Modelling the Relationship Between NBA Draft and the Career Longevity of Players Using Generalized Additive Models

Caio Garbo Miguel, Fabrφcio Joπo Milan, AndrΘ Luiz A. Soares, Ricardo Teixeira Quinaud, Luφsa Duarte Kós, Carlos Ewerton Palheta, Felipe Goedert Mendes & Humberto Moreira Carvalho[1]

Abstract

Draft is an annual event, which distributes players coming from college and overseas among a pool of eligible teams. Selecting a player in the draft is probably one of the most important decisions for NBA front offices. In this study we examined the relationships between NBA draft order of selection and career longevity between 1978 and 1998 using generalized additive models. Retrospective data of 1091 players selected in the NBA draft from 1978 to 1998 was considered. Data was extracted from the NBA official website (http://stats.nba.com/draft/history/) and from a specialized website in basketball (https://www.basketball-reference.com/). The variables considered for analysis were order of player's selection pick, years played in the league and the year of the NBA Draft. We assumed the possibility of non-linear patterns in the relations between Draft pick order, career longevity and years of draft selection, hence modeling these variables using a bivariate tensor spline. On average, players selected in the first five picks of the NBA had a longer career (~14 career years). Controlling for draft years, a non-linear trend of career longevity was observed with a decrease in the career longevity from the first pick until the 30th pick, leveling of thereafter. When accounting for draft pick, a non-linear trend was observed for career longevity per draft year. Career longevity increased between 1978 and 1985, leveling between 1985 and 1993 and increasing in the following years until 1998. Overall, the longevity of NBA careers has increased for players selected in the first picks of the NBA drafts of the most recent draft years.

Keywords: basketball, bayesian estimation, nonlinear analysis

Draft is an annual event, which distributes players coming from college and overseas among a pool of eligible teams. It is used in professional North American sports leagues, such as the National Basketball Association (NBA), National Football League (NFL), National Hockey League (NHL) and Major league baseball (MLB) (Koz, Fraser-Thomas and Baker, 2012). Selecting a player in the draft is probably one of the most important decisions for NBA front offices, not just for next season but for the longevity of the franchise (Staw and Hoang, 1995).

The draft order is based on the past season standings, which will determine the priority on the selection process (Berri, Brook and Fenn, 2011). Therefore, the weaker performance the team had in the last season standings, the higher it will be its chance to have a higher pick order, hence selecting a likely better player. In the NBA draft there is also the lottery, where the first three picks are drawn, and the weaker performance teams are more likely to get them, after that the order is based on the teams' standing, in the inverse order.

NBA is the most valuable and competitive professional basketball league in the world. To excel, teams´ perfor-

mance is dependent on several factors; being one of those having a deep, balanced and qualified roster. Hence, the draft is decisive to strengthen and rejuvenate a team´s roster. Besides the importance on the court, the draft has a great economic risk, given the rules restrictions on players' contracts (Arel and Tomas III, 2012). Nevertheless, the decision when selecting a player also needs to consider, for example, current team roster strengths and/or weakness, talent available on the draft, or the team strategies and game plan, among others (Abrams, Barnes and Clement, 2008; Wolff, 2001). Draft decisions in the NBA are complex and far from an exact science (Sailofsky, 2018).

Career longevity is an important factor to characterize a player's ability to thrive or not in the league (Abrams et al., 2008; Amico, 2001). Available data show a trend of association between career longevity with game statistics (traditional statistics) (Ballard, 2005). Nevertheless, available information is limited and there is a need to examine further the available game analysis data and possible non-linear trends, requiring statistical approaches beyond the traditional game statistics and data analysis approaches (mostly based on least squares estimations). In particular, we consider in the present study

1 Department of Physical Education, Sports Center, Federal University of Santa Caratina. Contact address: Humberto M. Carvalho. Departamento de Educação Física, Centro de Desportos, Universidade Federal de Santa Catarina - Campus Reitor João David Ferreira Lima - 88040-900, Florianópolis/Santa Catarina, Brasil. Tfno.: +55 48 3721-9926. E-mail: hmoreiracarvalho@gmail.com

the trends of association between career longevity and order of selection in the NBA draft between 1978 and 1998, using generalized additive models.

Methods

Data

For this retrospective study we considered data from 1091 players selected in the NBA draft, from 1978 to 1998. Data was retrieved from the official online website of the NBA (http://stats.nba.com/draft/history/), as well as from a specialized website in basketball (https://www.basketball-reference.com/). These sites contain official databases with information on sports statistics of professional players since the beginning of the NBA league (1948). Within the period considered in this study, changes in the number of teams and available draft picks for each team occurred, thus we right censored data at pick #50 for this study.

Variables

We retained for analysis each player's selection pick in their respective drafts year, and their years played in the league.

Statistical Analysis

Conditioned on the data, we examined initially the longevity of players per draft pick, accounting for year variation (at level-2) using Bayesian multilevel models (varying intercept model). We then predicted the players´ career longevity using the draft pick selection and draft year as explanatory variables. As the effect of the draft pick selection and draft year predictors is of unknown non-linear form, we used a bivariate tensor spline (Bürkner, 2018; Wood, Scheipl and Faraway, 2013).

Bayesian framework allows to derive probability statements for every quantity of interest or explicitly incorporate prior knowledge about parameters into the model (Bürkner, 2018; Gelman et al., 2013). Here we used default priors, which are non-informative (Bürkner, 2017, 2018).

The multilevel models and generalized additive models were implemented via Markov Chain Monte Carlo (MCMC) simulation using Hamiltonian Monte Carlo and its extension, the No-U-Turn Sampler. The MCMC simulations were implemented in using Stan probabilistic programming language (Stan Development Team, 2015), obtained using "brms" package (Bürkner, 2017) available as a package in the R statistical language (R Core Team, 2015). We ran four chains for 2000 interaction with a warm-up length of 1000 interaction to ensure convergence of the chains. We examined the convergence of the chains by visual inspection of the trace-plots.

Results

The figure 1 shows the mean between the picks selections order and the career longevity from each of those picks between the years 1978 to 1998 in the NBA draft.
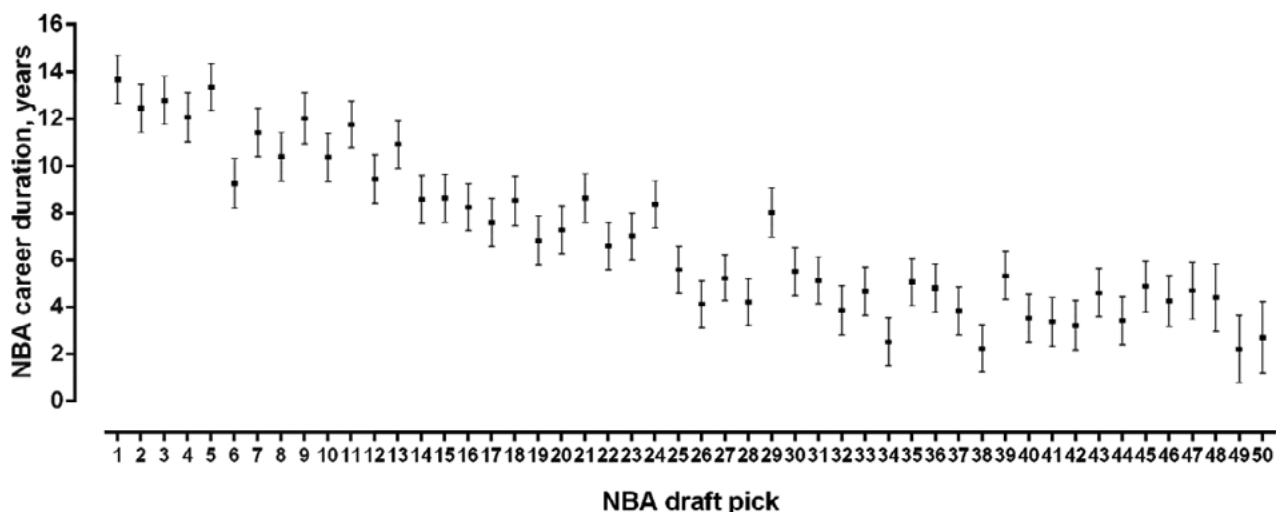

Figure 1. Mean and standard deviation of NBA career duration per draft pick

In Figure 2A, the marginal effect of each independent variable, i.e. draft picks and career years, is displayed with the other independent variable fixed at its mean. The credible intervals of the standard deviations of the coefficients forming the splines do not include zero indicating non-linearity in the combined effect of draft pick and draft year. The results showed that the athletes selected in the first five picks of the NBA had a longer career (13 to 14 career years). As pick selection increase there was a decrease in career longevity until about the 30th, leveling of thereafter, controlling. In Figure 2B, the combined effect of career years and draft year on players´ career longevity demonstrates the interaction between the variables. In particular, players´ career longevity was highest for higher picks in more recent draft years.
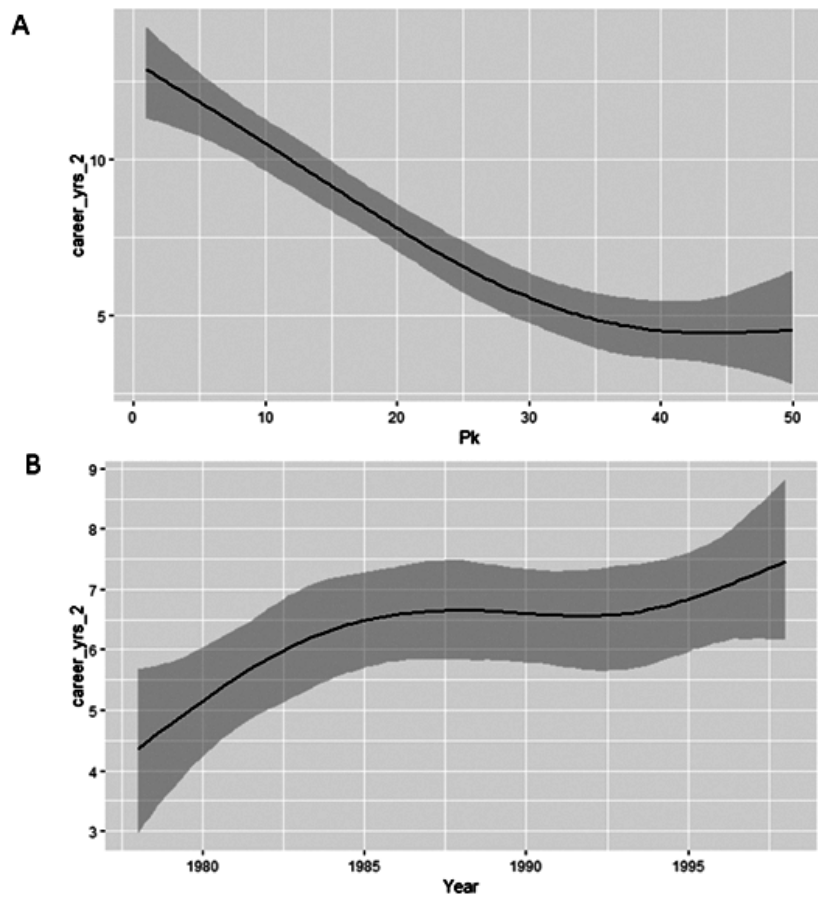
Figure 2 (A). Trend of NBA career duration by draft pick, controlling for variation between year (between 1978 and 1998)

Figure 2 (B). Trend of NBA career duration from 1978 to 1998, controlling draft pick
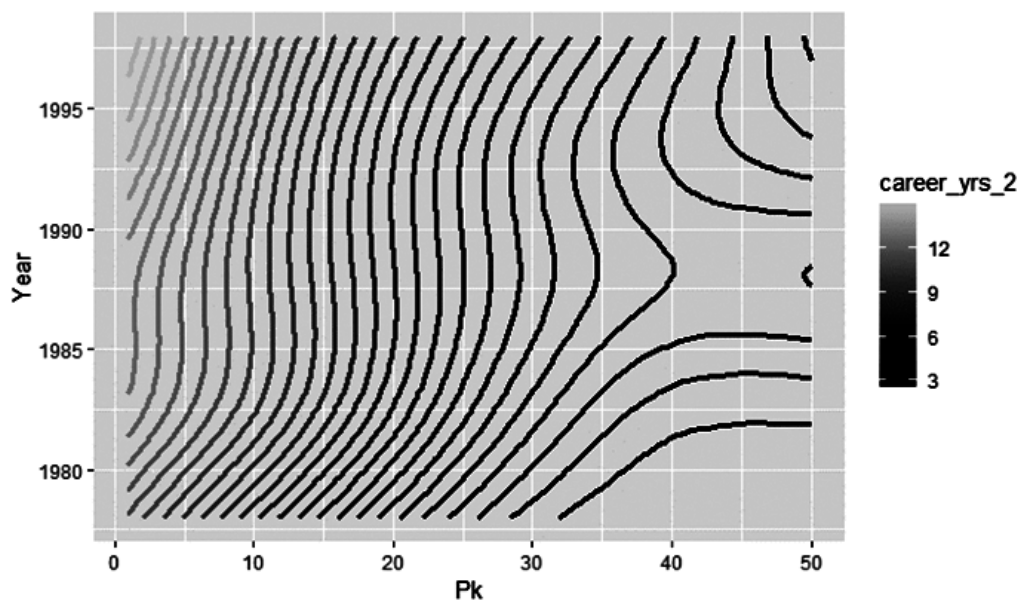

Figure 3. Surface plot of the generalized additive modeling of NBA career longevity by draft pick and draft year

Discussion

As career longevity is an important factor that characterizes the player's ability to thrive or not in the league (Amico 2001, Abrams, Barnes et al., 2008), we considered in the present study the associations between career longevity and selection order in the NBA draft between 1978 and 1998. The use of bayesian methods in sports and exercise research is limited (Leonardi et al., 2018; Mengersen, Drovandi, Robert, Pyne, & Gore, 2016). Bayesian methods allow a straightforward interpretation in terms of probabilistic statements (Mengersen, Drovandi, Robert, Pyne, & Gore, 2016), offering an alternative in the current debate of replication crisis, that has arrived also to sport and exercise research, in particular sport and exercise psychology (Schweizer & Furley, 2016). In the present analysis, we used multilevel modeling framework initially to explore the trends of career duration, acounting for possible variation by draft year. We further adopted generalized additive models allowing a smooth estimation of players´ career duration in funtion of the draft position and year of the draft. The results indicate that athletes from the first picks had longer career years compared to later picked players, regardless of draft year. On the other hand, NBA players' career longevity increased between 1978 to 1998 years, independent the order they were chosen in the draft. Lastly, players selected in lower draft position appear to have lower chance to have a long NBA career, particularly during the last decade of observation.

Between 1978 and 1998, players selected higher in the draft had, on average longer career than those selected at lower positions. These observations are similar to data examining draft in other professional sports (Sullivan, Kempton, Ward and Coutts, 2018).

Noteworthy, there was large uncertainty in players´ career longevity between picks (standard deviation at the first pick of 1.2 years, 95% credible interval 0.2 to 2.2 years, with small increase of 0.03 year per draft pick). This uncertainty estimates account for those players who were drafted in the first picks that had short NBA careers, as well as that despite being drafted in later positions had successful NBA careers. Considering the NBA draft, it has been suggested a positive relationship between athletes' longevity in some game specific positions (i.e. guard and forward). We did not consider players position, but it may be interesting in future analysis to consider whether the changes in the game across time influenced the pattern of selection by position and its implications on players´ career longevity. The present results demonstrate a clear link between early selection in the draft and career longevity, increasing between 1978 and 1998.

The present observations showed a trend of increase in NBA players´ career longevity between 1978 and 1998, particularly after about 1993. These observations suggest that the NBA franchises likely improved their selecting process in the draft by adding different athletes' information (Sailofsky, 2018; Sullivan et al., 2018), similar to other professional sports leagues selection process (e.g. *Australian Football League*) (Koz et al., 2012). The observed trends may also reflect the changes in the NBA structure (Amico 2001, Abrams, Barnes et al., 2008), such as travel or salary caps under the NBA collective bargaining agreement between players and owners), over the years of observation. On the other hand, player and team performance, and consequently the physiological demands of the game (Stojanovic et al., 2018), have developed as result of rule changes, and the increased emphasis on performance analytics to support both game-to-game decisions and roster composition (Sampaio et al., 2015).

As an exploratory study, we assume some limitations. We considered as a condition of success the longevity of athletes' career in years involved in the league. Also, we considered a period of 20 years, between 1978 and 1998. We selected 1998 since it is the first to have all players retired from active NBA career. Within the period of the study occurred four team expansions, from 22 to 27 teams in 1998. Hence, the number of available positions in the draft changed over the years. Also rules allowing high school players to enter the draft and rule of minimum age to enter the draft changed across the period of observation. Finally, we did not consider in this analysis the players who had successful NBA careers but were not selected or entered in the draft between 1978 and 1998. Hence, caution in warranted in the interpretation and generalization of the present observations. Further sources of variation may be considered in future research, such as individual players´ game performance indicators and contribution to team performance, position or level achieved in the career.

Achievement of the highest-level professional basketball is a highly selective process. Information about the trends of players´ career related to the NBA draft positions is of a practical relevance to inspect assumptions of the selection decisions, and decision making in the NBA, which conditioned on the data appears to be more effective as players selected in the first draft positions tend to have longer careers. Lastly, we ilustrate in this study a potential application of Bayesian methods to applied sport and exercise research.

Financial Support

Acknowledgments

Relación entre o Draft y el éxito de la carrera de los jugadores de la NBA

Resumen

Draft es un evento anual, que distribuye a los jugadores procedentes de la universidad y del extranjero entre un grupo de equipos elegibles. La selección de un jugador en el draft es probablemente una de las decisiones más importantes para los directoresde la NBA. El objetivo fue la asociación entre la longevidad de la carrera y el orden de selección en el draft de la NBA entre 1978 y 1998, adoptando un enfoque bayesiano multinivel. Para este estudio longitudinal retrospectivo, se recogieron datos de 1091 jugadores seleccionados en el draft de la NBA de 1978 a 1998, con edades comprendidas entre los 17.2 y los 27.3 años. A través del recurso oficial en línea de la liga profesional, el sitio (http://stats.nba.com/draft/history/) y el sitio web especializado en baloncesto (https://www.basketball-reference.com/). Las variables fueron el orden de selección del jugador, los años jugados en la liga y el año del Draft de la NBA. El promedio de años de los seleccionados en las primeras cinco selecciones de la NBA Draft tiene una carrera más larga (~ 14 años de carrera) y cuando las selecciones disminuyen hay una disminución en la longevidad de la carrera y desde la selección 25 los años de carrera se estabilizan (~ 4 años), este comportamiento se mantiene controlando el Año de Draft. Al controlar la selección del draft, podemos observar que hubo un crecimiento en el tiempo de la carrera a medida que pasaron los años, teniendo una estabilidad de los años 1990 a 1993. Por lo tanto, se vio que las primeras elecciones tienen más éxito en su carrera y los atletas tienen más longevidad en la carrera con el curso de los Draft anuales.

Palabras clave: prueba de rendimiento; NBA; baloncesto.

Relação entre o Draft da NBA e a longevidade da carreira dos atletas da NBA

Resumo

Draft é um evento anual, que distribui jogadores vindos da faculdade e do exterior entre um grupo de equipes elegíveis. Selecionar um jogador no draft é provavelmente uma das decisões mais importantes para os diretores das franquias da NBA. O objetivo do estudo foi a associação entre longevidade na carreira e ordem de seleção no draft da NBA entre 1978 e 1998, adotando uma abordagem Bayesiana em vários níveis. Para este estudo longitudinal retrospectivo, foram coletados dados de 1091 jogadores selecionados no draft da NBA de 1978 a 1998, com idade de 17.2 a 27.3 anos. Através do recurso online oficial da liga profissional do site (http://stats.nba.com/draft/history/) e do site especializado em basquete (https://www.basketball-reference.com/). As variáveis analisadas foram ordem de seleção de jogador, anos jogados na liga e ano do Draft. Os atletas selecionados nas primeiras cinco escolhas da NBA tem uma média de carreira mais longa (~ 14 anos de carreira) e conforme a ordem das escolhas regridem, há uma diminuição na longevidade da carreira, ocorrendo, a partir da escolha 25, há uma estabilização na média de anos de carreira (~ 4 anos), esse comportamento é observado quando controlado o Ano do draft. Ao controlar a ordem da escolha do draft, podemos observar que houve um aumento nos anos de carreira com o passar dos anos, tendo uma estabilidade dos anos de 1990 a 1993. Assim, foi visto que as primeiras escolhas têm mais sucesso em sua carreira e os atletas têm maior longevidade na carreira conforme o passar dos Drafts.

Palavras-chave: teste de performace; NBA; basquetebol.

References

Abrams, W., Barnes, J. C. and Clement, A. (2008). Relationship of selected pre-NBA career variables to NBA players' career longevity. *The sport journal, 11*(2).

Amico, S. (2001). *A basketball summer: The pro game as it's played in the off-season*: Protea Pub.

Arel, B., Michael, J. and Tomas III. (2012). The NBA draft: A put option analogy. *Journal of Sports Economics, 13*(3), 223-249.

Ballard, C. (2005). Measure of success. *Sports Illustrated, 103*(16), 78-82.

Berri, D. J., Brook, S. L. and Fenn, A. J. (2011). From college to the pros: Predicting the NBA amateur player draft. *Journal of Productivity Analysis, 35*(1), 25-35.

Bürkner, Paul-Christian (2017). brms: An R Package for Bayesian Multilevel Models using Stan. *Journal of Statistical Software, 80*, 1-28.

Bürkner, P. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal, 10*(1), 395-411.

Fynn, K. D. and Sonnenschein, M. (2012). An analysis of the career length of professional basketball players. *The Macalester Review, 2*(2), 3.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B. (2013). *Bayesian Data Analysis, Third Edition*. Boca Raton, FL: Chapman and Hall/CRC Press.

Koz, D, Fraser-Thomas, J and Baker, J. (2012). Accuracy of professional sports drafts in predicting career potential. *Scandinavian journal of medicine and science in sports, 22*(4), e64-e69.

Leonardi, T. J., Paes, R. R., Breder, L., Foster, C., Gonçalves, C. E. and Carvalho, H. M. (2018). Biological maturation, training experience, body size and functional capacity of adolescent female basketball players: A Bayesian analysis. International Journal of Sports Science & Coaching, 13(5), 713–722. doi: 10.1177/1747954118772489

Mengersen, K. L., Drovandi, C. C., Robert, C. P., Pyne, D. B. and Gore, C. J. (2016). Bayesian Estimation of Small Effects in Exercise and Sports Science. PLoS One, 11(4), e0147311. doi: 10.1371/journal.pone.0147311

R Core Team. (2015). *R: A Language and Environment for Statistical Computing* Retrieved from http://www.R-project.org/

Sailofsky, D. (2018). Drafting Errors and Decision Making Theory in the NBA Draft. Retrieved from http://www.sloansportsconference.com/content/drafting-errors-decision-making-theory-nba-draft/

Sampaio, J., McGarry, T., Calleja-Gonzalez, J., Jimenez Saiz, S., Schelling, I. D. A. X., and Balciunas, M. (2015). Exploring Game Performance in the National Basketball Association Using Player Tracking Data. *PLoS One*, 10(7), e0132894. doi: 10.1371/journal.pone.0132894

Schweizer, G. and Furley, P. (2016). Reproducible research in sport and exercise psychology: The role of sample sizes. Psychology of Sport and Exercise, 23, 114-122. doi: https://doi.org/10.1016/j.psychsport.2015.11.005

Stan Development Team. (2015). *Stan: A C++ Library for Probability and Sampling*. Retrieved from http://mc-stan.org/

Staw, B. M. and Hoang, H. (1995). Sunk costs in the NBA: Why draft order affects playing time and survival in professional basketball. *Administrative Science Quarterly*, 474-494.

Stojanovic, E., Stojiljkovic, N., Scanlan, A. T., Dalbo, V. J., Berkelmans, D. M. and Milanovic, Z. (2018). The Activity Demands and Physiological Responses Encountered During Basketball Match-Play: A Systematic Review. *Sports Medicine*, 48(1), 111-135. doi: 10.1007/s40279-017-0794-z

Sullivan, C., Kempton, T., Ward, P. and Coutts, A. J. (2018). Factors associated with early career progression in professional Australian Football players. *Journal of sports sciences*, 36(19), 2196-2201. doi: 10.1080/02640414.2018.1444277

Wolff, A. (2001). Your Lyin'Eyes. *Sports Illustrated, 94*(26), 82-93.

Wood, S.N., Scheipl, F. and Faraway, J.J. (2013). Straightforward intermediate rank tensor product smoothing in mixed models. *Statistics and Computing, 23*(3), 341–360. doi: https://doi.org/10.1007/s11222-012-9314-z