Estimating longitudinal emotional and self rated health trajectories in women undergoing menopause: A Bayesian latent factor approach

Irena Chen[1] and Susie D. Lee[2]

[1]*Department of Digital and Computational Demography, Max Planck Institute for Demographic Research (MPIDR), Rostock, Germany.*
[2]*Laboratory of Fertility and Well-Being Max Planck Institute for Demographic Research (MPIDR), Rostock, Germany.*

October 31, 2025

Abstract

Menopause is a significant life transition with implications for population pattern in fertility and health. To better understand the large heterogeneity in the timing and symptoms of menopausal transition, the present study proposes and tests a novel hypothesis that poor premenopausal health underpins faster ovarian aging, manifested through earlier menopause and severe symptoms. Using longitudinal data from the German Family Panel, we model latent health trajectories preceding menopause, and connect these to later menopausal outcomes self-reported by women. We find substantial individual variations in the evolution of premenopausal health. We also find that women who have a higher starting negative affect compared to the population tend to experience menopause earlier. We plan to further analyze the link between premenopausal health and symptoms, to be presented at the EPC next year. The present study contributes both theory and methods that are relevant for population research on menopause in particular and reproductive aging in general.

1 Introduction

Menopause, the permanent cessation of menstrual cycles, is a key life transition with implications for women's fertility and overall health. The rapid decline in the level of reproductive functions and hormone level during the menopausal transition impinge on not only fecundity but also health in several ways (Monteleone et al., 2018; Shuster et al., 2010). In this way, the menopausal transition demarcates a critical change point in the quality of life, of which significance has grown as more childbearing is being delayed and post-reproductive lifespan is getting longer.

There is large individual variation for both *when* menopause occurs and *how* is experienced. For instance, while final menstrual period (FMP) for women entering menopause naturally occurs between 45 to 55 years, it can be at younger age if

menopause occurs through surgical means or other conditions. Moreover, the transition to menopause, also known as perimenopause, and time afterwards are often accompanied by various menopausal symptoms experienced at different degree by individuals Gold and Greendale (2007); El Khoudary et al. (2019). The main drivers of these variations in the timing of menopausal transition and degree of symptoms remain poorly understood. This is an unfortunate gap in the literature, given that not only earlier onset of menopause but also severe menopausal symptoms have been shown to have negative impacts on the quality of life and post-menopausal health.

The present study proposes and tests a novel hypothesis that health comprises a pathway to individual differences in menopausal timing and symptoms. Specifically, we hypothesize that *poor premenopausal health predicts an earlier onset of menopausal transition and severe menopausal symptoms*. The hypothesis stems from converging theories that reproductive aging is underpinned by a broader aging process (Laven, 2022; Fraser et al., 2020), supported through evidence of genetic and molecular mechanisms linking general aging with ovarian aging Laven (2022).

Other studies have used indirect measures, such as indicators or predictors of health, to show their association with menopause. For example, women with chronic illness are more likely to experience earlier menopause Ramezani Tehrani and Amiri (2021); Zheng et al. (2024) and severe menopausal symptoms Waidyasekera et al. (2009); Miller et al. (2005). Moreover, life course predictors of chronic illness - such as low socioeconomic status, exposure to critical life events - have been shown to relate to the earlier onset of menopause and more severe symptoms Namazi et al. (2019); Schoenaker et al. (2014). While these studies use indirect measures of health, such as the presence of certain health condition or determinants of health, to provide evidence in support of the aforementioned hypothesis, there is notable lack of studies on the link between direct measures of premenopausal health and menopause. Our study aims to fill this gap, with a goal to help further develop theories for menopause in particular and reproductive aging in general.

We use the German Family Panel data that prospectively tracks individuals before and after they report menopausal transition. We first model health trajectory preceding the self-reported onset of menopause, and then test the link between health trajectory and two outcomes: age of self-reported onset of menopause and degree of discomfort -in terms of physical and mental discomfort- experienced during menopausal transition. We provide a detailed description of premenopausal health trajectories, based on two measures of health: self-rated general health and emotional health. Our latent factor model approach further allows characterizing premenopausal health trajectories in terms of both intercept ('level') and slope (1pace of change'). To the best of our knowledge, this is the first analysis that models latent longitudinal health trajectories that precede menopause, and connect these to later menopausal outcomes. Specifically, we test the following predictions:

1. Lower level of premenopausal health, measured in terms of self-rated general health and/or emotional health, is related to
   (a) earlier initiation of menopausal transition.
   (b) more severe discomfort during menopausal transition.

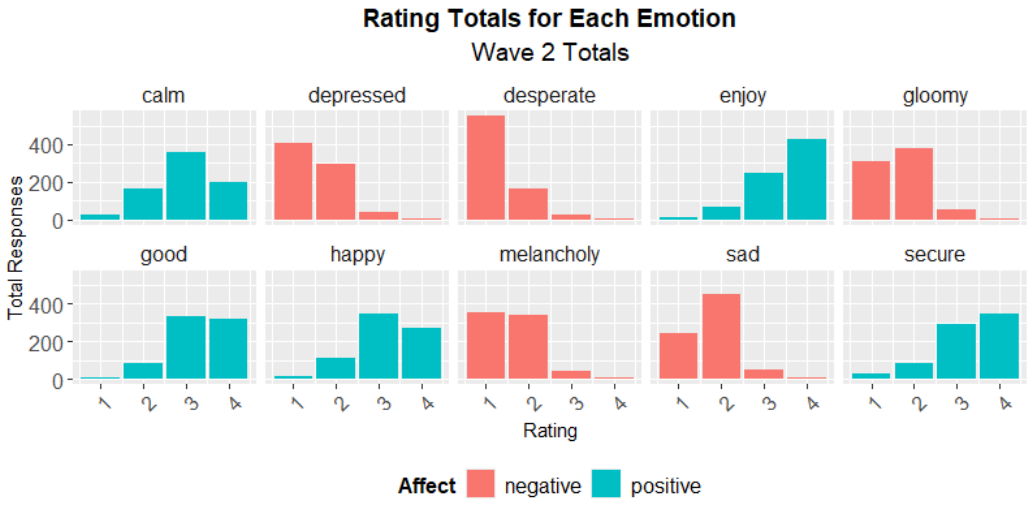

Figure 1: Response totals at wave 2 for each emotion.

2. Faster decline in premenopausal health, measured in terms of self-rated general health and/or emotional health, is related to

   (a) earlier initiation of menopausal transition.

   (b) more severe discomfort during menopausal transition.

1.1 Data

The German Panel Analysis of Intimate Relationships and Family Dynamics (PAIRFAM) is a longitudinal cohort study, which began in 2008 and 2009 (Brüderl et al., 2022). Participants were recruited from three birth cohorts: 1971–1973, 1981–1983, and 1991–1993 and have been surveyed annually since the start of the study. The focus of this study is women from the 1979-1973 cohort, because menopause-related questions were asked only to this group. We used data collected up to the most recent wave 14.

1.1.1 Predictors: Premenopausal health

We have two primary predictors of interest: self-rated general health and emotional health. First, self-rated general health is taken from 5-scale answers from "Bad" to "Very good" to the question "How would you describe your general health in the past 4 weeks?". The question was asked at every wave. Second, for emotional health, we used answers to the "State-Trait Depression Scales" (Spaderna et al., 2002). Respondents were asked to choose 4 out of the 10 emotions in the scales, which are comprised of 5 positive and negative emotions each. The self rated health question was asked in wave 1 while the STD-S questions were assessed first in wave 2. Both sets of questions were subsequently asked each year up to wave 14. Figure 1 displays the aggregated wave 2 responses for each emotion. We can see that generally, positive emotions were scored higher by participants and negative emotions tended to be scored lower by participants.

Variable	Statistic	Value	No. of Observations
Longitudinal Predictor	Count		
Emotional Health (10 variables)			8,887
Self Rated Health (1 variable)			9,179
Timing of Menopause	Count		
Left-censored		259	966
Interval-censored		279	966
Right-censored		428	966
Adjusted Covariates	Mean/SD		
Years of Education		13.5	
Number of Children		1.6	
Adjusted Covariates	Count		
Marital Status (Married)		621	966
Marital Status (Unmarried)		345	966
Ethnicity (German)		753	966
Ethnicity (Non-German)		213	966
Ever Smoked (No)		478	966
Ever Smoked (Yes)		488	966

Table 1: Descriptive statistics of the menopause dataset. For the self-rated health predictor, we had 9,179 observations from 966 women. For the emotional health predictors, we had 8,887 observations from 966 women. The adjusted covariate values were computed from each women's first observation in our final dataset.

1.1.2 Outcomes: Timing and symptoms of perimenopause

Starting in wave 10, three questions related to menopause experiences to the 1970-73 cohort. The first is about self-rated phase in menopausal transition, while the latter two are about self-rated discomfort experienced during menopause.

We used the first question to infer when perimenopause began in each women. The question lets each respondent to choose the statement that best describes her current situation among: "I have not yet entered menopause", "I am experiencing the first symptoms of menopause", "I am in the middle of menopause","I have already passed through menopause", "Don't know", and "I don't want to answer."

1.1.3 Final sample for analyses

We first restrict our dataset to women who were either left censored or interval censored (since right-censored women would not experience menopause-related symptoms). We then further subset the dataset to women who had at least one non-missing response to the self-rated physical or emotional health questions. Our final dataset for the longitudinal emotions comprised 966 women with 8,887 longitudinal observations and the final dataset for the self-rated physical health comprised the same 966 women with 9,179 observations.

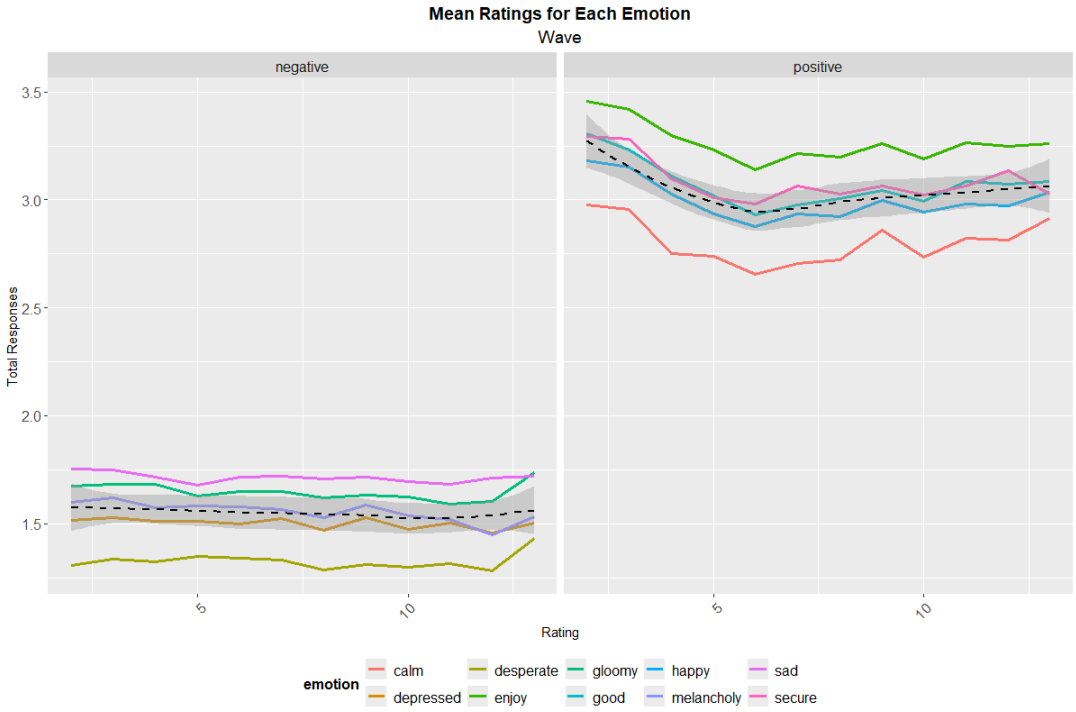

Figure 2: Average emotion rating for each wave, grouped by latent affect (negative and positive). A loess curve has been fitted to each affect group of emotions

2 Methods

2.1 Latent factor model for premenopausal health

Since self-rated general health and emotional health can be thought of as two different aspects of overall health, we chose to model each health trajectories and analyze them in relation to menopause separately. In the PAIRFAM data, we have self-rated emotional health questions for 10 different emotions corresponding to the "State-Trait Depression Scales". We summarize these emotions as being generated by two latent affects, one corresponding to positive emotions and the other corresponding to negative emotions, which we assign to each emotion before applying our model to the dataset.

Let $\boldsymbol{X}_{ij}$ be a $N \times P$ matrix of the observed P responses. We assume that $\boldsymbol{X}_{ij}$ are manifestations of a smaller number of K latent factors. We can then write $\boldsymbol{X}_{ij}$ as coming from a multivariate Normal distribution:

$$\boldsymbol{X}_{ij} \sim \mathcal{N}(\mu(\Lambda, \boldsymbol{\eta}_{ij}), \boldsymbol{\Omega}), j = 1, \ldots, n_i, \text{ independently for } i = 1, \ldots, N, \tag{1}$$

where Λ is the loadings $P \times K$ matrix, $\boldsymbol{\eta}_{ij}$ are the K latent trajectories and $\boldsymbol{S}$ is the diagonal variance matrix. $\boldsymbol{\eta}_{ij}$ is $N \times K$ matrix of unobserved latent trajectories; i.e. $\boldsymbol{\eta}_{ij} = [\eta_{ij1}, \ldots, \eta_{ijK}]$.

2.1.1 Latent trajectories

The η_{ijk} components are estimated with the following regression model:

$$\eta_{ijk} = \boldsymbol{Z}\boldsymbol{\beta}_k + \widetilde{\boldsymbol{Z}}_{ij}\boldsymbol{b}_{ik}, k = 1, \ldots, K,$$
$$\boldsymbol{\beta} \sim \mathcal{N}(0, \sigma_\beta), \boldsymbol{b}_i \sim \mathcal{N}(0, \Sigma_b),$$

where $\boldsymbol{Z}_{ij}, \widetilde{\boldsymbol{Z}}_{ij}$ are design matrices of covariates that may affect the latent trajectories, $\boldsymbol{\beta}_k$ are the global regression coefficients and b_{ik} are the random effects that describe individual deviations from the global trend. Σ_b is the matrix of random effects for all K factors. To ensure that the model is identifiable, we set the global intercept to 0. In practice, we set σ_β to be 1, i.e. a weakly informative prior.

2.2 Outcome model

2.2.1 Survival outcome

We chose to parameterize the survival outcome as an accelerated failture time (AFT) model with a Weibull distribution. We link the factor model and the hazard function of the outcome model via the random effects:

$$h_i(t) = e^{g(\boldsymbol{b}_i, \alpha)} h_0(t)(\frac{t}{e^{g(\boldsymbol{b}_i, \alpha)}}) \tag{2}$$

Our dataset comprises individuals who are subject to various forms of censoring. Due to the survey design, participants were not asked *at which age* they first experienced symptoms of menopause; rather only if they had started it. This corresponds to interval censoring, where we know the interval of ages where a woman began menopause, but we do not know exactly when she began. The menopause question was also first asked at wave 10, meaning that there may have been women who had already completed menopause in earlier waves. Finally, there was right censoring in the form of women who had not yet experienced menopause symptoms by wave 14 (the most recent wave of data collection). To account for these various types of censoring, we extended the survival model in 2.2.1 to allow for left, interval, and right-censoring.

2.3 Estimation

Since the joint distribution is intractable, we used MCMC sampling to estimate the posterior distributions of the unknown parameters via the Hamiltonian Monte Carlo algorithm using Stan and the rstan interface (Carpenter et al., 2017; Stan Development Team, 2020).

3 Results

3.1 Premenopausal health and the onset of perimenopause

Table 2 display the population level trend over time for the emotional affect and self rated health trajectories. All of the estimated coefficients were negative, meaning that both emotional and self rated health tends to decline over time.

Variable	Post. Mean	95% CrI
Negative Affect*	**-2.18**	(**-3.22, -1.12**)
Positive Affect*	**-6.30**	(**7.55, -5.10**)
Self Rated Health	**-0.17**	(**-0.38, -0.06**)

Table 2: Estimated population-level coefficients for latent trajectory outcome, fit with a linear time trend. The random effects in Tables 4 and 3 are estimated relative to these global estimates. * indicates that the posterior means and 95% CrI values have been multiplied by 10^2.

Variable	Post. Mean	95% CrI
Negative Affect Intercept	**0.28**	(**0.01, 0.56**)
Negative Affect Slope	**-2.32**	(**-4.25, -0.49**)
Positive Affect Intercept	-0.05	(-0.26, 0.15)
Positive Affect Slope	**-2.22**	(**-3.8, -0.66**)

Table 3: Estimated coefficients for the survival outcome model with the latent affect random effect predictors. All estimated posterior means and 95% CrI values have been multiplied by 10^2.

Tables 3 and 4 display the estimated regression coefficients for the one latent factor for self-rated health, and two latent factor (positive and negative affects) models. We chose to parameterize the outcome so that a positive regression coefficient indicates a decreased in time to event (i.e. shorter survival time) and a negative coefficient indicates an increased time to event (i.e. longer survival time), which better corresponds to the idea of accelerated failure.

For the emotional health models, we find that a higher negative affect intercept was associated with faster onset of perimenopause, compared to the population average. However, a higher slope value (i.e. faster decline) was associated with slower onset of menopause. As shown in Table 2, the estimated coefficient for the population-level negative affect trajectory was negative (see also Figure 2), which indicates that negative emotions on average tend to slightly decline over time. A higher individuak slope value in this case would mean that a woman's negative emotions are declining more slowly than the population average. This is significantly associated with later onset of perimenopause.

Variable	Post. Mean	95% CrI
Self-rated Health Intercept	-0.02	(-0.05, 0.01)
Self-rated Health Slope	-0.13	(-0.41, 0.14)

Table 4: Estimated coefficients for the survival outcome model with the self-rated health random effect predictors. All estimated posterior means and 95% CrI values have been multiplied by 10^2.

3.2 Planned Analysis: Sensitivity Analyses

We plan to conduct various sensitivity analyses, which will also be presented at EPC. Specifically, we will test to see if married and widowed status have different effects on emotional trajectories leading up to peri-menopause. In our current model, we specified marital status as binary, so that unmarried, divorced, and widowed were grouped in the same category.

3.3 Planned Analysis: Symptom Outcomes

After completing our analysis of the menopausal onset outcome, we will turn our focus to jointly modeling the menopausal symptom trajectories with the emotional and self rated health trajectories.

References

Brüderl, J., Drobnič, S., and Hank, K. (2022). The german family panel (pairfam). Backup Publisher: ZA5678 Data file Version 13.0.0.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76:1–32.

El Khoudary, S. R., Greendale, G., Crawford, S. L., Avis, N. E., Brooks, M. M., Thurston, R. C., Karvonen-Gutierrez, C., Waetjen, L. E., and Matthews, K. (2019). The menopause transition and women's health at midlife: a progress report from the study of women's health across the nation (swan). *Menopause*, 26(10):1213–1227.

Fraser, A., Johnman, C., Whitley, E., and Alvergne, A. (2020). The evolutionary ecology of age at natural menopause: implications for public health. *Evolutionary human sciences*, 2:e57.

Gold, E. B. and Greendale, G. A. (2007). Epidemiology of menopause: demographics, environmental influences, and ethnic and international differences in the menopausal experience. In *Treatment of the Postmenopausal Woman*, pages 77–96. Elsevier.

Laven, J. S. (2022). Early menopause results from instead of causes premature general ageing. *Reproductive BioMedicine Online*, 45(3):421–424.

Miller, S. A., Santoro, N., Lo, Y., Howard, A. A., Arnsten, J. H., Floris-Moore, M., Moskaleva, G., and Schoenbaum, E. E. (2005). Menopause symptoms in hiv-infected and drug-using women. *Menopause*, 12(3):348–356.

Monteleone, P., Mascagni, G., Giannini, A., Genazzani, A. R., and Simoncini, T. (2018). Symptoms of menopause—global prevalence, physiology and implications. *Nature Reviews Endocrinology*, 14(4):199–215.

Namazi, M., Sadeghi, R., and Behboodi Moghadam, Z. (2019). Social determinants of health in menopause: an integrative review. *International journal of women's health*, pages 637–647.

Ramezani Tehrani, F. and Amiri, M. (2021). The association between chronic diseases and the age at natural menopause: a systematic review. *Women & health*, 61(10):917–936.

Schoenaker, D. A., Jackson, C. A., Rowlands, J. V., and Mishra, G. D. (2014). Socioeconomic position, lifestyle factors and age at natural menopause: a systematic review and meta-analyses of studies across six continents. *International journal of epidemiology*, 43(5):1542–1562.

Shuster, L. T., Rhodes, D. J., Gostout, B. S., Grossardt, B. R., and Rocca, W. A. (2010). Premature menopause or early menopause: long-term health consequences. *Maturitas*, 65(2):161–166.

Spaderna, H., Schmukle, S. C., and Krohne, H. W. (2002). Bericht über die deutsche adaptation der state-trait depression scales (STDS). [report about the german adaptation of the state-trait depression scales (STDS).]. *Diagnostica*, 48(2):80–89. Place: Germany Publisher: Hogrefe Verlag GmbH & Co. KG.

Stan Development Team (2020). RStan: the R interface to Stan. WebPage. R package version 2.21.2.

Waidyasekera, H., Wijewardena, K., Lindmark, G., and Naessen, T. (2009). Menopausal symptoms and quality of life during the menopausal transition in sri lankan women. *Menopause*, 16(1):164–170.

Zheng, M., Li, J., Cao, Y., Bao, Z., Dong, X., Zhang, P., Yan, J., Liu, Y., Guo, Y., and Zeng, X. (2024). Association of different inflammatory indices with risk of early natural menopause: a cross-sectional analysis of the nhanes 2013–2018. *Frontiers in Medicine*, 11:1490194.