Parameter risk in time-series mortality forecasts

Kleinow, Torsten

T.Kleinow@hw.ac.uk Department of Actuarial Mathematics and Statistics and the Maxwell Institute for Mathematics Sciences Heriot-Watt University, Edinburgh

Richards, Stephen J.

stephen@longevitas.co.uk Longevitas Ltd, Edinburgh

Longevitas Seminar - 30 June 2017







- Projecting mortality rates is an essential part of valuing liabilities in life-insurance portfolios and pension schemes.
- An important tool for risk-management and solvency purposes is a stochastic projection model for mortality:

- Projecting mortality rates is an essential part of valuing liabilities in life-insurance portfolios and pension schemes.
- An important tool for risk-management and solvency purposes is a stochastic projection model for mortality:
 - we acknowledge that there is uncertainty about projected mortality rates coming from unpredicted future changes in mortality rates (volatility)

- Projecting mortality rates is an essential part of valuing liabilities in life-insurance portfolios and pension schemes.
- An important tool for risk-management and solvency purposes is a stochastic projection model for mortality:
 - ▶ we acknowledge that there is uncertainty about projected mortality rates coming from unpredicted future changes in mortality rates (volatility) ⇒ mortality scenarios

- Projecting mortality rates is an essential part of valuing liabilities in life-insurance portfolios and pension schemes.
- An important tool for risk-management and solvency purposes is a stochastic projection model for mortality:
 - we acknowledge that there is uncertainty about projected mortality rates coming from unpredicted future changes in mortality rates (volatility) => mortality scenarios and probabilities for scenarios

- Projecting mortality rates is an essential part of valuing liabilities in life-insurance portfolios and pension schemes.
- An important tool for risk-management and solvency purposes is a stochastic projection model for mortality:
 - we acknowledge that there is uncertainty about projected mortality rates coming from unpredicted future changes in mortality rates (volatility) => mortality scenarios and probabilities for scenarios
 - but there is also uncertainty about the stochastic model itself and its parameters.

- Projecting mortality rates is an essential part of valuing liabilities in life-insurance portfolios and pension schemes.
- An important tool for risk-management and solvency purposes is a stochastic projection model for mortality:
 - we acknowledge that there is uncertainty about projected mortality rates coming from unpredicted future changes in mortality rates (volatility) => mortality scenarios and probabilities for scenarios
 - but there is also uncertainty about the stochastic model itself and its parameters.



Stochastic models often include: age effects, cohort effects and period effects

- Stochastic models often include: age effects, cohort effects and period effects
- Time series models are commonly used to project period and cohort effects and generate mortality scenarios.
- In particular, ARIMA processes and random-walks with drift are often used to generate scenarios for the period effects.

- Stochastic models often include: age effects, cohort effects and period effects
- Time series models are commonly used to project period and cohort effects and generate mortality scenarios.
- In particular, ARIMA processes and random-walks with drift are often used to generate scenarios for the period effects.
- While random walk models are the most widely used models, projections based on ARIMA models can look very different.

- Parameters for those models need to be estimated and, therefore, parameter risk becomes an issue.
- We consider parameter risk from the point of view of an insurer using stochastic models for regulatory risk reporting.
- Decomposing overall risk into undiversifiable trend risk (parameter uncertainty) and diversifiable volatility.

Questions for this Presentation

- How should time series be projected for mortality forecasts?
- What is the importance of different sources of uncertainty?
- Is goodness of fit a reliable criterion for choosing forecasting models?
- What impact does parameter instability have on projected mortality rates and solvency capital requirements?
- How do central projections compare to the CMI model and how can we set the long-term rate in the CMI model?

$$D_{x,t} \sim \text{Poisson}\left(\mu_{x,t}E_{x,t}^{c}\right)$$

For each calendar year y and age x we observe

- $D_{x,t}$: Number of deaths,
- $E_{x,t}^c$: Central exposure-to-risk
- $\mu_{x,t}$: force of mortality

$$D_{x,t} \sim \text{Poisson}\left(\mu_{x,t}E_{x,t}^{c}\right)$$

For each calendar year y and age x we observe

 $D_{x,t}$: Number of deaths, $E_{x,t}^c$: Central exposure-to-risk $\mu_{x,t}$: force of mortality Model for the force of mortality μ :

$$\log \mu_{x,t} = \alpha_x + \beta_x \kappa_t$$

with age effects α_x and β_x , and period effect κ_t .

Future liabilities in year t + h will depend on the number of deaths $D_{x,t+h}$.

$$D_{x,t+h} \sim \text{Poisson} \left(\mu_{x,t+h} E_{x,t+h}^{c} \right)$$
$$\log \mu_{x,t+h} = \alpha_{x} + \beta_{x} \kappa_{t+h}$$

Future liabilities in year t + h will depend on the number of deaths $D_{x,t+h}$.

$$D_{x,t+h} \sim \text{Poisson} \left(\mu_{x,t+h} E_{x,t+h}^{c} \right)$$

og $\mu_{x,t+h} = \alpha_{x} + \beta_{x} \kappa_{t+h}$

Sources of Uncertainty (given information up to current calendar year *t*):

- Poisson noise in year t + h
- uncertainty about estimated age parameters α_x , β_x

Future liabilities in year t + h will depend on the number of deaths $D_{x,t+h}$.

$$D_{x,t+h} \sim \text{Poisson} \left(\mu_{x,t+h} E_{x,t+h}^c \right)$$

og $\mu_{x,t+h} = \alpha_x + \beta_x \kappa_{t+h}$

Sources of Uncertainty (given information up to current calendar year t):

- Poisson noise in year t + h
- uncertainty about estimated age parameters α_x , β_x
- uncertainty about future values κ_{t+h} of period effect κ

To investigate the risk coming from the period effect we distinguish between

• the realised period effect κ_{t+h} in year t+h

To investigate the risk coming from the period effect we distinguish between

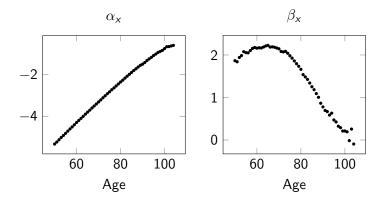
- the realised period effect κ_{t+h} in year t+h
- ▶ the predicted value of the period effect $\hat{\kappa}_t(h)$ for year t + h given information up to year t based on the estimated model for κ .

To investigate the risk coming from the period effect we distinguish between

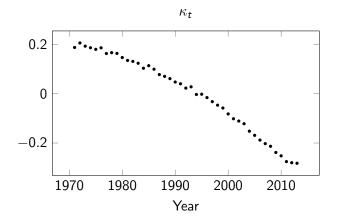
- the realised period effect κ_{t+h} in year t+h
- ▶ the predicted value of the period effect $\hat{\kappa}_t(h)$ for year t + h given information up to year t based on the estimated model for κ .

We will study the distribution of $\hat{\kappa}_t(h)$ and compare it to κ_{t+h} for an example data set (England & Wales).

Parameter estimates for Lee-Carter model fitted to mortality data for males in England & Wales aged 50–104 over the period 1971–2013.



Parameter estimates for Lee-Carter model fitted to mortality data for males in England & Wales aged 50–104 over the period 1971–2013.



Model for the period effect κ ($\epsilon_t \sim N(0, \sigma_{\epsilon}^2)$ i.i.d.):

$$\kappa_{t+1} = \kappa_t + \mu_0 + \epsilon_{t+1}$$

And the realised value h years ahead is given by

$$\kappa_{t+h} = \kappa_t + h\mu_0 + \sum_{j=1}^h \epsilon_{t+j}$$

What is unknown at time t?

Model for the period effect κ ($\epsilon_t \sim N(0, \sigma_{\epsilon}^2)$ i.i.d.):

$$\kappa_{t+1} = \kappa_t + \mu_0 + \epsilon_{t+1}$$

And the realised value h years ahead is given by

$$\kappa_{t+h} = \kappa_t + h\mu_0 + \sum_{j=1}^h \epsilon_{t+j}$$

What is unknown at time t?

 future error terms ε_{t+j}: Replaced by their expectation (Eε_{t+j} = 0) to obtain a central prediction for κ_{t+h} given information up to time t and the true drift parameter μ₀

Model for the period effect κ ($\epsilon_t \sim N(0, \sigma_{\epsilon}^2)$ i.i.d.):

$$\kappa_{t+1} = \kappa_t + \mu_0 + \epsilon_{t+1}$$

And the realised value h years ahead is given by

$$\kappa_{t+h} = \kappa_t + h\mu_0 + \sum_{j=1}^h \epsilon_{t+j}$$

What is unknown at time t?

- future error terms ε_{t+j}: Replaced by their expectation (Eε_{t+j} = 0) to obtain a central prediction for κ_{t+h} given information up to time t and the true drift parameter μ₀
- ► drift parameter μ_0 : Replaced with an estimate, $\hat{\mu}_0$ to obtain a forecast estimator *h* years ahead as

$$\hat{\kappa}_t(h) = \kappa_t + h\hat{\mu}_0 \tag{1}$$

We will use the standard estimator

$$\hat{\mu}_0 = \frac{1}{t-1} \sum_{i=2}^t (\kappa_i - \kappa_{i-1}) = \frac{\kappa_t - \kappa_1}{t-1}$$
(2)

with variance:

$$\operatorname{Var}(\hat{\mu}_0) = \operatorname{Var}\left(\frac{\kappa_t - \kappa_1}{t - 1}\right) = \frac{\sigma_{\epsilon}^2}{t - 1}$$
(3)

For our data we obtain $\hat{\mu}_0 = -0.011176$.

We will use the standard estimator

$$\hat{\mu}_0 = \frac{1}{t-1} \sum_{i=2}^t (\kappa_i - \kappa_{i-1}) = \frac{\kappa_t - \kappa_1}{t-1}$$
(2)

with variance:

$$\operatorname{Var}(\hat{\mu}_0) = \operatorname{Var}\left(\frac{\kappa_t - \kappa_1}{t - 1}\right) = \frac{\sigma_{\epsilon}^2}{t - 1}$$
(3)

For our data we obtain $\hat{\mu}_0 = -0.011176$. Our estimate of σ_{ϵ}^2 is the appropriate sample variance, σ_{ϵ}^2 : which gives $\hat{\sigma}_{\epsilon}^2 = 0.00011$ ($\hat{\sigma}_{\epsilon} = 0.010512$).

For risk management we are interested in

• the uncertainty about the realised κ_{t+h} , and

For risk management we are interested in

- ▶ the uncertainty about the realised κ_{t+h} , and
- ► the relative role of parameter uncertainty and uncertainty about future values e_{t+j} (volatility)

For risk management we are interested in

- the uncertainty about the realised κ_{t+h} , and
- ► the relative role of parameter uncertainty and uncertainty about future values e_{t+j} (volatility)

Projection error:

$$\mathsf{E}\left[\left(\hat{\kappa}_{t}(h) - \kappa_{t+h}\right)^{2}\right] = \underbrace{\frac{h}{t-1}h\sigma_{\epsilon}^{2}}_{\substack{\text{parameter}\\\text{uncertainty}}} + \underbrace{h\sigma_{\epsilon}^{2}}_{\text{volatility}}$$
(4)

where the parameter uncertainty is the variance of $h\hat{\mu}_0$, i.e. $h^2 \text{Var}(\hat{\mu}_0)$.

For risk management we are interested in

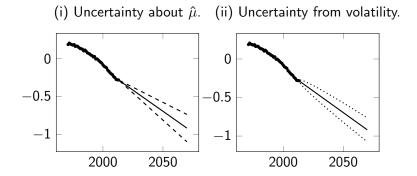
- the uncertainty about the realised κ_{t+h} , and
- ► the relative role of parameter uncertainty and uncertainty about future values ϵ_{t+j} (volatility)

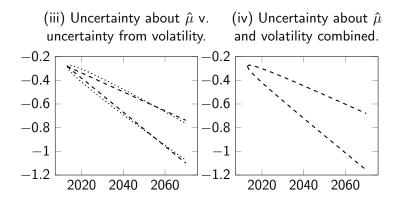
Projection error:

$$\mathsf{E}\left[\left(\hat{\kappa}_{t}(h) - \kappa_{t+h}\right)^{2}\right] = \underbrace{\frac{h}{t-1}h\sigma_{\epsilon}^{2}}_{\substack{\text{parameter}\\\text{uncertainty}}} + \underbrace{h\sigma_{\epsilon}^{2}}_{\text{volatility}}$$
(4)

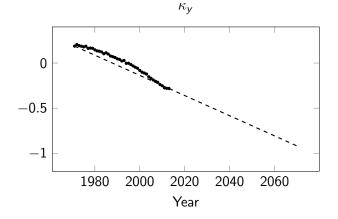
where the parameter uncertainty is the variance of $h\hat{\mu}_0$, i.e. $h^2 \text{Var}(\hat{\mu}_0)$.

Parameter uncertainty (variance of $h\hat{\mu}_0$) = $\frac{h}{t-1}$ Volatility

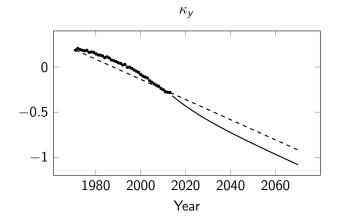




Predicted κ values $\hat{\kappa}_t(h)$ from RW model



Predicted κ values $\hat{\kappa}_t(h)$ from RW model and ARIMA(1,1,2) model.



Period Effect as ARIMA process

The structure of an ARIMA(p,1,q) process is just like the structure of a RW:

$$\kappa_{t+1} = \kappa_t + \mu + X_{t+1}^0$$

$$\kappa_{t+h} = \kappa_t + h\mu + \sum_{i=1}^h X_{t+i}^0$$

The structure of an ARIMA(p,1,q) process is just like the structure of a RW:

$$\kappa_{t+1} = \kappa_t + \mu + X_{t+1}^0$$

$$\kappa_{t+h} = \kappa_t + h\mu + \sum_{i=1}^h X_{t+i}^0$$

But with a different noise process:

$$X_t^0 = ar_1 X_{t-1}^0 + \ldots + ar_p X_{t-p}^0 + ma_1 \varepsilon_{t-1} + \ldots + ma_q \varepsilon_{t-q} + \varepsilon_t$$

where ε_t are i.i.d. normal.
In particular, an ARIMA(0,1,0) process is a random walk.

$$\kappa_{t+h} = \kappa_t + h\mu + \sum_{i=1}^h X_{t+i}^0$$

We define *h*-step ahead projections for κ as in the previous section, that is:

$$\hat{\kappa}_t(h) = \kappa_t + \sum_{i=1}^h \hat{X}_t^0(i) + h\hat{\mu}$$

h-step ahead projections for κ :

$$\hat{\kappa}_t(h) = \kappa_t + \sum_{i=1}^h \hat{X}_t^0(i) + h\hat{\mu}$$

For ARIMA(1,1,2) we obtain (all future noise terms ε are set to zero)

$$i = 1: \qquad \hat{X}_{t}^{0}(1) = \hat{a}r_{1}X_{t}^{0} + \hat{m}a_{1}\varepsilon_{t} + \hat{m}a_{2}\varepsilon_{t-1}$$

$$i = 2: \qquad \hat{X}_{t}^{0}(2) = \hat{a}r_{1}\hat{X}_{t}^{0}(1) + \hat{m}a_{2}\varepsilon_{t}$$

$$i > 2: \qquad \hat{X}_{t}^{0}(i) = \hat{a}r_{1}^{i-2}\hat{X}_{t}^{0}(2)$$

As for the Random Walk, we estimate μ with

$$\hat{\mu} = \frac{1}{t-1} \sum_{i=2}^{t} (\kappa_i - \kappa_{i-1}) = \frac{\kappa_t - \kappa_1}{t-1}$$
(5)

As for the Random Walk, we estimate μ with

$$\hat{\mu} = \frac{1}{t-1} \sum_{i=2}^{t} (\kappa_i - \kappa_{i-1}) = \frac{\kappa_t - \kappa_1}{t-1}$$
(5)

Since X^0 is a stationary ARMA process, we have $E[\hat{\mu}] = \mu$.

As for the Random Walk, we estimate μ with

$$\hat{\mu} = \frac{1}{t-1} \sum_{i=2}^{t} (\kappa_i - \kappa_{i-1}) = \frac{\kappa_t - \kappa_1}{t-1}$$
(5)

Since X^0 is a stationary ARMA process, we have $E[\hat{\mu}] = \mu$. Variance of $\hat{\mu}$:

$$\operatorname{Var}(\hat{\mu}) = \frac{\operatorname{Var}(X^0)}{t} + \frac{2}{t} \sum_{k=1}^{t-1} \gamma(k) \left[1 - \frac{k}{t} \right]$$
(6)

where $\gamma(k) = \text{Cov}(X_t^0, X_{t+k}^0)$ is the auto-covariance function of X^0 .

AICc (Akaike information criterion) values for various ARIMA(p, 1, q) models.

	q					
р	0	1	2	3		
0		-259.54				
1	-260.22	-257.88	-269.83	-267.14		
2	-258.10	-261.00	-267.14	-264.58		
3	-258.95	-262.60	-264.17	-261.29		

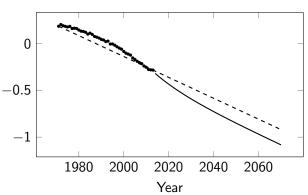
AICc (Akaike information criterion) values for various ARIMA(p, 1, q) models.

	q					
р	0	1	2	3		
0	-260.16	-259.54	-260.81	-262.78		
1	-260.22	-257.88	-269.83	-267.14		
2	-258.10	-261.00	-267.14	-264.58		
3	-258.95	-262.60	-264.17	-261.29		

Parameter estimates for ARIMA(1,1,2)

		Standard
Parameter	Estimate	error
ar_1	0.935	0.060
ma_1	-1.577	0.173
ma_2	0.815	0.149
σ_{ϵ}^2	0.000068	n/a
$\hat{\mu}$	-0.011	0.002

 κ values with RW and ARIMA(1,1,2) forecasts, $\hat{\kappa}_t(h)$.

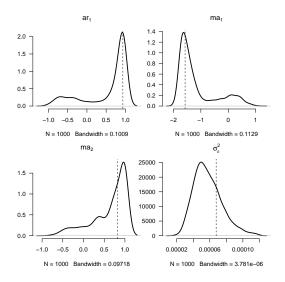


 κ_y

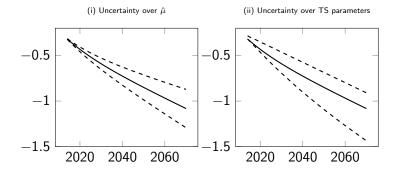
A Bootstrap method (Pascual et al. 2004)¹ is applied to study uncertainty about the Time series parameters. Idea:

- Simulate the past many times using estimated parameters and
- then re-estimate parameters for each simulated scenario.
- This gives an empirical distribution for the estimated parameters.

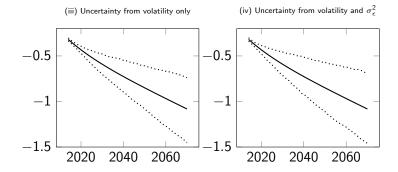
¹L. Pascual, J. Romo and E. Ruiz (2004): Bootstrap Predictive Inference for ARIMA Processes, Journal of Time Series Analysis 25(4)



Forecast κ values from ARIMA(1,1,2)



Forecast κ values from ARIMA(1,1,2)



Alternative ARIMA(1,1,0) process

Long term central projections for ARIMA(p,1,q) processes depend on AR terms and drift, but not on MA terms. But estimated parameter values and goodness of fit depend on

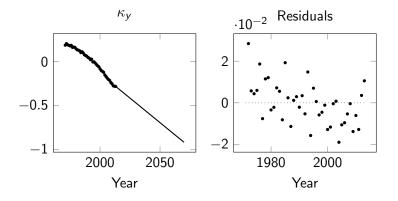
both, AR and MA terms.

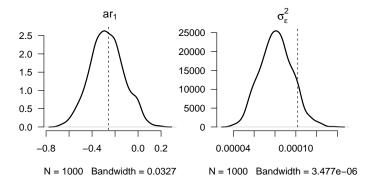
Estimated values for ARIMA(1,1,0):

	ARIMA(1,1,0)		ARIMA(1,1,2)		
Parameter	Estimate	Std. error	Estimate	Std. error	
ar_1	-0.259	0.166	0.935	0.060	
$\frac{\sigma_{\epsilon}^2}{\hat{\mu}}$	0.000102		0.000068		
$\hat{\mu}$	-0.011	0.002			

Impact on Central projections

 κ values with ARIMA(1,1,0) forecast and residuals from the ARIMA(1,1,0) fit





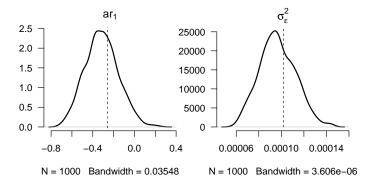


Figure: κ values with forecast from ARIMA(1,1,0) model with 95% bounds for various kinds of uncertainty.

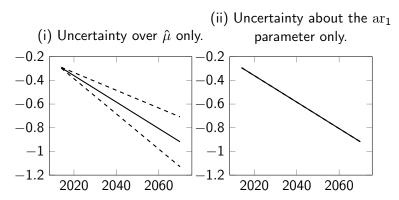
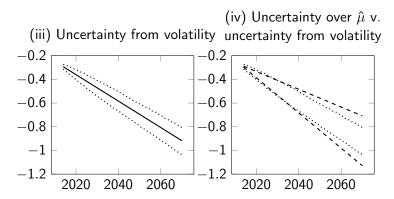


Figure: κ values with forecast from ARIMA(1,1,0) model with 95% bounds for various kinds of uncertainty.

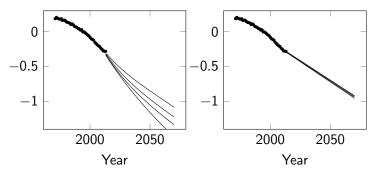


The importance of ar_1 can be seen when central projections are considered, i.e. where we set future error terms ε to zero.

h	RW and ARIMA(1,1,0)	ARIMA(1,1,2)
1	$X^0_t(1) = \mathit{ar}_1 X^0_t$	$X_t^0(1) = \operatorname{ar}_1 X_t^0 + \operatorname{ma}_1 \varepsilon_t + \operatorname{ma}_2 \varepsilon_{t-1}$
2	$X_t^0(2) = ar_1^2 X_t^0$	$X_t^0(2) = ar_1 X_t^0(1) + ma_2 \varepsilon_t$
> 2	$X_t^0(t)$	$h) = ar_1^{h-2}X_t^0(2)$

Random Walk: $ar_1 = 0$ ARIMA(1,1,0): $ar_1 = -0.259$ ARIMA(1,1,2): $ar_1 = 0.935$

Figure: Sensitivity of central projections when up to three years are removed from the end of the sample. Left panel: ARIMA(1,1,2) model. Right panel: ARIMA(1,1,0) model.



Capital Requirements

		Para.			VaR99.5	CTE99
Model	Volatility	uncert.	$ar{a}_{70:\overline{35}}^{50\%}$	$ar{a}^{99.5\%}_{70:\overline{35}}$	capital	capital
RW	Yes	No	12.50	12.70	1.62%	1.70%
	No	Yes	12.50	12.54	0.33%	0.34%
	Yes	Yes	12.49	12.72	1.79%	1.96%
(1,1,0)	Yes	No	12.51	12.68	1.36%	1.40%
	No	Yes	12.51	12.55	0.31%	0.34%
	Yes	Yes	12.51	12.69	1.43%	1.58%
(1,1,2)	Yes	No	12.59	12.87	2.25%	2.32%
	No	Yes	12.53	12.61	0.63%	0.64%
	Yes	Yes	12.53	12.87	2.70%	2.79%

interest rate: 2.5% p.a.

Capital Requirements

- Both Value-at-risk and CTE calculations are driven by the variability of mortality experience over a one-year horizon and how the model fit responds to this.
- Volatility makes the largest contribution: short time horizon for projections (1 year)
- The best fitting model, ARIMA(1,1,2), leads to
 - highest capital requirements
 - highest extra requirements for parameter uncertainty

Comparing Time Series Models

Random Walk

- standard model, most commonly used, very simple statistical methods required
- least flexible, few estimated parameters and strong assumptions (mortality improvements over time are i.i.d.)

Comparing Time Series Models

Random Walk

- standard model, most commonly used, very simple statistical methods required
- least flexible, few estimated parameters and strong assumptions (mortality improvements over time are i.i.d.)

ARIMA(1,1,2)

- best fitting model
- great uncertainty about parameters and projected rates leading to high capital requirements

Comparing Time Series Models

Random Walk

- standard model, most commonly used, very simple statistical methods required
- least flexible, few estimated parameters and strong assumptions (mortality improvements over time are i.i.d.)

ARIMA(1,1,2)

- best fitting model
- great uncertainty about parameters and projected rates leading to high capital requirements

ARIMA(1,1,0)

- similar goodness of fit as RW
- Iowest capital requirements
- weaker assumptions than RW allowing for structure in error terms

- Both, volatility and parameter uncertainty contribute to forecast uncertainty
- Parameter uncertainty is the driving force behind forecast uncertainty only for long forecast horizons

- Both, volatility and parameter uncertainty contribute to forecast uncertainty
- Parameter uncertainty is the driving force behind forecast uncertainty only for long forecast horizons
- The best fitting model, ARIMA(1,1,2), might not be the best model for projections since projected rates are not robust (one extra year of observed rates changes projected rates significantly)

- Both, volatility and parameter uncertainty contribute to forecast uncertainty
- Parameter uncertainty is the driving force behind forecast uncertainty only for long forecast horizons
- The best fitting model, ARIMA(1,1,2), might not be the best model for projections since projected rates are not robust (one extra year of observed rates changes projected rates significantly)

► ARIMA (1,1,0) vs. ARIMA(1,1,2): Although formulas (X_t⁰(h) = ar₁^{h-2}X_t⁰(2)) for projected mortality improvements are the same for projection horizons > 2, the inclusion of moving average terms leads to very different projected rates.

- Both, volatility and parameter uncertainty contribute to forecast uncertainty
- Parameter uncertainty is the driving force behind forecast uncertainty only for long forecast horizons
- The best fitting model, ARIMA(1,1,2), might not be the best model for projections since projected rates are not robust (one extra year of observed rates changes projected rates significantly)
- ► ARIMA (1,1,0) vs. ARIMA(1,1,2): Although formulas (X⁰_t(h) = ar^{h-2}₁X⁰_t(2)) for projected mortality improvements are the same for projection horizons > 2, the inclusion of moving average terms leads to very different projected rates.
- Choosing a model requires actuarial judgement taking objectives into account

Questions?