Can Sex-role Attitudes be Treated? Experiment in Schools

Kaire Põder, Niveditha Pankova[1]

[1]Estonian Business School*

Abstract

We study whether sex-role attitudes can be changed via experimental research design. Our data originate from a randomized control trial of 297 students aged 16 and up. After using randomization in the selection of (1) schools and (2) control and treatment groups within the school, we treated students of the treatment group with a short film lasting 20 minutes on family life. Our results show that the treatment does not affect traditional "male breadwinner," but does impact the respondents' egalitarian view of a "universal caregiver." In addition, we showed that certain family and background characteristics such as minority language, male, low education of parents correlate with more traditional male-breadwinner attitudes.

Keywords: work-life balance, attitudinal change, randomized control trial, exploratory factor analysis, universal caregiver, male breadwinner

1. Introduction

In families with children, the gender differences play out not only in career choices and wages earned in the labor market, but also at home in the form of unpaid care time. It could be argued that there are many institutional particularities, both formal and informal, behind it. Unequal distribution of responsibilities concerning childcare is one of the main nuisances within the complex set of problems that hinder work-life balance, i.e. balancing the professional and private lives, of families with under-aged children. Unequal division of care between women and men lies on a spectrum, with the extreme conventional "male breadwinner model," where the childcare responsibilities are allocated solely to the mother; this set-up is justified by the belief that mothers are better caregivers, and children will suffer under other care. Often these choices within families are considered rational or "economic". However, "rational" choices of parents are embedded in informal institutions such as traditional sex roles. However, a broad societal change and the feminist movement have challenged traditional—female homemaker and male breadwinner—sex roles; still, in some countries both evolving social changes and formal institutional support of the changes are more modest and slow than in others.

In this article, **we ask whether sex-role attitudes can be changed in young adults.** Our theoretical framework is based on the capabilities approach (Kremer 2006, Kurowska 2016, Lauri et al 2019) which **conceptualizes the work-life balance as the result of the capability to make choices between work and life**. In this approach sex-role attitudes interact with the legal and

economic institutional environment and mediate capabilities, such as both parents' capability to work and the capability to share childcare. Simply put, sex-role attitudes give incentives for families to "interact" with institutions, such as leave policies, childcare arrangements, legal rights for rest periods and holidays to name a few, and this interaction produces outcomes such as work-life balance.

We use an experiment, and our research design is a randomized control trial (RCT) based on measuring treatment effects on treated. Our research design has many elements: data collection, treatment, measurement of sex-role attitudes, and impact analysis. The treatment is a short film designed to address the work-life balance issues faced by full-time working parents. We measure the impact of a 20-minutes film on the sex-role attitudes expressed in two measures – **traditional (male breadwinner) and modern (universal caregiver) sex-role attitudes**. For the measurement of attitudes, we use exploratory factor analysis. In general, we apply the standard premises of RCT. The general idea of the impact analysis applying RCT is to randomly allocate subjects to two groups, treating them differently, and then comparing them to the measured response. The treatment group has the intervention being assessed, while the other—the control group—has no intervention. For analysis of treatment effect, we use multivariate regression including various controls of individual and family background.

We collect data in 10 schools in Estonia, and our sample consists of 297 students aged 16 and up. We rely on a survey instrument, which originates from Walter (2018) and is constructed for measuring attitudes about sex roles; we also include personal and family characteristics as control variables. As a particularity of the Estonian school system, it also allocates children to the Russian (minority) language and Estonian (majority) language schools, so we allow surveys to be run in both languages to control for cultural (language) effect.

We see the contribution of the current exercise in providing insights of measurement of sex-role attitudes and in the study of change of these. As stated, we rely on theoretical premises that individually shared attitudes about sex roles also have these so-called macro-consequences on social and economic institutionalized patterns, which constitute female employment, pay, career opportunities, and planning. Moreover, we test whether low-effort processes of persuasive communication can change sex-role attitudes. After valid measurement of sex-role attitudes, we find that the treatment moderately changed sex-role attitudes towards the "universal caregiver" end of the spectrum in a subset of the respondents; however, we were unable to break the attitude supporting the traditional "male breadwinner" family arrangement with our video. Thus, we consider our experiment a partial success in supporting the egalitarian modernization of family life, and it leaves the door open for further explorations on how the traditional mindset can be changed.

We continue as follows. First, we explain our theoretical approach. Second, we discuss our research design. Then in Section 4, we give descriptive statistics and the results of the impact analysis. Finally, we discuss our findings and give conclusions and limitations of the study.

2. [1]Theoretical Framework

2.1 The Capabilities Approach

Already Esping-Andersen (2014) argued that the changed balance in both paid and unpaid work is the key in explaining the changes of societies. Thus, there is an increasing understanding that women cannot achieve equality with men even if they work, as long as men do not share in the caring tasks. It is argued (Čermáková 1997), that the "double burden" of working and caring prevents most women from concentrating on their careers – women work, but only men have careers. So Esping-Andersen (2014) has argued that female employment per se is not decisive. His argument is based on empirical results which indicate that men's contribution to domestic tasks suddenly takes off and this take-off in terms of equal sharing occurs when the modal woman shifts from part-timer to full-timer status. So easy solution to transfer from male breadwinner to one-and-half breadwinners is the solution that finds most criticism because in such cases often part-time working is considered a forced choice by mothers, empirically contributing to the gendered wage and career gap. So, it has been argued for some time that the traditional mail-breadwinner type of family is about to change to cope with dual-earner families' (Korpi 2000; Sainsbury 1999).

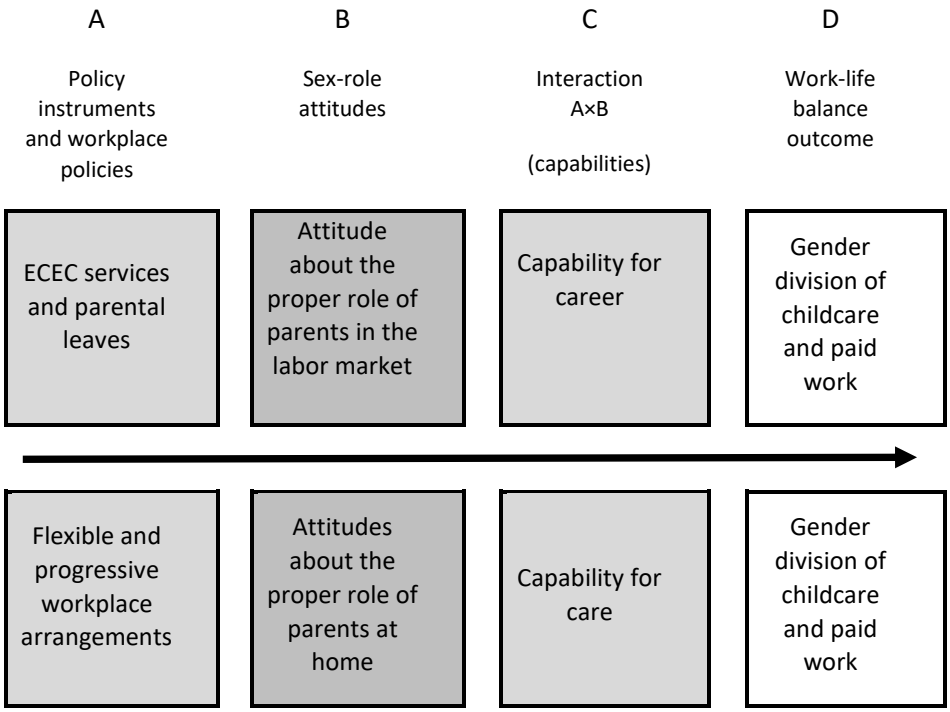

Figure 1: The Capabilities Approach (modified from Lauri et al. 2019)

However, this change has been slow. One explanation of the "unfinished gender revolution" (Gerson 2010) is that even individuals holding gender-egalitarian ideas, their choices (of work,

*kaire.poder@ebs.ee; niveditha.prabakaran@ebs.ee.

career, and children), either being men or women, are often not that egalitarian in their behavioral choices. Thus by empirical studies of choices, the effects of constraints (institutions and policies) are difficult to subtract from the effect of preferences or attitudes. Indeed there are also arguments that gendered aspects of individual identities (internalized believes and attitudes about caregiving and earning) operate alongside gendered institutions to maintain patterns of inequality (Feree, Lober and Hess 1999). So the question of whether preferences (or attitudes) are endogenous to formal institutions or formal institutions are endogenous to gendered attitudes, the reverse causality issue, that often remains both a theoretical and empirical puzzle. Thus, we limit ourselves to studying attitudes, not behavior.

Our theoretical underpinnings rely on premises of the capabilities approach (Kremer 2006, Kurowska 2016, Lauri et al. 2019). In general, we see that a certain mix of legal and economic constraints with "right" attitudinal factors can bring along good outcomes in terms of equal capabilities of fathers and mothers to proceed with childbearing and job market careers. It is also indicated by Lauri et al. (2019) that certain policy mixes measuring legal and economic constraints produce different outputs, e.g. moderately generous leaves and father quotas work well in Finland, while Austria and Latvia follow the same policy mix with no success. Our theoretical argument is that the capabilities of families differ not only in different institutional context but even in the same institutional context (see Figure 1).

Figure 1 illustrates the capabilities approach and explains capabilities as the interaction of sex-role attitudes and policies, such as affordability, quality, and availability of public early childhood education and care (ECEC); various leave policies (including targeted quotas or leaves for fathers); progressive workplace policies. It is theorized by this approach that capabilities, neither attitudes nor institutions or policies alone, have outcomes for the labor market which can be operationalized by the gender wage gap, female employment, gendered care time, or similar. Moreover, we don't rely on "defamilisation" premise as conceptualized by McLaughlin and Glendinning (1994), meaning that care is took away from the responsibilities of family, instead caretime is considered valuable both for the society and for the individual.

We operationalise the attitude about the proper role of parents in the labor market and at home by agreement of statements (see Appendix 1 as our measurement model and Appendix 2 for survey instrument) about how traditionally or universally are arranged gendered roles – whether males have comparative advantage in the labour market and females in care. Thus, we measure how traditional are the attitudes regarding roles in labour market, this indicates male breadwinner attitude; and how modern are the roles by sharing caregiving at home, this indicates universal caregiver attitude. These family constructs are largely debated in literature and referred to as Fraser's (1994) ideal types. Fraser's ideal types debate that universal caregiver, or so called Nordic utopia, triggers gender equality by men carework instead of female equal division of labour. Thus, symmetry of carework is considered a main source of equality by feminist literature, and universal caregiver type is considered as ideal for sex roles in modern societies (e.g. Ciccia & Verloo, 2012; Norchost & Siim, 2008).

In our research design, we target sex role attitudes in particular, and our attitudinal factors – attitude about the proper role of parents in the labor market (attitude of traditional male breadwinner) and at home by providing care (attitude of equal gender roles not as female as homemaker) – are at the core of our study.

2.2 Channels for Attitudinal Change

Informal institutions are considered reluctant to change and gender equality promoting formal institutional changes (e.g. father quotas or public childcare arrangements) are not bringing outcomes, that often are hoped for (Wayen 2014). This is common knowledge that attitudinal regularities persist over time. Different trails of literature, however, discuss different channels of the change of attitudes.

First, attitudes are considered endogenous to social change. From this perspective, Thijs et al. (2017) show from longitudinal studies that more egalitarian attitudes emerge over time. This also implies that attitudes can change throughout the life course (Brooks & Bolzendahl, 2004). This change is often explained by modernization theory (Inglehart & Baker, 2000) which assumes that industrialization produces pervasive social and cultural changes which also lead to changes in gender roles. So social changes are assumed to affect gender role attitudes (Brajdić-Vuković et al., 2007). Hence, these approaches would assume that as institutional contexts change, gender role attitudes should also change. However, we are interested in "treating" the attitudes, thus instead of slow social change, we seek the channels for "manipulating" attitudes.

So second, it can be assumed that attitudes can be reconstructed. The debate of whether and how attitudes can be changed is at the core of social psychology. In this literature, the arguments to support or contradict the idea of attitudinal change can be broadly divided into two extremes – theories that rely on the idea that attitudes are stored in memory or theories that argue that attitudes are constructed on the spot (Bohner & Dickel 2011). And there is some middle ground, arguing that current evaluations are constructed from relatively stable representations (Cunningham et al. 2007). Still, the battle goes between constructivist view and stable-entity view, where a radical construction view (Conrey & Smith 2007) emphasizes that attitudes are "time-dependent states of the system" rather than "static 'things' that are 'stored' in memory". Empirical research has shown that stability of attitudes comes with the repetition in the stable environment (Bohner & Dickel 2011), and thus it can be argued that young people (with less repetition in life) are more open to attitudinal change, while older age groups are prone to it.

The mechanism of change can be a persuasive communication (see discussion in Petty et al. 2006). Petty et al. (2003) give an overview of mechanisms where persuasion can bring along attitudinal change. In general, they (ibid.) distinguish between low- and high-effort processes. These include some largely automatic associative processes as well as simple inferential processes. From this debate we make a theoretical assumption that sex-role attitudes can be changed by following low effort processes:

- **Heuristics.** Heuristics are simple decision rules based on prior experiences or observations. We aim to show in our treatment that males can perform well in traditional women roles, for example as caregivers at home or as care-workers in general.
- **Attributional processes**. Attribution concerns the inferences that people make about themselves and others after witnessing a behavior and the situation in which it occurred. We aim to show that witnessing males and females relaxed and happy in transformed and more equal gender roles can adjust the self-perception and produce attitudinal change, making the attitudes toward equal sex roles more positive.
- **Cognitive balance**. Balance is achieved when we agree with those we like and disagree with those we dislike. Simply we aim to present the main characters of film pleasant and

kind, so it makes easy to agree with their life choices. Also, film characters are played by popular actors, which presumably makes it easier to like them.

- **Classical conditioning**. Classical conditioning is the positive and repetitive stimulus related to an attitude object. In our case we try to reflect the main male character in a positive light regarding his behavior at home or work, he is always calm and oriented to solutions of equal care, also the child is relaxed and able to perceive transformed gender roles as new normality.

We utilize these channels of attitudinal change by the design of the treatment, See section 3.2. According to the cognitive dissonance theory (Festinger 1957), people are motivated to hold consistent attitudes, otherwise, they can feel unpleasant physiological arousal. To resolve this dissonance, they are ready to change their attitude. For example, inconsistency between the belief of gendered ableism (e.g. mother is better caregiver) and modern labor market institutional pressures (income, including retirement earnings, are personalized and based on merit and effort) force change the attitudes of gendered ableism.

3. Methodology

3.1 Sampling, Survey Design, and Data Collection

To execute RCT the following procedure was followed. Sampling unit is a school, in total there are 58 gymnasiums, in each year approximately 7000 students graduate secondary education. So with a confidence level of 95% and an error margin of 5% and with a population size of 7000, we are targeting 365 students as our optimal sample size. Taking account of the average size of the school we were using a cluster of 10 schools as our main sampling units. The list of 10 schools was randomly drawn from the population of schools. Some schools refused to participate in the experiment (3 schools), these were substituted by a new selection of random schools.

The data collection was done using a survey instrument from Walter (2018). There were 10 questions related to sex-role attitudes (see Appendix 1). Walters (2018) argues that social developments create the need to phrase items differently from many explicit measures used in research of sex-role attitudes, e.g. Old-Fashioned Sexism (OFS) and Modern Sexism (MS) scales (Swim et al., 1995; Swim and Cohen, 1997), since, for example, the employment of mothers and fathers meant something different back then. The modern scales should include the items to measure the attitudes toward the division of labor within the family and its deviation from the male breadwinner model, as well as how these deviations have been perceived with regards to child welfare. So, items can reflect two scales – one for measuring attitudes of sex roles regarding labor market (male breadwinner) and the other regarding child welfare in respect of equality of childcare (see Appendix 2). All 10 items in the survey are measured in Likert 6-point scales.

RCT design of the experiment was theoretically "Post-test-only control group design" ($R\frac{C_O}{X_O}$, where O is observation, X is treatment, and C = control, while R indicates randomization, i.e. random assignment of subjects into treatment and control groups). All data were collected in May 2021 during the Covid-19 school closedown. So all sessions of the experiment were held in the Zoom, where after a short introduction of the experiment the class was divided randomly into two breakup rooms. The control group was sent to fulfill the online survey (by 20 minutes),

simultaneously treatment group was assigned to watch the film (see for the description of the film in the next section). After this treatment group is asked to fulfill the survey, while the control group was able to see the film after the survey.

3.2 Treatment Design

Our treatment is a 20-minute film, produced by Kuukulgur Film company. It consists of short sketches of family life, where the full-time working mother (IT specialist) and full-time working father (specialist in a logistic company) are balancing their work and life as parents of two children (a school-age child and a toddler), while also caring for a grandfather. It includes 8 sketches, starting from a stressful, late-evening scenario (see Figure 2), and guides the audience through multiple scenes of the father's and mother's workplace conflict and school issues of the child, seeking the solutions to each, such as more flexible work schedules, rest periods, holidays, and bigger involvement of the father in caregiving. Also, it presents the evolution of the workplace culture to support the mother's career, and also workplace acceptance of the father as an equal caregiver of the children.

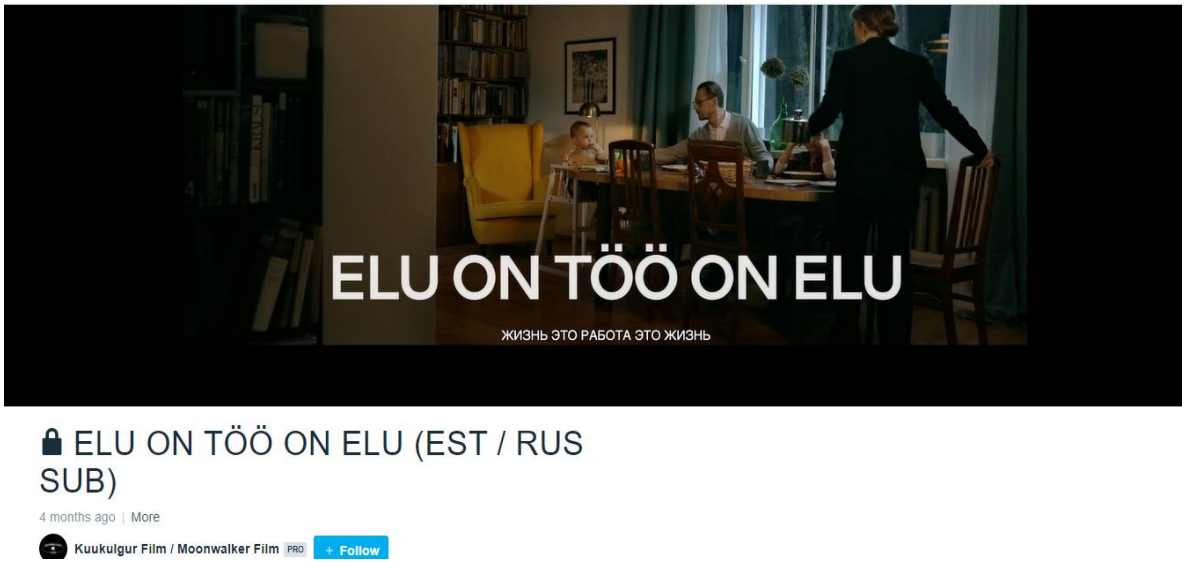

Figure 2: Starting scene of the film "Life is work is life"

The film takes into account the channels of attitudinal change, as our core design principles of the treatment, discussed in Section 2.2. In the design of the treatment, we see attitudinal changes as an inferential process (first three items in the list in Section 2.2) rather than an associative process (last item from the list). To strengthen the effect, the treatment also includes guiding questions within video clips to push the audience towards a higher cognitive effort; for example, we ask "Is it the best way to organize family and work-life for both partners, so that they both work full-time and look after the home and children equally?" or "Do you think that the child benefits if his or her mother has a job rather than just concentrating on the home?". We assume that this cognitive effort will trigger the following responses: (a) change the ratio of positive to negative thoughts about equal sex roles; (b) more thinking will help attitudinal change; (c) increase confidence that discrete thoughts are associated with the systematic thinking about sex-roles. The latter is also related to the dissonance theory introduced in Section 2.2.

3.3 Analysis Methods

For constructing a scale for measuring sex-role attitudes, we rely on the theoretical model laid out in Appendix 1. As Appendix 1 indicates, there are two scales for measuring sex-roles attitudes: traditional attitudes and modern (egalitarian) attitudes; we denote these as the latent variables: *male breadwinner* and *universal caregiver*. For constructing a measurement scale for both, we first use Cronbach's alpha, which indicates the reliability of the aggregate scales. If the correlations are above the conventional threshold of Cronbach's alpha of 0.7, we can calculate the composite scores of the ten items. However, Cronbach's alpha is not a test of one-dimensionality, and further, has the underlying assumption of tau-equivalence, i.e., the same true score for all test items, which implies equal factor loadings of all items in the factorial model. This is a requirement for Cronbach's alpha to be equivalent to the reliability coefficient (Cronbach, 1951). If the assumption of tau-equivalence is violated, the true reliability value will be underestimated (Raykov, 1997; Graham, 2006). Moreover, aggregate measures, such as the sum of the items, do not account for measurement errors. In order to account for the possible violation of tau-equivalence, we employ exploratory factor analysis (EFA) using the R-package Lavaan ((Rosseel, 2012) and estimate our measurement model M0, while allowing for the violation of tau-equivalence. Finally, we specify a revised model using the modification indices, an estimation option in the R-package Lavaan (Rosseel, 2012). Modification indices help to answer 'what if?' questions, such as whether freeing parameter constraints or adding paths to the model will improve the model. The final model M1 is based on the model-fit characteristics, which we explain in sections below.

Following measurement, we construct scales for sex-role attitude measures and use multivariate regression analysis, clustering the standard errors at the school-level. Standard errors need to be clustered at the school-level due to unobserved components in outcomes for students within schools which can be correlated. Even though our design is RCT, our randomization in the first stage is a subset of schools, and then in the second stage, students within schools are randomly divided into treatment and control groups, so clustering of standard errors is justified, particularly when we include school-level variables.

The idea of the specification of the model is to isolate the treatment effects. We use a pooled dataset, and our model specification is as follows:

$$y_i^j = \beta_0 + \beta_1 T_i + \beta_k \Sigma_k c_{k_i} + \epsilon_{is} \qquad (1)$$

where y_i^j is the measure of attitudes of individual i, where j indicates either a **traditional (male breadwinner) or modern (universal caregiver) attitude measure,** and s indicates individual i's school. We are mainly after the treatment effect β_1, where T_i equals 1 if the individual i is assigned to the treatment group and 0 for control. Other background covariates (c_k) are added in the regression using the step-wise modeling technique.

4. Results

In total, our sample consists of 297 students from 10 schools, and per our RCT research design, these students were split randomly into two groups: treatment group (n=147) and control group (n=150). In our descriptive analysis we ask three main questions: (a) whether there are differences in the background characteristics of these two groups; (b) whether we can compose

scales using 10 survey items; and finally, (c) whether there are distinct patterns in individual survey items when comparing treatment and control groups.

4.1 Descriptive Analysis: Student Demographics

Table 1 displays the following demographics of our study. In general, the average age of students was 17 years, ranging from 15 to 19 years. 47% of the students are males. Most of the students' language at home is the same as the language in which they took the survey (observable variable *LanguageHome*). The variable *Language* indicates whether the student comes from an Estonian or Russian language school (1=Estonian, 2=Russian). In general, 44% of the observations are from Russian-language schools. The average number of siblings (variable *Siblings*) is slightly below 1, indicating that 2-child families are the norm. In addition, we have five types of families, encoded by the unordered variable *Family*; thus, the means and standard deviations don't give us much insight. Instead, Figure 3 reveals the pattern for this variable, indicating that most of the students either come from traditional or single-parent family. Figure 3 demonstrates that there are no significant differences in family types by school languages, and traditional and single-parent family types are prevalent.

Table 1: Demographic characteristics of students

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Age	297	17.077	0.868	15	16	18	19
Gender	297	0.465	0.500	0	0	1	1
LanguageHome	297	0.899	0.302	0	1	1	1
Family	297	1.673	1.039	1	1	2	5
Siblings	297	0.919	1.070	0	0	1	8
MotherEducation	297	4.845	1.626	1	3	6	8
Language	297	1.444	0.498	1	1	2	2

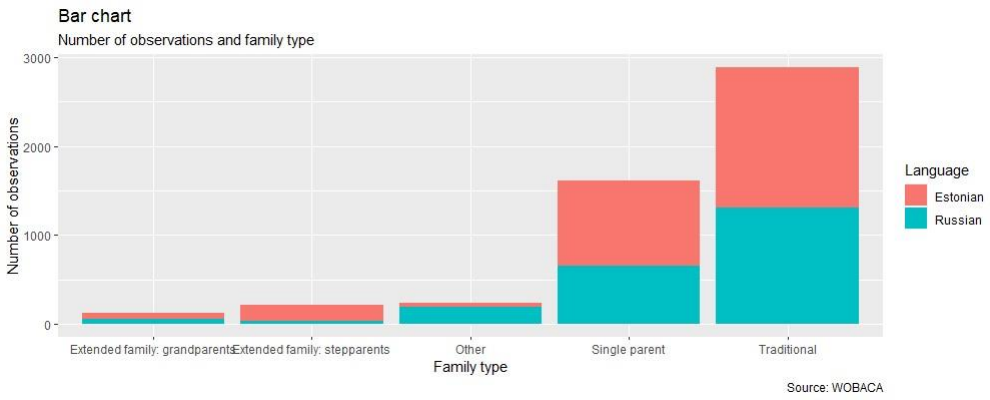

Figure 3: Family types and the language of the students

Furthermore, in our dataset, we have some background characteristics that are nominal (meaning that there is no intuitive order), and some that are measured in an ordered, categorical

scale. For example, *MotherEducation* is categorical and ordered in the following way indicating the education of the students' mothers:

1 Primary education
2 Lower secondary education
3 Secondary education
4 Vocational (secondary) education
5 Unfinished bachelor education (2 or more years)
6 Bachelor's degree
7 Master's degree
8 Doctoral degree

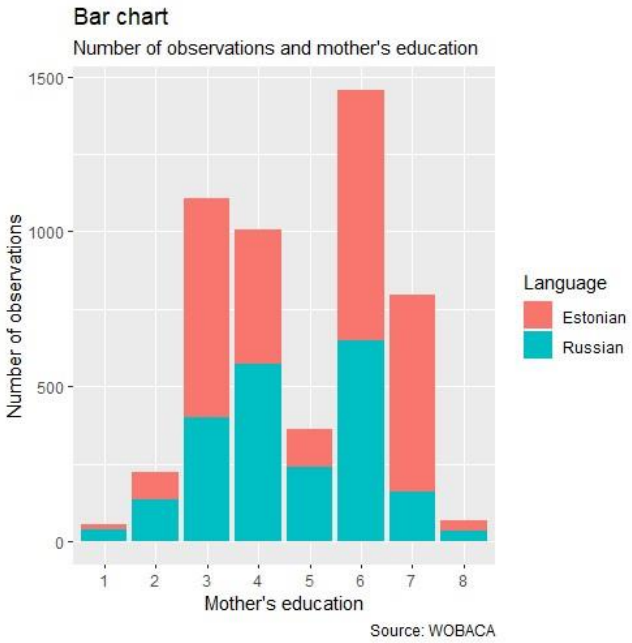

Figure 4: Mother's education and the language of the students

Table 2: Demographic Characteristics of the Treatment Group

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Age	147	17.082	0.856	15	16	18	19
Gender	147	0.435	0.498	0	0	1	1
LanguageHome	147	1.912	0.285	1	2	2	2
Family	147	1.558	1.021	1	1	2	5
Siblings	147	1.014	1.126	0	0	1.8	8
MotherEducation	147	4.748	1.659	1	3	6	8
Language	147	1.476	0.501	1	1	2	2

Table 3: Demographic Characteristics of the Control Group

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Age	150	17.073	0.883	15	16	18	19

Gender	150	0.493	0.502	0	0	1	1
LanguageHome	150	1.887	0.318	1	2	2	2
Family	150	1.787	1.046	1	1	2	5
Siblings	150	0.827	1.008	0	0	1	8
MotherEducation	150	4.940	1.594	2	3.2	6	8
Language	150	1.413	0.494	1	1	2	2

Figure 4 above shows that most students have relatively highly educated parents, and furthermore, significantly more students from Estonian-language schools have mothers with a Master's degree, while Russian-school students in general have mothers with vocational education.

Finally, Tables 2 and 3 indicate that there are no statistically significant differences in any of the background characteristics of the control and treatment group. This shows that our randomization was successful.

4.2 Attitudinal Measures: Factor Analysis

Tables 4 and 5 present the summary statistics of the ten sex-role attitudes in the treatment and control groups. The lowest mean-score values occur in the variables MBREAD, FULDBAD, WRKMHOME (Likert-scale category 2 indicates disagreement with the statement and category 3 slight disagreement). MBREAD tests the "male breadwinner" sentiment by stating: "It is much better for everyone if the man goes to work and the woman stays at home and looks after house and children." FULDBAD indicates that a father who works full-time cannot care for his children properly, while WRKMHOME states that in case of both parents working full-time, the mother must take full responsibility for the home and children.

In broad terms students are, on average, relatively egalitarian in their sex-role attitudes; they rather weakly prefer full- or part-time work of both parents and equal caregiving at home. We see some minor differences between the control and treatment groups, but these descriptive insights are inconclusive. Thus, after constructing the composite scales for attitudes, we can test whether the differences between the two groups are significant. As stated earlier, we have 10 attitudinal items in our survey, some of which indicate modern sex-role attitudes, while others traditional.

Table 4: Descriptives of the Attitudes: Control Group

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
FULMW	150	3.853	1.172	1	3	5	6
BOTHFUL	150	4.447	1.014	2	4	5	6
CHLSUF	150	3.000	1.087	1	2	4	6
MBREAD	150	2.673	1.184	1	2	3	6

CHLBEN	150	4.120	1.055	1	4	5	6
BOTHPART	150	4.047	1.228	1	3	5	6
FULDBAD	150	2.373	1.150	1	2	3	6
WRKMHOME	150	2.667	1.109	1	2	3	6
FULDW	150	4.147	1.276	1	3	5	6
ROLECHA	150	4.800	0.990	2	4	6	6

Table 5: Descriptives of the Attitudes: Treatment Group

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
FULMW	147	3.762	1.268	1	3	5	6
BOTHFUL	147	4.476	1.036	1	4	5	6
CHLSUF	147	3.293	1.166	1	3	4	6
MBREAD	147	2.653	1.317	1	2	4	6
CHLBEN	147	3.864	1.231	1	3	5	6
BOTHPART	147	4.449	1.130	1	4	5	6
FULDBAD	147	2.395	1.214	1	1	3	6
WRKMHOME	147	2.531	1.273	1	1.5	4	6
FULDW	147	3.952	1.401	1	3	5	6
ROLECHA	147	4.857	1.034	1	4	6	6

The results of the reliability analyses (Appendix 3) lead us first to confirmatory factor analysis, and further, to the exploration of a good-model-fit-based optimal factor structure through the use of modification indexes. We initiate our model estimation from M0 (Figure A1 in Appendix 1) and report the model fit characteristics in Table 6. The model-fit statistics in Table 6 for model M0 suggest revising and modifying the model. We report the following model-fit characteristics: Chi-square (cut-off value $p < 0.05$); the Comparative Fit Index (CFI), which compares the fit of a target model to the fit of the null model (the target cut-off value $CFI > 0.90$); and the Root Mean Square Error of Approximation (RMSEA), where values close to 0 represent a good fit.

Table 6: Model Fit Statistics: Comparison of CFA (M0) and EFA (M1)

Model	$Chi^2(df,p)$	CFI	TLI	RMSEA (CI_{90})	$\Delta Chi^2(df)$
M0: theoretical	186.23(32,.000)	0.628	0.476	0.127(0.11-0.15)	
M1: final	15.64(10,.002)	0.984	0.967	0.044(0.000-0.083)	170.58***(22)

The Chi-Square indicates not that good fit, but this statistic may not discriminate between good and poor fitting models (Kenny & McCoach, 2003). Due to the restrictiveness of the Chi-Square

statistic, we rely on CFI and RMSEA. RMSEA is now regarded as one of the most informative fit indices (Hooper et al., 2008; Diamantopoulos & Siguaw, 2000) because it favors parsimony by choosing the model with the lower number of parameters. Recommendations for RMSEA cut-off points have been reduced considerably in the past fifteen years. However, most of the authors state a cut-off value close to 0.06 (Hu & Bentler, 1999) or a stringent upper limit of 0.07 (Steiger, 2007) as a consensus amongst experts in this area (Hooper et al., 2008). The confidence interval around the RMSEA value indicates that in a well-fitting model the lower limit is close to 0, while the upper limit should be lower than 0.08.

So, M0 generally performs poorly, and we use the modification indexes for a better model fit. Our final model M1 is shown in Figure 5, and its fit characteristics, which meet the above-described threshold criteria, are given in Table 6.

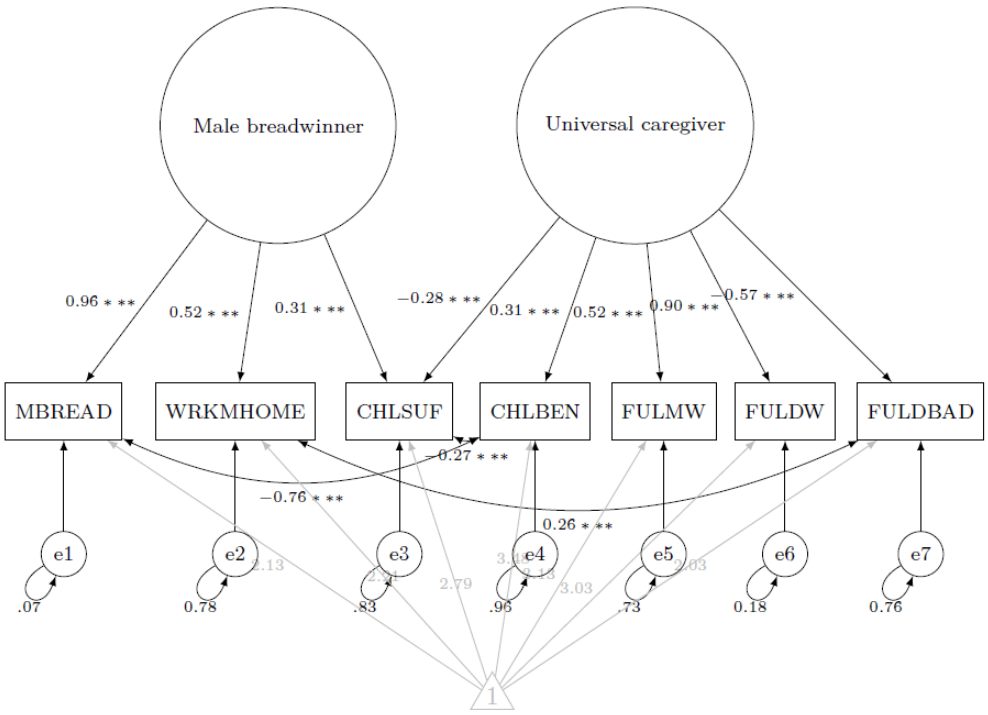

Figure 5: Final Factor Model
Notes: All variable names are given in Section 4.2.4.

In the final model, we see that the latent variable *Traditional sex-role attitudes* has 3 observable items—MBREAD, WRKMHOME, and CHLDSUF—all with positive factor loadings. The highest loading is on MBREAD, similar to Walter (2018). The other latent variable *Modern sex-role attitudes* has 5 observable items: CHLDSUF, CHILDBEN, FULMW, FULDW and FULDBAD, the first and last of which have negative factor loadings. FULDW (A full-time working father can normally establish just as close a relationship with his small child as a father who doesn't work) has the highest factor loading. Seemingly, egalitarian (modern) sex-role attitudes are not only measuring the feminine aspect and mothers' equal rights/obligations of work and care, but also consider males'/fathers' roles as vital in family affairs.

In Figure 6 we report the distribution of the scores for both factors by control and treatment group. It is worth mentioning that the distribution of the traditional attitudes is right-skewed, i.e. the mean is higher than the median, while modern attitudes are normally distributed. In both cases we can visually see that there are some differences between the treatment and control group, and so, we now proceed to test whether these differences are statistically significant.

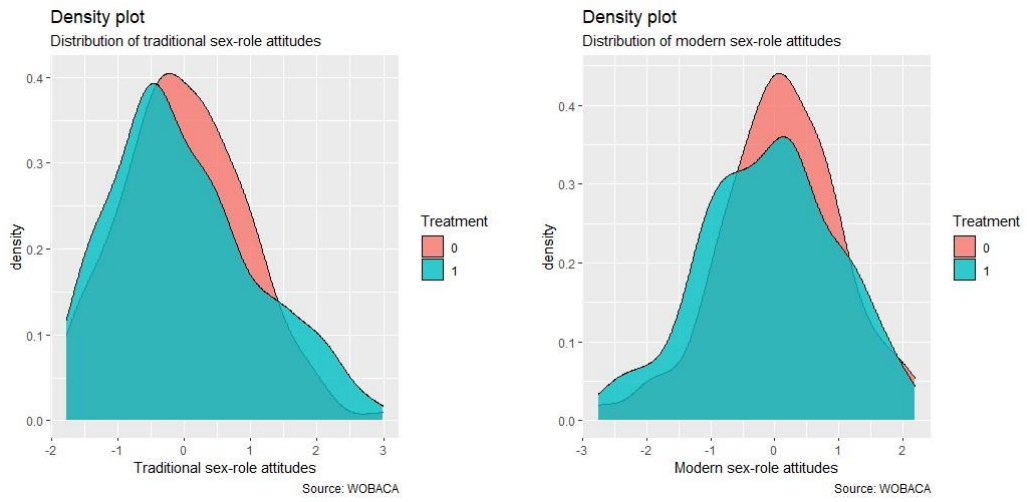

Figure 6: Distributional Properties of the Attitudinal Factors by the Treatment (=1) and Control Group (=0)

4.3 Treatment Effects

We are interested in whether we can find any treatment effects in the two attitudinal factors (measures for traditional and modern sex-role attitudes) indicated in the previous section. Our first step of the analysis is to standardize the factors scores. For testing H1 (whether there are average treatment effects), as indicated in Section 3.4, we run a Welch two-sample t-test to see whether we can find evidence to support our hypothesis that the treatment has a negative effect on traditional sex-role attitudes and a positive effect on modern. The test results show that the null hypothesis (there is no difference between the control and treatment group) can be rejected in both cases with a p-value smaller than 0.000. The treatment group average score on traditional sex-role attitudes was 0.500 compared to -0.005 in the control group, suggesting that, contrary to our expectation, the treatment made the students *more* traditional in their attitudes on average. Simultaneously, modern sex-role attitude scores also increase on average: treatment group average was 0.500 and control group -0.000, meaning that the treatment had the intended effect of increasing the modern sex-role attitudes as well.

In Figure 6, we show visual evidence by groups. In the upper panels, Estonian students are presented and in the lower panels, Russian students; right and left panels represent the different genders. Simple lines with a negative slope indicate that, as expected, more modern/egalitarian students are also less traditional (on average). We also gain a descriptive insight that only Estonian female students with highly modern sex-role attitudes were pushed to be less modern,

while highly traditional Estonian males became less traditional. We see no effect in Russian male students, but Russian female students with highly modern attitudes became less modern (as in the case of Estonian females), while also extremely traditional became less traditional.

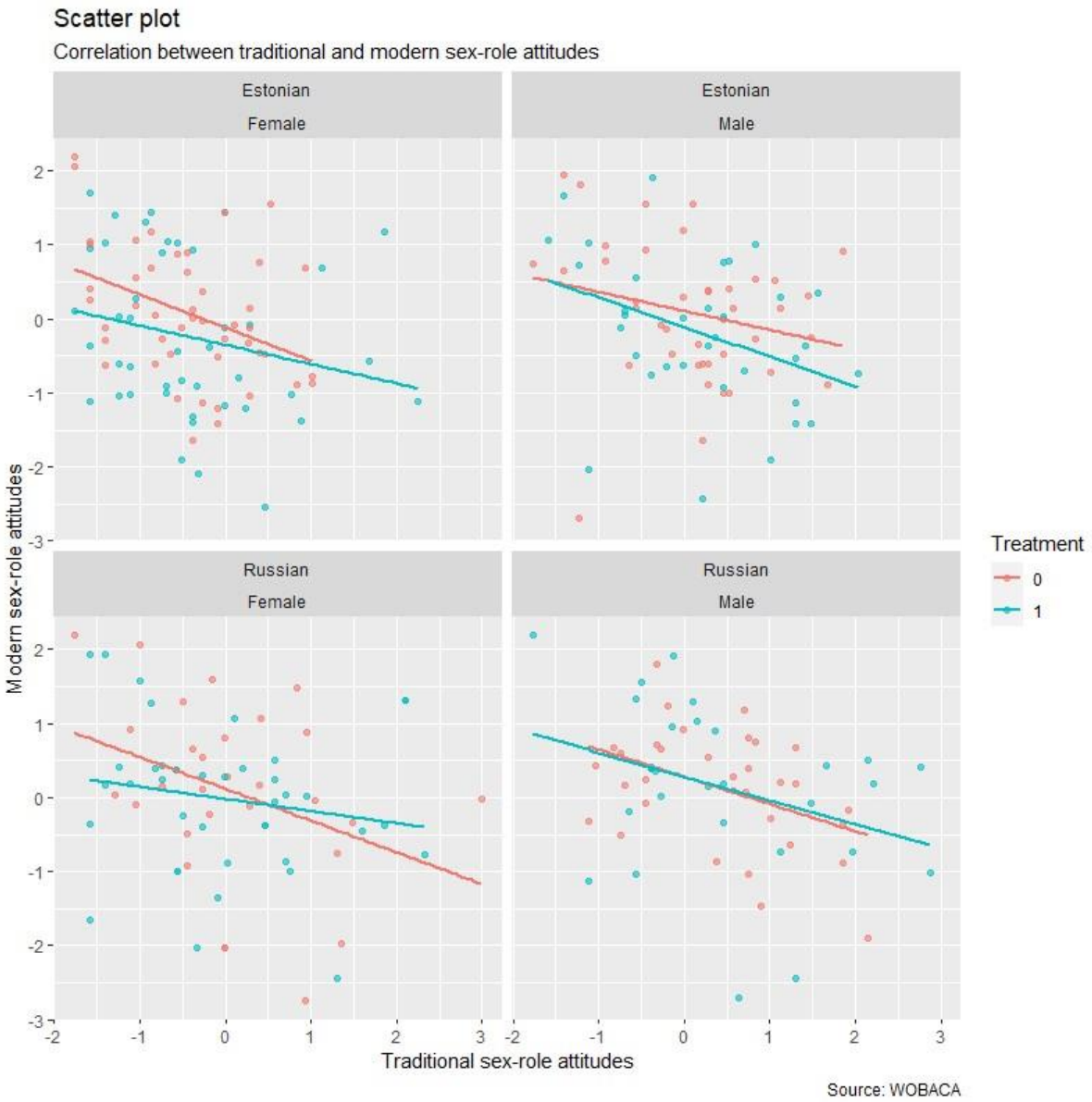

Figure 6: Correlations between Traditional and Modern sex-role attitudes

4.3.1 Traditional Sex-role Attitudes

Now we turn to the stepwise multivariate regression models to estimate the treatment effects. The model specification is given by Equation 1, and we first estimate *Traditional sex-role attitudes* and then *Modern sex-role attitudes*. In both cases, we use the same covariates on the right-hand side of the equation.

Table 7 indicates that age, number of siblings, and mother's education do not affect traditional sex-role attitudes. However, we see that males have much higher traditional sex-role attitudes when compared to females (approximately +0.5 s.d). Russian school students (LanguageQuestionnaire = Russian) also have higher traditional sex-role attitudes when compared to their Estonian counterparts. Unfortunately, we are not able to identify treatment effects.

Table 7: Demographic Components of Traditional Sex-role Attitudes

	Dependent variable: Traditional.z			
	(1)	(2)	(3)	(4)
Age	0.031	0.03	−0.013	0.025
	(0.906)	(0.112)	(0.097)	(0.071)
SexMale	0.495***	0.498***	0.510***	0.495***
	(0.172)	(0.176)	(0.172)	(0.178)
Siblings		0.016	0.016	0.041
		(0.074)	(0.072)	(0.073)
MotherEducation 2			−0.568	−0.521
			(0.404)	(0.265)
MotherEducation 3			−0.169	−0.024
			(0.227)	(0.174)
MotherEducation 4			−0.492	−0.407
			(0.356)	(0.243)
MotherEducation 5			0.172	0.214
			(0.175)	(0.115)
MotherEducation 6			−0.592**	−0.470**
			(0.284)	(0.185)
MotherEducation 7			−0.704**	−0.487**
			(0.31)	(0.221)
MotherEducation 8			−0.685	−0.561
			(0.864)	(0.849)
LanguageQuestionnaire				0.369***
				(0.103)
Treatment	0.029	0.026	0.007	−0.018
	-(0.107)	-(0.108)	-(0.085)	-(0.083)
Observations	297	297	297	297
R^2	0.062	0.062	0.125	0.154
Adjusted R^2	0.052	0.049	0.091	0.118
Residual Std. Error	0.974	0.975	0.953	0.939
F Statistic	6.436***	4.834***	3.695***	4.310***

*Note: * p < 0.1; ** p < 0.05, *** p < 0.01*

In Table 8 we ask whether specific groups can be treated: (a) males and (b) Russian language students. We see that all interaction effects with treatment are insignificant, and so we were unable to treat the individuals in the more traditional groups of male students and Russian language students.

Table 8: Treatment Effects of Males and Russian School Students

	Dependent variable: traditional.z	
	(1)	(2)
Age	0.025	0.025
	(0.071)	(0.07)
SexMale	0.440**	0.439**
	(0.177)	(0.177)
Siblings	0.04	0.04
	(0.073)	(0.072)
MotherEducation 2	−0.527	−0.527
	(0.285)	(0.284)
MotherEducation 3	−0.027	−0.029
	(0.177)	(0.178)
MotherEducation 4	−0.417	−0.419
	(0.265)	(0.264)
MotherEducation 5	0.209	0.206
	(0.13)	(0.133)
MotherEducation 6	−0.478**	−0.480**
	(0.198)	(0.197)
MotherEducation 7	−0.494**	−0.495**
	(0.233)	(0.263)
MotherEducation8	−0.541	−0.544
	(0.842)	(0.846)
LanguageQuestionnaire	0.371***	0.382***
	(0.101)	(0.127)
Treatment	−0.071	−0.039
	(0.116)	(0.302)
SexMale:Treatment	0.113	0.114
	(0.159)	(0.157)
LanguageQuestionnaire:Treatment		−0.022
		(0.169)
Observations	297	297
R^2	0.155	0.155
Adjusted R^2	0.116	0.113

Note: *p<0.1; **p<0.05; ***p<0.01

4.3.2 Modern sex-role attitudes

Similar to the previous section, we estimate a step-wise model for examining the effect of treatment on *Modern sex-role attitudes*. Table 9 indicates that there are no group differences by

any other demographic characteristic other than mother's education. Recall that the mother's education was an ordered, categorical variable with the following structure: 1) primary education, 2) lower secondary education, 3) secondary education, 4) vocational (secondary) education, 5) unfinished Bachelor education (2 or more years), 6) Bachelor's education, 7) Master's degree and 8) Doctoral degree.

Table 9: Demographic Components of Modern Sex-role Attitudes

	Dependent Variable: Modern.z			
	(1)	(2)	(3)	(4)
Age	−0.105	−0.105	−0.071	−0.052
	(0.073)	(0.073)	(0.07)	(0.06)
SexMale	0.067	0.067	0.069	0.061
	(0.092)	(0.091)	(0.091)	(0.086)
Siblings		0.0001	0.004	0.017
		(0.026)	(0.028)	(0.028)
MotherEducation 2			1.469*	1.493*
			(0.824)	(0.76)
MotherEducation 3			0.697	0.771
			(0.786)	(0.749)
MotherEducation 4			0.884	0.927
			(0.666)	(0.618)
MotherEducation 5			0.377	0.398
			(0.787)	(0.738)
MotherEducation 6			1.061	1.123*
			(0.813)	(0.762)
MotherEducation 7			1.08	1.191*
			(0.72)	(0.679)
MotherEducation 8			1.667*	1.730*
			(0.993)	(0.934)
LanguageQuestionnaire				0.188
				(0.14)
Treatment	−0.172	−0.172	−0.142	−0.154
	(0.145)	(0.145)	(0.129)	(0.129)
Constant	1.843	1.843	0.332	−0.332
	(1.143)	(1.145)	(1.332)	(1.397)
Observations	297	297	297	297
R2	0.017	0.017	0.087	0.095
Adjusted R2	0.007	0.004	0.052	0.057
Residual Std. Error	0.996	0.998	0.974	0.971

F-Statistic	1.712	1.279	2.477***	2.480***

Note: *p<0.1; **p<0.05; ***p<0.01

So, Table 9 reveals that students with more modern sex-role attitudes come from families where the mother either has a lower secondary education or a higher post-secondary education (Bachelor's or higher). All other demographic background characteristics are insignificant. Unfortunately, the treatment also does not have a significant effect in increasing modern sex-role attitudes in students.

However, upon interacting the treatment with mother's education variable, to see whether we influenced the students with more egalitarian attitudes to be even more modern, we find some interesting results (see Table 10).

Table 10: Treatment Effects: Interaction with Mother's Education

	Dependent variable: Modern.z	
Age	−0.055	(0.069)
SexMale	0.085	(0.084)
Siblings	0.025	(0.025)
MotherEducation 2	3.700***	(1.103)
MotherEducation 3	2.636**	(1.077)
MotherEducation 4	2.722***	(0.977)
MotherEducation 5	2.077	(1.097)
MotherEducation 6	2.976***	(1.092)
MotherEducation 7	3.267***	(0.989)
MotherEducation 8	2.673***	(0.801)
LanguageQuestionnaire	0.21	(0.147)
Treatment	1.713**	(0.774)
MotherEducation 2:Treatment	−2.584***	(0.884)
MotherEducation 3:Treatment	−1.852**	(0.793)
MotherEducation 4:Treatment	−1.720**	(0.679)
MotherEducation 5:Treatment	−1.461	(0.921)
MotherEducation 6:Treatment	−1.832***	(0.564)
MotherEducation 7:Treatment	−2.291***	(0.684)
Observations	297	
R^2	0.123	
Adjusted R^2	0.066	
Residual Std. Error	0.966 (df = 278)	
F-Statistic	2.168*** (df = 18; 278)	

Note: *p<0.1; **p<0.05; ***p<0.01

First, we find that the average treatment effect becomes positive and weakly significant, i.e. on average the video led to an increase in modern sex-role attitudes. However, students whose mother's have secondary, vocational, or unfinished bachelor education experience no (positive) treatment effect, while there are relatively large (by magnitude), positive, and statistically significant effects for students whose mothers have a lower secondary school education, i.e. the video increased modern sex-role attitudes to a greater extent among these students than on average. Moreover, it is worth mentioning that in case of students whose mothers hold Bachelor's and/or Master's degrees, the treated students became less modern in their sex-role attitudes.

5. Discussion

We showed that treatment by work-life balance promotion film does not affect the support for traditional male-breadwinner attitudes while it can change the attitude toward more egalitarian views on the childcare of the respondents. All effects we measure were short-run, meaning just after treatment, and we cannot argue about the persistence of the effects (or no effects). While we are not aware of any other experiment that treats sex-role attitudes by film, there is some supporting evidence in neighboring fields, e.g. attitudes toward safer sex (Kyes 1990), climate change (Howell 2011; Beatie et al. 2011).

In addition, we showed that certain family and background characteristics correlate with more traditional male-breadwinner attitudes. Namely males and Russian speakers have more traditional sex-role attitudes, which were not open to treatment. Similarly, Thijs, Te Grotenhuis, and Scheepers (2017) find the change in attitudes of men is much harder to explain than the change in women's attitudes. Similar to Pampel (2011) who argues that compositional changes of the population change attitudes of innovative, high-status groups first before egalitarian attitudes spread to other groups, we also demonstrate that support for male breadwinners is lower and support for equality in childcare (universal caregiver) is higher among highly educated. Pepin and Cotter (2018) also show that an increase in egalitarian attitudes depends on what measure you look at. They also found that attitudes regarding the family are less egalitarian than attitudes towards public roles of women. We did not measure the latter directly, but we showed that male breadwinner attitudes are less prevailing.

Even though our theoretical frame was arguing that more modern sex-role attitudes in interaction with existing institutional policies will bring along work-life reconciliation, methodologically, it is difficult to ascertain to what extent sex-role attitudes have an impact on fertility levels or labor market outcomes. Still, some of the scholars (Gornick and Meyers 2009; Esping-Andersen 2009) have proposed a set of policies designed to foster more egalitarian family divisions of labor by strengthening men's ties at home and women's attachment to paid work. Mainly inspired by the experience of the Nordic countries this template is based on three sets of policies: 1) individual rights of mothers and fathers to moderately paid leaves; 2) high-quality, publicly financed, universally accessible early childcare services to emphasize the investment-related policy approach (see Morgan (2005) for instance); 3) working-time regulations and more flexible work schedules that limit full-time work hours and increase the availability and quality of part-time jobs. However, part-time work arrangements are recently contested by Esping-Andersen (2015) who emphasize that the prevalence or recommendations of part-time working among females are at the same time the main obstacle of improving the bargaining power (pp. 175). Still, the aforementioned proposed set of policies indicates that the

quality of work-family policy in terms of influencing the behavior of females is dependent on the interplay of attitudes and work-family policy instruments. This interplay is converted to capabilities through sex-role attitudes. We cannot test this premise empirically, rather we take it as an underlying assumption of our study and argue that more universal caregiving and less traditional male-breadwinner attitudes can produce more gender-balanced labor market outcomes if conditioned to a "good policy mix". Latter includes not only public financing of childcare and leave policies but also supporting organizational practices of more flexible work schedules, rest periods, and holidays for parents.

Besides social structural developments, the institutional context can also influence gender role attitudes, and cultural norms can have an effect on the formulation, institutionalization, and efficacy of work-family policies (Budig, Misra, & Boeckmann, 2012). It can be argued that family policies influence gender role attitudes by signaling what is defined as appropriate behavior and by shaping the choices which are available to individuals (Jakobsson & Kotsadam, 2010). Grunow, Begall, and Buchler (2018) refer to policy feedback theory to explain the effect of family policies on gender ideologies. Here the idea is that "interests, beliefs, and ideologies held by citizens…feedback into the policy-making process" (p.47). Work-family policies on the other hand may influence individual gender role attitudes through role exposure and norm-setting. So policies could serve as cultural and normative reference points. Therefore, the complexity of interaction effects of policies and informal institutions, including attitudes, can make our experiment as a tool for manipulating attitudes more valuable due to its direct effect on attitudes.

6. Conclusion

We were studying the impact of the treatment on sex-role attitudes using the RCT design. Our treatment was a short movie that promoted modern and egalitarian sex roles, especially father in the caregiving role of parenting. In total, we randomized 16-year and older students from 10 schools to treatment and control groups, with 297 observations in total.

Our theoretical approach was the so-called capabilities approach, which argues that informal institutions (attitudes in our case) interact with work-family policies and generate capabilities of the parents, that determine macro consequences such as labor participation, the gender wage gap, and distribution of care-work within the family. Also, our measuring exercise was resting on the premise that attitudes can be measured, and each individual has a set of traditional (male breadwinner) and egalitarian (universal caregiver) parental role attitudes, which are highly (negatively) correlated. Whether informal institutions such as attitudes are persistent or open to change? On this question we relied on the classical arguments of social psychology, stating that low-effort processes of persuasive communication can alter attitudes.

We show that attitudes can be measured in a two-factor scale, where one factor is measuring the traditional roles of parents in the labor market (male breadwinner) and another is measuring the egalitarian role of parents at home (universal caregiver). Our experimental results show that males and Russian speakers have more traditional male breadwinner attitudes, while highly educated mothers are positively correlated with the supporting attitude about egalitarian parental roles at home – universal caregiving. Similarly, in literature (citation is missing) it was argued that males and East Europeans are less egalitarian regarding attitudes capturing mother/child relation and less traditional regarding male-breadwinner. This empirical insight is to a large extent in line with our results.

Our main results are related to the experiment of studying the treatment effects of the short film, promoting modern sex roles at work and home. Our results show that traditional sex-role attitudes of a male breadwinner cannot be treated by the short film, while a more egalitarian universal caregiving attitude can be promoted by it. However, the positive effects are significant statistically and in size only in the case of students with low or unfinished bachelor education mothers. In the case of students with master's degree mothers, the treatment effects of the latter were negative.

The main limitation of the study is originated mainly from the data collection instrument. The rationale behind self-report scales of attitudes is that people are both willing and able to accurately report their attitudes. It can be argued that these conditions are not always met; people can behave strategically or have other incentives to withhold information, or just present them in a more positive light (see Krosnick et al. 2005, Schwarz 2008). One limitation regarding our choice of explicit instrument is the possibility that the topic of sex-role attitudes is considered to be socially sensitive. If people feel that it is not right to have traditional gender roles (or that they would be perceived badly for having such attitudes), there would be a bigger tendency not to reveal true attitudes. Then, our approach to go for the explicit measure (self-reported instrument) can be criticized by the insight that in sensitive domains implicit measures predict behavior better than explicit ones (e.g. intergroup prejudice and discrimination (Greenwald et al. 2009).

The other limitation of the study regarding external validity can originate from two sources – treatment can be case-specific or treatment effects can be nonlinear, e.g. effects are significant only in case of relatively moderate pre-treatment factor scores. What concerns the first limitation, then we see some channels for attitudinal change (e.g. actors are well known only locally) that might hinder external validity. What concerns second, then it is empirically shown that Europeans in Continental and Anglo-Saxon countries have the most traditional attitudes, especially regarding a joint contribution to the income (Walter, 2018) and most East Europeans are less egalitarian and more traditional than West Europeans (Dotti Sani & Quaranta 2017, Tanaka & Lowry 2011). At the same time, it is also shown that East Germans are more egalitarian than West Germans (Bauernschuster & Rainer 2012; Lee et al. 2007). So studies about East Europe have mixed results. We may argue that due to Nordic influence in our case students are relatively less traditional and more egalitarian than most post-communist cases. Still, it could be that only relatively low scores on egalitarian universal-caregiver attitudes are open to treatment.

7. Bibliography

Beattie, G., Sale, L., & Mcguire, L. (2011). An inconvenient truth? Can a film really affect psychological mood and our explicit attitudes towards climate change? Semiotica, 2011(187). doi:10.1515/semi.2011.066

Beere, C.A. 1990. Gender roles: A Handbook of Tests and Measures. Greenwood Press: New York.

Beere, C.A. 1990. Gender roles: A Handbook of Tests and Measures. Greenwood Press: New York.Behr, D., Braun, M., Kaczmirek, L., Bandilla, W., 2012. Testing the Validity of Gender Ideology Items by Implementing Probing Questions in Web Surveys. Field Methods 0, 1-18.

Borchorst, A., & Siim, B. (2008). Woman-friendly policies and state feminism: Theorizing Scandinavian gender equality. Feminist theory, 9(2), 207-224.

Braun, M., 2008. Using Egalitarian Items to Measure Men's and Women's Family Roles. Sex Roles 59, 644-656.

Budig, M. J., Misra, J., & Boeckmann, I. (2012). The motherhood penalty in cross-national perspective: The importance of work–family policies and cultural attitudes. Social Politics, 19(2), 163-193.

Ciccia, R., & Verloo, M. (2012). Parental leave regulations and the persistence of the male breadwinner model: Using fuzzy-set ideal type analysis to assess gender equality in an enlarged Europe. Journal of European Social Policy, 22(5), 507-528.

Conrey, F.R., Smith, E.R. 2007. Attitude representation: attitudes as patterns in a distributed, connectionist representational system. Soc. Cogn. 25:718–35.

Crompton, R., Brockmann, M., & Lyonette, C. (2005). Attitudes, women's employment and the domestic division of labour: a cross-national analysis in two waves. Work, employment and society, 19(2), 213-233.

Cronbach, L.J. 1951. Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334.

Cunningham W.A., Zelazo, P.D., Packer, D.J., van Bavel, J.J. 2007. The iterative reprocessing model: a multilevel framework for attitudes and evaluation. Soc. Cogn. 25:736–60.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., Moors, A. 2009. Implicit measures: a normative analysis and review. Psychol. Bull. 135:347–68.

Fazio, R.H., Jackson, J.R., Dunton, B.C., Williams, C.J. 1995. Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? J. Personal. Soc. Psychol. 69: 1013–27.

Fazio, R.H, Olson, M. A. 2003. Implicit measures in social cognition research: their meaning and use. Annual Review of Psychology. 54: 297–327.

Festinger, L. (1957). A Theory of cognitive dissonance. Stanford, CA: Stanford University Press.

Fraser, N. 1994. After the family wage: gender equity and the welfare state. Political Theory. 22 (4): 591–618.

Graham, J. M. 2006. Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability. Educational and Psychological Measurement 66:6, pages 930-944.

Greenwald, A.G., McGhee, D.E., Schwartz, J.L.K. 1998. Measuring individual differences in implicit cognition: the implicit association test. J. Personal. Soc. Psychol. 74:1464–80.

Greenwald, A.G., Poehlman, T.A., Uhlmann, E.L., Banaji, M.R. 2009. Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. J. Personal. Soc. Psychol. 97:17–41

Gerson, K. 2010. The Unfinished Revolution: Coming of Age in a New Era of Gender, Work, and Family. Oxford University Press.

Hooper, D., Coughlan, J., Mullen, M. 2008. Equation Modelling: Guidelines for Determining Model Fit. Electronic Journal of Business Research Methods, 6(1), 53-60.

Howell, R. A. (2011). Lights, camera … action? Altered attitudes and behaviour in response to the climate change film The Age of Stupid. Global Environmental Change, 21(1), 177–187. doi:10.1016/j.gloenvcha.2010.09.004.

Jakobsson, N., & Kotsadam, A. (2010). Do attitudes toward gender equality really differ between Norway and Sweden?. Journal of European Social Policy, 20(2), 142-159.

Knight, C.R., Brinton, M.C., 2017. One Egalitarianism or Several? Two Decades of Gender-Role Attitude Change in Europe. American Journal of Sociology 122, 1485-1532.

Kremer, M. 2006. The Politics of Ideals of Care: Danish and Flemish Child Care Policy Compared. Soc. Polit. Int. Stud. Gend. State Soc. 13, 261–285.

Krosnick, J.A., Judd, C.M., Wittenbrink B. 2005. The measurement of attitudes. In The Handbook of Attitudes, ed. D Albarracin, BT Johnson, MP Zanna, pp. 21–76. Mahwah, NJ: Erlbaum

Kurowska, A. 2016. (De)familialization and (De)genderization – Competing or Complementary Perspectives in Comparative Policy Analysis? Social Policy & Administration, 52(1): 29-49.

Kyes, K. B. (1990). The effect of a "safer sex" film as mediated by erotophobia and gender on attitudes toward condoms. Journal of Sex Research, 27(2), 297–303. doi:10.1080/00224499009551559

Lauri, T., Poder, K., Ciccia, R. 2019. Pathways to Gender Balance: A Configurational
Analysis of Childcare Instruments and Outcomes in 21 European Countries. Social Policy & Administration, 54 (5): 646-665.

Likert, R. 1932. A Technique for the Measurement of Attitudes. Archives of Psychology, 140, 1–55.

McLaughlin, E. and Glendinning, C. (1994), "Paying for care in Europe: is there a feminist approach?", in Hantrais, L. and Mangen, S. (Eds), Family Policy and the Welfare of Women, University of Loughborough, Loughborough, pp. 52-69.

Morgan, K. J. (2005). The "production" of child care: How labor markets shape social policy and vice versa. Social Politics: International Studies in Gender, State & Society, 12(2), 243-263.

Petty, R.E., Krosnick, J.A. 1995. Attitude Strength: Antecedents and Consequences. Mahwah, NJ: Erlbaum

Petty, R. E., Wheeler, S. C., & Tormala, Z. L. 2003. Persuasion and attitude change. In T.
Millon & M. Lerner (Eds.), Handbook of psychology: Personality and social psychology (pp. 353-382). Hoboken, NJ: Wiley.

Petty, R.E., Tormala, Z.L., Brinol, P., Jarvis, W.B.G. 2006. Implicit ambivalence from attitude change: an exploration of the PAST model. J. Personal. Soc. Psychol.
90:21–41.

Raykov, T. 1997. Scale Reliability, Cronbach's Coefficient Alpha, and Violations of Essential Tau-Equivalence with Fixed Congeneric Components, Multivariate Behavioral Research, 32:4, 329-353.

Richetin, J., Perugini, M., Prestwich, A., O'Gorman, R. 2007. The IAT as a predictor of food choices: the case of fruits versus snacks. Int. J. Psychol. 42:166–73.

Rosseel, Y. 2012. Lavaan: An R Package for Structural Equation Modeling, Journal of Statistical Software, Journal of Statistical Software, Vol 48, 2: 1-36.

Schwarz, N. 2008. Attitude measurement. See Crano & Prislin 2008, pp. 41–60.

Spearman, C. 1904. "General intelligence," objectively determined and measured. American Journal of Psychology, 15, 201–293.

Walter, J. 2018. Measures of gender role attitudes over time. Doctoral disstertation, Inauguraldissertation zur Erlangung des akademischen Grades einer Doktorin der Sozialwissenschaften, Universität Mannheim.

Waylen, G. (2014). Informal Institutions, Institutional Change, and Gender Equality. Political Research Quarterly, 67(1), 212-223. http://www.jstor.org/stable/23612047

Ziegert JC, Hanges PJ. 2005. Employment discrimination: the role of implicit attitudes, motivation, and a climate for racial bias. J. Appl. Psychol. 90:553–62.

Appendix 1: Measurement model

For our measurement, we use the model which specifies the structure between sex-role attitudes from Walter (2018). In this framework (see Figure 7) there are two strongly correlated latent variables—traditional and modern (egalitarian) sex-role attitudes—and 10 observable items. Even though in the original framework (ibid.) the model is built on the assumption that egalitarianism is not merely the reverse of traditionalism (e.g. Behr et al., 2012; Braun, 2008), we see the first factor as a traditional male-breadwinner measure and second as a universal or egalitarian-caregiver measure. Indeed egalitarianism is a multifaceted concept (Knight & Brinton, 2017), so some observables might contribute to both latent variables, but are expected to have negative factor loadings in one latent variable compared to the other (e.g. λ_5 and λ_7 both point on CHBEN – A child actually benefits if his or her mother has a job, rather than just concentrating on the home – but the first factor loading is negative while the second is positive). Also, Walter (2018) shows that the biggest positive factor loading in the latent variable indicating traditional sex roles attitudes (*Male breadwinner*) is λ_1 and λ_2, and in the latent variable measuring *Modern sex roles attitudes* (Universal caregiver) it is λ_9 (see Figure below). In addition, the estimate for the correlation between the latent variables is -0.60 (ibid). The figure below constitutes our measurement model (M0), where all abbreviations of the variables can be found in Appendix 2.

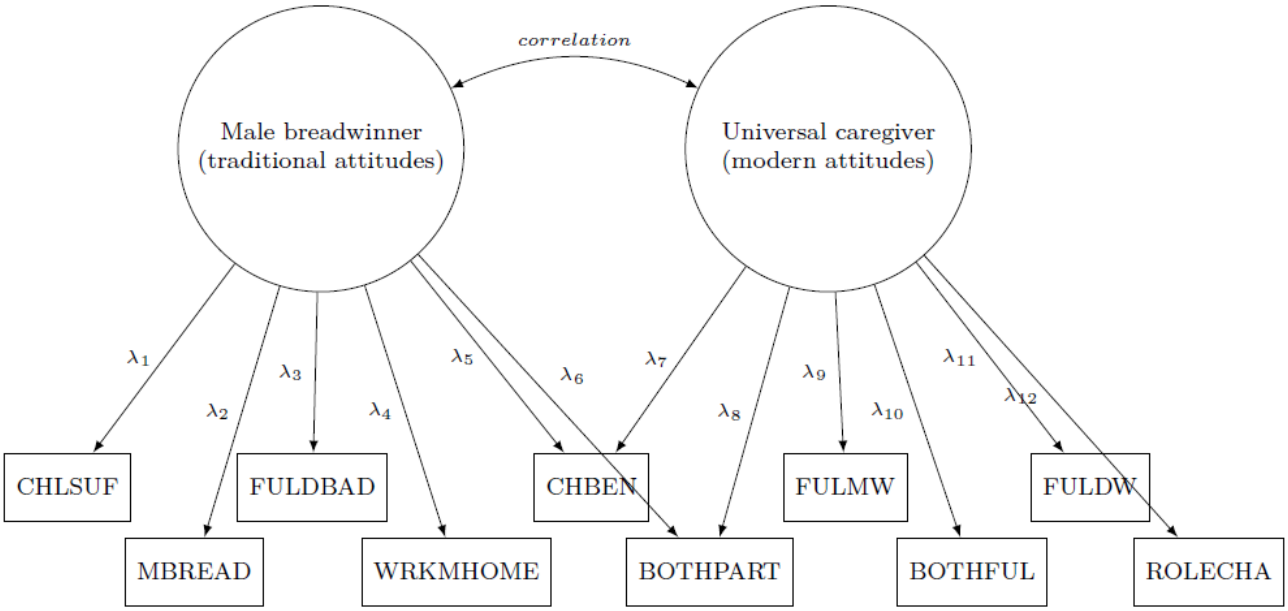

Figure A1: Initial Model M0

Appendix 2: Sex-role Attitudes Scale

Totally Disagree	Disagree	Slightly disagree	Slightly agree	Agree	Totally agree
1	2	3	4	5	6

CHLSUF A small child is bound to suffer if his or her mother goes out to work.

MBREAD It is much better for everyone concerned if the man goes out to work and the woman stays at home and looks after the house and children.

FULDBAD A father who works full-time cannot care for his children properly.

WRKMHOME Even if both parents work full-time, it is still better if the mother has the main responsibility of looking after the home and children.

CHLBEN A child actually benefits if his or her mother has a job, rather than just concentrating on the home.

BOTHPART The best way to organise family and work life is for both partners to work part-time and to look after the home and children equally.

FULMW A full-time working mother can normally establish just as close a relationship with her small child as a mother who doesn't work.

BOTHFUL The best way to organise family and work life is for both partners to work full-time and to look after the home and children equally.

FULDW A full-time working father can normally establish just as close a relationship with his small child as a father who doesn't work.

ROLECHA A man can be responsible for looking after the home and children just as well while the woman works full-time.

Appendix 3: Reliability of the Scales (Online Appendix)

We use Cronbach's alpha (Cronbach 1951) to test the internal reliability of the scales. The threshold for reliability is conventional, meaning that Cronbach's alpha should be bigger or equal to 0.7 (see Table 3a). The columns in Table 3a report raw alpha – the correlation of the item with the entire scale; standard alpha – the correlation of the item with the entire scale in case each item is standardized; and average r – the average inter-item correlation. We also report Guttman's Lambda 6 (G6), which considers the amount of variance in each item that can be accounted for in the linear regression of all of the other items (the squared multiple correlation or SMC), or more precisely, $\lambda_6 = 1 - \frac{\sum e_j^2}{V_x}$, where e_j^2 is the variance of the errors and V_x is the amount of variance in each item that can be accounted for by the linear regression of all of the other items. For tests with equal factor loadings, alpha is greater than G6, but if the loadings are unequal or if there is a general factor, G6 is greater than alpha, which we find to be the case for both factors (latent variables).

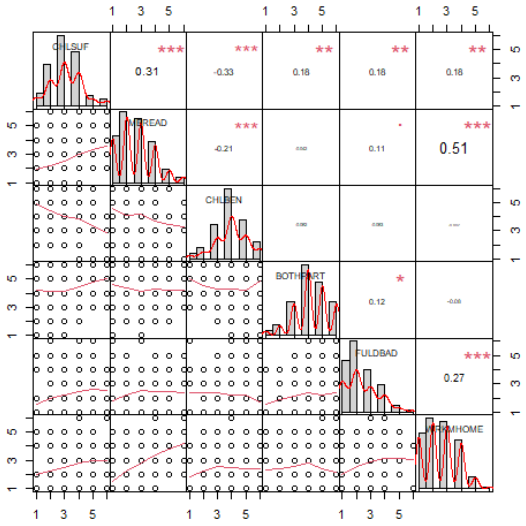

Figure A3.1: Correlation of observable items for latent variable *Universal caregiver*

Traditional sex-role attitudes (*Male breadwinner)* has the following observable items: CHLSUF, MBREAD, FLDBAD, WRKMHOME, CHBEN, BOTHPART. In Figure A3.1 we show that CHBEN is not significantly correlated with other items, and BOTHPART has a very weak correlation only with CHLSUF and FULDBAD. So, as expected, the Cronbach's raw alpha of the composite score of the latent is 0.32 and the standardized measure is 0.31. This low value is not only due to the issue that some items are measured in reverse scale (e.g. CHLBEN) for the traditional sex-role attitudes, but seemingly, some items (e.g. BOTHPART) reflect more the economic dimension of families than sex-roles attitudes per say. Also, Table A3.1 reports the measures if the item is dropped from the composite score, e.g. CHILDBEN increases the Cronbach's alpha, but not enough to meet the threshold.

Table A3.1: Reliability of the scale in measure of the latent "Male breadwinner" (if the item is dropped)

	raw alpha	std.alpha	G6	average r
CHLSUF	0.225	0.216	0.31	0.052
MBREAD	0.165	0.159	0.23	0.036
CHLBEN	0.515	0.515	0.53	0.175
BOTHPART	0.357	0.344	0.43	0.095
FULDBAD	0.204	0.190	0.32	0.045
WRKMHOME	0.083	0.076	0.16	0.016

The following observable items indicate the latent variable Modern sex-role attitudes (*Universal caregiver*): CHLBEN, BOTHPART, FULMW, BOTHFUL, FULDW, and ROLECHA. As Figure A3.2 shows, ROLECHA is not (or very weakly) correlated with the rest of the items. However, some items are moderately correlated, e.g. FULMW – a full-time working mother can normally establish just as close a relationship with her small child as a mother who doesn't work – with FULDW (same statement regarding father). In both cases (Figure A3.1 and A3.2) we see that some items—CHILDBEN, BOTHPART, and ROLECHA—are weakly or weakly and insignificantly correlated with the rest of the items, hindering the reliability.

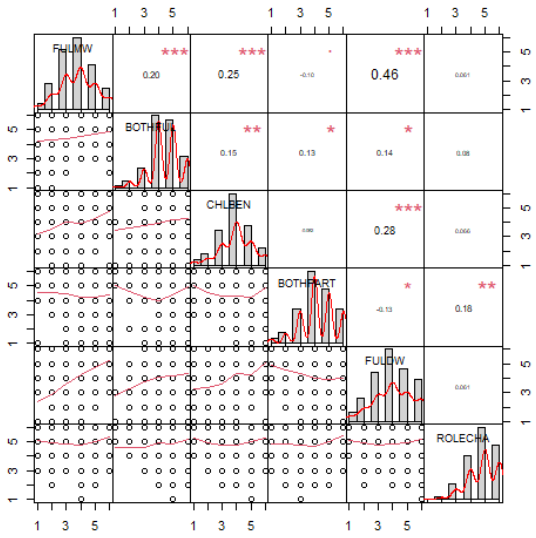

Figure A3.2: Correlation of observable items for latent variable *Universal caregiver*

Similar to the reliability of the scale of the previous latent variable, the raw Cronbach's alpha is relatively row (0.45), and some items do not match with the scale (see Table A3.2). Table A3.2 also indicates that dropping BOTHPART will increase Cronbach's alpha.

Table A3.2: Reliability of the scale in measure of the latent "Universal caregiver" (if the item is dropped)

	raw alpha	std.alpha	G6(smc)	average r
FULMW	0.31	0.33	0.32	0.088
BOTHFUL	0.38	0.37	0.39	0.105
CHLBEN	0.38	0.38	0.39	0.108
BOTHPART	0.53	0.52	0.49	0.176
FULDW	0.33	0.34	0.33	0.094
ROLECHA	0.43	0.43	0.44	0.131

In conclusion, our calculations demonstrate that not all items can be included in calculating the measures for Male breadwinner and Universal caregiver. The analysis also suggests that the tau-equivalence is violated in the case of measuring both latent variables.