- I have tried to use a half-Normal prior on the "standard deviation" of the spline coefficients, $\lambda^{-1/2}$, but it failed to converge even for Zimbabwe

- all hyperpriors are now common between sexes, except for the dispersion parameters for the Negative Binomial (maybe will fix that too)

- simulating the posterior of the hyperparameters took extremely long time, given up on that at the moment

- Explored the PC priors and trimmed down the number of hyperparameters, as well as added spikes for TiPS at 5 and 10 that were previously missing, will add plots visualising the marginal prior distribution of the parameters soon (at the moment fitting the gumbel too all countries and don't want R to explode)

  – For the mortality model parameters, these are previously given GMRF with precision matrix $\boldsymbol{Q} = \lambda_1(\boldsymbol{D_2}'\boldsymbol{D_2} + \lambda_2 \boldsymbol{I})$. The ratio between λ_1 and λ_2 controls how much the linear extrapolation is shrunk towards 0. Writing the penalty (precision matrix) as $\boldsymbol{Q} = \tau(\boldsymbol{D_2}'\boldsymbol{D_2} + c\boldsymbol{I}) = \tau\boldsymbol{R}$, under the PC priors, we elicit information on the effective degrees of freedom (EDF) and translate this into a prior for λ. Following Ventrucci and Rue (2016), the EDF is approximated by the trace of the hat matrix under the classical linear regression model. Consider:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{B\beta} + \boldsymbol{\varepsilon} \\ \text{where} \quad & \varepsilon \sim N(0, \tau_\varepsilon^{-1}) \\ \text{with penalty} \quad & \tau\boldsymbol{\beta}'\boldsymbol{R}\boldsymbol{\beta}, \end{aligned}$$

  the hat matrix is $(\boldsymbol{B}'\boldsymbol{B} + \frac{\tau}{\tau_\varepsilon}\boldsymbol{R})^{-1}\boldsymbol{B}'\boldsymbol{B}$ and hence the EDF is $d(\tau) = tr(\boldsymbol{I} + \frac{\tau}{\tau_\varepsilon}\boldsymbol{R}(\boldsymbol{B}'\boldsymbol{B})^{-1})^{-1} = \sum_k (1 + \frac{\tau}{\tau_\varepsilon} v_k)^{-1}$. By setting a upper bound of the EDF U such that $P(d > U) = \alpha$ for some small probability α, the PC prior for τ can be obtained as a Gumbel Type 2 $(\frac{1}{2}, \theta)$ where $\theta = -\log(\alpha)\sqrt{d^{-1}(U)}$.

  There are several tuning parameters for this PC hyperprior, $\alpha, U, \tau_\varepsilon$ and c. I set $\alpha = 0.01$ and $U = 1$ for $\{\phi, A\}$, $U = 1.5$ for $\{\psi, B\}$ and $U = 5$ for $\{\lambda, \delta, \epsilon\}$, meaning that we don't really believe the EDF of the splines would be higher than 1, 1.5 or 5 for the respective parameters, which correspond to a constant ($U = 1$) and somewhere between a constant and a linear trend ($U = 1.5$).

  c is also estimated as a parameter but now I have decided to fix it as a constant to remove the computational burden. A higher value of c allows a lower τ in order to reach a specific EDF. In the limiting case when $c \to \infty$, the penalty degenerates to a i.i.d precision matrix. Conventionally in Bayesian P-splines, c is chosen to be some small constant so that it completes the rank of the penalty matrix but leaves the null space minimally affected. I have chosen c to be a small value $1e-3$ so that smoothness is prioritised before shrinking the spline to 0 or a constant, but thinking about increasing this value to allow more conservative priors on τ. Alternatively, I can increase U to be at least 2 so that c would have minimal effect on the limiting case under the 2nd order difference penalty.

  τ_ε is the precision of the 'observed data', here I interpret this as the precision of the 'observed' parameters that we expect from the data. I have set it to $(\log(1.3)/1.96)^{-2}$ for the parameters of the child and old age components and $(\log(3)/1.96)^{-2}$ for the hump components, roughly meaning that we expect majority of the mass would lie within 30% of the IGME derived estimates (or 300% for the hump components from the baseline estimates in 1960). I am thinking to increase this value to give a more conservative bound on τ, i.e. allowing more flexibility.

  Alternatively c and τ_ε could also be given priors and together form a joint prior with τ, i.e. $f(\tau, c, \tau_\varepsilon) = f(\tau|c, \tau_\varepsilon)f(c)f(\tau_\varepsilon)$ so that it would lower the risks of mis-specifying their values, but this would require numerical methods within *TMB* arising from calculating $d^{-1}(U)$, i.e. the τ implied by the values of c and τ_ε at each iteration.

A compromise would then be to use a discrete mixture of the Gumbel priors evaluated at a range of pre-specified values for c and τ_ε. Would this be an overkill? Or should I simply increase c or τ_ε to have a more conservative prior on τ?

- For f_{xt}, the spline coefficients are simply given i.i.d MVN with $U = 4$ (approximately 2 in each direction, age and time), $\alpha = 0.01$ and $\tau_\varepsilon = (\log(1.5)/1.96)^{-2}$
- For g_{xt}, previously it is estimated as a 2D tensor P-splines with penalty on the coefficients in the form $\boldsymbol{Q} = \lambda_{1\boldsymbol{x}}\boldsymbol{D_1}'_{\boldsymbol{x}}\boldsymbol{D_1} + \lambda_{2\boldsymbol{t}}\boldsymbol{D_1}'_{\boldsymbol{t}}\boldsymbol{D_1} + \lambda_{3\boldsymbol{xt}}\boldsymbol{D_1}'_{\boldsymbol{xt}}\boldsymbol{D_1} + \lambda_4\boldsymbol{I}$. I have swapped to using 2nd order differences penalty for smoothness, removed penalty for age-time and trimmed down the number of estimable parameters to $\boldsymbol{Q} = \tau(0.5_{\boldsymbol{x}}\boldsymbol{D_2}'_{\boldsymbol{x}}\boldsymbol{D_2} + 0.5_{\boldsymbol{t}}\boldsymbol{D_2}'_{\boldsymbol{t}}\boldsymbol{D_2} + c\boldsymbol{I})$. For some small constant c, this would mean that the smoothness is split equally in the age and time direction. Priors for τ is obtained using $\tau_\varepsilon = (0.08/1.96)^{-2}$, $c = 1e-3$, $U = 7$ and $\alpha = 0.01$. $U = 7$ means that the upper bound is approximately having 2-3 EDF in each direction, age and time.

- The PC priors do seem to work well in preventing under-smoothing, especially in B that was previously problematic, probably because of the potentially overly tight priors given? However, the D_2 penalties still imply a strong linear trend in the hump parameters outside the data range as the given upper bound of EDF is 5. I have also used a D_1 penalty on the mortality model parameters.
  - in the following, the results using D_2 penalty will be shown first, then D_1
  - estimated smoothing parameters
  - parameters using D_2 penalties are sometimes insensible, especially for ϵ (location of the hump), which went down to 0 in the earliest period due to the linear smoothness induce by the D_2 penalty. Variance estimates are also implausible for λ (level of the hump). Uncertainty around other parameters, especially A maybe be too low? Also for f_{xt} and possibly g_{xt}.
  - when using D_1 penalty, parameter estimates seem to be more sensible, however, uncertainty around the child/old age component parameters are again seem to be too low, as well as f_{xt} and possibly g_{xt}. In addition, asymmetric uncertainty range around the mode is obtained for females ${}_5q_0$ and the population counts (will check it on log-scale).
    * when using D_1 penalty, it is possible to elicit information on the expected variation in $\boldsymbol{B\beta}$ for each parameter as the variance structure is much more uniform than the D_2 penalty, hence I can consider acting directly on the expected variance for $\boldsymbol{B\beta}$ to derive the PC prior, removing the need to estimate precision of the pseudo data τ_ε
    * Tried to do this last night, uncertainty around parameters are still similar, showing that the priors are kind of robust

Zimbabwe

```
## [1] "Census Females"
```

```
## # A tibble: 86 x 6
##      age `1969`  `1982`  `1992`  `2002`  `2012`
##    <dbl>  <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
##  1     0   215. 137199. 169638. 170997. 215623.
##  2     1   200. 134655. 159190. 172117. 211071.
##  3     2   219. 133815. 158420. 168708. 198666.
##  4     3   229. 131120. 158291. 164925. 190736.
##  5     4   240. 128619. 160494. 162274. 183611.
##  6     5   251. 126617. 162425. 159507. 177197.
##  7     6   258. 125956. 164145. 156778. 173472.
##  8     7   261. 125351. 166412. 155377. 171651.
##  9     8   262. 124105. 167841. 153895. 169450.
## 10     9   259. 118378. 165018. 153789. 171475.
## # ... with 76 more rows
```

```
## [1] "Census Females 5-year"
```

```
## # A tibble: 18 x 2
##      age  `1969`
##    <dbl>   <dbl>
##  1     0 1100.
##  2     5 1252.
##  3    10 1181.
##  4    15  935.
##  5    20  709.
##  6    25  541.
##  7    30  437.
##  8    35  379.
##  9    40  303.
## 10    45  227.
## 11    50  168.
## 12    55  124.
## 13    60   95.4
## 14    65   67.1
## 15    70   35.5
## 16    75   17.0
## 17    80    9.26
## 18    85   13.4
```

```
## [1] "Census Males"
```

```
## # A tibble: 86 x 6
##      age `1969`  `1982`  `1992`  `2002`  `2012`
##    <dbl>  <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
##  1     0   238. 133357. 168079. 170637. 213895.
##  2     1   232. 129754. 158096. 172216. 210140.
##  3     2   231. 129163. 156962. 168433. 197420.
##  4     3   235. 127343. 157057. 164960. 189835.
##  5     4   242. 125798. 159247. 162197. 182613.
##  6     5   248. 124587. 161170. 159217. 176094.
```

```
## 7      6   255. 124714. 162716. 156209. 172218.
## 8      7   265. 124146. 164082. 154419. 170262.
## 9      8   271. 122838. 164791. 152264. 167883.
## 10     9   271. 117620. 161620. 152041. 170380.
## # ... with 76 more rows
```

```
## [1] "Census Males 5-year"
```

```
## # A tibble: 18 x 2
##      age  `1969`
##    <dbl>   <dbl>
##  1     0 1156.
##  2     5 1278.
##  3    10 1207.
##  4    15  931.
##  5    20  636.
##  6    25  459.
##  7    30  389.
##  8    35  341.
##  9    40  280.
## 10    45  229.
## 11    50  186.
## 12    55  142.
## 13    60   97.1
## 14    65   60.0
## 15    70   32.4
## 16    75   15.2
## 17    80    8.24
## 18    85    7.21
```

Thiele log-Normal Hump Spline

```
## [1] "relative convergence (4)"
```

```
##    log_tau2_logpop    log_tau2_logpop    log_tau2_logpop    log_tau2_logpop
##        4.188311709        6.003080052        4.348179163        5.757824367
##      log_lambda_fx      log_lambda_gx     log_dispersion     log_dispersion
##       10.634214243       10.881303496        1.188040609        1.182731942
##      log_lambda_tp           tp_slope        tp_params_5       tp_params_10
##        3.589509963       -0.006255899        0.225700924        0.421151124
##     log_lambda_phi     log_lambda_psi       log_lambda_A       log_lambda_B
##       15.275057021       14.654590261       15.252392419       14.444144593
##  log_lambda_lambda   log_lambda_delta log_lambda_epsilon
##        2.922373903        5.739305891        3.906946997
```

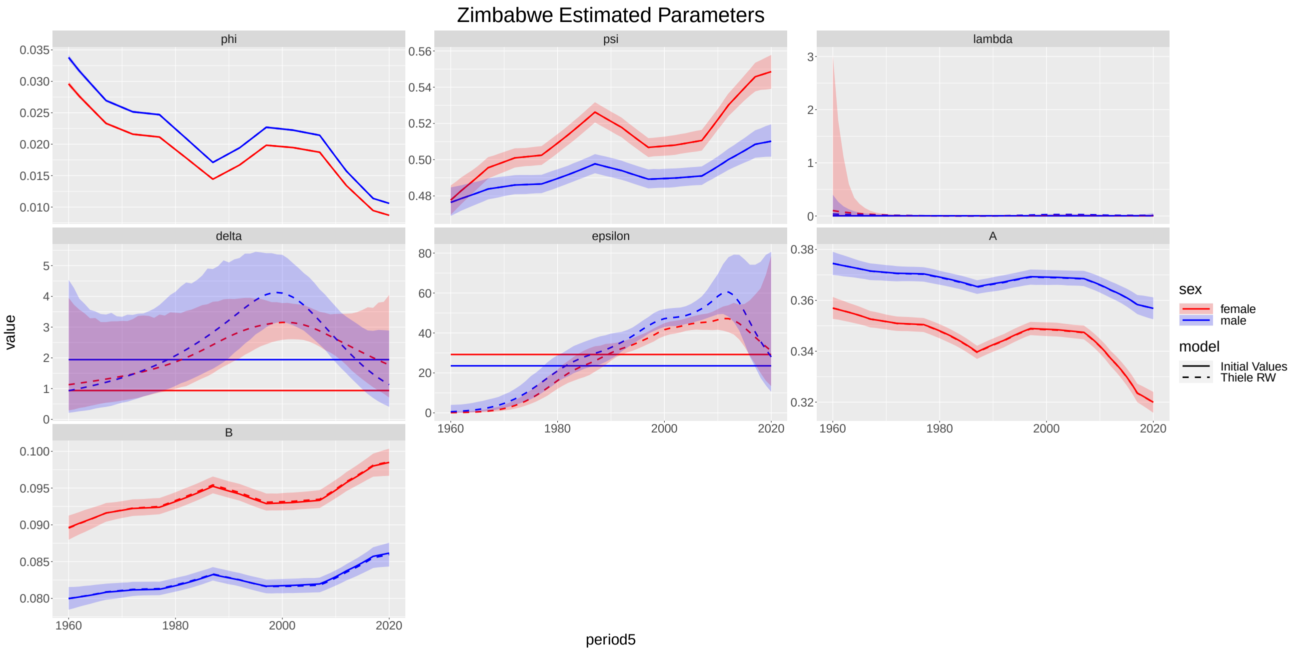

Figure 1: Estimated parameters

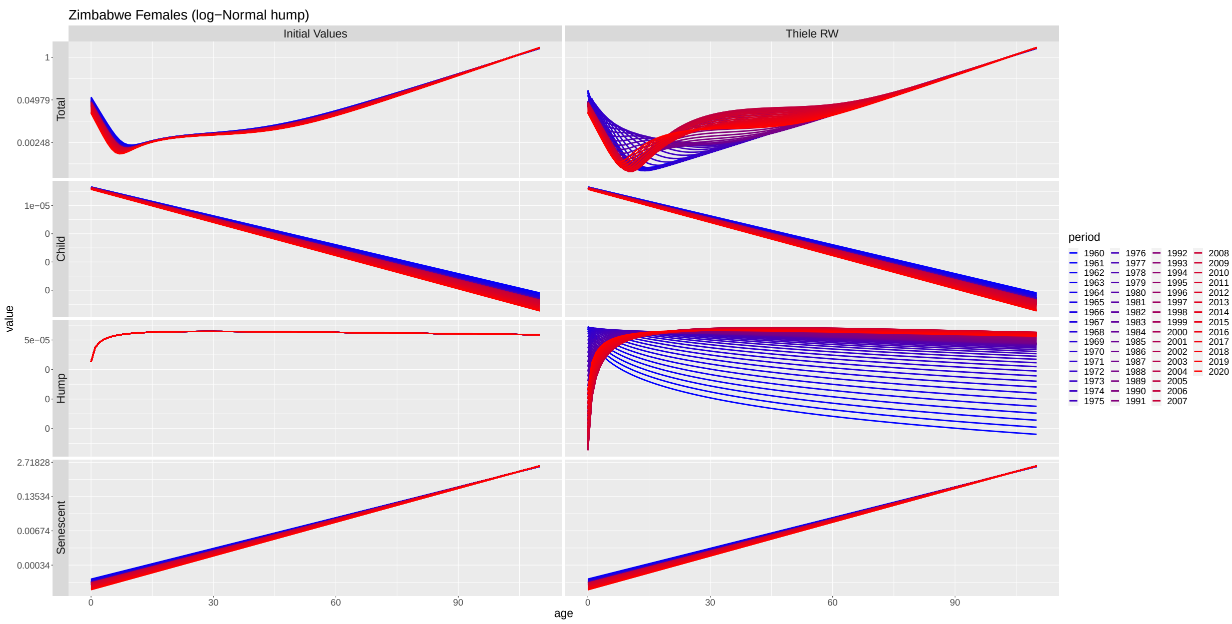

Figure 2: Thiele Decomposed

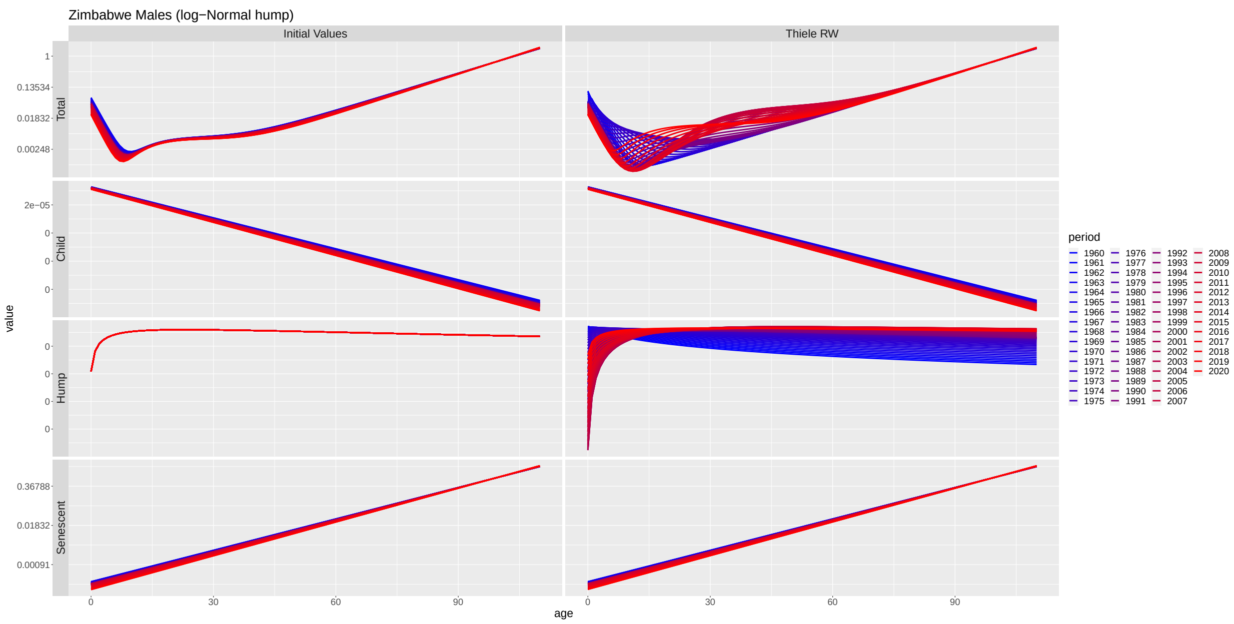

Figure 3: Thiele Decomposed

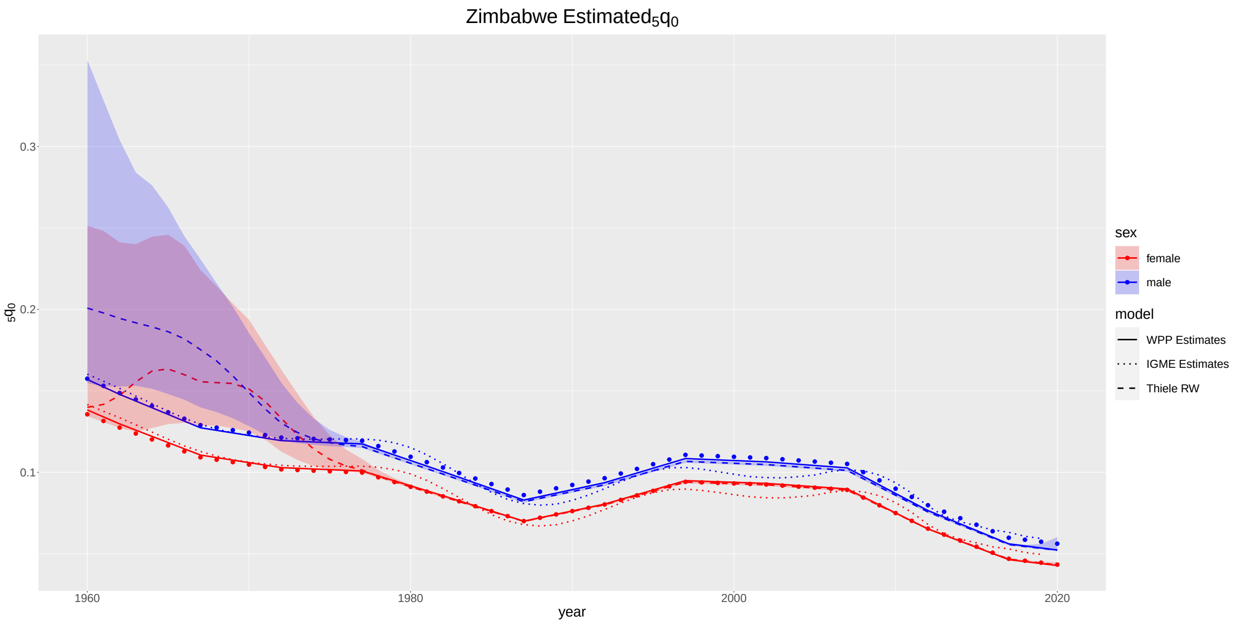

Figure 4: Estimated $_5q_0$

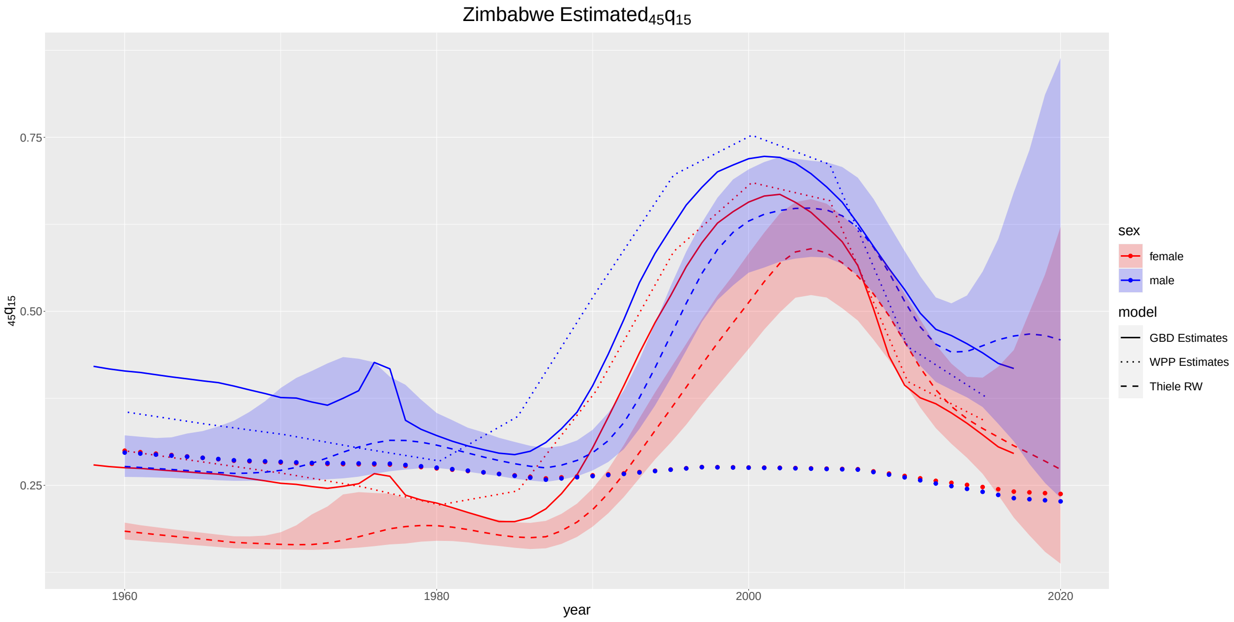

Figure 5: Estimated $_{45}q_{15}$

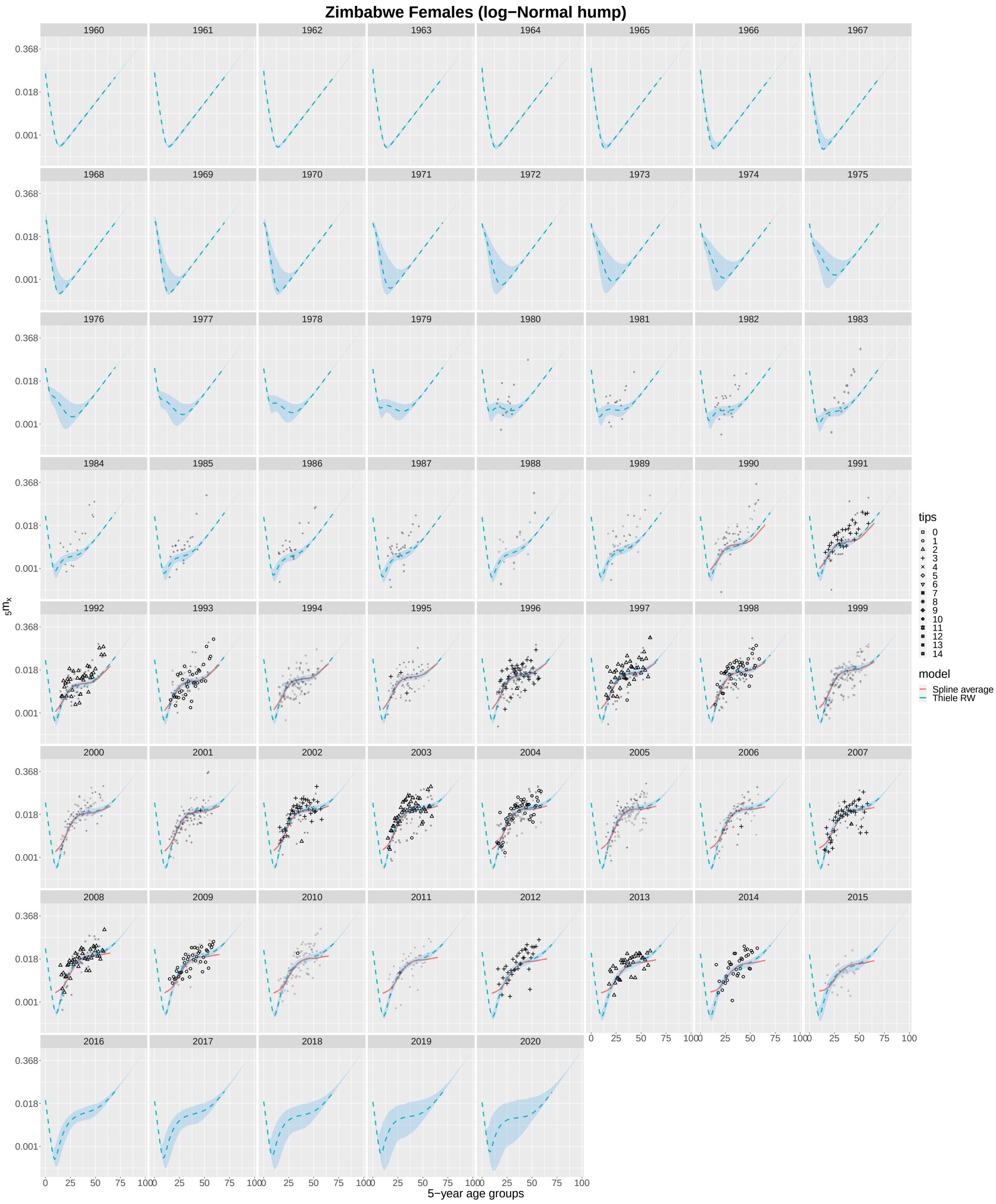

Figure 6: Mortality Schedules

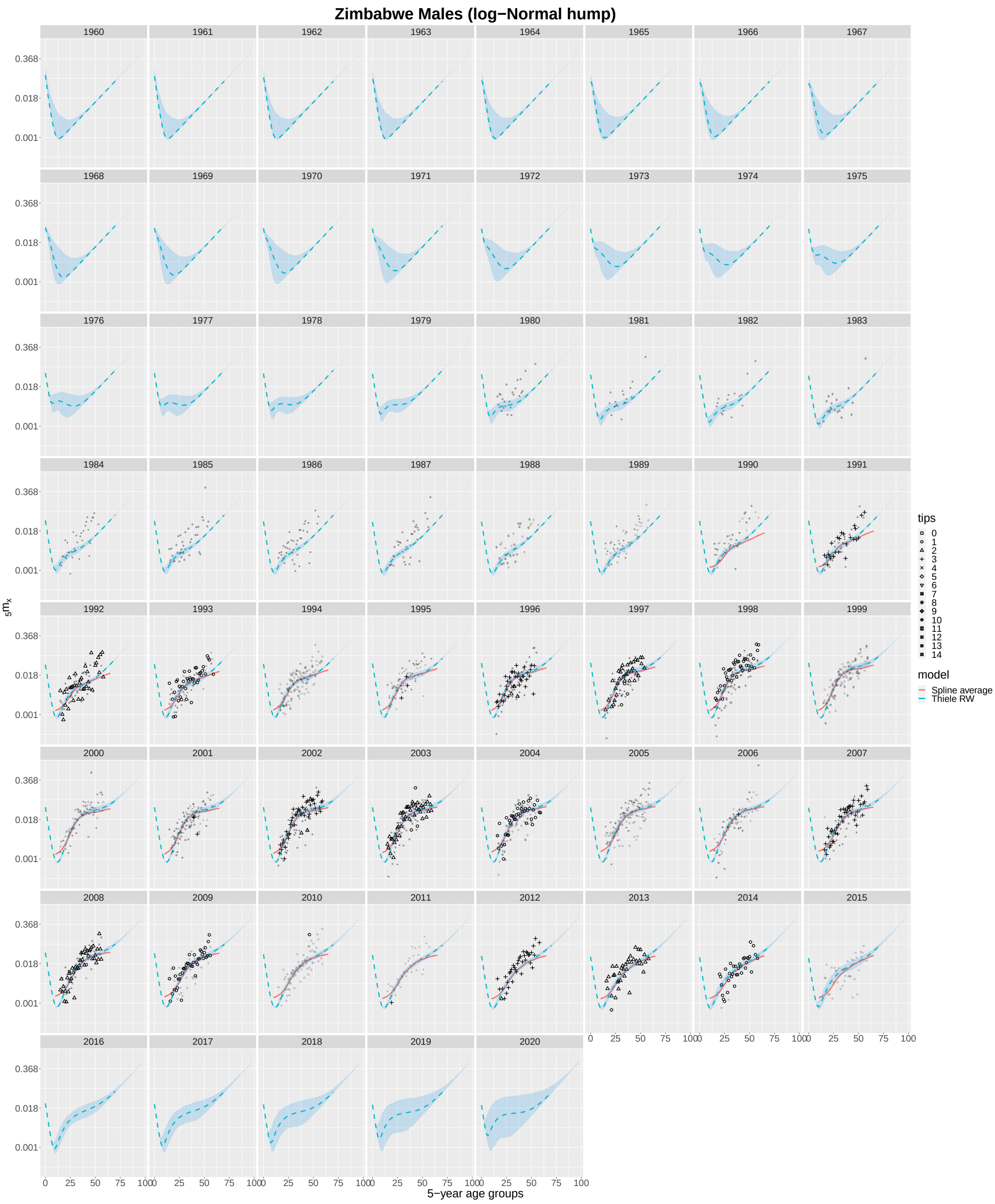

Figure 7: Mortality Schedules

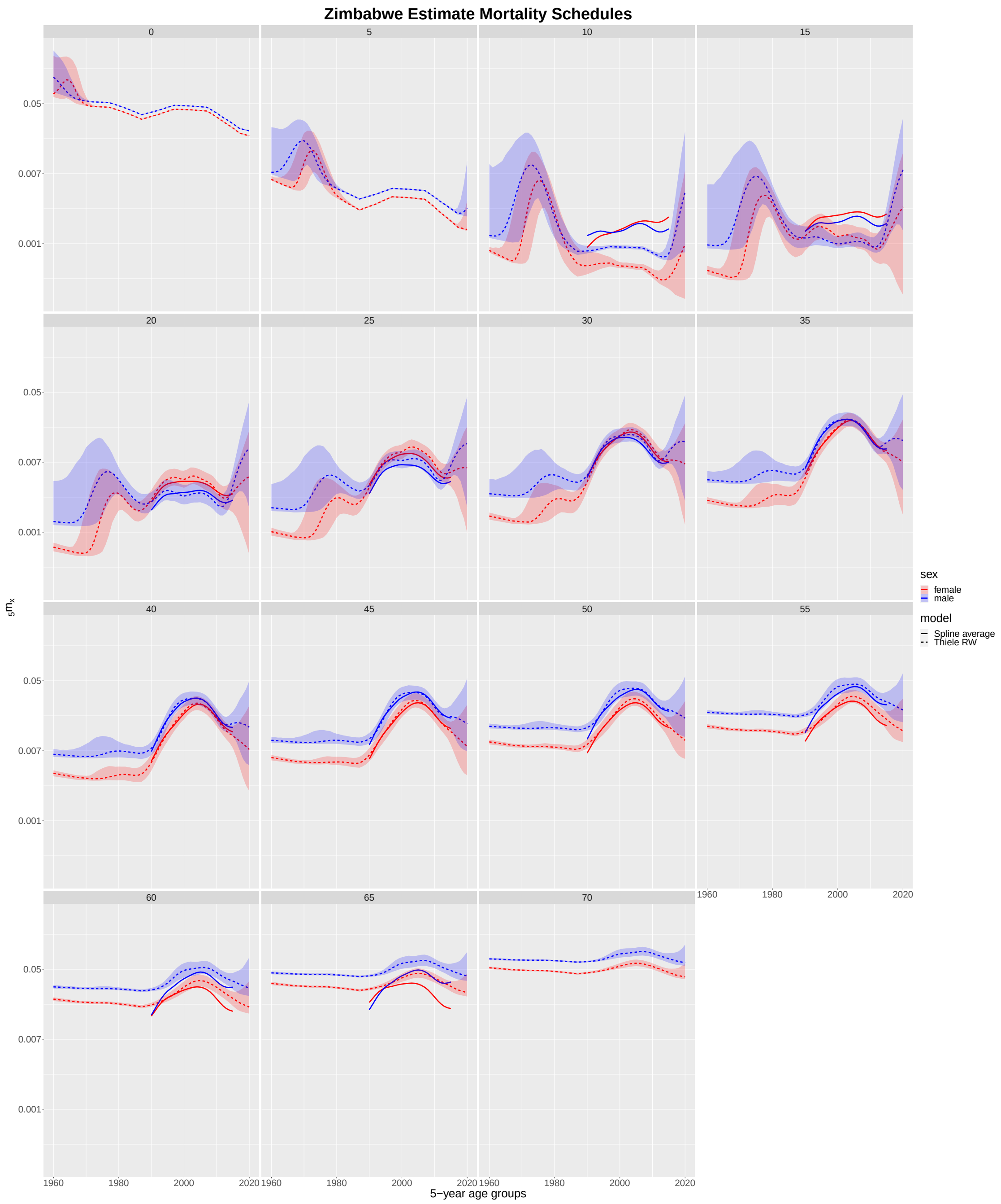

Figure 8: Mortality Schedules

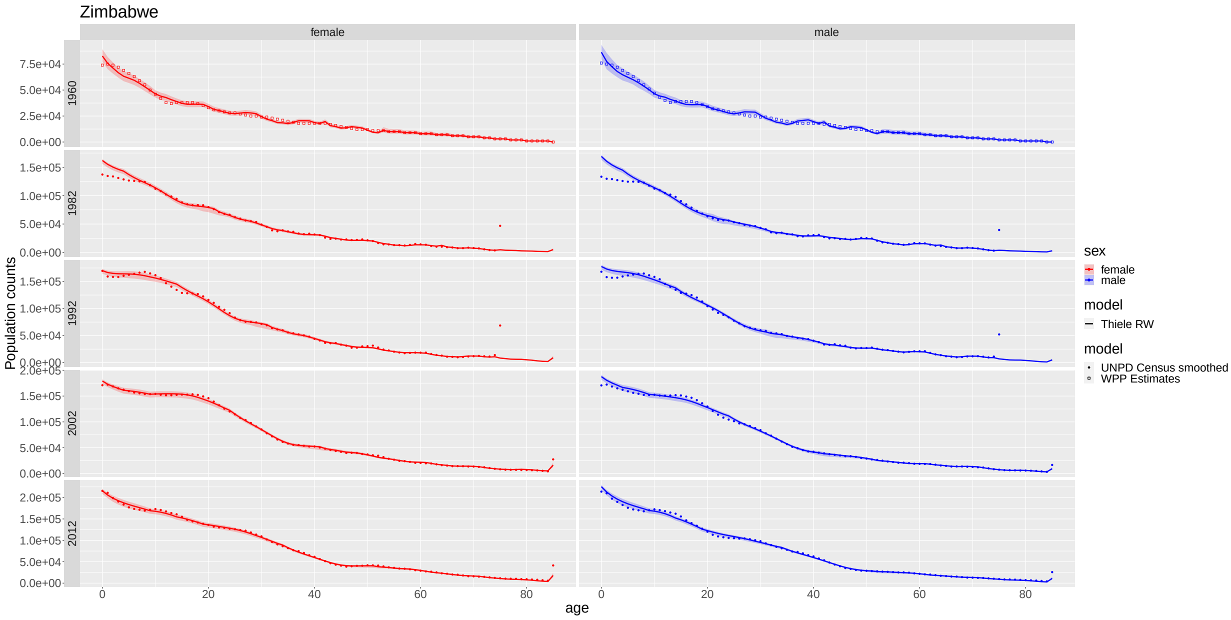

Figure 9: Population

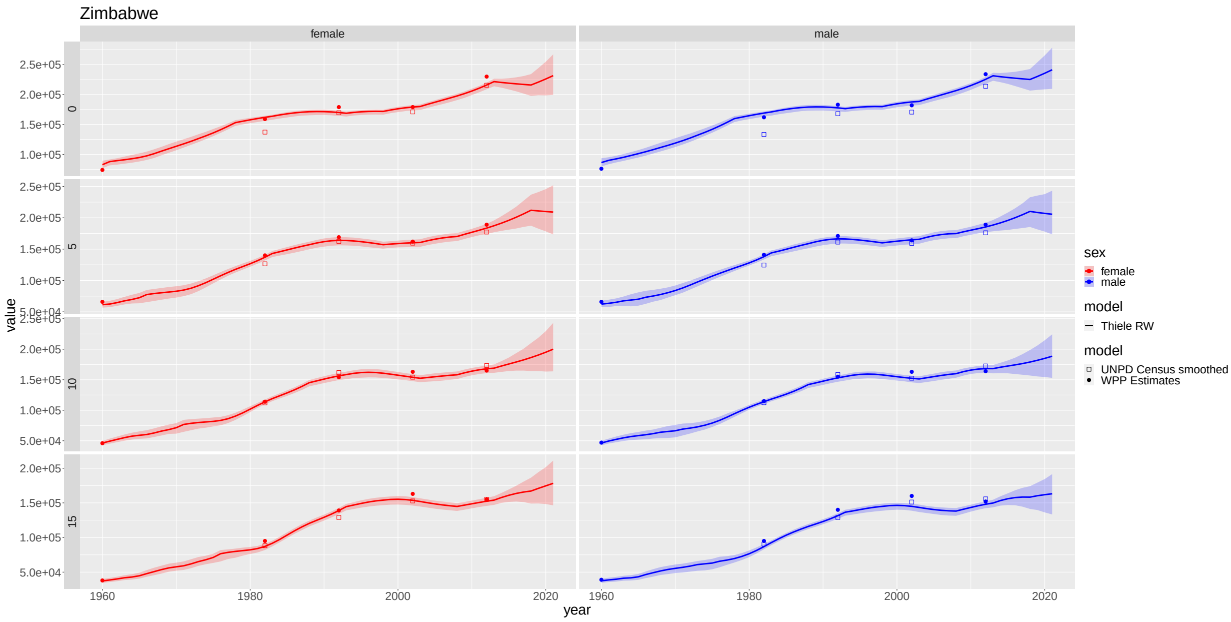

Figure 10: Population

Figure 11: Population

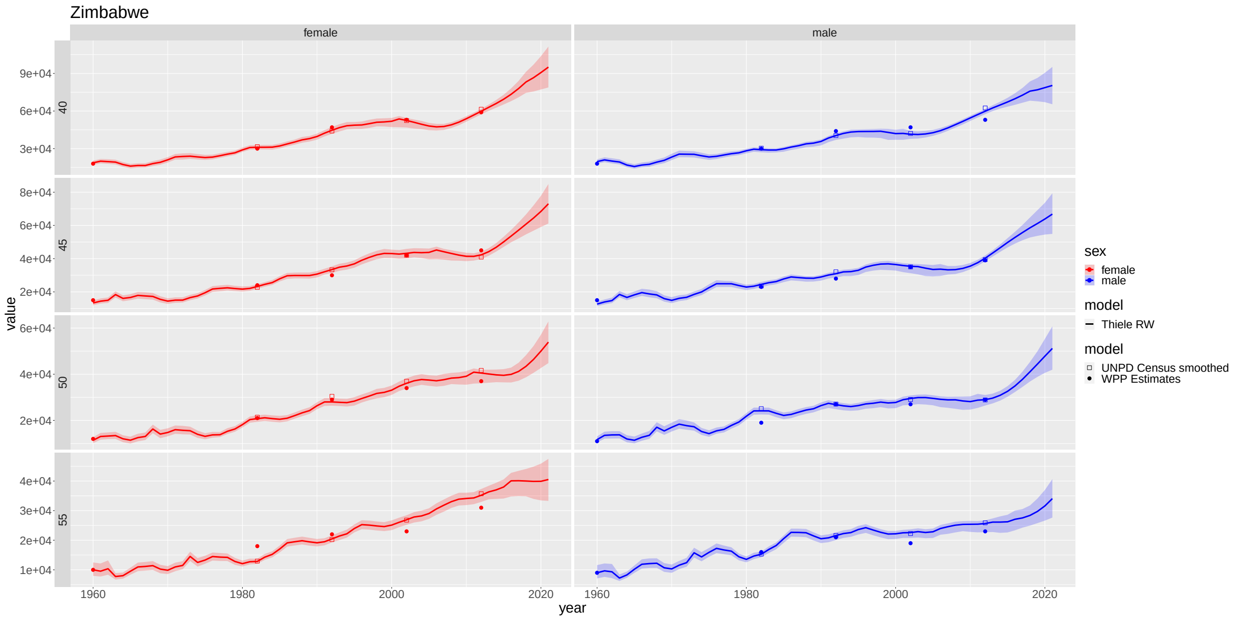

Figure 12: Population

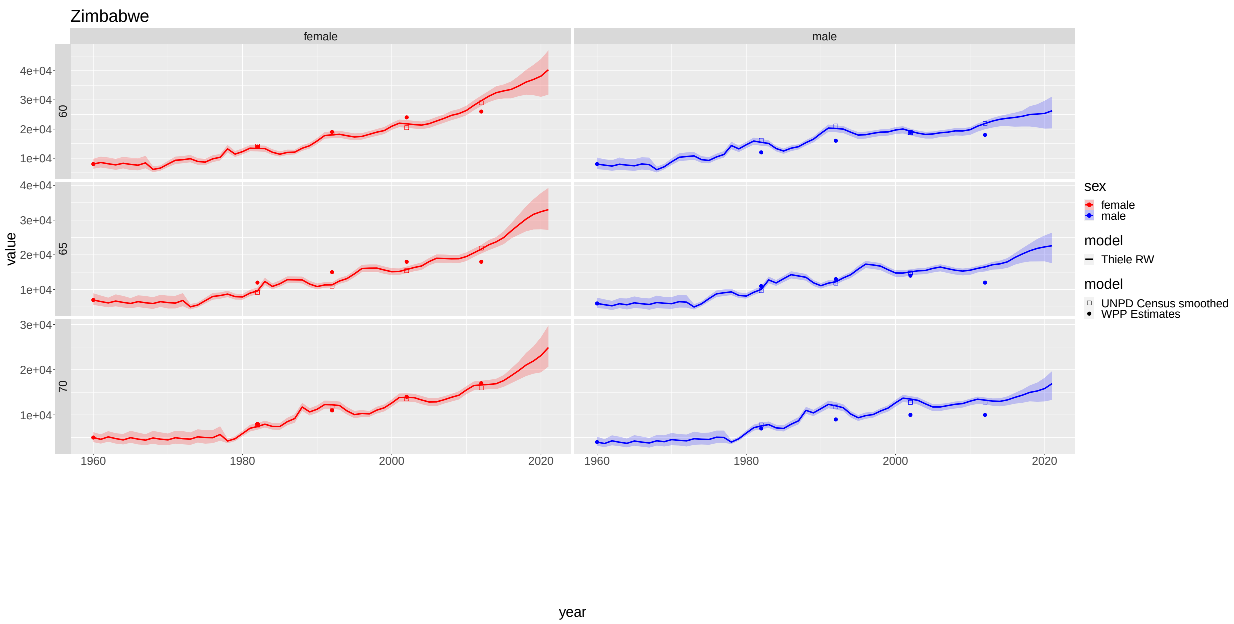

Figure 13: Population

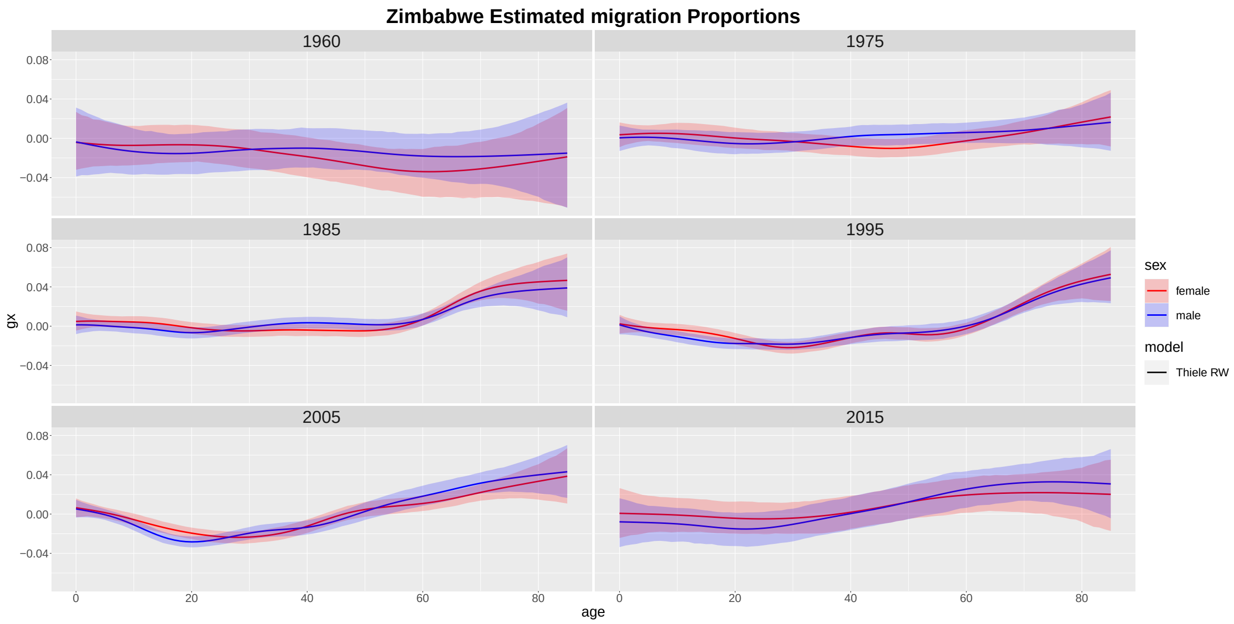

Figure 14: Migration

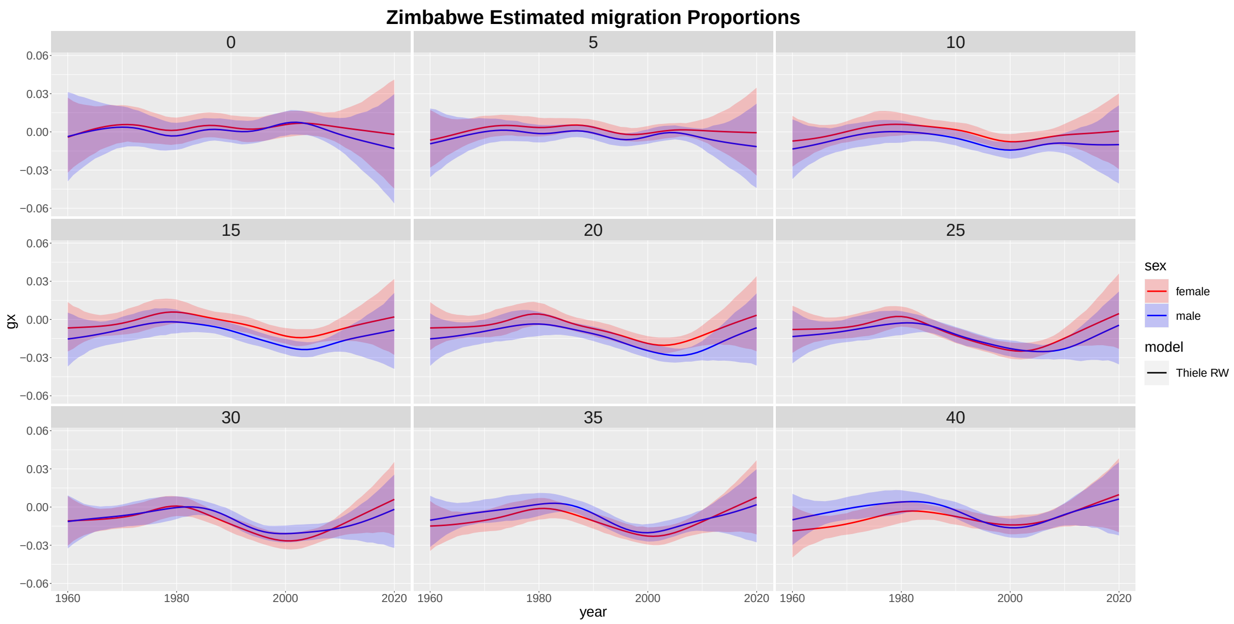

Figure 15: Migration

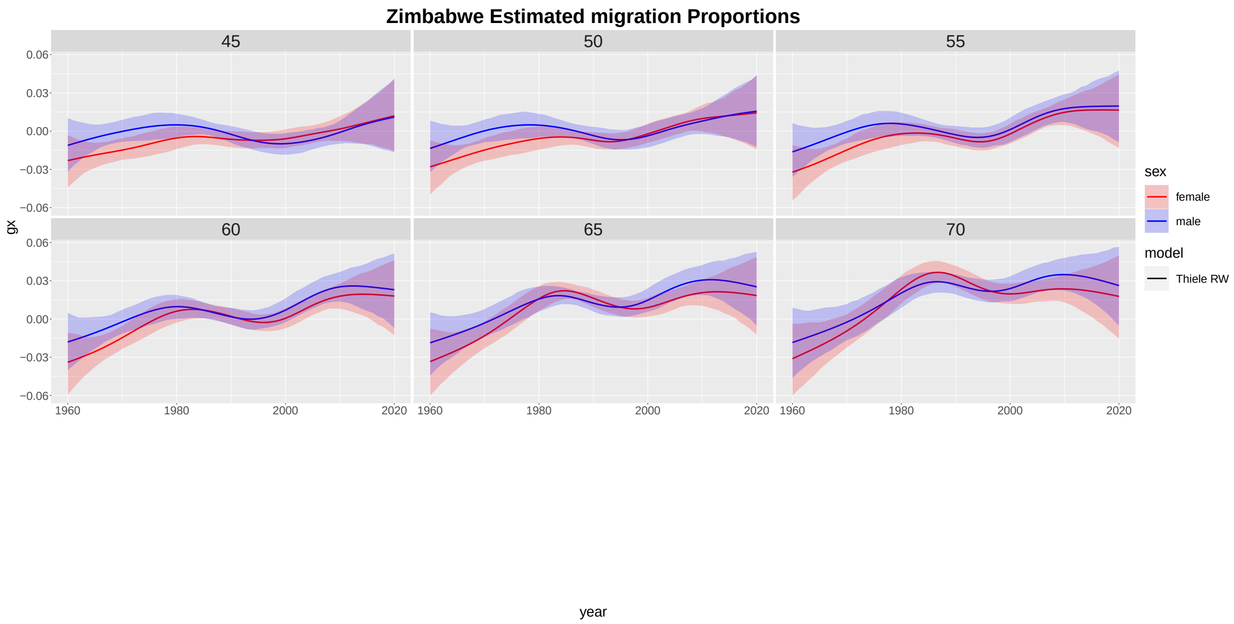

Figure 16: Migration

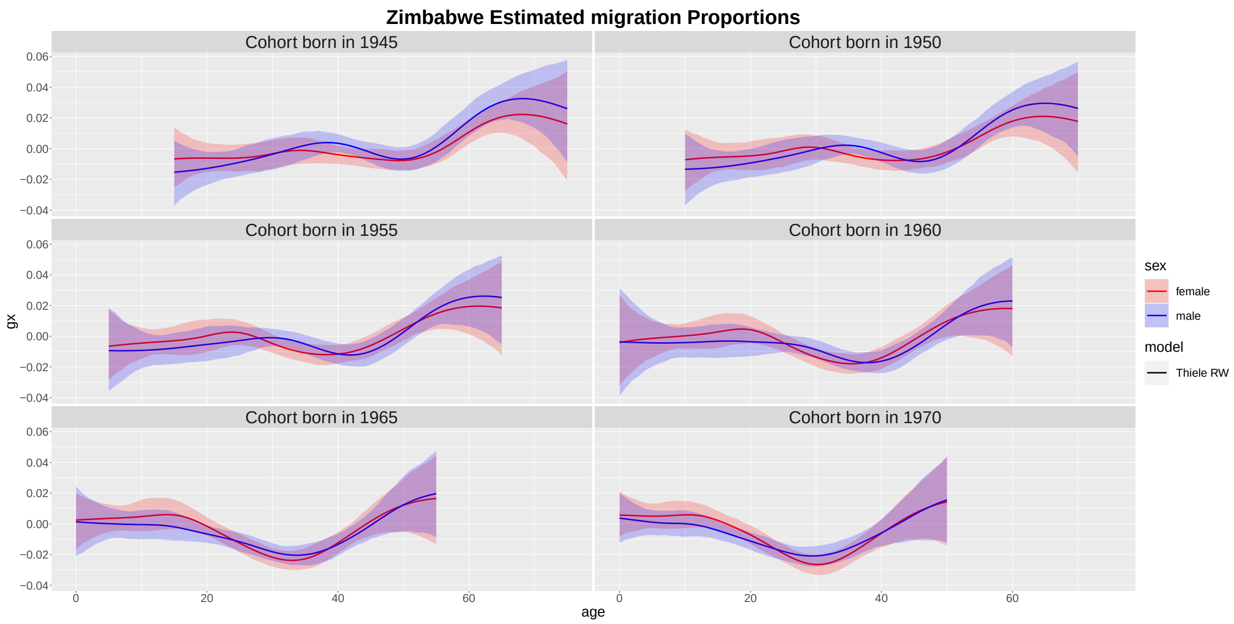

Figure 17: Migration

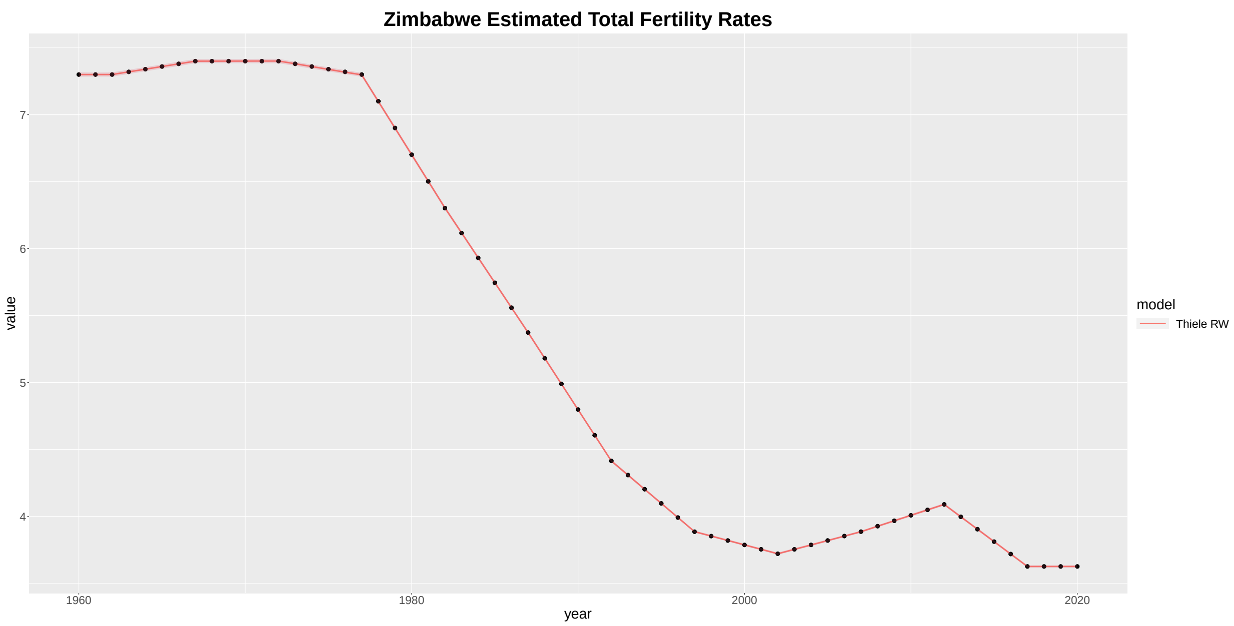

Figure 18: Total Fertility

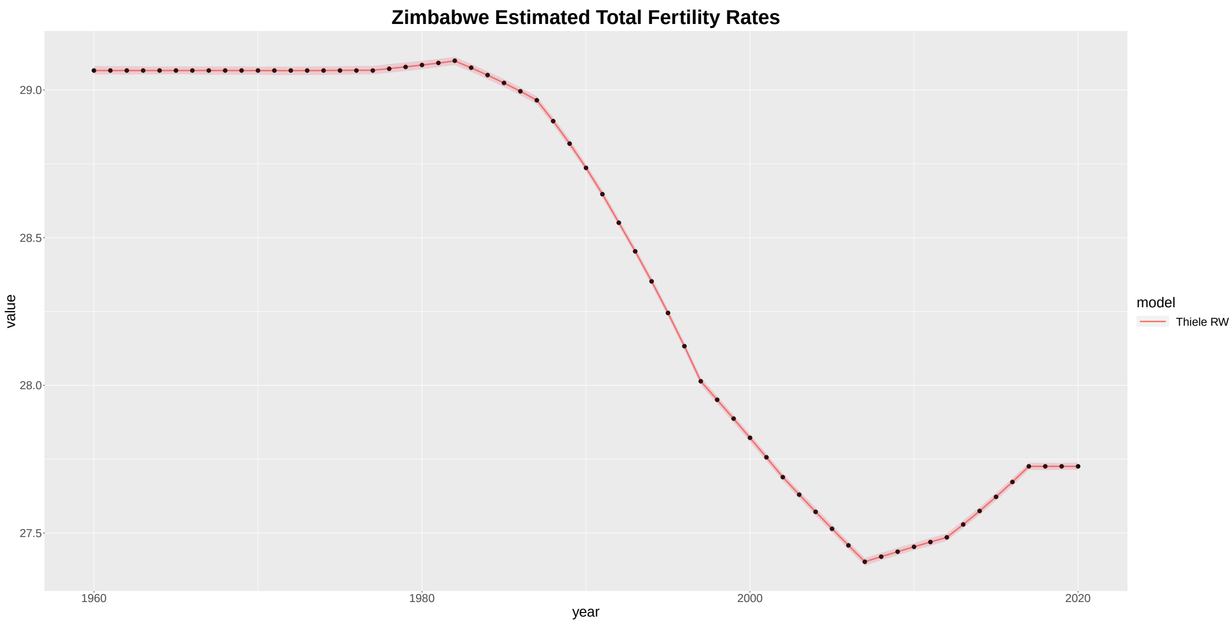

Figure 19: Mean age at births

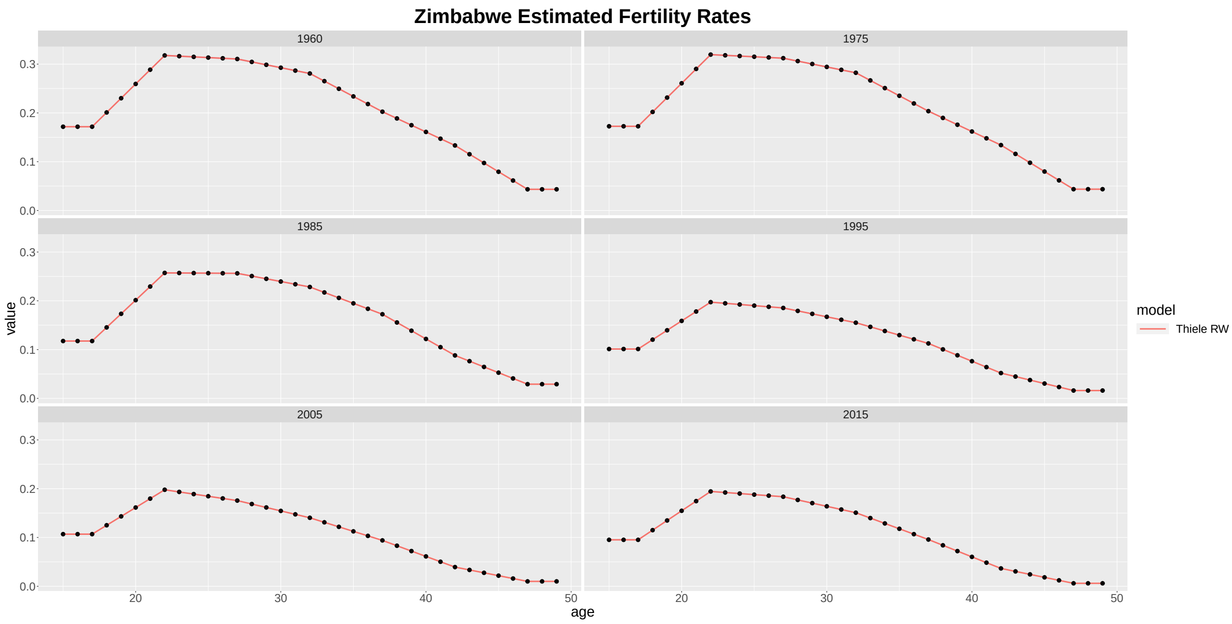

Figure 20: Fertility

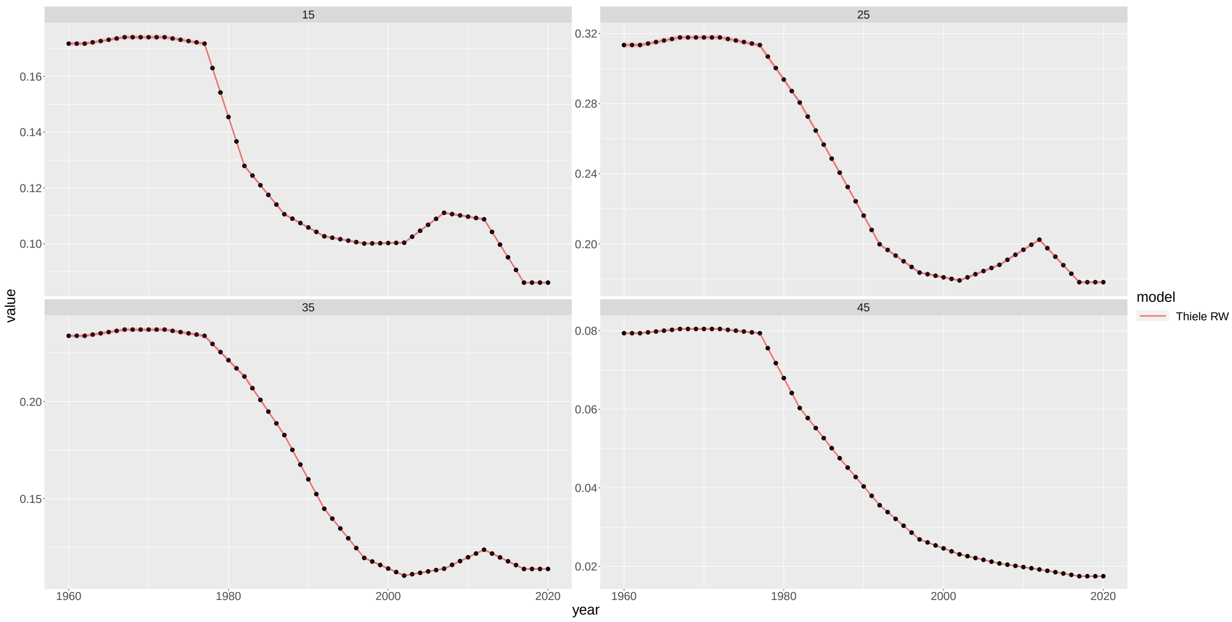

Figure 21: Fertility

Zimbabwe

```
## [1] "Census Females"
```

```
## # A tibble: 86 x 6
##      age `1969`  `1982`  `1992`  `2002`  `2012`
##    <dbl>  <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
##  1     0   215. 137199. 169638. 170997. 215623.
##  2     1   200. 134655. 159190. 172117. 211071.
##  3     2   219. 133815. 158420. 168708. 198666.
##  4     3   229. 131120. 158291. 164925. 190736.
##  5     4   240. 128619. 160494. 162274. 183611.
##  6     5   251. 126617. 162425. 159507. 177197.
##  7     6   258. 125956. 164145. 156778. 173472.
##  8     7   261. 125351. 166412. 155377. 171651.
##  9     8   262. 124105. 167841. 153895. 169450.
## 10     9   259. 118378. 165018. 153789. 171475.
## # ... with 76 more rows
```

```
## [1] "Census Females 5-year"
```

```
## # A tibble: 18 x 2
##      age  `1969`
##    <dbl>   <dbl>
##  1     0 1100.
##  2     5 1252.
##  3    10 1181.
##  4    15  935.
##  5    20  709.
##  6    25  541.
##  7    30  437.
##  8    35  379.
##  9    40  303.
## 10    45  227.
## 11    50  168.
## 12    55  124.
## 13    60   95.4
## 14    65   67.1
## 15    70   35.5
## 16    75   17.0
## 17    80    9.26
## 18    85   13.4
```

```
## [1] "Census Males"
```

```
## # A tibble: 86 x 6
##      age `1969`  `1982`  `1992`  `2002`  `2012`
##    <dbl>  <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
##  1     0   238. 133357. 168079. 170637. 213895.
##  2     1   232. 129754. 158096. 172216. 210140.
##  3     2   231. 129163. 156962. 168433. 197420.
##  4     3   235. 127343. 157057. 164960. 189835.
##  5     4   242. 125798. 159247. 162197. 182613.
##  6     5   248. 124587. 161170. 159217. 176094.
```

```
##  7     6   255. 124714. 162716. 156209. 172218.
##  8     7   265. 124146. 164082. 154419. 170262.
##  9     8   271. 122838. 164791. 152264. 167883.
## 10     9   271. 117620. 161620. 152041. 170380.
## # ... with 76 more rows
```

```
## [1] "Census Males 5-year"
```

```
## # A tibble: 18 x 2
##      age  `1969`
##    <dbl>   <dbl>
##  1     0 1156.
##  2     5 1278.
##  3    10 1207.
##  4    15  931.
##  5    20  636.
##  6    25  459.
##  7    30  389.
##  8    35  341.
##  9    40  280.
## 10    45  229.
## 11    50  186.
## 12    55  142.
## 13    60   97.1
## 14    65   60.0
## 15    70   32.4
## 16    75   15.2
## 17    80    8.24
## 18    85    7.21
```

Thiele log-Normal Hump Spline

```
## [1] "relative convergence (4)"
```

```
##    log_tau2_logpop    log_tau2_logpop    log_tau2_logpop    log_tau2_logpop
##        4.206964445        6.003090142        4.362942390        5.740807419
##      log_lambda_fx      log_lambda_gx     log_dispersion     log_dispersion
##       10.545033588       10.874594772        1.192539799        1.184173835
##      log_lambda_tp           tp_slope        tp_params_5       tp_params_10
##        3.591096962       -0.004455636        0.226154565        0.426490737
##     log_lambda_phi     log_lambda_psi       log_lambda_A       log_lambda_B
##       13.740352556       12.768610775       13.693942532       12.041326087
##  log_lambda_lambda   log_lambda_delta log_lambda_epsilon
##        1.828135883        4.703624972        3.376302654
```

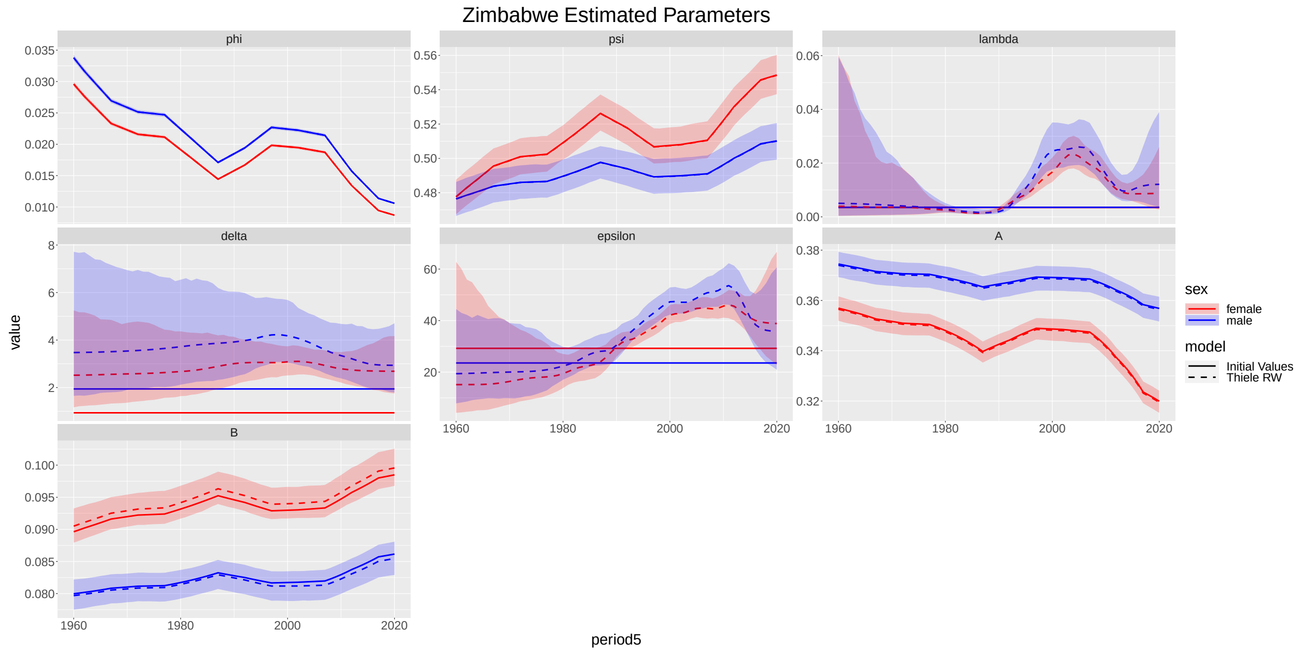

Figure 1: Estimated parameters

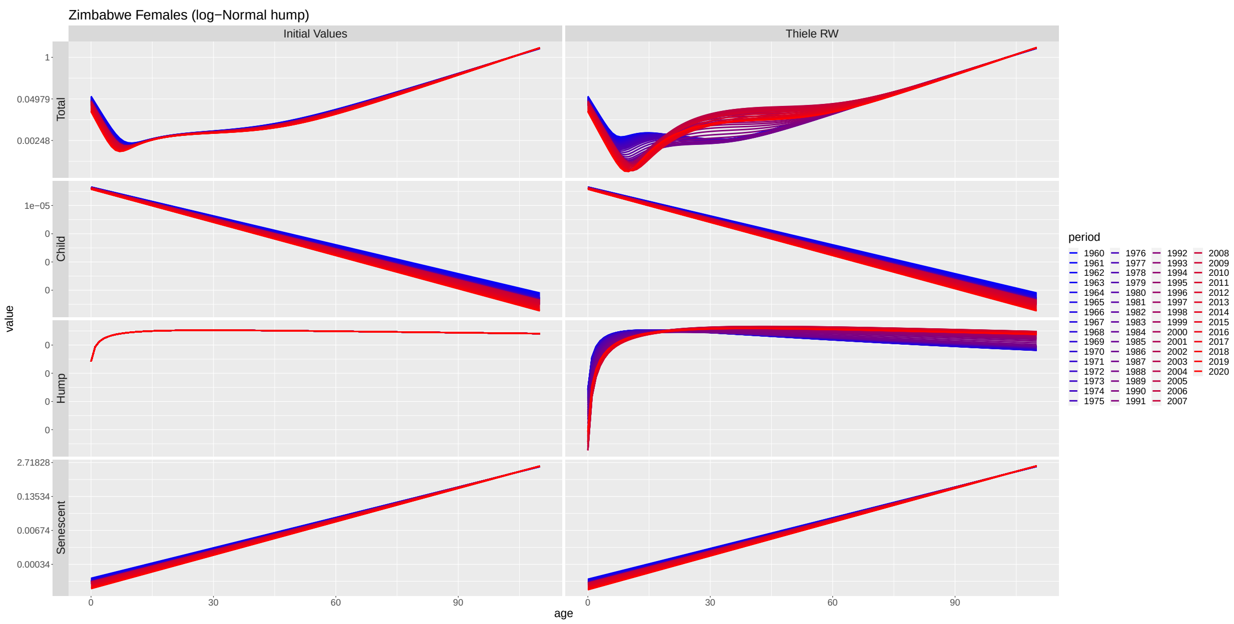

Figure 2: Thiele Decomposed

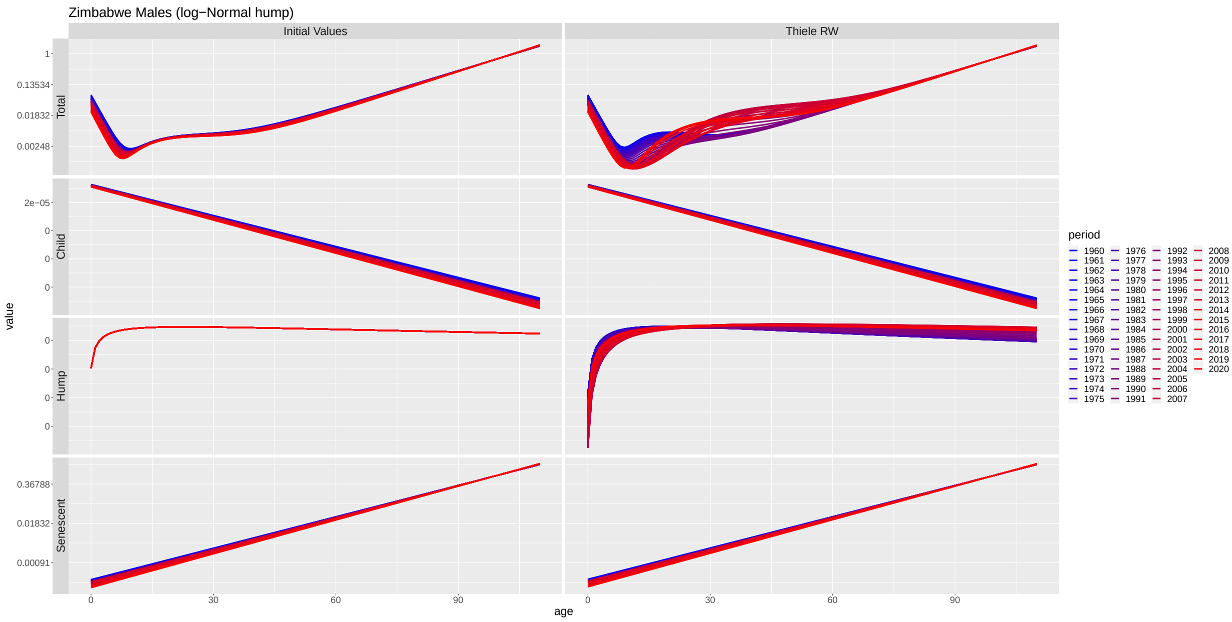

Figure 3: Thiele Decomposed

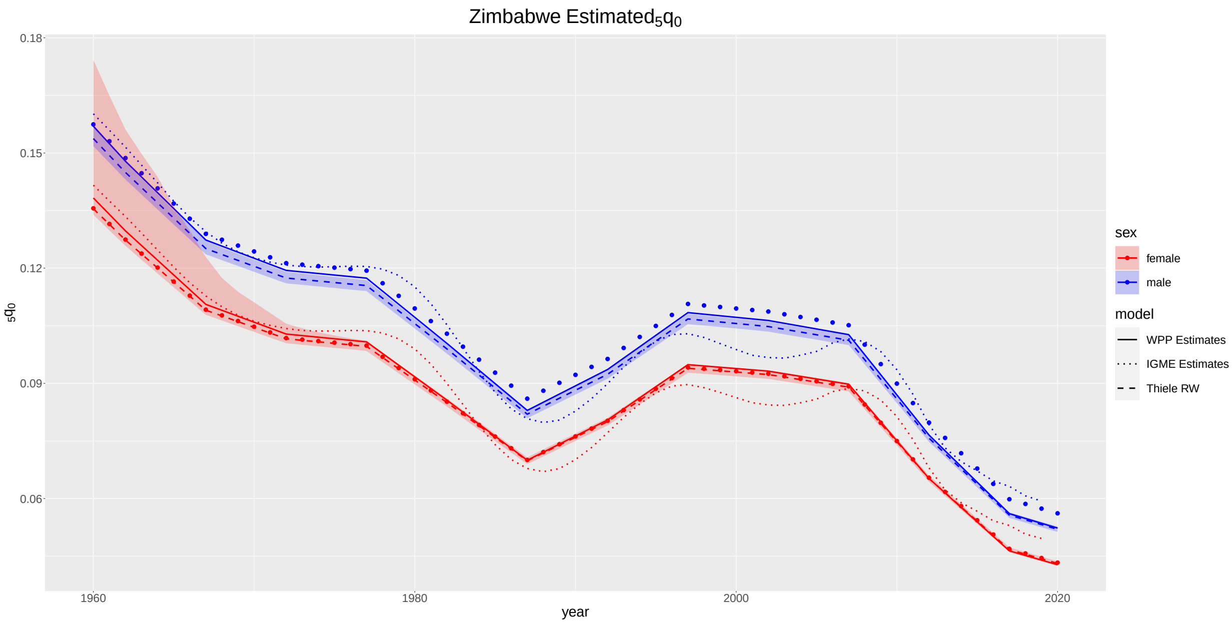

Figure 4: Estimated $_5q_0$

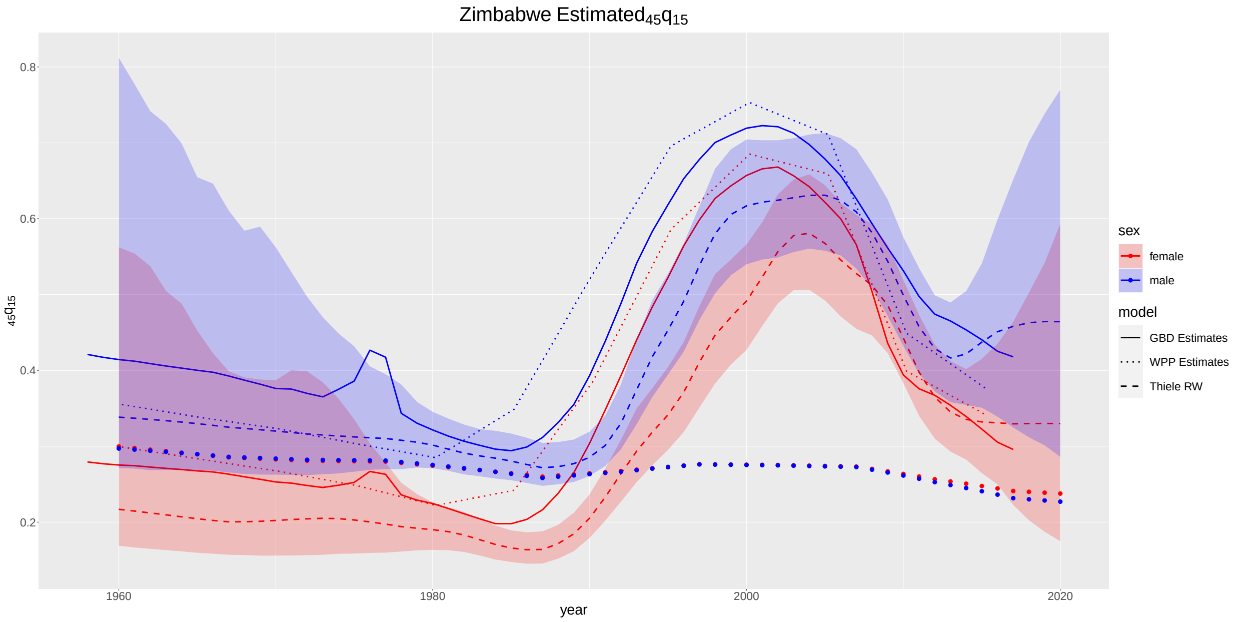

Figure 5: Estimated $_{45}q_{15}$

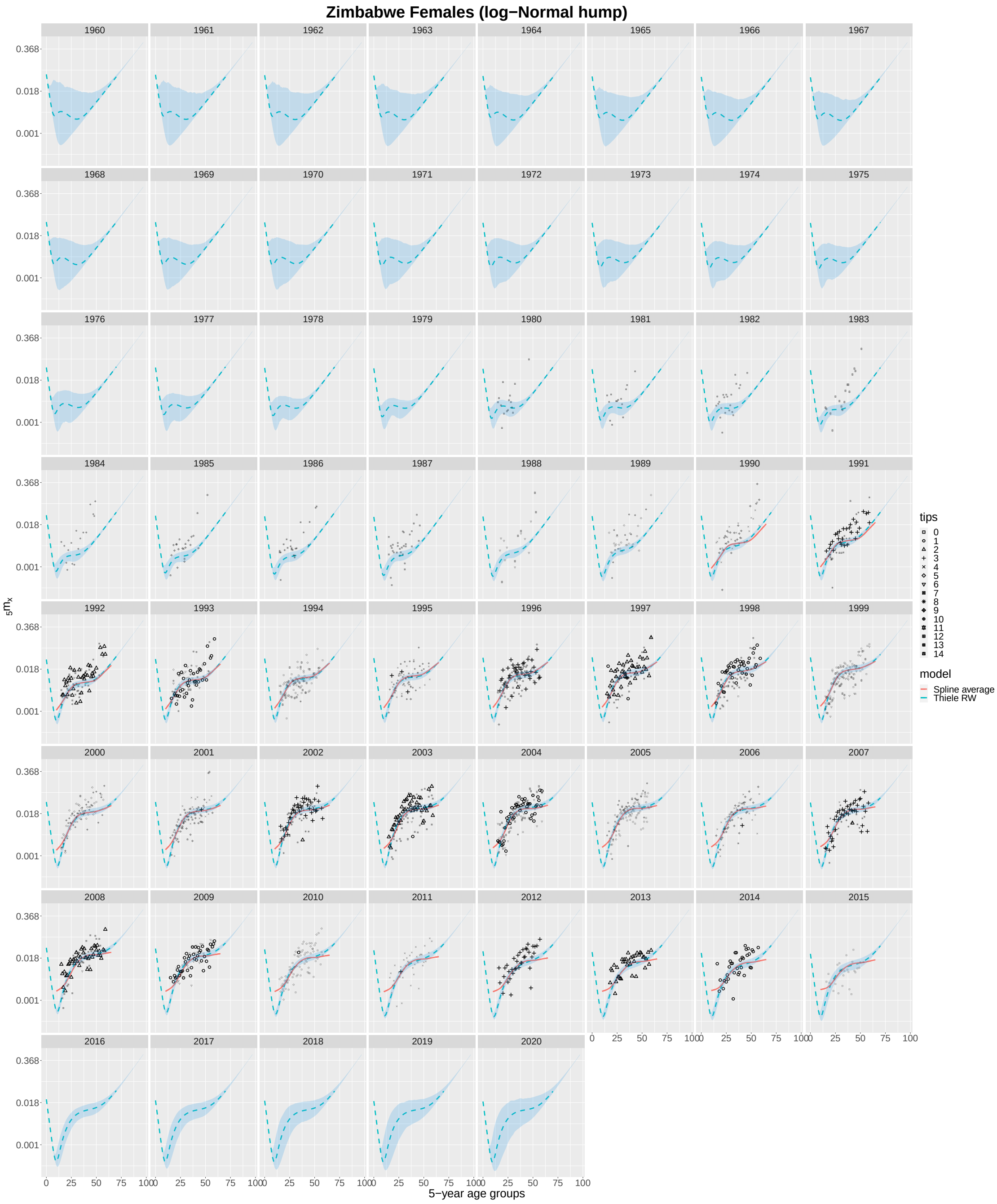

Figure 6: Mortality Schedules

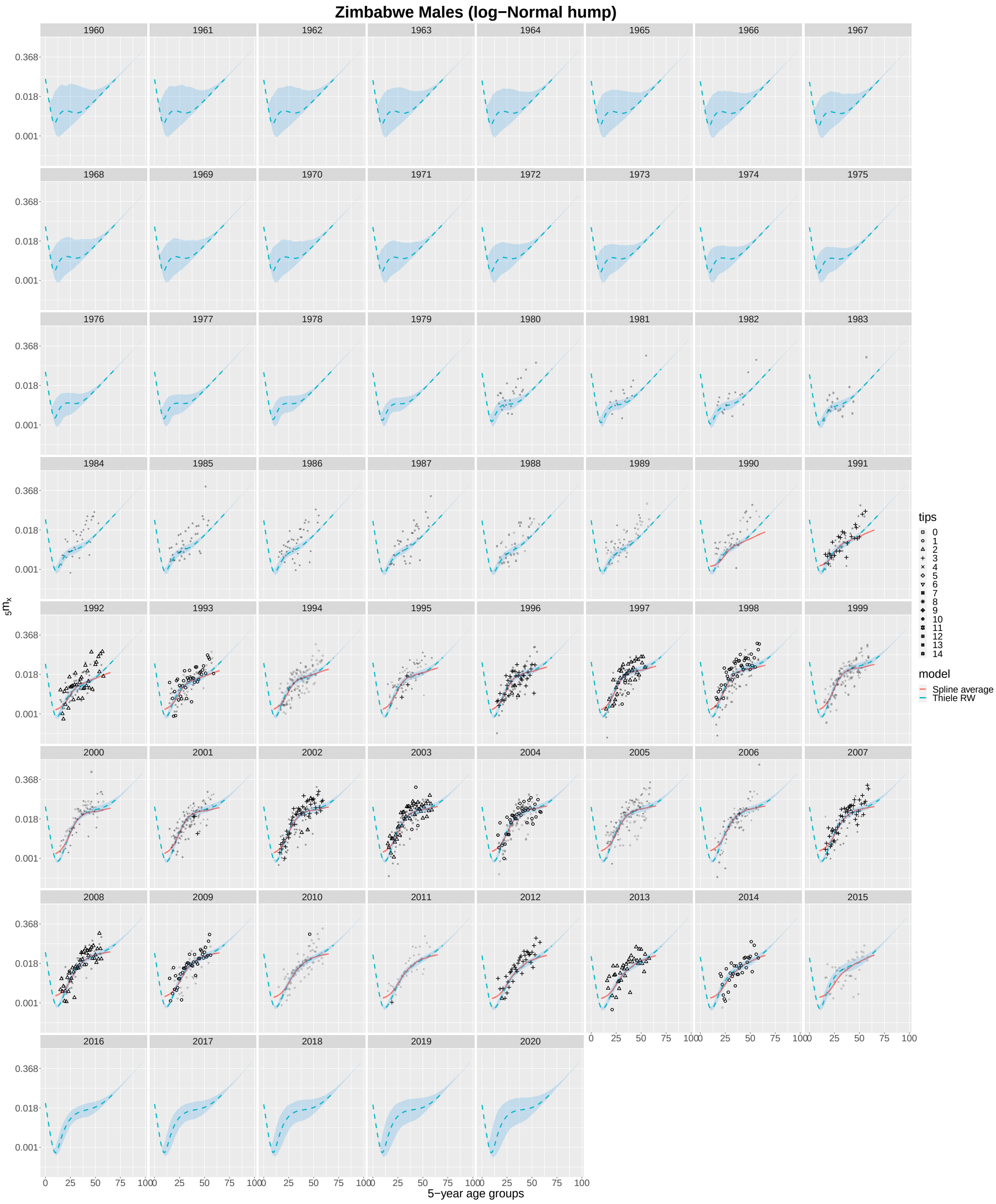

Figure 7: Mortality Schedules

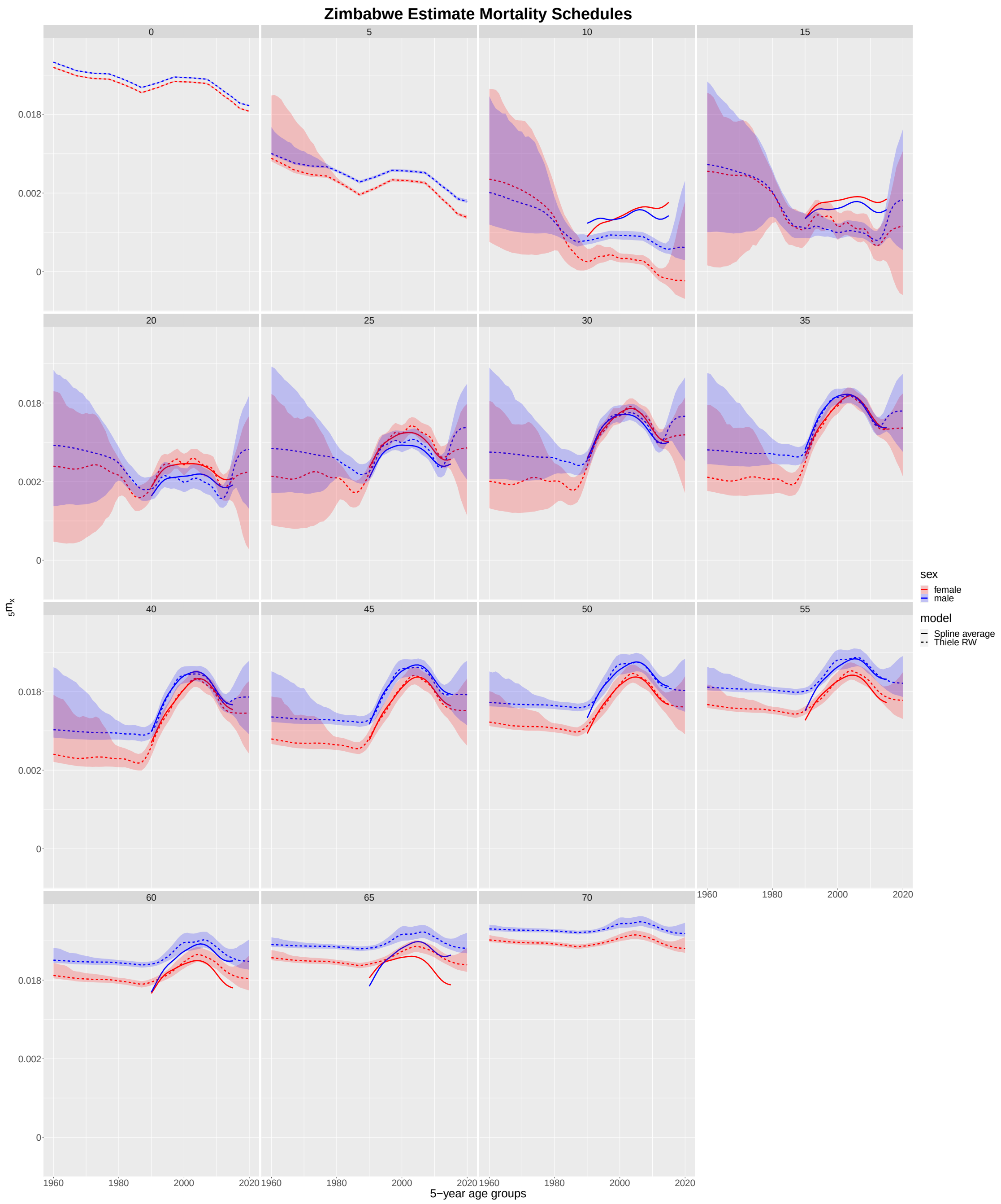

Figure 8: Mortality Schedules

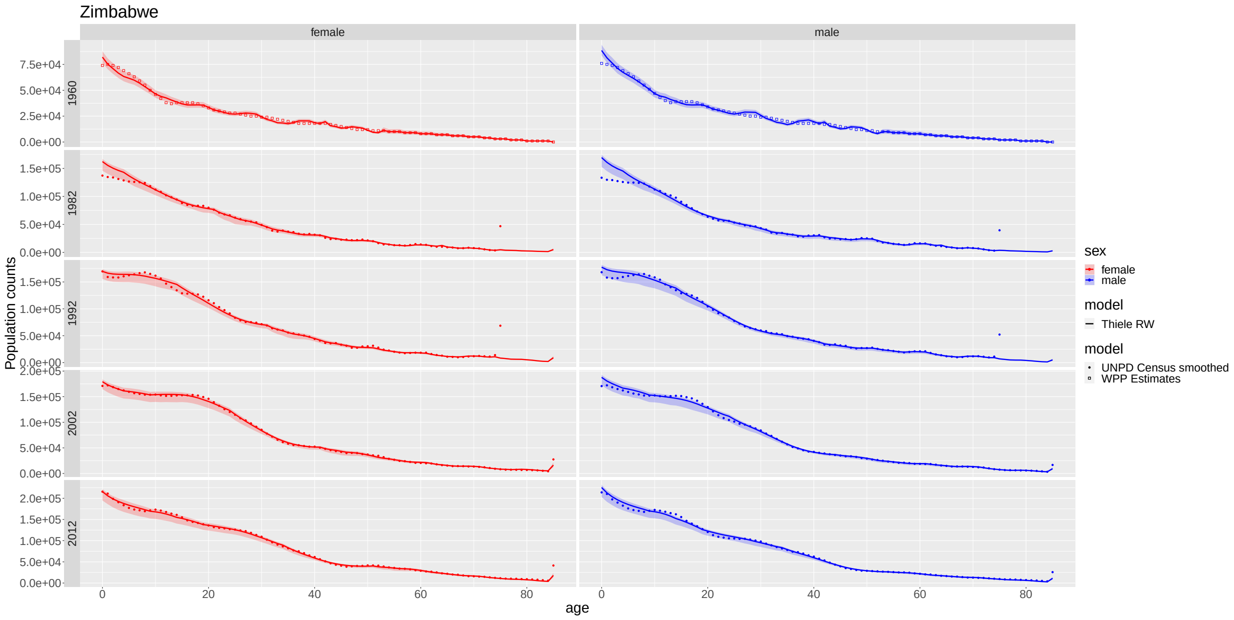

Figure 9: Population

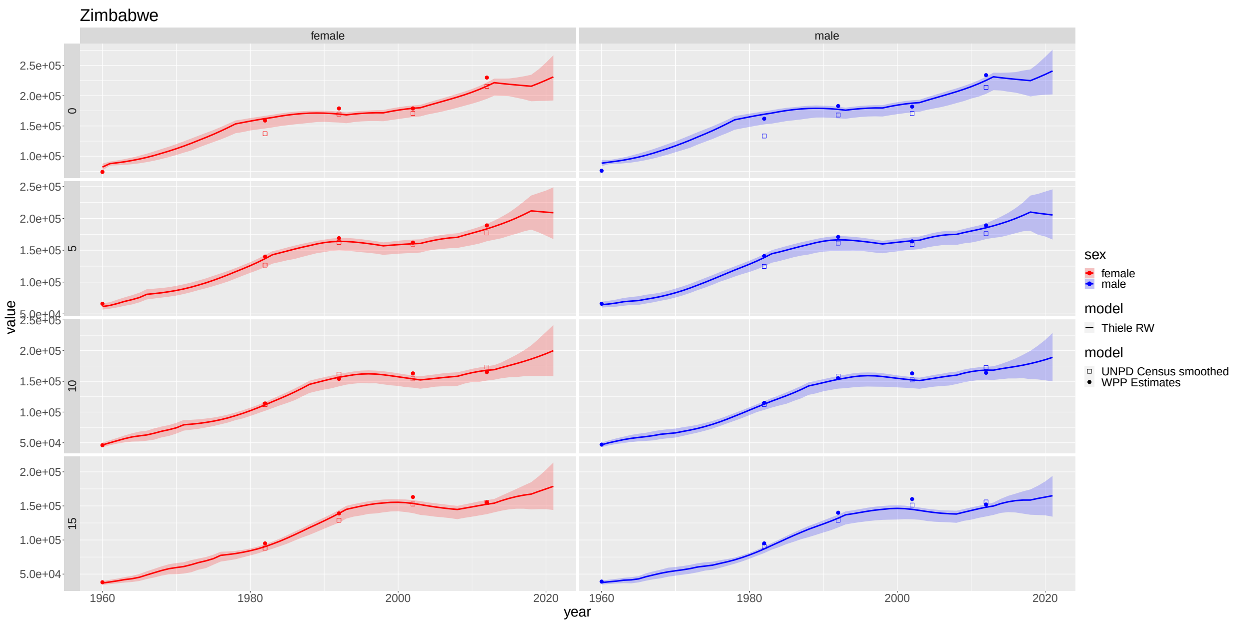

Figure 10: Population

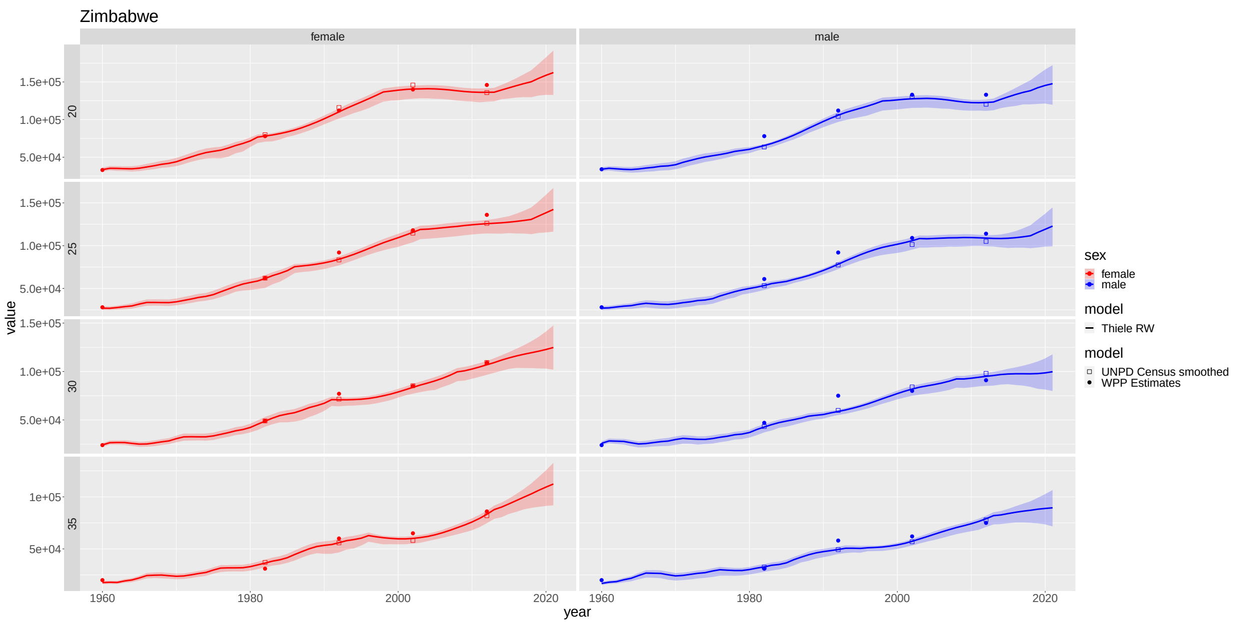

Figure 11: Population

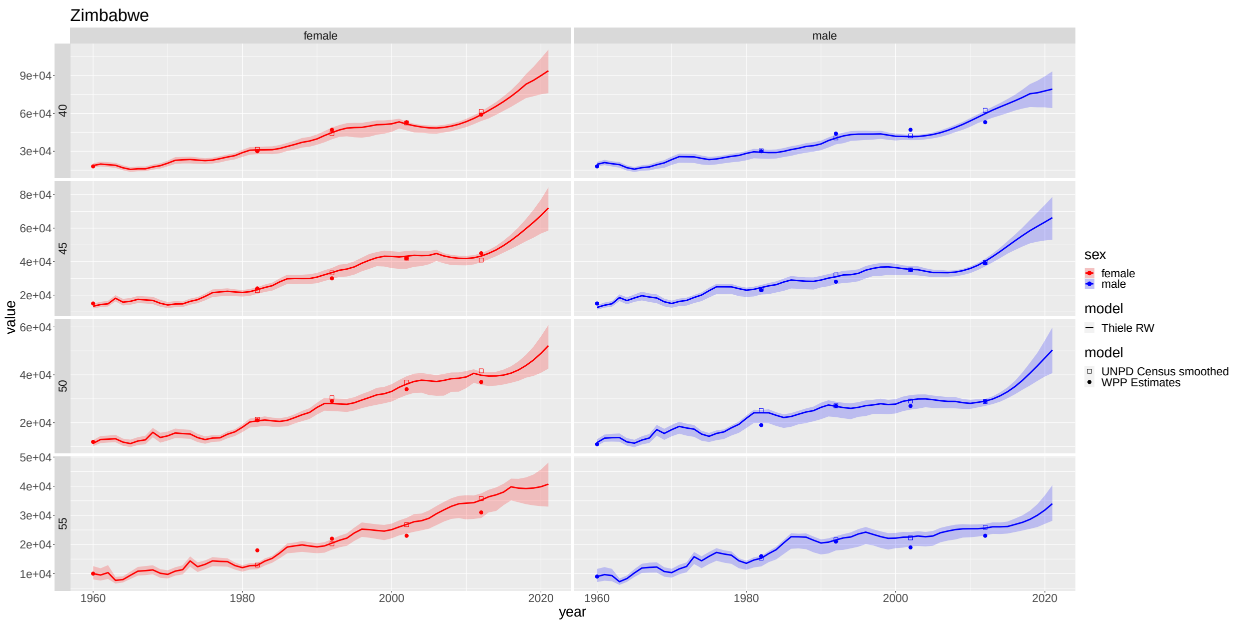

Figure 12: Population

Figure 13: Population

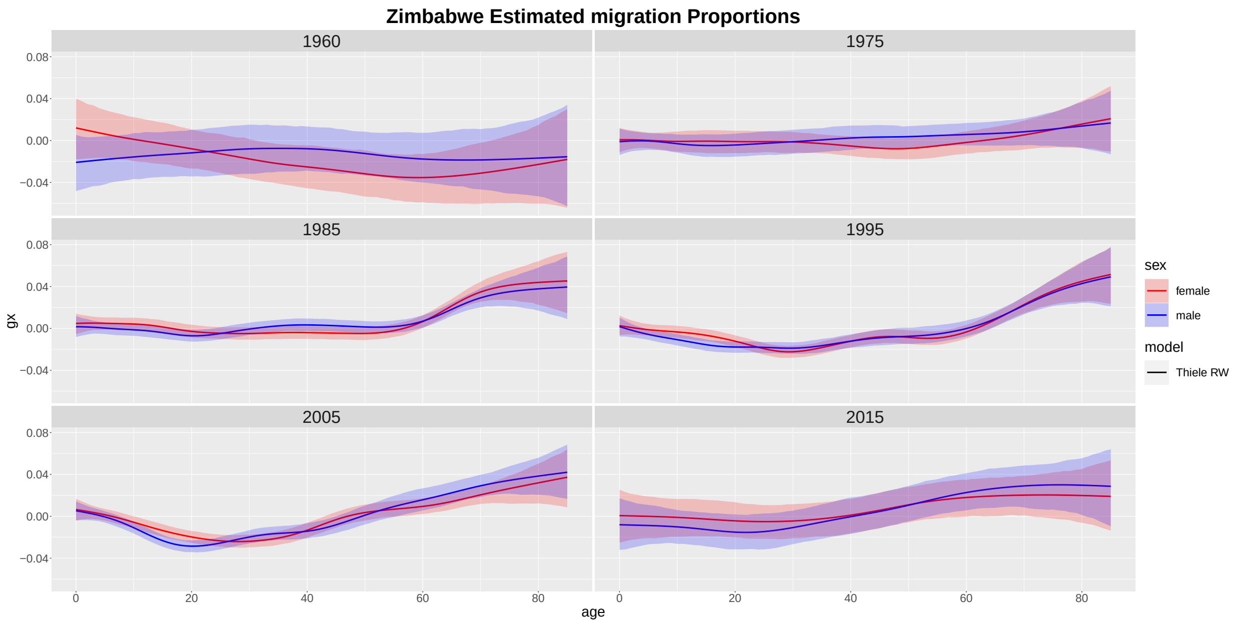

Figure 14: Migration

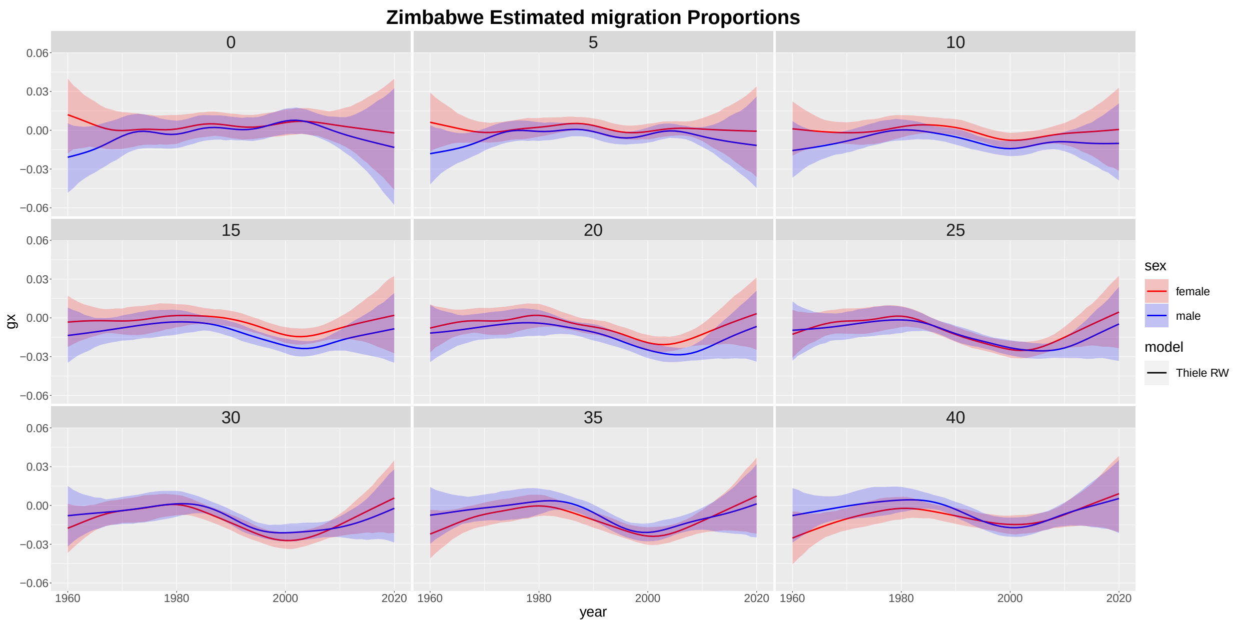

Figure 15: Migration

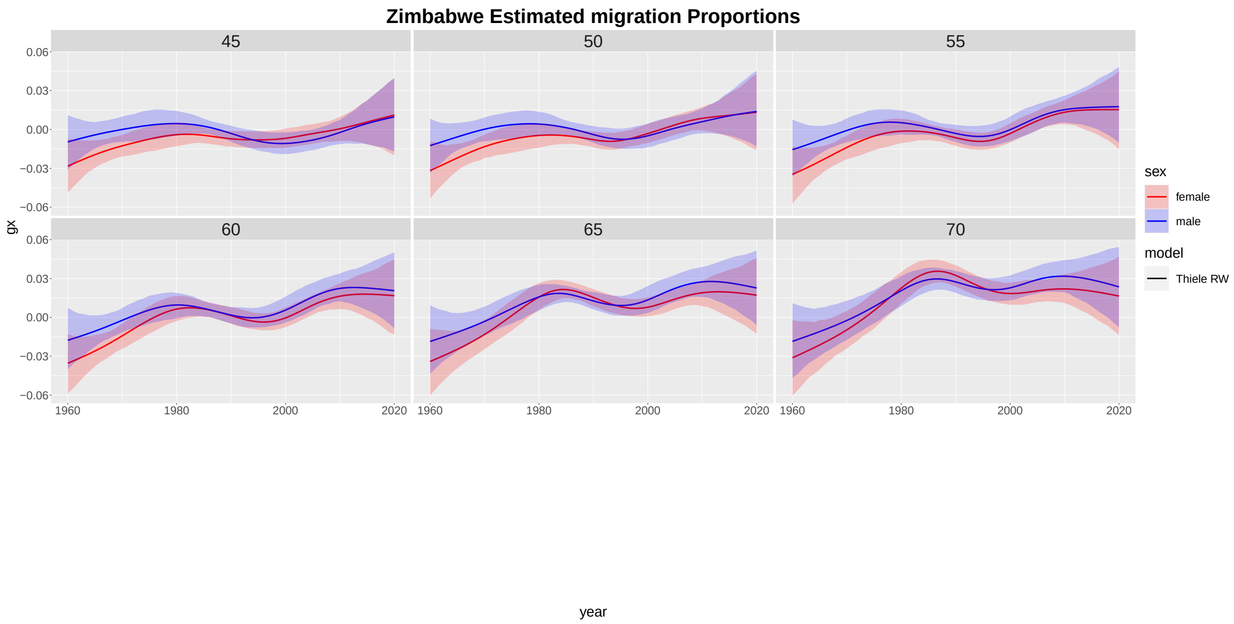

Figure 16: Migration

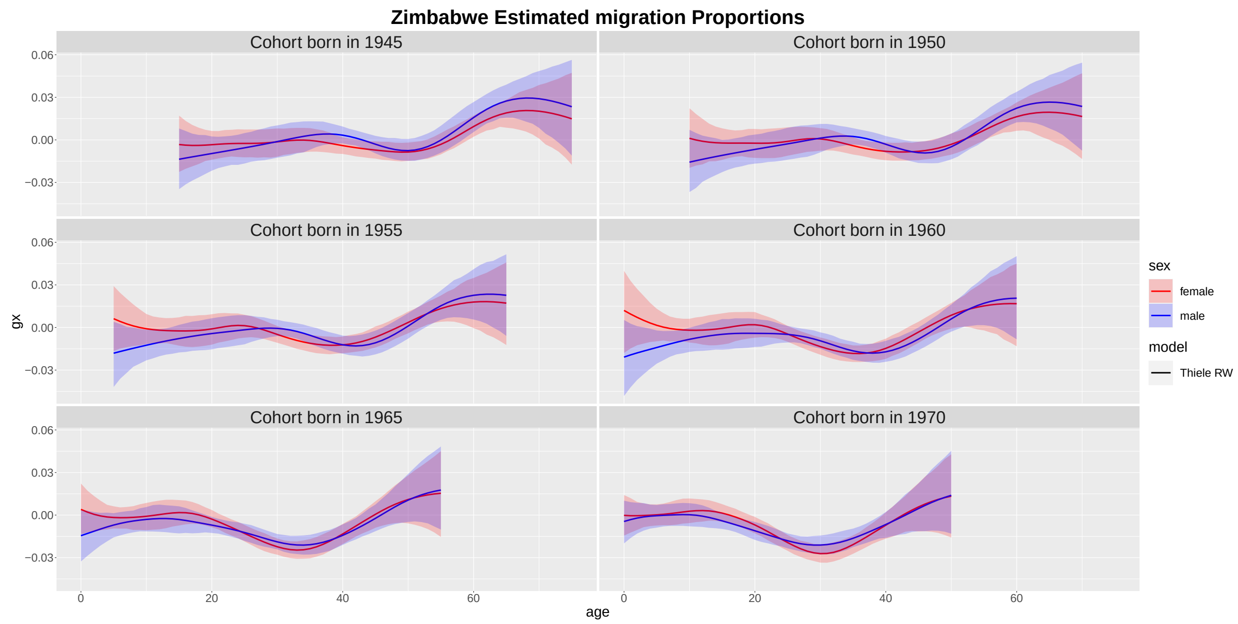

Figure 17: Migration

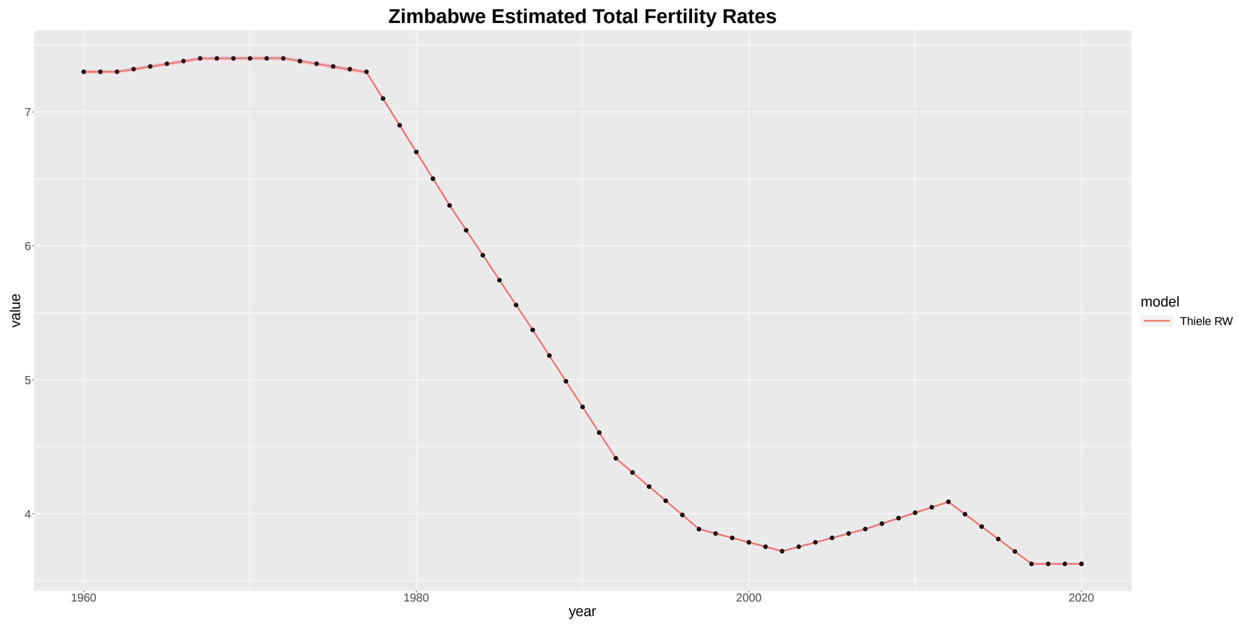

Figure 18: Total Fertility

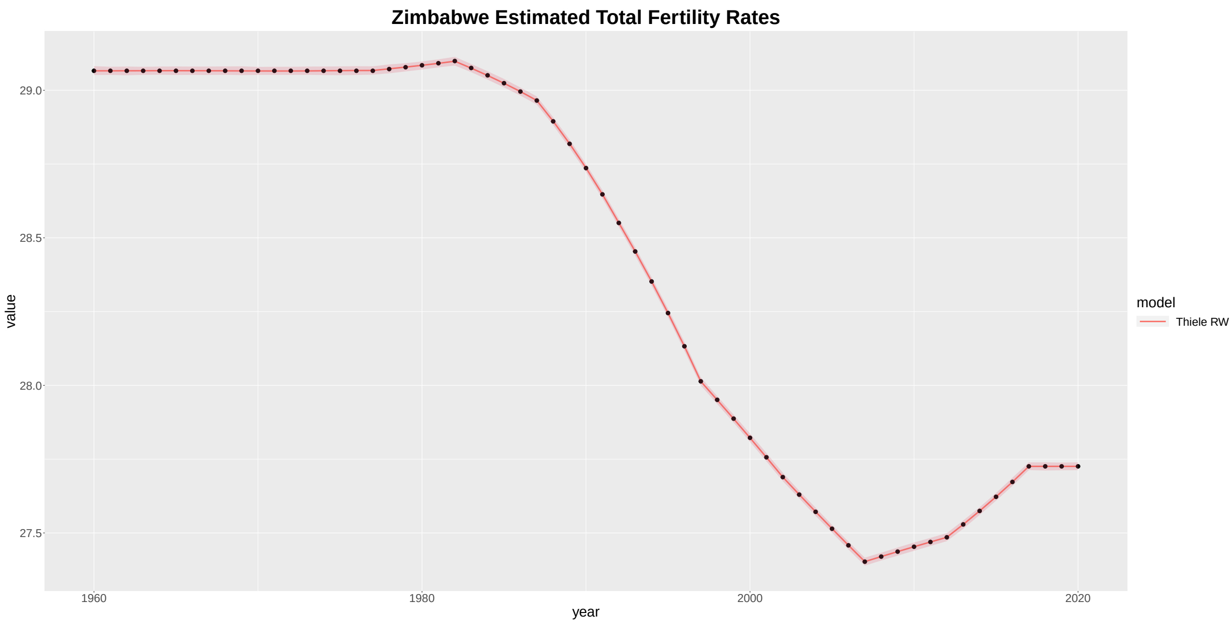

Figure 19: Mean age at births

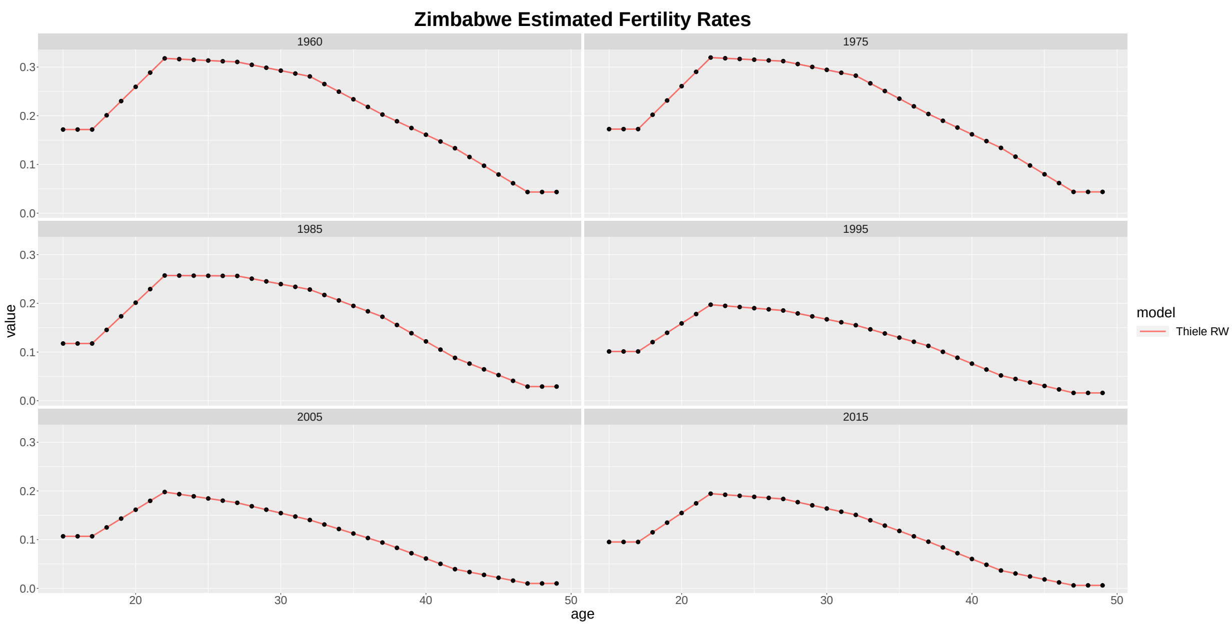

Figure 20: Fertility

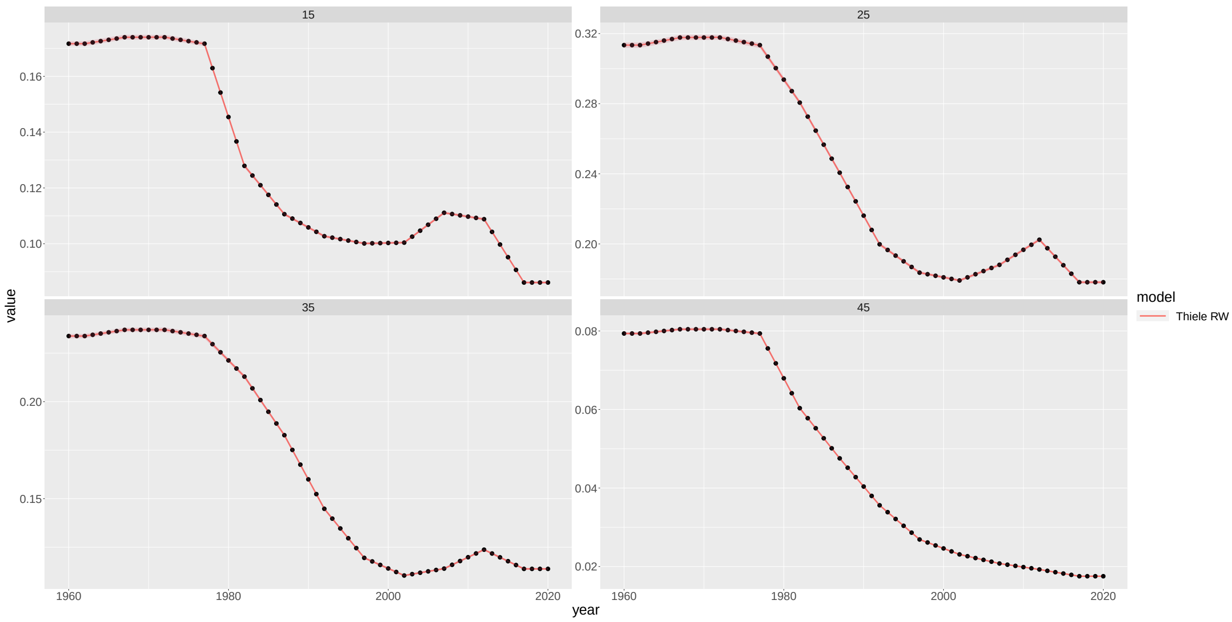

Figure 21: Fertility