

Research Article

Open Access

Fulvia Pennoni* and Miki Nakai

A latent class analysis towards stability and changes in breadwinning patterns among coupled households

<https://doi.org/10.1515/demo-2019-0012>

Received December 18, 2018; accepted June 12, 2019

Abstract: A latent class model is proposed to examine couples' breadwinning typologies and explain the wage differentials according to the socio-demographic characteristics of the society with data collected through surveys. We derive an ordinal variable indicating the couple's income provision-role type and suppose the existence of an underlying discrete latent variable to model the effect of covariates. We use a two-step maximum likelihood inference conducted to account for concomitant variables, informative sampling scheme and missing responses. The weighted log-likelihood is maximised through the Expectation-Maximization algorithm and information criteria are used to develop the model selection. Predictions are made on the basis of the maximum posterior probabilities. Disposing of data collected in Japan over thirty years we compare couples' breadwinning patterns across time. We provide some evidence of the gender wage-gap and we show that it can be attributed to the fact that, especially in Japan, duties and responsibilities for the child care are supported exclusively by women.

Keywords: Akaike information criterion, Expectation-Maximization algorithm, gender inequality, household income composition, weighted log-likelihood

MSC: 62C12, 62D05, 62H12, 62P25, 91B40

1 Introduction

A clear division of paid and unpaid work along gender lines in households is found in most countries, but the recent trend in several advanced societies is away from the male breadwinner model towards a dual-earner family. However, despite the increase in the paid employment of married women in Japan as well as Western industrialized societies, the notion of gendered division of labour, that is men should be the family's primary economic provider and women should stay at home to focus mainly on domestic labour and care for family members, has been accepted as a tradition. While recently the number of households with wives entirely dependent on their spouses' income has dramatically declined, most women in dual-earner households still earn much less than their spouses and the households where wives earn equal to or more than their husbands are very few.

The increasing rate of labour force participation of women does not necessarily mean that economic inequality within couples is expected to have changed. Instead, dual-earner households does not liberate women from their traditional gender role, especially in the strong male breadwinner model like Japan. The Japanese tax and pension systems are legislated regarding the male breadwinner household as "standard

***Corresponding Author: Fulvia Pennoni:** Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Via Bicocca degli Arcimboldi, 8, 20126 Milano, E-mail: fulvia.pennoni@unimib.it

Miki Nakai: College of Social Sciences, Ritsumeikan University, 56-1 Tojiin Kitamachi, Kyoto 603-8577 Japan, E-mail: mnakai@ss.ritsumei.ac.jp

household". The Japanese government policies carried out over time have encouraged married women to stay at home and to be dependent on their husbands, to work part-time or in non-regular jobs and to have broken career patterns. Work and family reconciliation is difficult and this perpetuates women's dependence on a male breadwinner.

As gender inequalities in the division of labour within marriage/family are closely related to gender inequalities in other spheres of life, particularly in the labour market as well as political and other societal arenas, understanding what determines gendered intra-household inequality is essential for researchers who want also to understand other aspects of gender stratification as well as for politicians interested to adopt laws favouring an active women role also in the labour market. Many studies argued that women's economic dependency on men is an important attribute of stratification systems and essential force in the maintenance of gender inequality (see, among others [26]). Recent studies aim at exploring if the gender pay inequalities are due to the perception of both man and woman that they are fair (see [3]). However, as women's educational attainment has been rising, more and more women might be part of the dual-career couple and other types of non-traditional couples in contemporary Japan. Having said that, some people may still rely on and find value in the conventional gender roles as recently stressed by [27].

We aim to examine couples' breadwinning patterns such as, male-breadwinner or equal-income among Japanese couples according to a latent class model [14] and study how they have changed over the past four decades, during which we have seen an increase in the women's labour force participation and a rising of educational attainment. As well as, after identifying distinct latent classes of breadwinning types, we aim to explore the associations with their socio-economic statuses to detect the main determinants of this classification. In Japan, heterosexual marriage has been considered as the only way to form a family until recently and marriage must be officially registered at a local government. Lately there are some people who choose alternatives to marriage such as consensual unions and registered partnerships, but the way people live together has not significantly changed (1.1 percent lived in a consensual union in 2015). It is not only because married couples receive many legal and economic benefits not available to unmarried couples, but also because marriage has been necessary for women's financial survival, social interaction and personal well-being until recently [22].

Many studies concerns United States and Australia and investigate what differentiates dual-income couples from husband sole provider couples, for example [21] shows what differentiates equal-provider couples from male-breadwinner couples among dual income households and their changes over time. Few studies have been conducted in Europe, for example [30] investigate women as main earners by comparing European countries according to social policies. Previous research on this topic in Japan is scarce. Recently, [17] attempts to explore by means of a multinomial logit model the determinants of the factors that differentiate equal-provider couples in Japan, where wives share the breadwinning responsibility equally with their partners, from male-breadwinner couples among dual-income households. However, it focused only on dual-earner families and the respondents in sole-breadwinning families were omitted from the analysis. Furthermore, the breadwinning typology commonly used in those studies are usually operationalized based on a priori percentages of wife's income of the total household income. Therefore, breadwinning categories of "male breadwinner" or "equal" might arbitrary be constructed.

Our analysis is conducted by employing a latent class model and relying on the declared incomes of respondent and spouse to national surveys carried out along 40 years. We properly define an ordinal response able to disentangle the earning proportions derived by comparing two underlying continuous variables and we account for missing responses as well as for sampling probabilities. In this way, we get consistent estimates of the parameters of interest, reliable standard errors adjusting for possible failures of the sample in covering the target population. The latent class model has been considered for the analysis of data arising in different contexts. Within survey data it is a flexible tool to account for the heterogeneity among responses. The proposed model-based approach enables us to investigate the associations with the couple's characteristics through a suitable parameterization as well as to dispose of data driven couple's typologies (see, among others, [19] and [4]). Maximum likelihood estimation of the model parameters is well established and it is carried out through the Expectation-Maximization algorithm. However, even if missing observations arise frequently the use of weighted methods for the estimation of the parameters with missing responses has been recently

employed. Instead of performing case deletion we retain the missing responses under the missing-at-random assumption and we impute the missing values on the concomitant variables. In this work we extend the proposal of [20] by considering comparisons across waves.

The paper is organized as follows: in Section 2 we describe the surveys and the data on which the empirical analyses are carried out. In Section 3 we discuss the latent class model with sample weights and maximum likelihood estimation of the model parameters. In Section 4 we report and compare the results across waves. Finally, a discussion concludes the paper.

2 Survey data

Data are obtained from the past three decades of Japanese cross-sectional data of the Social Stratification and Social Mobility (SSM) surveys, and the last decade of the Japanese Stratification and Social Psychology (SSP) survey. The 1985 SSM survey consists of three surveys: a) for men (N=1,239, response rate is around 61 percent), b) for men (N=1,234, response rate is around 61 percent), and f) for women (N=1,474, response rate is around 68 percent). Each survey used different questionnaires, but some questions overlapped and it is possible to dispose of almost the same variables. The 1995 survey consists of two surveys: a) for men and women (1,248 men and 1,405 women, response rate is around 66 percent) and b) for men and women (1,242 men and 1,462 women, response rate is around 67 percent). Each survey used different questionnaires, but again, some questions overlapped and we combine the two surveys.¹ The respondents are interviewed and asked a wide range of questions such as respondents' socio-economic background, education, social consciousness, and the spouse's information when a respondent has a spouse. The surveys have been conducted with face-to-face interviews with a special focus on social stratification and inequality in contemporary Japan. They are based on selected national representative respondents through a multiple-stage sampling design. The respondents were aged between 20 and 69 for 1985, 1995, and 2005, and between 20 and 64 for the 2015 SSP survey. Therefore, we use the data collected from 2,473 men and 1,474 women in 1985, 2,490 men and 2,867 women in 1995, 2,660 men and 3,082 women in 2005, and finally 1,644 men and 1,931 women in 2015. The response rates are 63.3% in 1985, 66.4% in 1995 (response rate for men and women are 61.8% and 71.1%, respectively), 44.1% in 2005 (response rates for men and women are 40.9% and 47.2%, respectively), and 43.0% in 2015.² The 2005 and the 2015 survey data are provided with individual sampling weights, not available with the 1985 and 1995 waves. The weights are obtained with respect to the following auxiliary variables: gender, age group, and region of residence.

The couple's income provision-role type is obtained by the declared incomes of both survey respondent and spouse. Incomes include earned and investment incomes and all the other additional incomes such as pensions and dividend on stock shares. The measurement of couple's income provision-role type is based on who may be a dominant provider and the share of contribution of wife's income for the household income. The relative importance of this concept is also highlighted by [18] and [21]. Five ordinal categories of the key variable are considered: (1) husband sole provider, which consists of couples where only husband is employed; (2) husband provides majority, which consists of couples where husbands' earnings represent 60 percent or more of the sum of the spouses' income; (3) equal providers, which identifies couples where wife's earnings represent somewhere from 40 percent to 60 percent, meaning that each partner contributes between 40 and 60 percent of total household income; (4) wife provides majority, which consists of couples where wives' earnings represent 60 percent or more of the combined total income of the husband and wife; (5) wife sole provider, which consists of couples where only wife is employed.

Table 1 shows the weighted frequencies of the response variable for each year. We notice some changes of breadwinning patterns among married couples over the past four decades. The proportion of the households with husbands as sole provider has declined from 42.9% in 1985 to 22.9% in 2015. However, despite the continuing rise in Japanese women's participation in the economy, husbands until recently have been the sole or the primary breadwinner in 65% of the couples and equal-provider couples have been only 11.8% up to 2015. The highest shares of item nonresponse for household income and/or spouse income is observed in 2005.

Table 1: Observed and missing frequencies of the couple's income provision-role type (weighted with survey weights for years 2015 and 2005): (1) husband sole provider, (2) husband provides majority, (3) equal providers, (4) wife provides majority, (5) wife sole provider.

Years	Income provision-role categories (%)					
	1	2	3	4	5	Missing
2015	22.9	42.2	11.8	5.3	0.6	17.2
2005	33.6	28.1	6.9	1.6	1.0	28.7
1995	38.2	37.8	6.2	1.2	0.7	16.0
1985	42.9	36.7	6.7	1.2	0.1	12.4

Table 2: Classes and categories for the continuous and categorical covariates defined according to the survey wave.

Covariates	Categories									
	1	2	3	4	5	6	7	8	9	10
Preschool children	1									
Relative education	1	2								
Wife education	1	2	3	4						
Size of the place	1	2	3	4	5					
Number of children	0	1	2	≥ 3						
2015										
Husband age	≤ 34	(34, 39]	(39, 43]	(43, 46]	(46, 50]	(50, 54]	(54, 58]	(58, 61]	(61, 64]	> 64
Wife age	≤ 32	(32, 37]	(37, 40]	(40, 44]	(44, 47]	(47, 51]	(51, 55]	(55, 58]	(58, 61]	> 61
Husband income	≤ 175	(175, 275]	(275, 325]	(325, 375]	(375, 425]	(425, 500]	(500, 600]	(600, 700]	(700, 900]	> 900
2005										
Husband age	≤ 34	(34, 40]	(40, 45]	(45, 50]	(50, 54]	(54, 57]	(57, 61]	(61, 64]	(64, 68]	> 68
Wife age	≤ 32	(32, 37]	(37, 42]	(42, 47]	(47, 51]	(51, 55]	(55, 58]	(58, 61]	(61, 65]	> 65
Husband income	≤ 175	(175, 225]	(225, 275]	(275, 375]	(375, 425]	(425, 500]	(500, 600]	(600, 700]	(700, 900]	> 900
1995										
Husband age	≤ 33	(33, 38]	(38, 42]	(42, 45]	(45, 48]	(48, 52]	(52, 56]	(56, 60]	(60, 65]	> 65
Wife age	≤ 31	(31, 35]	(35, 40]	(40, 43]	(43, 46]	(46, 49]	(49, 53]	(53, 57]	(57, 62]	> 62
Husband income	≤ 150	(150, 250]	(250, 350]	(350, 400]	(400, 500]	(500, 600]	(600, 700]	(700, 800]	(800, 900]	> 900
1985										
Husband age	≤ 32	(32, 36]	(36, 39]	(39, 42]	(42, 46]	(46, 49]	(49, 53]	(53, 57]	(57, 62]	> 62
Wife age	≤ 29	(29, 33]	(33, 36]	(36, 39]	(39, 44]	(44, 48]	(48, 51]	(51, 56]	(56, 61]	> 61
Husband income	≤ 130	(130, 190]	(190, 250]	(250, 310]	(310, 350]	(350, 430]	(430, 480]	(480, 550]	(550, 650]	> 650

The presence of a preschooler is coded as a binary variable with respect to children's age 0-6. Couples' relative education is measured by two categories: 1 husband has higher education than wife, 2 wife has higher education than husband. Wife's education: 1 less than high school, 2 high school, 3 two-years college degree, 4 four-year college degree or higher. Size of the place of living: 1 major cities; 2 > 200,000; 3 (100,000, 200,000]; 4 ≤ 100,000; 5 small towns and villages. Husband's income is measured in ten thousand yen a year.

Concomitant variables are selected according to previous research among the available survey measurements on the basis of their possible influence on the division of labour within households. They are shown in Table 2 according with the employed categorization. Educational level is considered as a proxy of human resources and we expect a relative importance of this variable on married couples' choice of breadwinning patterns. Couples' relative education-level between spouses has been considered to connect to values and culture which may influence division of labour within household. Whether the marriage is homogamous, hypergamous or hypogamous may be associated with patriarchal culture. More patriarchal households, which may be associated with female hypergamous couples, may prefer traditional breadwinning type, as well as their relative power within the marriage, see, among others [24]. In addition, when available, we consider the variable related to locality of the respondents. The continuous covariates such as wife's and husband's age, husband income, are considered according to the quantiles illustrated in Table 2 and included as dummy variables to adjust distributions throughout the thirty-year period. We define age and income bands since their

effects on the couple's provision-role type might not be linear and they can be due to the impact of various life events and career stages. Descriptive statistics for these covariates are presented in Table 3.

Table 3: Descriptive statistics for the categorical covariates by survey year (weighted frequencies with survey weights for years 2015 and 2005). The categories are defined in Table 2.

Covariates (%)	Categories									
	1	2	3	4	5	6	7	8	9	10
2015										
<i>Husband age</i>	13.6	12.1	11.9	9.1	10.5	10.8	10.6	8.6	8.2	5.2
<i>Wife age</i>	12.6	12.1	10.6	11.5	9.4	10.9	9.9	7.5	8.6	6.9
<i>Husband income</i>	10.1	11.3	9.7	8.1	9.6	13.3	10.1	8.1	10.1	9.5
<i>Relative education</i>	32.6	16.5								
<i>Wife education</i>	4.1	56.5	24.4	18.3						
<i>Preschool children</i>	25.2									
<i>Number of children</i>	87.4	24.3	43.6	18.0						
<i>Size of the place of living</i>	28.9	24.8	17.2	20.5	8.7					
2005										
<i>Husband age</i>	15.0	13.3	10.3	10.7	9.0	9.6	10.1	7.0	9.1	5.9
<i>Wife age</i>	14.8	12.2	10.8	12.0	9.1	10.8	9.3	6.5	8.6	6.0
<i>Husband income</i>	10.9	6.9	7.8	16.3	11.4	12.1	11.1	6.8	8.3	8.5
<i>Relative education</i>	28.3	14.3								
<i>Wife education</i>	12.9	57.0	16.6	11.3						
<i>Preschool children</i>	20.3									
<i>Number of children</i>	90.1	16.6	48.3	24.6						
<i>Size of the place of living</i>	20.0	22.3	18.7	25.3	13.8					
1995										
<i>Husband age</i>	11.2	9.3	10.0	10.0	10.2	10.3	9.1	10.3	10.7	8.9
<i>Wife age</i>	11.8	8.4	12.5	9.0	10.3	8.6	9.7	10.0	11.1	8.6
<i>Husband income</i>	4.9	7.2	13.5	14.1	14.5	12.3	10.1	7.0	6.4	10.0
<i>Relative education</i>	29.1	12.6								
<i>Wife education</i>	23.3	55.8	13.2	7.2						
<i>Preschool children</i>	18.3									
<i>Number of children</i>	92.3	15.0	51.6	25.7						
1985										
<i>Husband age</i>	11.1	11.0	9.4	9.1	11.4	8.8	10.8	9.7	10.3	8.6
<i>Wife age</i>	11.3	9.7	10.5	9.3	11.4	8.8	9.3	11.9	8.9	8.7
<i>Husband income</i>	7.8	8.1	13.1	10.5	9.5	11.1	6.7	12.6	8.5	12.1
<i>Relative education</i>	28.0	12.3								
<i>Wife education</i>	34.3	52.3	7.4	5.0						
<i>Number of children</i>	98.0	6.0	16.4	11.3						

3 Latent class model

The latent class model belongs to the class of finite mixture models [16]. It includes an unobserved random variable with random parameters. The advantages of using latent variable models in social research are il-

illustrated among others, by [25] and [5] Chapters 1 and 2. With reference to a random unit drawn from the population of interest let Y_{ij} be the derived response variable observed for individual i , $i = 1, \dots, n$ on the j , $j = 1, \dots, r$ ordered category. Its values are obtained by comparing two continuous variables therefore we assume that it is derived by an underlying unobserved latent variable denoted as U_i for i , $i = 1, \dots, n$. This latent variable has a discrete distribution left unspecified with k support points having finite discrete values. In this way a semi-parametric model results [15]. As in the standard latent class model formulation the observed responses are assumed independent one another conditionally to the latent variable. The latter is named local independence assumption.

The parameters are referred to the measurement and to the latent model. Those of the measurement model are the conditional probabilities of the response variable given the latent variable

$$\phi_{j|u} = p(Y_j = y|U = u), \quad u = 1, \dots, k, \quad j = 1, \dots, r.$$

The latent model is defined by the probability of belonging to each latent class. Considering a set of covariates arranged in the vector \mathbf{X} , where \mathbf{x} is the corresponding realization, they enter into the model through the following parameterization

$$\log \frac{p(U = u|\mathbf{X} = \mathbf{x})}{p(U = 1|\mathbf{X} = \mathbf{x})} = \log \frac{\pi_{u|\mathbf{x}}}{\pi_{1|\mathbf{x}}} = \beta_{0u} + \mathbf{x}'\boldsymbol{\beta}_u, \quad u = 2, \dots, k, \quad (1)$$

where β_{0u} is an intercept specific of each latent class and $\boldsymbol{\beta}'_u$ is the vector of regression parameters. Note that, in the parameterization above, the first latent class is taken as a reference category and has a special role when the latent classes are ordered according to response categories. It allows us to compare the patterns with respect to the first latent class and this choice makes the interpretation of the results easier. Other link functions can be chosen such as those presented in [1]. Concerning the missing responses we assume that the response variable is independent from the missing responses given the observed covariates \mathbf{X} and the latent variable U . This is a less stringent assumption with respect to the missing at random assumption.

To account for individual sampling weights denoted as w_i , $i = 1, \dots, n$ as those provided within the surveys and related to the inverse of the selection probability we estimate the model through a weighted log-likelihood. The latter is determined given a sample of n independent individuals for which we observe the responses y_1, \dots, y_n as

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n w_i \ell_i(\boldsymbol{\theta}), \quad \ell_i(\boldsymbol{\theta}) = \log p(y_1, \dots, y_n|\mathbf{x}_i),$$

where $\boldsymbol{\theta}$ denotes the overall vector of free parameters arranged in a suitable way and \mathbf{x}_i denotes the vector of the observed covariates configuration for individual i , $i = 1, \dots, n$. The above quantity is maximized through the Expectation-Maximization (EM) algorithm [7] and [10] representing the main tool to estimate the models with latent variables. It is based on the *complete data log-likelihood*, for more details see [5].

We apply a two-step approach to avoid bias for the estimated coefficients referred to the covariates. First, we estimate the model with the missing responses and sampling weights and we perform a model selection strategy to choose the suitable number of latent classes. This consists to estimate the model with each k several times and to check for local maxima. We rely on the Akaike Information Criterion (AIC, [2]) and on the Bayesian Information Criterion (BIC, [23]) to guide the selection. They are measures of the relative goodness of fit of a model, accounting simultaneously for its accuracy and complexity. The AIC is defined on the basis of the following index

$$\text{AIC}_k = -2 \hat{\ell}_k(\boldsymbol{\theta}) + 2\#\text{par},$$

where $\hat{\ell}_k$ denotes the maximum of the log-likelihood of the model with k latent classes and $\#\text{par}$ denotes the number of free parameters in the model. The BIC differs from the previous by the penalized term that also includes the sample size through the term $\log(n)$. Then, by fixing the parameters of the measurement model we estimate the remaining parameters by adding the full set of covariates as in Equation (1). Standard errors for the parameter estimates are obtained according to the observed information matrix computed through numerical methods.

By updating the probabilities of belonging to each latent class through the Bayes' theorem [8] we compute the estimated posterior probability to be assigned to each latent class determined as

$$\hat{q}_u = \frac{\prod_{j=1}^r \hat{\phi}_{j|u} \hat{\tau}_{u|x}}{\sum_{u=1}^k \prod_{j=1}^r \hat{\phi}_{j|u} \hat{\tau}_{u|x}}, \quad u = 1, \dots, k, \quad j = 1, \dots, r. \quad (2)$$

In this way, we dispose of a suitable allocation rule for each respondents and they are profiled according to the latent class composition.

4 Results

In the following we show the results of the models estimated on the data collected in 2015, 2005, 1995 and 1985. First, we present the results for the model selection and the estimates of the parameters of the measurement model. Then, we show the estimated parameters of the latent model and we compare results across waves. Lastly, we discuss the predicted latent class membership.

4.1 Latent class model

We performed the imputation for the missing values of age and husband's income by using the weighted mean matching method, as explained in the previous section, by considering sampling weights and covariates as predictors [29]. In this way, it is possible to allocate to a latent class also respondents with missing provision-role behavior. Under the missing-at-random assumption [28] show that multiple imputation methods perform equally well as the latent class imputation, see also [31] for a comparison on the performance of the imputation strategies with an ordinal response variable.

The number of latent classes is selected for each wave by estimating the models for a number of classes ranging from 1 to 4 with different initialisations of the EM algorithm. The model is estimated by adapting the functions of the R package `LMest` [6]. The highest maximum log-likelihood at convergence, the number of free parameters and the corresponding AIC and BIC values are reported in Table 4. The selected number of latent classes is identical across waves. Table 5 shows the estimated conditional probabilities of the response for each wave.³

The latent classes are characterized by qualitatively different couple's typologies and the results allow us to label the first latent class as that of *traditional couples* (U_T) and the other one as that of *new couples* (U_N) since clearly the first consists exclusively of couples where the husband is the unique provider and the second is a more heterogeneous class with prevalence of couples where the husband is provider of the majority. The latent class of traditional couples (U_T) is characterized by a high degree of gender role specialization, strong gender based division of work where the husband specialize in market-work and wife in domestic work and caregiver. In 2015 we notice that (U_N) is composed of 58% of couples where the husband provides majority, 16% of couples where each component is equal provider and 7% of couples where the wife provides the majority. The latter percentage is higher than that estimated for the latent class model with data from the first wave.

Table 4: Maximum log-likelihood, number of parameters, values of the Akaike and Bayesian information criteria of the latent class models with two latent classes selected for the data collected at each wave.

Years	$\hat{\ell}$	#par	AIC	BIC
2015	-2,456.7	11	4,935.4	4,999.4
2005	-3,421.2	11	6,864.4	6,934.6
1995	-3,653.5	11	7,329.1	7,397.3
1985	-2,737.5	11	5,497.1	5,562.4

Table 5: Estimated conditional probabilities of the responses given the latent classes ($\hat{\phi}_{j|w}$) under the selected model for each year (U_T for latent class labeled as traditional couples and U_N for new couples, categories are defined in Table 1).

Income provision-role categories		1	2	3	4	5
Latent classes						
2015	U_T	1.000	0.000	0.000	0.000	0.000
	U_N	0.179	0.578	0.161	0.074	0.008
2005	U_T	1.000	0.000	0.000	0.000	0.000
	U_N	0.214	0.586	0.145	0.033	0.022
1995	U_T	1.000	0.000	0.000	0.000	0.000
	U_N	0.312	0.568	0.093	0.018	0.010
1985	U_T	1.000	0.000	0.000	0.000	0.000
	U_N	0.359	0.526	0.096	0.018	0.002

4.2 Latent class model with covariates

Concomitant variables enter into the model as in Equation (2). The maximum log-likelihood of the four models are: $\hat{\ell}_{15} = -2,366$ and $\hat{\ell}_{05} = -3,293$ with 51 free parameters, respectively, $\hat{\ell}_{95} = -3,456$ with 47 free parameters and $\hat{\ell}_{85} = -2,659$ with 46 free parameters. Table 6 reports the estimates of the logit regression parameters affecting the transition from the latent class of traditional couples (U_T) to the new couples (U_N) for each wave. The estimated intercepts are positive and significant indicating a general tendency towards the latent class of new couples (U_N) especially in 2015.

The estimated coefficients can be compared across waves. We observe that having preschool children shows the highest estimated coefficient whose negative sign indicates that couples with babies tend to belong to the cluster of traditional couples (U_T) with respect to new couples (U_N). In fact, the estimated odds ratio of belonging to (U_N) for the couples with preschool children is $\exp(\hat{\beta}) = 0.08$ the estimated odds for those not having preschool children. The negative effect of having preschool children decreases in 2005 with respect to 1995, but in 2015, increases again, meaning that having preschool children implies to belong to (U_T). This result might suggest continued, or even strengthened, gender differentiation in contemporary Japanese society. Although mothers in many countries experience strong normative pressure to fulfill the norms of motherhood, Japanese mothers hold extremely high standards of mothering that lead them to perceive it as gruelling and difficult to manage successfully [11].

As for the wife's education, a significant effect of women's tertiary education on the hypothesized positive effect is found. Compared with high school graduates, women who have completed four years of college or more are expected more likely to belong to the class (U_N), since they may have less traditional gender-role attitudes. However, significant effects of women's tertiary education are not found in other periods than 2005. Interestingly, we observe a negative effect of the lower level of education to transit to new couples (U_N) in 2015. Having said that, the pattern of the effects of couple's relative educational level seems constant and significant through the four waves: hypergamy, meaning that husband has higher education than his spouse, implicates less likely to be new couple, whereas hypogamy, meaning that wife has higher education than her spouse, implicates more likely to be new couple. The results suggest that hypogamy is linked to women's greater bargaining power within households. In addition, it contributes to more gender-egalitarian attitudes inside and outside the house.

The effect of the husband age is positive for the transition to (U_N) only in 1985 and it is restricted to middle ages (from 39 to 53). On the other hand, in 1995 wife ages from 46 and above have a great impact on the transition to the new type of family (U_N). In 2005, living in small towns contributes to this transition. The results can be understood in the light of the situation in the Japanese society: the labour force participation rate of women is generally lower in urban areas than in smaller or rural cities. In large cities, commuting time to work tends to be longer than in rural areas. Moreover, in more urbanised regions, working hours are longer than in rural areas. Working parents who have young children often rely on grandparents for childcare,

Table 6: Estimates of the regression parameters ($\hat{\beta}$) for each wave and percentages of predicted couples in U_T (categories are defined in Table 2).

		1985	1995	2005	2015
	$\hat{\beta}_0$	1.129**	2.506**	1.377*	3.777**
<i>Husband age</i>	1	-0.081	-0.180	0.106	0.349
	3	0.147	-0.324	0.846**	0.613
	4	1.324**	-0.373	0.441	0.776
	5	1.234**	-0.668	0.434	-0.599
	6	1.223**	1.018	-0.314	-0.331
	7	1.730**	-2.512**	-0.630	0.599
	8	0.612	-2.571**	-0.480	-1.299
	9	0.030	-3.784**	-0.305	-1.412
	10	-0.124	-4.280 [†]	0.167	0.428
	<i>Wife age</i>	1	-0.709 [†]	-0.462	-0.451
3		0.414	0.452	-0.272	-0.437
4		0.238	0.729	0.334	-0.877
5		0.413	2.723	0.133	-0.498
6		0.198	2.738**	0.221	0.271
7		-0.508	2.945**	-0.176	-0.411
8		-0.338	2.133**	-0.819	1.284
9		-0.920**	1.998**	-1.077	-1.045
10		-0.995**	1.841**	-1.473*	-1.425
<i>Wife education</i>		1	0.550**	0.309	0.034
	3	-0.399	-0.011	-0.131	0.115
	4	-0.407	0.515	0.050*	0.471
<i>Relative education</i>	1	-0.459**	-0.476**	-0.609**	0.328
	2	0.785 [†]	1.051**	0.323	0.887*
<i>Preschool children</i>	1	-	-2.114**	-0.970**	-2.485**
<i>Number of children</i>	1	0.639	-0.621	0.133	-0.464
	2	-0.216	-0.559**	-0.331**	-0.479*
	4	0.444	0.707**	0.149	0.149
<i>Size of the place of living</i>	1	-	-	0.048	-0.164
	3	-	-	0.129	-0.379
	4	-	-	0.223	0.236
	5	-	-	0.438*	0.853
<i>Husband income</i>	1	-0.019	0.256	0.380	0.772
	2	0.877	0.300	0.110	0.952
	4	0.291	-0.253	0.184	0.333
	5	0.255	-0.061	0.131	0.072
	6	-0.241	-0.673**	0.061	-0.547
	7	0.030	-0.434**	-0.272	-0.427
	8	-0.684*	-1.038	-0.657**	-1.479**
	9	-0.810*	-0.078	-0.671**	-1.221**
	10	-0.975**	-0.538	-0.632**	-1.265**
	<i>Couples predicted in U_T</i>		21.5%	22.2%	31.0%

Significance at 10%([†]), 5%(*), 1%(**). The missing category is taken as baseline.

but it is difficult for working mothers to ask parents to look after their children in big cities because many grandparents live in another area of the city or they can still work.

Concerning husband's incomes we note that the median income of Japanese households is decreasing from 4,620,000 yen in 2005 to 4,270,000 yen in 2015, probably due to the global economic crisis. Husband's high incomes in 2015 and 2005 determine a lower probability to belong to the new family (U_N), meaning that couples with husband having relatively high income, for example over 6 million yen (approximately 44,600

euro), are less likely to be new families (U_N). The income effects have a more clear cut after 2000, although having upper middle level income in 1985 and 1995 implies still to belong to traditional couples.

4.3 Posterior probabilities

The spouses's allocation to a latent class is performed through the estimated maximum posterior probability as in Equation (2). The last row of Table 6 reports the estimated percentages of couple's predicted in (U_T) for each wave. The lowest percentage is estimated for the 2015 wave and it is equal to 11.21%. Table 7 reports the composition of this subpopulation according with the covariates for each wave in order to compare the values with those reported in Table 3 related to the observed sample proportions.

Traditional families are mainly prevalent when the husband and the wife are young. Categories from 5 to 8 corresponding to wife's age from 44 to 58 (see Table 2) are less represented in 2015, meaning that especially recently, some wives are induced to get back to work during their middle ages. We expected that the younger couples support gender egalitarian values more and this would be reflected in gender equality in couples earnings structures. The twentieth century gave rise to profound changes in traditional gender roles as argued by [12], although the force of this "rising tide" has varied among rich and poor societies. The authors demonstrate that richer, postindustrial societies support the idea of gender equality more than agrarian and industrial societies and intergenerational differences in values are largest in postindustrial societies and relatively minor in agrarian societies, suggesting that the former are undergoing intergenerational changes in values. They also argue that cohort change in gender-role attitudes in postindustrial societies is unidimensional, with newer cohorts consistently more egalitarian than older cohorts. However, we found an association between age and the probability of being in (U_N) since households with wife in younger birth cohorts are more likely to be traditional families (U_T). It is still not normative for young married women to share equal financial responsibilities within household. This is partly due to the chronic shortages of regular childcare arrangements.

Considering the last survey none of the couples is predicted in the traditional family if the husband's income is less than 1,750,000 yen a year (approximately 13,000 euro) and that 61% of the couples predicted as traditional family has preschool children. We observe that (U_T) is mainly characterized by relative young households living in big cities, generally with husband having income of more than 4,250,000 yen a year, wives having high school education and husband's education that is higher than that of the wife, with one or more preschool children.

5 Discussion

About 20-30% of the couples are predicted in the latent class labeled as "Traditional couple" where only husbands undertake the financial responsibility for the family and wives bear domestic responsibility. These percentages show an increasing trend up to 2005. However, in 2015 we estimate a relevant increase of couples predicted in the latent class labeled as "New couple" that are non-traditional families. One of the reasons why the new families have been more represented in the past decade is that being a conventional single-income household is becoming more difficult due to the recent economic crisis causing income stagnation and decline. Over the past two decades, the average male annual income fell steadily. Therefore, in cases where husbands do not have relatively profitable and economically sound jobs, wives tend to have shared breadwinner roles. According to a survey carried out at national level in 2018 females employment rate in Japan is higher than that of United States working women but their entry-level positions and wages are lower than their male counterpart, demonstrating that high qualification occupations are often precluded for women.

We disentangle how the couple's typologies are related with the covariates and we highlight some considerations. Having young children has been constantly related to the greater likelihood that the respondents are in couples where husband is the sole provider. It is partly because of gender stereotype on wife's employment

Table 7: Frequencies of the covariates (%) for couples predicted in U_T (weighted with survey weights for years 2015 and 2005). The categories are defined in Table 2.

		Survey years			
	Categories	1985	1995	2005	2015
<i>Husband age</i>	1	19.1	27.9	18.9	25.5
	2	19.3	15.1	17.0	19.5
	3	15.9	8.7	5.6	10.1
	4	4.0	3.5	4.3	3.2
	5	1.0	1.4	2.6	9.9
	6	3.2	0.1	8.3	4.7
	7	2.2	1.1	13.2	1.0
	8	6.6	2.4	10.3	1.0
	9	13.6	18.2	13.1	14.7
	10	10.1	21.6	6.7	1.5
<i>Wife age</i>	1	17.3	30.3	20.7	29.4
	2	6.1	16.1	12.4	11.6
	3	5.8	8.4	9.6	11.9
	4	4.3	2.6	3.3	11.9
	5	5.1	0.0	4.6	5.5
	6	4.5	0.1	7.2	1.1
	7	11.1	0.1	10.1	5.8
	8	12.6	6.2	10.5	0.0
	9	14.5	17.1	12.7	11.3
	10	18.5	19.0	8.8	11.6
<i>Husband income</i>	1	4.3	4.6	7.6	0.0
	2	0.8	5.1	7.4	5.3
	3	9.5	16.9	8.3	10.2
	4	7.6	20.1	16.4	5.8
	5	6.1	14.4	9.6	10.0
	6	10.1	14.0	9.3	18.1
	7	4.7	6.9	12.0	11.0
	8	19.9	6.8	8.3	12.5
	9	13.3	3.2	10.2	11.9
	10	23.6	7.9	10.9	15.2
<i>Wife education</i>	1	15.6	20.2	14.5	8.4
	2	61.1	55.6	53.8	58.1
	3	13.1	17.4	20.4	20.4
	4	9.0	6.0	10.8	20.9
<i>Relative education</i>	1	46.0	44.8	43.3	34.4
	2	4.7	5.3	6.4	10.0
<i>Preschool children</i>	1	–	51.0	35.5	61.1
<i>Number of children</i>	0	6.2	6.2	4.9	4.9
	1	7.3	28.7	24.7	40.9
	2	16.1	51.5	50.7	38.9
	3	4.7	13.6	19.7	14.9
<i>Size of the place of living</i>	1	–	–	22.7	33.1
	2	–	–	26.7	27.0
	3	–	–	19.8	23.4
	4	–	–	21.7	13.6
	5	–	–	9.1	2.9

and on family roles in Japan where women have to stay at home, cook, clean, and care for children when having little babies. Another challenge for the families with preschoolers is the chronic shortages of regular child-care arrangements. Childcare has become an important policy issue in Japan as well as in many countries, but still there is a large number of children on waiting lists for childcare centers. Other childcare services such as home-based group care by family day care providers or babysitters who watch the children in either their

own homes or the children's home are not very popular in Japan and they are expensive. Moreover, when the employed women become pregnant and hope to take some maternity leave, it happens that employers place pressure on women to quit. This leads to the coining of the phrase "maternity harassment". These conditions have precluded mothers of preschoolers from staying in the workforce. At the same time, Japanese mothers, not fathers, are subjected to an extraordinary degree of excoriation by physicians, educators, preschool directors, and even high-level of government officials. In recent years, municipal governments in Japan have initiated various types of support programs. However, these programs might promote the notion that child rearing should be standardized, and make it extremely difficult for mothers to fulfil the requirements of being "good wives, wise mothers". Moreover, the popularity of social media might exacerbate normative pressures to perform well because mothers can look at what other mothers are doing, judge themselves in comparison, and feel inadequate in their mothering abilities. Consequently, deep sense of guilty might help mothers with young children withdraw from the labour market more in recent years.

High wage husbands imply more chance to build a traditional family. The result can be better understood in the light of the following arguments. When the levels of husband's income are enough to maintain a family and run their home, wives tend to accept to be economically dependent and stay at home. If husband's earnings are not enough to take care of all family financial obligations, wives tend to enter the labour force to supplement the husband's income. Under the current Japanese social security system, there is a tax reduction for married couples if one person make less than 1.5 million yen a year and dependent spouses of regular employees are automatically entitled to the basic pension without making any payments. Critics say that this entitlement deters full-time housewives from participating into the labour market.

Concerning the relative educational level of the couple we notice that hypergamous couples where husband is educated more than his wife heighten a couple's likelihood of having traditional, non-egalitarian marriage practices. This suggests that power imbalances within relationships may be related to the couple's choice of income provision-role type. This asymmetric gender relation within families seem to be associated with a patriarchal culture inducing to prefer traditional marriage practices. Having tertiary education does not necessarily show a higher probability towards a new type of couple, than lower educational levels. This leads to low economic returns to investment in education for women in Japan.

A worthwhile direction for future research would be to impute the missing responses by using a full information maximum-likelihood method in order to compare the results with those obtained with the current approach. Another research aspect would be to apply the proposed method to make comparative analysis on breadwinning patterns beyond Western societies along the research lines proposed by [9] and [13]. Comparing countries where breadwinning and caregiving expectations are institutionalized differently would contribute to improve further understanding of the women's opportunities in the various institutional contexts.

Notes

¹The Social Stratification and Social Mobility Survey data up to 2005 are available from the Center for Social Research and Data Archives, Institute of Social Science <https://csrda.iss.u-tokyo.ac.jp/>.

²In recent years, many countries are facing the increasing rate of missingness in surveys as well as Japan. Substantial declines in survey response rates after 2005 are mainly related to enactment of Act on the Protection of Personal Information in 2003 in Japan which evoked growing awareness of privacy or confidentiality issues.

³The R code is available from <https://github.com/penful/breadwinners>

Acknowledgments: Fulvia Pennoni acknowledges the financial support of the grant "Finite mixture and latent variable models for causal inference and analysis of socio-economic data" (FIRB – Futuro in ricerca) funded by the Italian Government (RBFR12SHVV). Miki Nakai acknowledges the financial support of the JSPS Grant-in-Aid for Scientific Research (No. 26380658, No. 17K04103 and No. 16H02045 as part of the Social Stratification and Social Mobility project). She thanks 2015 SSM Survey Management Committee for the permission to use the SSM data.

References

- [1] Agresti, A. (2019). *An Introduction to Categorical Data Analysis*. Third edition. John Wiley & Sons, Hoboken NJ.
- [2] Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. In B. N. Petrov and F. Csaki (Eds.), *Proceedings of the 2nd International Symposium on Information Theory*, pp. 267–281. Akademiai Kiado, Budapest.
- [3] Auspurg, K., T. Hinz, and C. Sauer (2017). Why should women get less? Evidence on the gender pay gap from multifactorial survey experiments. *Am. Sociol. Rev.* 82(1), 179–210.
- [4] Bartolucci, F., S. Bacci, and M. Gnaldi (2016). *Statistical Analysis of Questionnaires: A Unified Approach Based on R and Stata*. CRC Press, Boca Raton FL.
- [5] Bartolucci, F., A. Farcomeni, and F. Pennoni (2012). *Latent Markov Models for Longitudinal Data*. CRC Press, Boca Raton FL.
- [6] Bartolucci, F., S. Pandolfi, and F. Pennoni (2017). LMest: An R package for latent Markov models for longitudinal categorical data. *J. Stat. Softw.* 81(4), 1–38.
- [7] Baum, L. E., T. Petrie, G. Soules, and N. Weiss (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Statist.* 41, 164–171.
- [8] Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philos. Trans. R. Soc. Lond.* 53, 370–418.
- [9] Crompton, R. and C. Lyonette (2006). Work-life 'balance' in Europe. *Acta Sociol.* 49(4), 379–393.
- [10] Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B* 39(1), 1–38.
- [11] Holloway, S. D. (2010). *Women and Family in Contemporary Japan*. Cambridge University Press.
- [12] Inglehart, R. and P. Norris (2003). *Rising Tide*. Cambridge University Press.
- [13] Korpi, W., T. Ferrarini, and S. Englund (2013). Women's opportunities under different family policy constellations: gender, class, and inequality tradeoffs in Western countries re-examined. *Soc. Politics* 20(1), 1–40.
- [14] Lazarsfeld, P. F. and N. W. Henry (1968). *Latent Structure Analysis*. Houghton Mifflin, Boston MA.
- [15] Lindsay, B., C. C. Clogg, and J. Grego (1991). Semiparametric estimation in the Rasch model and related exponential response models, including a simple latent class model for item analysis. *J. Am. Statist. Assoc.* 86(413), 96–107.
- [16] McLachlan, G. and D. Peel (2000). *Finite Mixture Models*. John Wiley & Sons, New York.
- [17] Nakai, M. (2017). Changes in couples' breadwinning patterns and wife's economic role in Japan. In F. Greselin, F. Mola, and M. Zenga (Eds.), *Book of Short Papers (CLADAG 2017)*. Universitas Studiorum, Mantova.
- [18] Nock, S. L. (2001). The marriages of equally dependent spouses. *J. Fam. Issues* 22(6), 755–775.
- [19] Pennoni, F. (2014). *Issues on the Estimation of Latent Variable and Latent Class Models*. Scholars' Press, Saarbrücken.
- [20] Pennoni, F. and M. Nakai (2018). A latent variable model for a derived ordinal response accounting for sampling weights, missing values and covariates. In S. Capecchi, F. Di Iorio, and R. Simone (Eds.), *Proceedings of the International Conference on Advances in Statistical Modelling of Ordinal Data*, pp. 155–162. FedOAPress, Naples.
- [21] Raley, S. B., M. J. Mattingly, and S. M. Bianchi (2006). How dual are dual-income couples? Documenting change from 1970 to 2001. *J. Marriage Fam.* 68(1), 11–28.
- [22] Saito, M. (2000). *Modern Girl*. Magazine House, Tokyo (in Japanese).
- [23] Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Statist.* 6(2), 461–464.
- [24] Simpson, I. H. and P. England (1982). Conjugal work roles and marital solidarity. In J. Aldous (Ed.), *Two Paychecks: Life in Dual-Earner Families*, pp. 147–172. Sage, Beverly Hills CA.
- [25] Skrondal, A. and S. Rabe-Hesketh (2004). *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. CRC Press, Boca Raton FL.
- [26] Sørensen, A. and S. McLanahan (1987). Married women's economic dependency, 1940-1980. *Am. J. Sociol.* 93(3), 659–687.
- [27] Steel, G. (2019). *Beyond the Gender Gap in Japan*. University of Michigan Press, Ann Arbor MI.
- [28] Sulis, I. and M. Porcu (2017). Handling missing data in item response theory. Assessing the accuracy of a multiple imputation procedure based on latent class analysis. *J. Classif.* 34(2), 327–359.
- [29] Van Buuren, S. (2018). *Flexible Imputation of Missing Data*. Second edition. CRC Press, Boca Raton FL.
- [30] Vitali, A. and D. Mendola (2014). Women as main earners in Europe. Available at <https://ssrn.com/abstract=2555172>.
- [31] Wu, W., F. Jia and C. Enders (2015). A comparison of imputation strategies for ordinal missing data on likert scale variables. *Multivar. Behav. Res.* 50(5), 484–503.