Can Gender Role Items Improve the Prediction of Income? Insight from Machine Learning

Klara Raiber

University of Mannheim
Graduate School of Social Science

klara.raiber@gmx.de

ESRA Conference
July 18, 2019

Overview

General Issue

Interest / General question:

- There is a gender pay gap; Women earn less than men (Blau & Kahn, 2017; Weichselbaumer & Winter-Ebner, 2005)
- Why is that the case? What influences income?
- Different underlying mechanisms are not sufficiently disentangled so far

General Issue

- One line of argumentation is that the attitudes and behavior of a person related to their sex can explain differences (Mueller & Plug, 2006; Drydakis et al., 2018)
- We only measure sex in surveys
- Sex is then used as a proxy for gender

General Issue

Sex and gender are not the same!

- Sex is a categorization upon socially agreed on biological attributes (West & Zimmermann, 1987)
- Gender is the situational social behavior related to sex
- An understanding of the gender pay gap might not be possible without a better measurement for gender

Research Question

- Gender role items (GRI) might capture the concept of gender related to wage
- GRI measure attitudes towards gender ideologies and labor market participation vs. family (Braun, 2013)
- Some scholars find a relationship between GRI and income (Stickney & Konrad, 2007; Grove, 2011)
- How big is the impact of GRI on income compared to sex?
- Can GRI improve the prediction of income?
- Prediction can give insights on how important a variable is

Research Strategy

Comparing the results of four different models:

1. Empty model with only controls
2. Model including sex and control variables
3. Model including gender roles items and control variables
4. Full model with sex, gender roles items and control variables

Applying random forest to evaluate prediction performance and variable importance

Data

- International Social Survey Program (ISSP) from 2012
- Topic in 2012 was "Family and Changing Gender Roles"
- 40 countries (Turkey excluded) - 44 country variables
- 40.051 observations

Variables

- Income: 1. annualized, 2. log transform, and 3. standardized
- Controls: Age, country, working hours, legal partnership, health, location place of living, education, employment status, employment type, supervisor, organization type, and union membership

Variables

GRI (5-point scale + can't choose) (GESIS, 2012):

1. All in all, family life suffers when the woman has a full-time job.
2. Both the man and woman should contribute to the household income.
3. A man's job is to earn money; a woman's job is to look after the home and family.
4. Still thinking about the same couple, if both are in a similar work situation and are eligible for paid leave, how should this paid leave period be divided between the mother and the father?

mother entire, father not any / mother most, father some / mother and father half / father most, mother some / father entire, mother not any / can't choose

Random Forest

- 80% of the data as training data
- 500 regression trees per model
- Grid: mtry $= 4$
- Caret package in R (Kuhn, 2017)
- R^2, prediction performance, and variable importance to compare models

Results

	Empty Model	Sex Model	GRI Model	Full Model
R^2	0.468	**0.474**	0.444	0.445
Prediction	1.62	**1.61**	1.69	1.68

Table: Evaluation of the predictions

Results

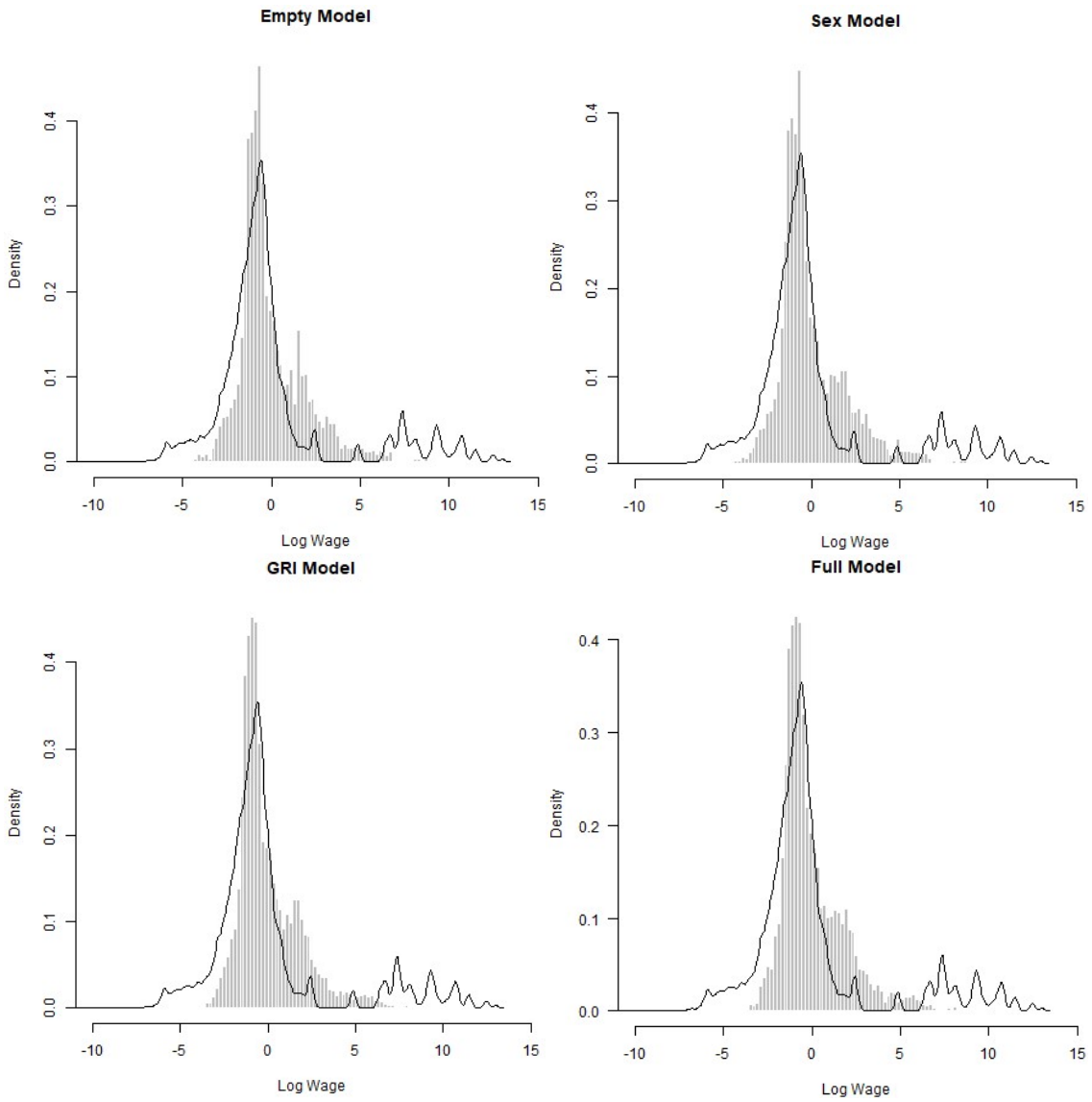

Figure: Prediction of income - in gray - compared to actual values - the black line

Results

Variable importance

Empty Model:

- Age
- Country
- Unemployed

Sex Model:

- Sex 10th most important variable (value 52)

GRI Model:

- No GRI in the first 20 variables
- Small importance (values from 0 to 35)
- Most important cut point: A mans job is to earn money; a womans job is to look after the home and family - strongly disagree

Results

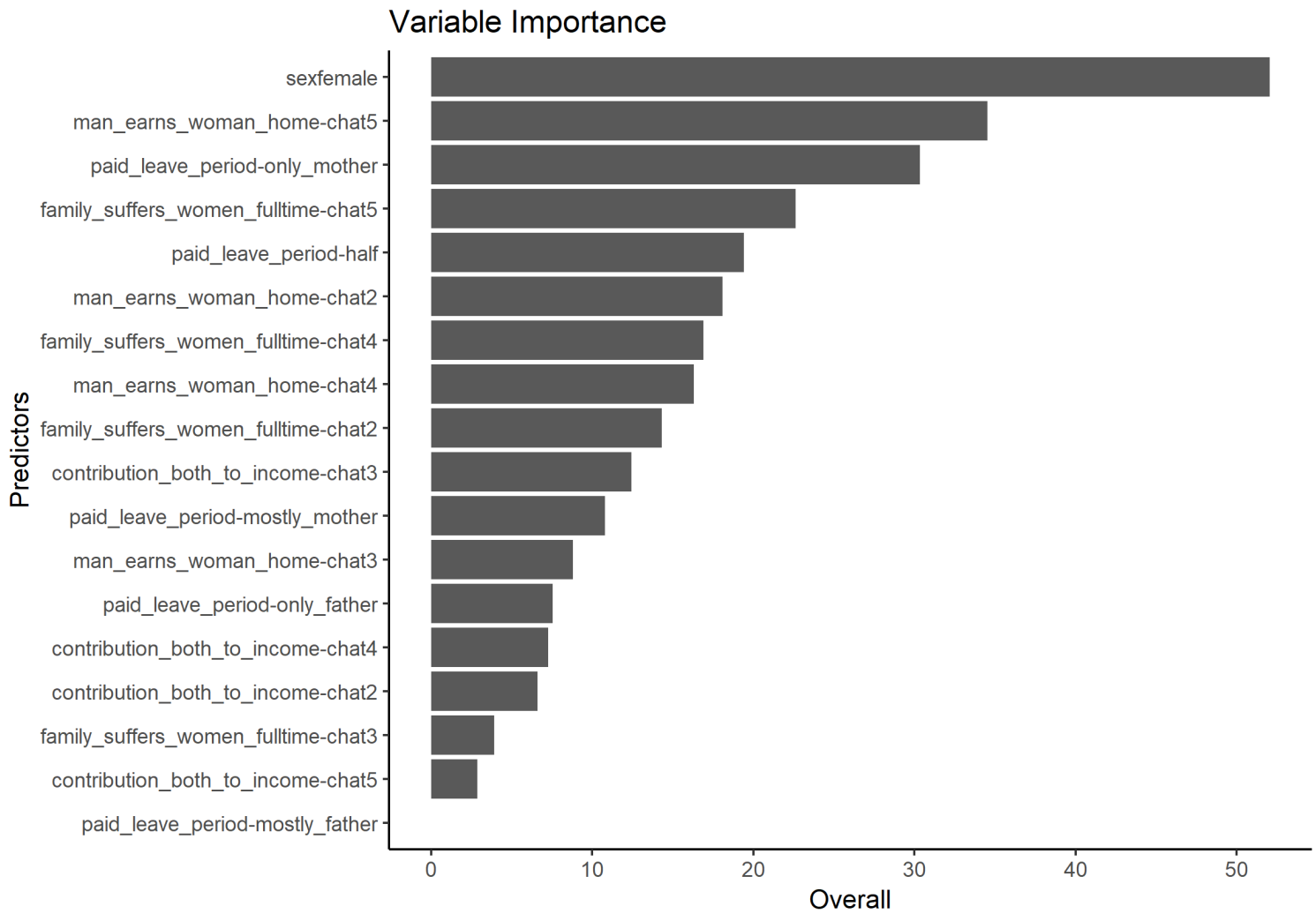

Figure: Variable importance of sex and GRI in the Full Model

Conclusion

- GRI cannot improve the prediction of income
- Sex is an important predictor for the income

What does that mean substantially?

- GRI might be correlated with income, but they do not have a big impact on income
- Attitude towards gender roles and family vs. labor market participation might not (fully) capture the concept of gender
- A better operationalization for gender is needed
- Sex is used by research to categorize, but maybe also by others like employers and firms

Thank you!

References

- Blau, F. D., & Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3), 789865.
- Braun, M. (2013). Gender roles. *In Comparative politics: The problem of equivalence* (pp. 111134). Colchester: ECPR Press.
- Drydakis, N., Sidiropoulou, K., Bozani, V., Selmanovic, S., & Patnaik, S. (2018). Masculine vs feminine personality traits and womens employment outcomes in Britain: A field experiment. *International Journal of Manpower*, 39(4), 621630.
- GESIS. (2012). International social survey programme: Family and changing gender roles ivissp 2012. gesis data archive, cologne. za3880 data file version 4.0.
- Grove, W. A., Hussey, A., & Jetter, M. (2011). The gender pay gap beyond human capital heterogeneity in noncognitive skills and in labor market tastes. *Journal of Human Resources*, 46(4), 827874.

References

- Kuhn, M. (2017). *Caret: Classification and regression training*. R package version 6.0-78. Retrieved from https://CRAN.R-project.org/package=caret
- Mueller, G., & Plug, E. (2006). Estimating the effect of personality on male and female earnings. *ILR Review*, 60(1), 322.
- Stickney, L. T., & Konrad, A. M. (2007). Gender-Role Attitudes and Earnings: A Multinational Study of Married Women and Men. *Sex Roles*, 57(11-12), 801811.
- Weichselbaumer, D., & Winter-Ebmer, R. (2005). A meta-analysis of the international gender wage gap. *Journal of Economic Surveys*, 19(3), 479511.
- West, C., & Zimmerman, D. H. (1987). Doing gender. *Gender & society*, 1(2), 125151.