 and WAI2 versus number of children
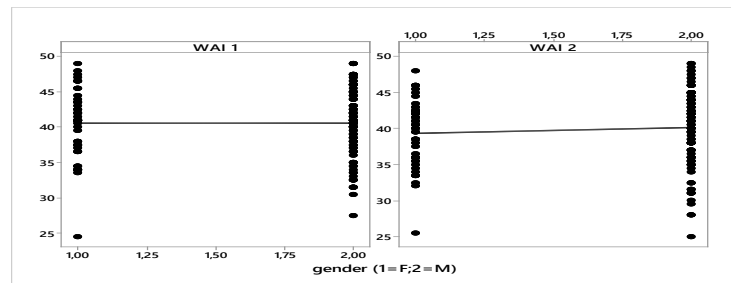

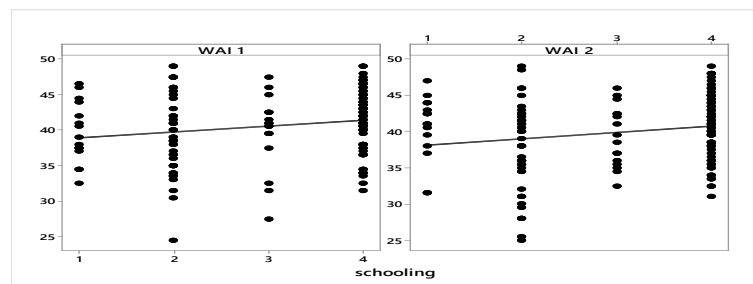Figure 4: Scatterplot of WAI1 and WAI2 versus gender

Figure 5. Scatterplot of WAI1 and WAI2 versus level of education

From these results, it is possible to see that the use of different statistical analysis assumed in this study were very important to detect the factors (covariates) affecting the work capacity of workers in the agriculture company in both times where WAI index was evaluated for each worker.

It is also important to point out that the use of beta regression models under a Bayesian approach could be a suitable alternative in the data analysis of longitudinal WAI data in presence of covariates, especially considering the introduction of a latent variable which

482

captures the possible dependence for the repeated data (longitudinal data). The reparametrized form of the beta distribution (4) (Ferrari & Cribari-Neto, 2004; Jorgensen, 1997) gives easy interpretations. Posterior summaries of interest can be easily obtained using MCMC (Markov Chain Monte Carlo) methods especially from the free available software OpenBugs which only requires the definition of the likelihood function and the prior distributions for the parameters of the model.

REFERENCES

Ahlstrom, L., Grimby-Ekman, A., Hagberg, M., & Dellve, L. (2010). The work ability index and single-item question: associations with sick leave, symptoms, and health – a prospective study of women on long-term sick leave. **Scandinavian Journal of Work, Environment & Health**, 36(5), 404-412. DOI: https://www.ncbi.nlm.nih.gov/pubmed/20372766

Box, G. E. P., & Tiao, G. C. (1973). **Bayesian inference in statistical analysis**, Addison-Wesley Publishing Co, Reading, Mass. London-Don Mills.

Cepeda-Cuervo, E., Achcar, J. A. & Lopera, L. G. (2013). Bivariate beta regression models: joint modeling of the mean, dispersion and association parameters. **Journal of Applied Statistics**,41, 1-11. DOI: https://ideas.repec.org/a/taf/japsta/v41y2014i3p677-687.html

Cepeda-Cuervo, E., & Gamerman, D. (2005). Bayesian methodology for modeling parameters in the two parameter exponential family, **Estadística**, 57, p. 93-105. DOI: https://mathscinet.ams.org/mathscinet-getitem?mr=2406467

Chib, S., Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm, **The American Statistician**, 49(4), 327-335.

Cordeiro, T. M. S. C., & Araujo, T. M. (2016). (in portuguese) Capacidade para o trabalho entre trabalhadores do Brasil. **Revista Brasileira de Medicina do Trabalho**, 14(3), 262-274.

Draper, N.R., & Smith, H. (1981). **Applied Regression Analysis**. 2nd Edition, John Wiley & Sons, New York.

Ferrari, S., & Cribari-Neto, F. (2004) Beta regression for modeling rates and proportions, **Journal of Applied Statistics**, 31(7), 799-815.

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities, **Journal of the American Statistical Association,** 85(410), 398-409. DOI: https://www.tandfonline.com/doi/abs/10.1080/01621459.1990.10476213

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). **Bayesian Data Analysis** (second ed.). Boca Raton, Florida: CRC Press. P. 4–5. ISBN 1-58488-388-X.

Ilmarinen, J. (2006). **Towards a longer worklife!** Finnish Institute of Occupational Heath, 467 p.

Ilmarinen, J. (2007). The Work Ability Index (WAI). **Occupational Medicine,** 57(2), 160.

Ilmarinen, J. (2009). Work ability: a comprehensive concept for occupational health research and prevention . **Scandinavian Journal of Work, Environment & Health**, 35(1),1-5.

Jorgensen, B. (1997). Proper dispersion models (with discussion), **Brazilian Journal of Probability and Statistics**,11,89-140.

Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). OpenBugs: a Bayesian modelling framework: concepts, structure, and extensibility. **Statistics and Computing,** 10(4), 325-337. DOI: https://link.springer.com/article/10.1023/A:1008929526011

Martelin, T., Sainio, P., Koskinen, S., & Goul, R. (2008) **Education**. In: Gold, R., Ilmarinen, J., Jarvisalo, J., & Koskinens, S. **Dimensions of work ability.** Finnish Centre for Pensions, 42-44.

Martinez, M. C., & Latorre, M. R. D. O. (2009) (in portuguese) Fatores associados à capacidade para o trabalho de trabalhadores do Setor Elétrico. **Cadernos de Saúde Pública**, 25(4), 761-772.

Martinez, M. C., Latorre, M. R. D. O., & Fischer, F. M. (2010) (in portuguese) Capacidade para o trabalho: revisão de literatura. **Ciência & Saúde Coletiva**,15(1), 1553-1561.

Montgomery, D. C., & Runger, G. C. (2010) **Applied statistics and probability for engineers**, 5nd Edition, Wiley & Sons.

Moreira, P. S. V. (2013) (in portuguese) **Aplicação do índice de capacidade para o trabalho na equipe de enfermagem*:* estudo descritivo**. 2013. 102 f. Dissertação (Mestrado Profissional em Enfermagem Assistencial). Universidade Federal Fluminense, Niterói.

Pohjonen, T. (2001) Perceived work ability of home care workers in relation to individual and work-related factors in different age groups. **Occupational Medicine,** 51(3), 209-217.

Renosto, A., Biz, P., Hennington, E. A., & Pattussi, M. P. (2009) (in portuguese) Confiabilidade teste-reteste do Índice de Capacidade para o Trabalho (ICT) em trabalhadores metalúrgicos do Sul do Brasil. **Revista Brasileira de Epidemiologia**, 12(2), 217-225.

Seitsamo, J., Tuomi, K., & Martikainen, R. (2007) Activity, functional capacity and well-being in ageing Finnish workers. **Occupational Medicine**, 57(2), 85-91.

Silva Junior, S. H. A., Vasconcelos, A. G. G., Griep, R. H., & Rotenburg, L. (2013) (in portuguese) Confiabilidade teste-reteste do Índice de Capacidade para o Trabalho (ICT) em trabalhadores de enfermagem. **Revista Brasileira de Epidemiologia**, 16(1), 202-209.

Tuomi, K., Huuhtanen, P., Nykyri, E., & Ilmarinen, J. (2001) Promotion of work ability, the quality of work and retirement. **Occupational Medicine**, 51(5), 318-324.

Tuomi, K., Ilmarinen, J. ; Jahkola, A., Katajarinne, L., & Tulkki, A. (2010) (in portuguese) **Índice de capacidade para o trabalho**. Translation of Frida Marina Fischer (Coord.). São Carlos: EdUFSCar, 59 p.

APPENDIX 1 (LEVELS OF THE COVARIATES USED IN THE STUDY)

- Gender: 1-Female ; 2 - Male

- Education degree: 1 - Elementary School; 2 - High School; 3 – Undergraduate; 4 - Graduate

- Age: in years

- Categorized Age: 1 - 31 to 40 years; 2 - 41 to 50 years; 3 - 51 to 60 years old; 4 - Over 60

- Marital status: 1 - Single; 2 - Married

- Number of children: 0 - None; 1 – one; 2 – two; 3 - Three or more

484

- Physical Activity: 1 - Every day; 2 - Every week; 3 - A few times a month; 4 - Rarely; 5 - Do not practice physical activity

- Nature of the task: 1 – Research; 2 – Administrative; 3 - Research Support; 4 - Field

- Position: 1 - Researcher; 2 - Analyst; 3 – Technical; 4 - Assistant

- Employment Bond: 1 - Effective; 2 - Retired by the National Security Institute

- Company Time: 1 - Up to 10 years; 2 - 11 to 20 years; 3 - 21 to 30 years; 4 - Over 30 years

485