The long term cost of partnership and fertility trajectory: later life labour market income of women across Europe

Joanne Sophie Muller, Nicole Hiekel, & Aart C. Liefbroer

The increase of female employment was the most significant change in labour markets during the past century (Goldin, 2006). However, a woman's earnings remain closely related to her changing family role over the life course. Mothers' wages lag behind those of childless women and men, even after controlling for individual work experience. This so called "motherhood penalty" is a well-established finding in many Western countries (e.g. Harkness & Waldfogel, 2003). More specifically, research shows that female earnings relate to the number- and timing of children and presence of a partner (Budig & England, 2001; Miller, 2011; Pienta, 1999). Also country-comparative research suggests that motherhood effects on income are shaped by contextual factors, most importantly family polices and gender attitudes (Budig, Misra, & Boeckmann, 2012).

Against this background, the present study will address the research question: what is the association between family trajectory, i.e. combined partnership and fertility history, and later life (50-65) labour market income among women? Also, we ask whether this association differs between countries. We contribute to the literature in three main ways: we take a holistic life course approach, we focus at later life outcomes and we provide a cross-national perspective.

First, previous research mainly focused on the effects of single events in the family trajectory on later life, for instance mother's age at first birth (e.g. Gough & Noonan, 2013; Pienta, 1999). The interplay between events in the fertility and partnership trajectories is understudied. Combining full fertility and union histories, we assess their combination, i.e. joint timing of partnership unions, dissolutions and child births, and take into account the interdependencies between the two. This approach is relevant, because events are highly likely to be related (e.g. the older the age at first birth, the lower the number of children is likely to be) and because the effect of events on female income might depend on other events or states (e.g. whether there are children present or not in case of divorce). We create a typology of family trajectories using sequence analysis (Cornwell, 2015).

Second, we focus on later life outcomes. Most studies regard only short-term income effects of motherhood. However, young mother's decision to quit their job or reduce working hours does not only lower their current income, but also compromises future earnings. Personal earnings after age fifty are the sum of cumulative work experience over the life course. Moreover, mothers might have lost attachment to the labour market because their live was devoted to care work. We wonder whether mothers stick to their custom even when children left the home and their direct care-related reason to stay out of the labour force is over. These are reasons we focus at personal labour market income rather than household income. In addition, earnings in later life are of interest because they reflect economic independence of women and the (under)use of their skills. On average mothers have less opportunity to provide a financial buffer for themselves which could make them vulnerable, especially at later ages. Furthermore, care duties might withhold mothers from reaching their full productive potential, reflected by their lower labour market incomes.

Third, this study attempts to provide a cross-country perspective. We expect that the effect of the family trajectory on later life income may differ across countries. Relevant macro-level aspects are family policies, cultural attitudes and their combination. If a country offers generous family policies

we would expect higher maternal incomes, because such a supportive environment enables women (and men) to combine care and labour market investments, rather than specialize in one of them. However, Budig et al. (2012) found some evidence that cultural attitudes moderate this effect. Parental leave and public childcare are associated with higher earnings for mothers when cultural support for maternal employment is high. Yet where cultural attitudes support the male breadwinner-female caregiver model, the effect of parental leave and public childcare is not clear-cut. We contribute to the literature by assessing this relation in a large number of countries, and focus at cross-country differences in later life outcomes.

Research design

Data

Our study uses data from the first wave of the Generations and Gender Programme (GGP). These surveys contain retrospective fertility and partnership histories as well as information about current employment and income. The data were collected around 2005. We restrict our sample to women aged 50 to 65. Our total analytical sample consists of 21.748 women from 14 European countries which are Belgium, Bulgaria, Czech Republic, Estonia, France, Georgia, Germany, Hungary, Lithuania, Netherlands, Norway, Poland, Romania and Sweden.

Dependent measure

The outcome of interest is personal net labour market income. We created two variables; a dummy indicating whether a woman is employed or not, and among employed women we calculated the labour market income percentile rank per country.

In all countries income information was collected with survey questions, except Sweden and Norway where register data was used. The questionnaire asked what the average monthly net earnings from a job were during the last 12 months. By multiplying we estimated the yearly net labour market income. If respondents were not able or willing to indicate an exact amount, they were asked in which band (range) their income lays. We replaced such indicated bands by the median of the income of respondents who did indicate an exact value within that band. We used the median rather than the mean, because income is not normally distributed within bands but skewed positively.

In five countries the GGP did not provide monthly income information, but the net total annual income received. This measure also includes other income sources than earnings from a job, such as social benefits. In order to minimize measurement difference we made two adjustments. Using additional survey questions we first checked whether respondents were employed. These questions regard current activity status (employed or not), working hours per week and income payment types received. Second, for employed respondents only we calculated the percentile rank of their income among women aged 50-65 within their country, ranging from 0 to 100. This improved comparability between the two income measures (total vs. labour market income). We assume that the lion's share of the total income of women with earnings from a job depends on their earnings. Therefore we expect that the ranking of the two income definitions is roughly the same. However we do realize that especially at the bottom of the income distribution, the ordering between the net labour market income and total net income could be different due to social benefits. The first robustness tests were promising; the pooled models with and without the five 'alternatively' measured countries showed similar results. However, we plan to do more extensive robustness tests in the nearby future.

Using percentile ranks has some additional advantages. Countries used different currencies in their income questions. By standardizing units of measurement we make country results easier to

compare. Also results are easier to interpret since we have a scale from 0 to 100. For instance a percentile rank of 80 is clearly a relatively high income, while 20 is relatively low.

We imputed missing labour market incomes. Percentage missings ranged from 2.49% in France to 24.83% in Czech Republic. In the two countries using register data the percentage of missings was nearly zero. In the multiple imputation model a large selection of control variables was used covering personal information (gender, age), educational level, family characteristics (e.g. young children), house characteristics (e.g. house size), health information, possessions, job characteristics (e.g. work hours per week) and financial indicators (e.g. 'difficulty to meet ends').

Independent measure

The main independent variable of interest is a woman's family trajectory. Family trajectory is measured as a sequence of monthly intervals from age 18 to 50. This implies that we created a sequence of 384 chronological states for every woman in the dataset. We specified eight possible states based on a combination of the age of the youngest child and partnership status: (1) No child, no partner; (2) No child with partner; (3) Youngest child 0-3, no partner; (4) Youngest child 0-3 with partner; (5) youngest child 4-11, no partner; (6) youngest child 4-11 with partner; (7) youngest child 12+, no partner; (8) youngest child 12+ with partner. We consider the youngest child only because it is known that the care burden is larger for younger children. The states "with partner" includes both cohabiting and married partnerships.

Controls

In addition to family trajectory we included education, age, working hours and country as control variables to our models. Education is measured with a dummy for being higher educated based on ISCED level (1 = ISCED 5 or higher; 0 = ISCED 4 or lower). Age is measured in months at the time of interview. Working hours are the typical number of hours worked per week. Last we added country as dummy in the pooled models, to adjust for country composition differences.

Analytical strategy

Our analytical approach consisted of two parts. First we created a typology of family trajectories using sequence cluster analysis. Second we used this typology in regression models to assess whether it explains variance in labour market incomes of women.

The software we used to perform the sequence cluster analysis is the TraMineR package in R (Gabadinho, Ritschard, Müller, & Studer, 2011). First we calculated the distance matrix which includes the distance between all possible pairs of individual sequences, i.e. a matrix of 21,748 by 21,748. We used Optimal Matching with a constant substitution matrix of 1 and indel cost of 1 (Studer & Ritschard, 2014). We plan to experiment with other distance measures in future analyses.

Next we used the distance matrix to do a Ward's cluster analysis. This means that sequences with the smallest distance to each other were clustered together. To determine the most appropriate number of clusters, we considered several cluster cut-off criteria, including the average silhouette Widths (ASW) and point biserial correlation (PBC) (Studer, 2013), and theoretical considerations.

The next step in our analytical strategy was assessing the variance in labour market incomes explained by the family trajectory typology across countries. However we only observe earnings among employed women. Yet employed women are not a random subgroup of the population of women. Therefore, estimating the predictors of women's incomes from the subpopulation of employed women might introduce bias. To control for this potential bias, we added a Heckman sample selection correction to our model. This is a two-step approach, including a selection equation that predicts the likelihood of employment and the regression equation of our interest. We added

two extra predictor variables to the selection equation, namely presence of an under aged child in the household and presence of a partner. However, the results we found using the Heckman approach were similar in terms of effect size and significance to a simpler two-part model, i.e. separate logistic and regression model. Therefore at this point we report the results of this simpler model.

First results

Typology

Figure 1 (Appendix) shows the sequence index plots of the eight family trajectory clusters we found. In these graphs every horizontal line represents the family trajectory of a woman in the data. Every colour represents a state. We labelled each cluster based on its characteristics. First, we identified four types of 'traditional' family clusters. Women in these four clusters went through the same sequence: they spent some time on their own, they met a partner, had one or more children, and stayed together with their partner. The difference between the four clusters was the timing of these events. Women in the "late family" cluster had their first child later than women in the "early family" cluster, while women in the "mid family" were in between. The fourth type is called "stretched family" because women in this cluster had many children or a large time gap between births. Typically, they had their first child at young age and their last child at older age and thus, spent much of their adult life with the rearing of young children.

The other four family trajectory types were less 'traditional'. We identified two clusters of childless women, who either spent the majority of their life with a partner ("No child with partner") or without ("No child, no partner"). The other two clusters comprise women who experienced single motherhood. Women in the "dissolution" cluster typically started their trajectory as a 'traditional' family; they lived with a partner and had a child. However, after union dissolution they continued living with a child and without a partner. Lastly, women in the "single mother" category spent the majority of their life without a coresiding partner and raised their child(ren) on their own.

Pooled model

Our first result is that we found some evidence that the family trajectory typology is related to both the odds of being employed and the labour market income of women at later stages of their life course. Table 1 and 2 (Appendix) show respectively the results of the logistic regression and linear regression. The logistic model has the binary employment variable as dependent: whether women are employed (1) or not (0). Next the regression model assessed the percentile rank of labour market incomes among women that have a job. In both models we found a significant effect of family trajectory. We used a Wald's test that assessed the effect of all family trajectory dummies at once. The effect of family trajectory remained significant after controlling for education, age and country. In the regression model we also controlled for the number of hours women are working.

Looking at differences between clusters we found most interesting results in the linear regression model. We present the predicted income ranks by family cluster in figure 2 (Appendix). This graph shows the marginal effects controlled for age, education, working hours and country. In general we can see that women who experienced 'traditional' family trajectories, i.e. spent the majority of their adult life with children and a partner, on average have lower incomes than women in less traditional trajectories, i.e. who are childless and/or without a partner. This divide suggests that women with children and a partner tend to withdraw from the labour market to specialize in housework which may imply that they never fully re-integrate into the labour market – leading to lower earnings still much later in life. Therefore our second finding is that being a mother not always gives an income

penalty (i.e. the "motherhood penalty" found in previous literature), rather being a mother *and having a partner* 'penalizes' women in terms of income.

This is in line with our finding that women who experienced the "stretched family" trajectory have the lowest incomes; in fact their earnings are significantly lower than all other clusters. We could say that women in this trajectory experienced the most extreme version of the 'traditional' family trajectory: they spent lots of time with young children in combination with having a partner.

Also we can see within the 'traditional' family clusters that the earlier a mother gets her first child, the lower her earnings in later life. "Early family" has lowest earnings of the family clusters, "late family" is highest and "mid family" is in between. This finding corresponds to the well-known effect of mother's age at first birth on income in the existing literature.

At the other end of the spectrum it seems that the absence of a partner, rather than the absence of children is associated with higher earnings later in life. The two top earning clusters spent the majority of their lives without a cohabiting partner. This is both the case for childless women ("No child no partner") and mothers ("Single mother"). These two clusters are both significantly higher than all four 'traditional' family clusters. It suggests that women without a partner simply do not have the option to specialize in housework, attaching them strongly to the labour market throughout their life, whether they have children or not.

Country differences

Our third main result is that we found country differences in the effect of family trajectory on later life income. In this extended abstract we present findings on two countries representing extreme cases in terms of inequality between the family trajectory clusters: the Netherlands and Czech Republic. Figure 3 and 4 (Appendix) are the graphs showing the predicted income rank per family trajectory cluster for these two countries. In the Netherlands we find a similar pattern as in the overall pooled model (including all 14 countries): a clear distinction between the 'traditional' family clusters and the less 'traditional' clusters. On the other hand, in Czech Republic we did not find such a clear distinction, rather we did not find any effect of family trajectory on later life earnings at all. This suggests that the Czech context in some way supported or motivated women with a child and partner to stick to the labour market and/or that Czech women had the opportunity to re-integrate into the labour market in later life.

Our future analyses will focus on these cross-national differences. We will try to explain them by both cultural attitudes and family policies. Also we aim to include more countries in the analysis. We are currently working on the harmonization of SHARELIFE data, which will expand our GGP sample by another seven countries. This is an important step for our project, because it enables us to include Southern European countries as well.

Literature

Budig, Michelle J., & England, Paula. (2001). The Wage Penalty for Motherhood. *American Sociological Review, 66*(2), 204-225.

Budig, Michelle J., Misra, Joya, & Boeckmann, Irene. (2012). The Motherhood Penalty in Cross-National Perspective: The Importance of Work-Family Policies and Cultural Attitudes. *Social Politics, 19*(2), 163-193.

Cornwell, Benjamin. (2015). *Social Sequence Analysis*. Cambridge (UK): Cambridge University Press.

Gabadinho, Alexis, Ritschard, Gilbert, Müller, Nicolas S., & Studer, Matthias. (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software, 40*(4), 1-37.

Goldin, Claudia. (2006). The Quiet Revolution that Transformed Women's Employment, Education, and Family. *American Economic Review, 96*(2), 1-21.

Gough, Margaret, & Noonan, Mary. (2013). A Review of the Motherhood Wage Penalty in the United States. *Sociology Compass, 7*(4), 328-342.

Harkness, Susan, & Waldfogel, Jane. (2003). The Family Gap in Pay: Evidence from Seven Industrialized Countries. *Research in Labor Economics, 22*, 369-414.

Miller, Amalia R. (2011). The effects of motherhood timing on career path. *Journal of Population Economics, 24*, 1071-1100.

Pienta, Amy. (1999). Early Childbearing Patterns and Women's Labor Force Behavior in Later Life. *Journal of Women & Ageing, 11*(1), 69-84.

Studer, Matthias. (2013). WeightedCluster library manual: A practical guide to creating typologies of trajectories in the social sciences with R (Vol. Lives Working Paper). Geneva, Switserland: University of Geneva Institute for Demographic and Life Course Studies.

Studer, Matthias, & Ritschard, Gilbert. (2014). A comparative review of sequence dissimilarity measures (Vol. Lives Working Paper). Geneva, Switzerland: Swiss National Centre of Competence in Research.

Appendix

Figure 1. Sequence index plots of family trajectory clusters, women aged 50-65 from 14 European countries (GGP data)

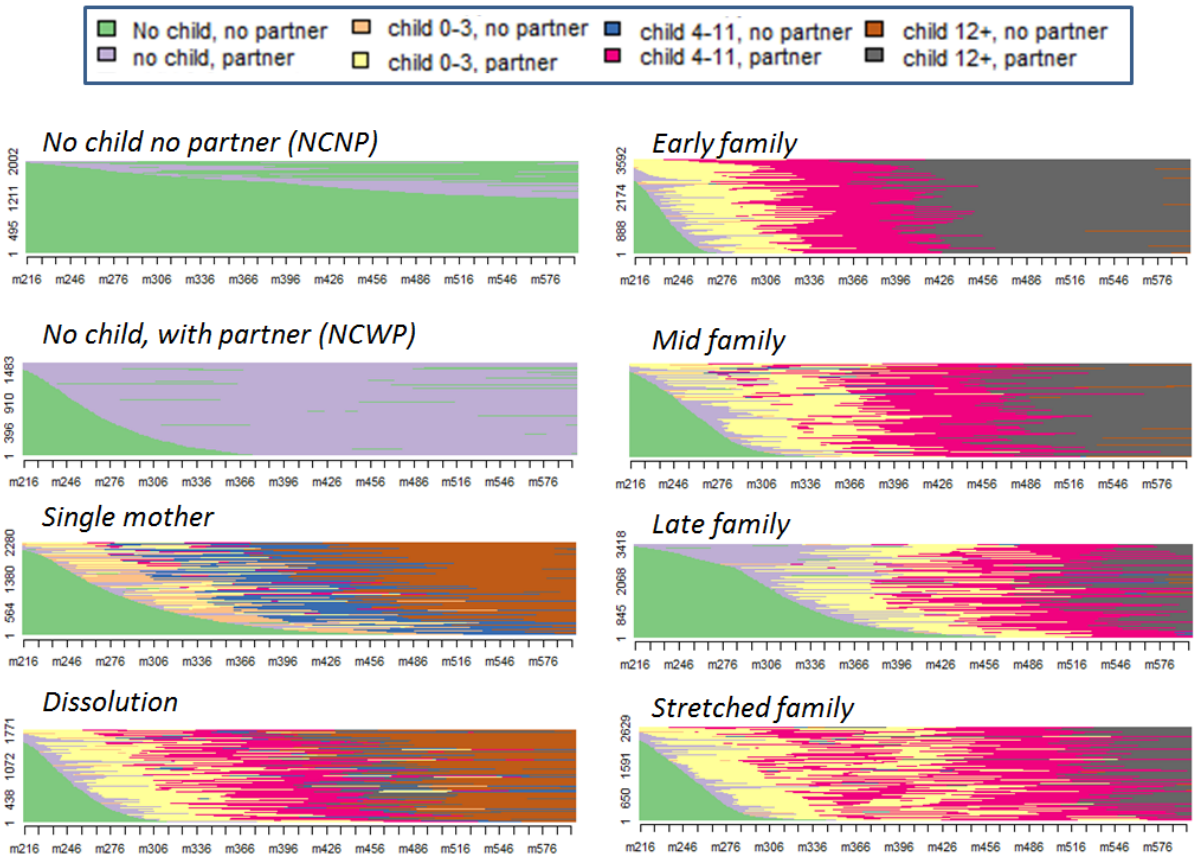


Table 1. Logistic regression model
Dependent: labour market participation
(1: earnings from a job,
0: no earnings from a job)

Late family	**1.36*** **
Mid family	**1.13***
Dissolution	1.11
No child, with partner	1.11
Single mother	1.09
No child, no partner	1.03
Early family	0.93
Higher Educated	**2.87*** **
Age	**0.98*** **
Belgium	**1.48*** **
Bulgaria	**1.50*** **
Czech Republic	**2.11*** **
Estonia	**3.30*** **
France	**2.27*** **
Georgia	**1.32*** **
Germany	**3.96*** **
Hungary	**4.14*** **
Lithuania	**2.90*** **
Netherlands	**1.47*** **
Norway	**13.41*** **
Romania	1.01
Sweden	**20.31*** **
Constant	**77384.14*** **

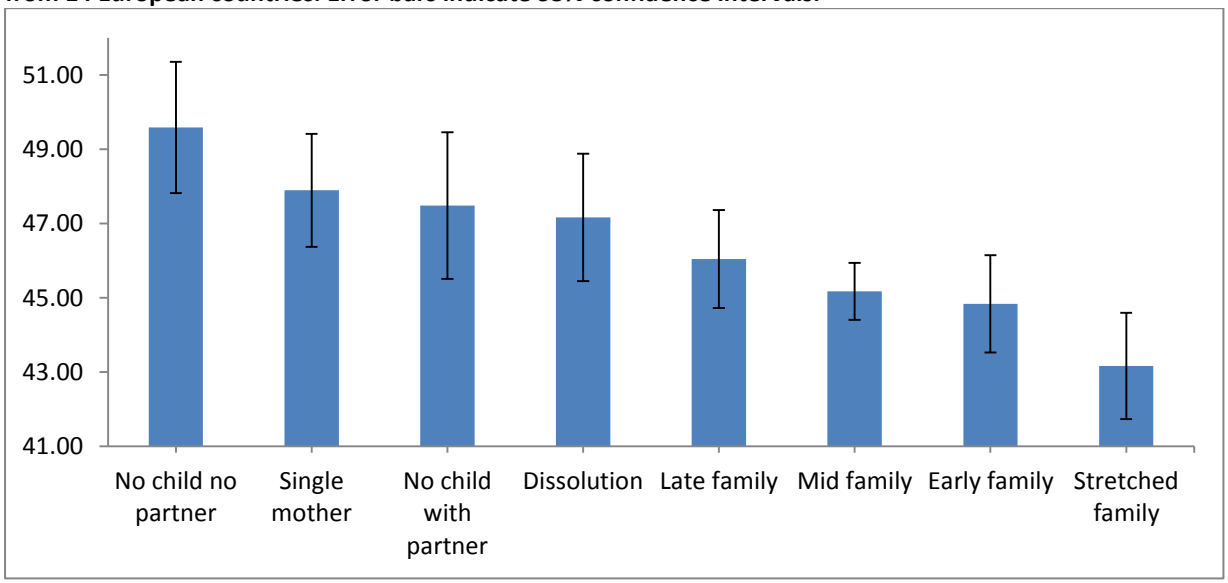
Note. Significant **odds ratios** bold (*** p < .001, ** p < .01, * p < .05, $^{+}$ p < .10), two-sided test. Reference categories are respectively “stretched family” and Poland. Wald's test all family trajectory clusters is significant (chi2(7)= 40.51, p< 0.001).

Table 2. Linear regression model
Dependent: Income percentile rank

No child, no partner	**6.19*** **
Single mother	**4.76*** **
No child, with partner	**4.08*** **
Dissolution	**3.92*** **
Late family	**2.72***
Mid family	**1.99***
Early family	1.45
Higher educated	**19.10*** **
Age	**0.02*** **
Working hours	**0.64*** **
Belgium	1.26
Bulgaria	**7.16*** **
Czech Republic	**2.48***
Estonia	**1.90**$^{+}$
France	**10.43*** **
Georgia	1.31
Germany	**9.35*** **
Hungary	**17.69*** **
Lithuania	0.29
Netherlands	**10.21*** **
Norway	**10.69*** **
Romania	-0.59
Sweden	**6.28*** **
Constant	-5.48

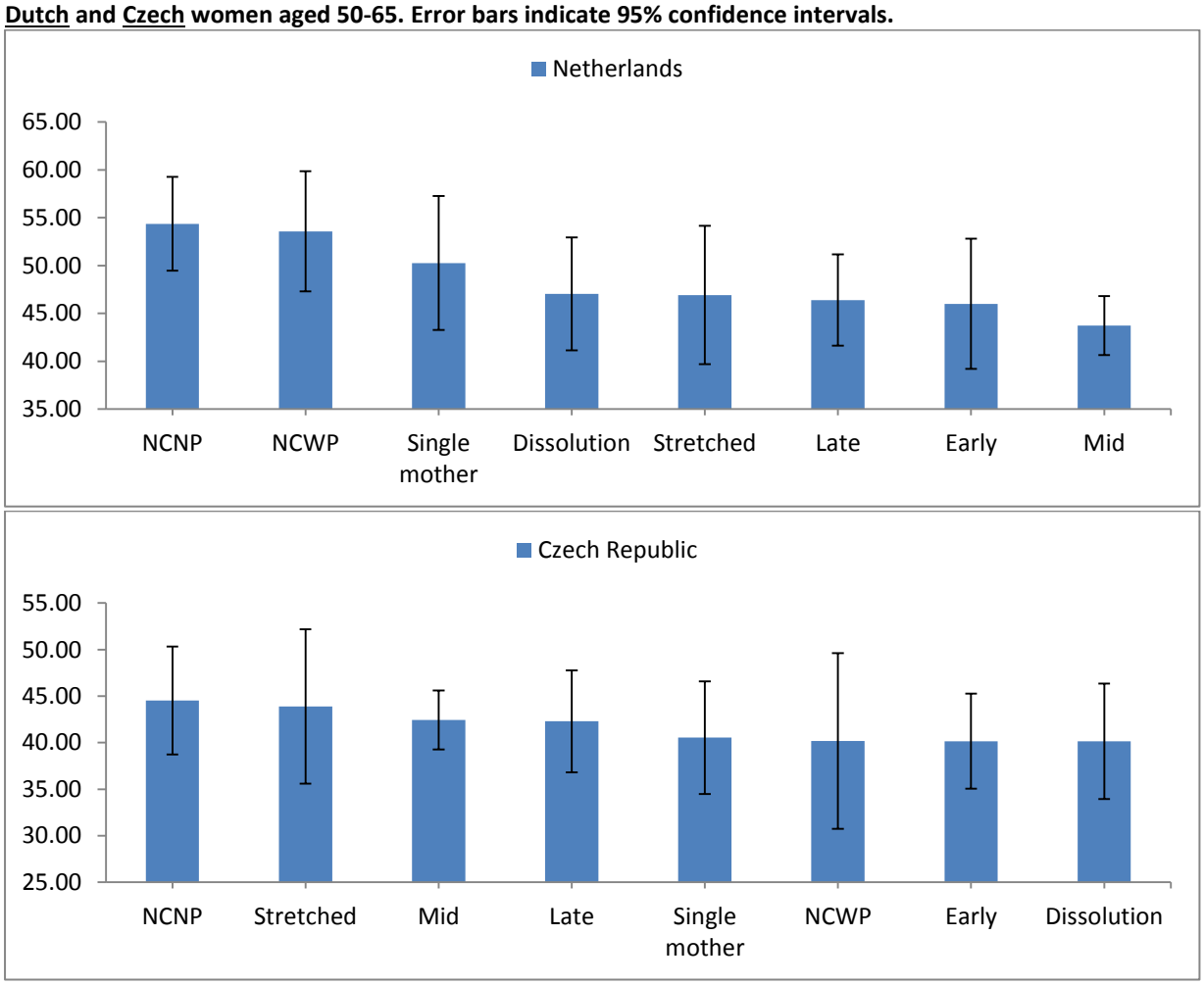
Note. Significant coefficients bold (*** p < .001, ** p < .01, * p < .05, $^{+}$ p < .10), two-sided test. Reference categories are respectively “stretched family” and Poland. Wald's test all family trajectory clusters is significant (F(7, 471.5) = 5.99, p< 0.001).

Figure 2: Predicted income (percentile rank) according to family trajectory cluster of women aged 50-65 from 14 European countries. Error bars indicate 95% confidence intervals.



Note. Margins holding all other variables at their mean values. Result of linear regression model. The model also controlled for age, education, working hours and country. GGP data.

Figure 3 and 4: Predicted income (percentile rank) according to family trajectory cluster of respectively Dutch and Czech women aged 50-65. Error bars indicate 95% confidence intervals.



Note. Margins holding all other variables at their mean values. Result of linear regression model. The model also controlled for age, education, working hours and country. GGP data.