**1 Title Slide**

Hello, my name is Ricarda Duerst, I am PhD student at the University of Helsinki and the Max Planck Institute for Demographic Research.

I am going to present ongoing work on forecasts of Nordic period fertility, with empirical prediction intervals. This is joint work with Julia and my supervisor Jonas Schöley from the Max Planck Institute.

**2 TFR**

To start off, a little reminder what the so-called Nordic Fertility Regime looks like. This is the Total Fertility rate over time in Finland, Denmark, Sweden, and Norway. We have high variability during the first half of the 1900s, strong decreases during the 1970s. Followed by less variability under the replacement level with moderate increases until 2010. And then, another drop in TFR to all-time lows, at least for Finland and Norway, during the 2010s until now.

**3 Aims**

Our aims for this project are forecasts of Nordic fertility and sensible prediction intervals. On the other hand, what we can work with and choose from is a range of forecast models and model-based prediction intervals that are usually too narrow.

**4 Research Plan**

Therefore, we came up with a research plan. … (see slide)

**5 Forecast Models**

Our selection of forecast models consists of a scenario-based model, an extrapolative model and a naïve model, so far.

**6 PPS**

The Postponement scenario was developed by Jessica, Julia, Pekka and Mikko. The main assumption of this scenario-based forecast model is, that the fertility postponement that has verifiably happened and that caused the decrease of the TFR in the last decade, will continue, but slow down. And eventually it will stop. That is why in this forecast from 2020 to 2040, the TFR is increasing and then levelling off.

**7 Lee-Carter**

Second model is the lee-carter model for fertility. This is an extrapolation-type forecast model and a demographic classic. It is based on the same method as the lee-carter model for mortality. Using singular value decomposition, age-specific and time-specific components are estimated and the time-specific components are then forecast using a random walk with drift.

And we also included a naïve forecast approach as reference. We simply fix the last observed value of the TFR.

**9 Empirical Prediction Intervals**

Now, for the exciting part. What are empirical prediction intervals?

**10 Principle of Empirical Prediction Intervals**

The idea or principle behind them is simple: A forecast is only as precise as similar forecast in the past turned out to be. Which means that we need historical forecasts to calculate them. The steps to derive empirical prediction intervals can be divided into steps of the calibration, the validation, and the application. The application in this case are real forecasts for Nordic fertility into the future.

The idea of empirical prediction intervals is not a new one. It first appeared 1971 in economics and made its first appearance in demography in the 80s. Ronald Lee, one of the authors of the Lee-Carter method, puts it like this 1998: “[Empirical prediction intervals are] a very valuable line of research that has been largely neglected since the publication of [the demographic paper in the 80s].” But then in 2004 Nico Keilman and Dinh Pham did something very similar to what we are doping: They applied empirical prediction intervals to fertility forecasts for the Nordic countries.

We are not only using them because a few leading scientists think they are great. But because they have advantages for our case. First of all, history has shown that model-based prediction intervals for population forecasts are too narrow. Second, you have to be prepared for exceptions and that is why in risk management you need to make probabilistic decisions in the face of risks. And history has shown us, that there is a non-neglectable risk to be very wrong with our population forecasts.

**12 Empirical Prediction Intervals**

Let’s get to business. How do we derive empirical prediction intervals for the Nordic countries? First, we take the available time series of fertility for each country and cut them up into smaller pieces. These are our cross-validation series. Then each series is divided into a training period, blue, and a testing period, pink. To the training period we apply a forecast method, and the forecast results will then be compared to the observed data in the testing period.

**13**

When we do this for all the cross-validation series, we can plot the forecast error distribution over the forecast length. This can now be either modelled by taking the raw quantiles of the distribution or by more fancy things like modelling with a skew-normal distribution. If we then apply this empirical forecast error distribution to the point forecasts of our actual forecast, we get empirical prediction intervals.

**14 EMP PIs for Scenarios**

For extrapolative data-driven forecast models, this is very intuitive. But we also chose a scenario-based approach to forecast Nordic fertility. So, how do we derive empirical prediction intervals for models, that only work under specific circumstances? The assumption of the PPS model is that the fertility postponement that has happened so far, will continue but will slow down and stop. In hindsight, it is easy for us to detect other periods where the same has happened and apply our forecast model to them. The fertility postponement already began in the 1970s and there was consensus among researchers that this will slow down and some point. Therefore, we can derive empirical prediction intervals for scenario-based models via cross-validation. That is one way.

However, there is a limitation. In hindsight, it is easy to apply a model to a time period, where it is certain that it will work. The resulting errors will be small. But this is not reflecting the reality of forecasting. We don’t know, whether our assumption is right and that should be reflected in the prediction intervals!

**16 Publications**

That is why we will also use another approach to retrieve empirical prediction intervals. If we use, like Keilman and Pham also did, published historical forecasts and see how wrong they are we get a different empirical error distribution. One that does reflect the real uncertainty of forecasting, because the published fertility forecasts are the expert consensus of their time.

**19 Preliminary Results**

I will know show some preliminary results. Because of the time, I will focus on Finland. Further, we don’t have results for all forecast models and the different types of prediction intervals yet. That’s why I will focus on the PPS forecast and the empirical prediction intervals from cross-validation.

**20 TFR PPS**

**21 Finland**

**24 End**

Thank you very much for your attention! I am now happy to receive your questions, and comments.