



ARC Centre of Excellence in Population Ageing Research

Working Paper 2016/04

The Impact of Systematic Trend and Uncertainty on Mortality and Disability in a Multi-State Latent Factor Model for Transition Rates

Zixi Li¹, Adam W. Shao² and Michael Sherris³

¹School of Risk and Actuarial Studies and Centre of Excellence in Population Ageing Research (CEPAR), UNSW Business School, UNSW Australia

²CEPAR, UNSW Business School, UNSW Australia, email:
wenqiang.shao@unsw.edu.au

³CEPAR, School of Risk and Actuarial Studies, UNSW Business School, UNSW Australia, email: m.sherris@unsw.edu.au

This paper can be downloaded without charge from the ARC Centre of Excellence in Population Ageing Research Working Paper Series available at www.cepar.edu.au

The Impact of Systematic Trend and Uncertainty on Mortality and Disability in a Multi-State Latent Factor Model for Transition Rates

Zixi Li, Adam W. Shao and Michael Sherris*

School of Risk and Actuarial Studies and ARC Centre of Excellence in Population Ageing Research (CEPAR), UNSW Australia.

February 8, 2016

Abstract

Multiple state functional disability models do not generally include systematic trend and uncertainty. We develop and estimate a multi-state latent factor intensity model with transition and recovery rates depending on a stochastic frailty factor to capture trend and uncertainty. We estimate the model parameters using U.S. Health and Retirement Study (HRS) data between 1998 and 2012 with Monte Carlo maximum likelihood estimation method. The model shows significant reductions in disability and mortality rates during this period and allows us to quantify uncertainty in transition rates arising from the stochastic frailty factor. Recovery rates are very sensitive to the stochastic frailty. There is an increase in expected future lifetimes as well as an increase in future healthy life expectancy. The proportion of lifetime spent in disability on average remains stable with no strong support in the data for either morbidity compression or expansion. The model has widespread application in costing of government funded aged care and pricing and risk management of LTC insurance products.

Keywords: long term care, systematic trend and uncertainty, disability, multi-state transitions, latent factor

JEL Classifications: C15, C23, G22, H51, I11

*[Corresponding author]. Email: m.sherris@unsw.edu.au; Postal address: School of Risk and Actuarial Studies, UNSW Business School, UNSW Australia, Kensington, NSW 2052, Australia; Phone: +61-2-9385 2333.

1 Introduction

Aged care, also referred to as long term care, is a significant and increasing cost as life expectancy increases. According to Australia's 2015 Intergenerational Report (Treasury of the Commonwealth of Australia, 2015), the number of Australians aged 65 and over will be more than doubled by 2055; aged care cost as a percentage of GDP will rise from 0.9% to 1.7% over the same period. By 2055, the number of people aged 15-64 for every person aged 65 and over in Australia will halve from 4.5 people today to 2.7 people. These demographic trends are also found in many other countries. Mortality improvement will have a significant impact on long term care costs, reflecting the extent of morbidity compression or expansion (e.g. Gruenberg, 1977, Fries, 1980, Manton, 1982).

Increasingly individuals are recognizing the importance of funding their own health care costs, as government budgets come under pressure. This highlights the important role that private Long Term Care (LTC) insurance can play in financing these costs. A viable long term care insurance market requires sophisticated models to project health care costs, to fairly price the products and to manage the associated risks. The aim of this paper is to propose and estimate a multiple state functional disability model with transition rates that include systematic trend and uncertainty, suitable for long term costing and assessing risks for LTC insurance.

Systematic improvement trends in mortality have been modeled at an aggregate population level. Among these models, the most widely used include the Lee-Carter model and its extensions (Lee, 2000; Lee and Carter, 1992), the Cairns-Blake-Dowd model and its variations (Cairns et al., 2006, 2011, 2009), the affine mortality models (Biffis, 2005; Dahl, 2004; Schrage, 2006), and the subordinated Markov model (Liu and Lin, 2012). There is evidence that systematic mortality trends vary by individual characteristics. Xu et al. (2015) model systematic mortality improvement with the Lee-Carter model using U.S. Health and Retirement

ment Study (HRS) data and consider a number of risk factors by classifying individuals into sub-populations. Mortality also varies by level of functional disability as shown in Fong et al. (2015) and Kwon and Jones (2008). Mortality improvement trends are expected to vary depending on level of functional disability. The incidence, recovery and mortality of disability will be impacted by systematic uncertainties which cannot be eliminated by pooling.

Multi-state Markov Chain models are widely used for LTC modeling. Olivieri and Pitacco (2001) have a single level of disability and Rickayzen and Walsh (2002) include two levels of disability. Fong et al. (2015) include recovery rates and Shao et al. (2015) apply the model to estimate premiums and solvency capital requirements for a wide range of LTC insurance products. Many prior studies estimate health transition rates based on cross-sectional prevalence data (Olivieri and Pitacco, 2001; Rickayzen and Walsh, 2002). As individual level data become available, recent studies calibrate their models to these data using longitudinal methods (e.g. Fong et al., 2015; Stallard, 2007, 2011). None of these studies, however, include systematic trend and volatility in health status transitions. Majer et al. (2013) develop a multi-state model with transition probabilities that depend on age and calendar time. They apply the Lee-Carter method to forecast future transition probabilities using Dutch population data. The model does not directly include recovery and is calibrated to aggregate population level data.

We develop a multi-state model for health state transitions incorporating systematic improvement and estimate transition rates using individual level U.S. Health and Retirement Study (HRS) data. We adopt the multi-state latent factor intensity (MLFI) model with a common stochastic factor, referred to as a frailty, and time trend. This has been used to model credit rating migrations, which are analogous to transitions among multiple health states, in Koopman et al. (2008). The MLFI model is a more parsimonious and consistent model than that used in Majer et al. (2013). It directly allows for a range of characteristics, or covariates, such as age and sex and allows estimation of the improvement trend

simultaneously with the transition rates using individual data. The model readily allows for the inclusion of other covariates. The Markov assumption can also be relaxed in the MLFI model which is relevant for LTC model development. We estimate the MLFI model using the U.S. Health and Retirement Study (HRS) data.

Simulations are used to investigate the effect of systematic uncertainties on distributions of healthy life expectancy and time spent in disability. We obtain some interesting results. We quantify the significant improvement in healthy-dead transition (mortality) rates for both genders, which is not found in mortality rates for disabled lives. Recovery rates are shown to be highly sensitive to the stochastic frailty factor. The systematic improvement in transition rates, after including a time drift term and a stochastic frailty factor and taking uncertainties into account, produces substantial increases in the older-age survival probabilities, and increases total and healthy life expectancy, for both genders. As total life expectancy increases, the number of old-aged disabled individuals also increases, resulting in a stable expected proportion of time spent in disability.

This paper is structured as follows. In Section 2, the models and estimation methodology are explained. In Section 3, we provide a brief description of the HRS data used. Section 4 presents the model parameter estimation results including the frailty factor using the HRS data. Section 5 presents transition rate estimates and uncertainties along with simulated survival curves. Section 6 shows the simulated distributions of disability. In Section 7, an analysis of future expected lifetimes along with expected time spent in disability is provided. We conclude the paper in Section 8.

2 Model Framework

Based on Fong et al. (2015) the multi-state LTC model contains 3 states: healthy, disabled and dead (absorbing state), as shown in Figure 1. The model allows for recovery from

disability. There are four types of transitions - healthy to disabled (*inc*), disabled to healthy (*rev*), healthy to dead (*hd*), and disabled to dead (*dd*).

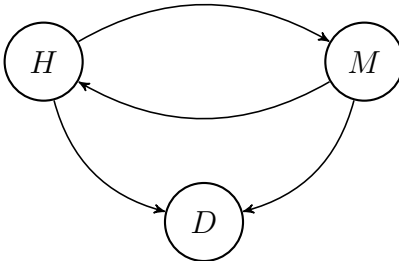


Figure 1. Simple three-state LTC transition model allowing for recovery; $H = \text{Healthy/abled}$, $M = \text{Morbid/disabled}$, and $D = \text{Dead}$.

2.1 Model Specification

We adopt a proportional hazard specification, similar to that used in Koopman et al. (2008) for credit rating transitions. The transition intensity for transition type s for an individual k at time t is assumed to be of the form:

$$\lambda_{sk}(t) = \exp[\beta_s + \gamma'_s w_k(t) + \alpha_s \psi(t)] \times H_{sk}(t), \tag{1}$$

where β_s is the baseline log-intensity for transition type s , with $s \in \{inc, rev, hd, dd\}$, independent of time and common across all individuals. The vector $w_k(t)$ contains the observed covariates for each individual k , and we restrict our covariates to gender and age. We assume the transition rates are piecewise constant for integer ages. $\psi(t)$ is the unobserved stochastic process that drives systematic uncertainties, also referred to as a frailty. The parameter vector γ'_s and scalar α_s measure the sensitivities of logarithm of $\lambda_{sk}(t)$ with respect to $w_k(t)$ and $\psi(t)$. Lastly, the scalar function $H_{sk}(t)$ is the baseline hazard function introduced by Koopman et al. (2008) to allow for duration dependence. If a semi-Markov process is assumed, then $H_{sk}(t)$ depends on the duration spent in the current state up to time t ; whereas

$H_{sk}(t) = 1$ in case of the Markov assumption. Note that both transition intensities and the latent process change at discrete time points.

We adopt the Markov assumption, i.e., $H_{sk}(t) = 1$ for computational reasons and to reflect the available data. Prior studies have shown evidence of a non-Markovian property in disability dynamics (Hardy and Gill, 2004) and a number of studies have used semi-Markov models for LTC (e.g. Biessy, 2015; Lepez, 2006; Tomas and Planchet, 2013). Given the widely spaced measurement intervals in almost all population-based panel studies, Markov models are widely used in LTC (e.g. Leung, 2006; Levantesi and Menziatti, 2012; Pritchard, 2006; Brown and Warshawsky, 2013; Shao et al., 2015). Since we have a two-year window between two consecutive waves in the HRS, accurate measurement of durations in the disabled state is not possible, limiting us to the Markov model assumption.

The latent dynamic process $\psi(t)$ is modeled with a simple random walk process. The latent process produces uncertainty in the transition intensities with the direction and magnitude varying for different transition types. The magnitude of the effect for type s is given by the coefficients α_s .

We adopt three models for estimation: no-frailty model, no-frailty model with linear time trend, and the frailty model with time trend:

1. In the “no frailty” model, the transition rate λ_{skx} is assumed to be dependent on age and sex only. For the simplest case, the logarithm of λ_{skx} is linear in age, the specification under the Markovian assumption is as follows:

$$\ln\{\lambda_{skx}\} = \beta_s + \gamma_s^{age} \cdot x + \gamma_s^{female} \cdot F. \quad (2)$$

The coefficients γ_s^{age} and γ_s^{female} describe the sensitivity of $\ln\{\lambda_{skx}\}$ to age x and to the female indicator variable F . The coefficient β_s is the reference level of $\ln\{\lambda_{skx}\}$

and varies by transition type.

2. To allow the transition rates to change over time, we add a linear time index into the no-frailty model. The no-frailty model with linear time trend is

$$\ln\{\lambda_{skx}(t)\} = \beta_s + \gamma_s^{age} \cdot x_t + \gamma_s^{female} \cdot F + \phi_s t, \quad (3)$$

where t indicates the time period, and ϕ_s measures the slope of change in $\ln\{\lambda_{skx}(t)\}$ with respect to the time index t .

3. The systematic latent factor is then included to give the frailty model with time trend:

$$\ln\{\lambda_{skx}(t)\} = \beta_s + \gamma_s^{age} \cdot x_t + \gamma_s^{female} \cdot F + \phi_s t + \alpha_s \psi(t), \quad (4)$$

where α_s measures the sensitivity of the log transition rates to the common latent factor or frailty $\psi(t)$. The frailty factor $\psi(t)$ can be interpreted as a mortality index with respect to time, similar to the κ_t in the Lee-Carter model, which is usually modeled as a simple random walk with drift term (e.g. Majer et al., 2013). With the linear time variable t to capture the deterministic time drift, the latent factor $\psi(t)$ is modeled as a simple random walk:

$$\psi(t) = \psi(t-1) + \epsilon_t, \quad \epsilon_t \sim NIID(0, \sigma^2). \quad (5)$$

We restrict σ to be independent of time. Not all α_s parameters and σ can be identified simultaneously, so without loss of generality we assume $\sigma = 1$.

2.2 Parameter Estimation

Maximum likelihood is used for estimation of the no-frailty models. Let θ denote the parameters of interest, then the likelihood functions for the no-frailty model and the no-frailty

model with time trend are

$$L(\theta|\mathcal{F}_T)_{\text{no frailty}} = \prod_{t=1}^T \prod_{k=1}^K \prod_{s=1}^S \exp \left\{ Y_{sk}(t) \ln[\lambda_{skx}] - R_{sk}(t) \int_{t-1}^t \lambda_{skx} du \right\}, \quad (6)$$

and

$$L(\theta|\mathcal{F}_T)_{\text{no frailty with time}} = \prod_{t=1}^T \prod_{k=1}^K \prod_{s=1}^S \exp \left\{ Y_{sk}(t) \ln[\lambda_{skx}(t)] - R_{sk}(t) \int_{t-1}^t \lambda_{skx}(u) du \right\}, \quad (7)$$

respectively, where dummy variables $Y_{sk}(t) = 1$ if individual k experiences a transition type of s at time t and $R_{sk}(t) = 1$ if individual k is exposed to transition type s between time $t - 1$ and t .

For the frailty model, the likelihood function is given by:

$$L(\theta|\mathcal{F}_T) = \int L(\theta|\mathcal{F}_T, \Psi) p(\Psi) d\Psi, \quad (8)$$

where the likelihood function conditional on Ψ , which denotes the complete path of $\psi(t)$, is:

$$L(\theta|\mathcal{F}_T, \Psi_T) = \prod_{t=1}^T \prod_{k=1}^K \prod_{s=1}^S \exp \left\{ Y_{sk}(t) \ln[\lambda_{skx}(t)] - R_{sk}(t) \int_{t-1}^t \lambda_{skx}(u) du \right\}. \quad (9)$$

The high-dimensional integral makes maximum likelihood computationally intensive. Monte Carlo techniques are used where we simulate N paths of Ψ from $p(\Psi)$ denoted by $\Psi^{[1]}, \dots, \Psi^{[N]}$, for a large number N . We then use the Monte Carlo estimator of Equation (8) computed as

$$\hat{L}(\theta|\mathcal{F}_T) = N^{-1} \sum_{j=1}^N L(\theta|\mathcal{F}_T, \Psi^{[j]}), \quad (10)$$

for parameter estimation.

Koopman et al. (2008) propose an alternative approach that is a combination of importance

sampling and the Kalman filter and smoother. This involves constructing an approximating linear Gaussian state space model, according to criteria in Durbin and Koopman (1997). In the approximating linear Gaussian state space model, the observation is the indicator variable $Y_{sk}(t)$ and the hidden state is a vector v_t that contains the frailty $\psi(t)$. The observation $Y_{sk}(t)$ is then linked to the state v_t with a linear equation. The Kalman filter and smoother algorithm is used to retrieve the conditional distributions of v_t , from which the conditional sample of the frailty is simulated. A more detailed explanation of this approach can be found in Koopman et al. (2008).

This alternative approach is not used for estimation for a number of reasons. It is more complex to implement, requiring smoothing and simulation in each step of the maximisation process, leading to a reduction in computational speed. The efficiency of the approach depends on how effective the approximating linear Gaussian model recovers the unobserved $\psi(t)$ process from the observed data. Koopman et al. (2008) perform a simulation experiment to show that for their problem the algorithm is adequate. In their simulation experiment, log transition rates of each type are restricted to the sum of a constant term and the frailty effect, whereas in our health state transitions the constant term is replaced by an age-dependent variable. Note that in the approximating Gaussian model, the observation (an indicator variable, Y_{skt} that equals one when a transition occurs) is the sum of $\log(\lambda_{skt})$ and a random observation noise. The observation noise in our case has a higher variance than the one in the simulation experiment carried out by Koopman et al. (2008). A more volatile observation noise reduces the effectiveness of the Kalman filter and smoother to track the latent $\psi(t)$.

To demonstrate this, we consider a simple linear state space model with a random walk state process plus observation noise. We apply the Kalman filter and smoother to the models with $\sigma = 1$ and $\sigma = 10$, where σ is the standard deviation of the observation noise. Both state space models are based on the same latent state process, and the standard deviation of the

state noise is 1. As shown in Figure 2, the approach performs less adequately than suggested by Koopman et al. (2008) when the observation noise has a larger variance relative to the state noise.

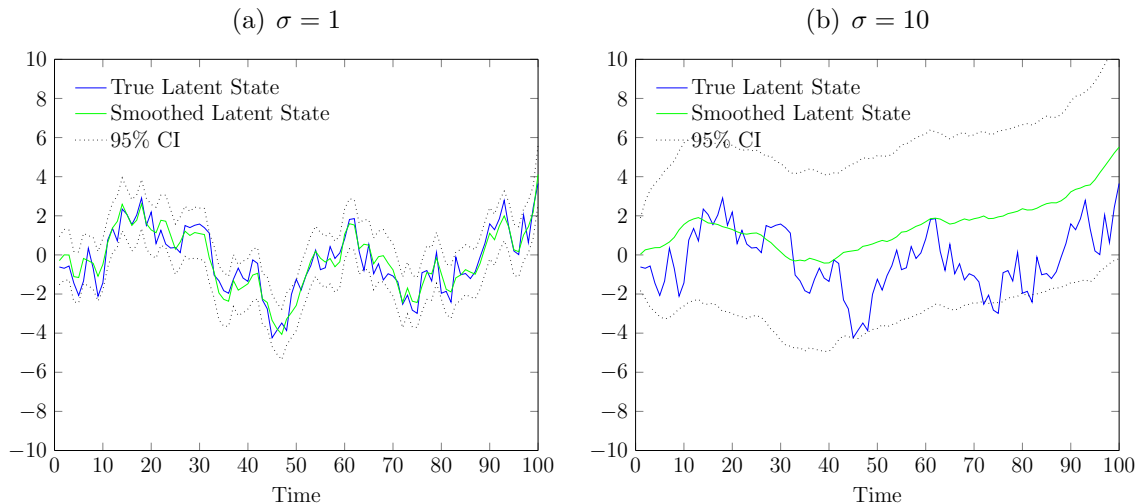


Figure 2. True vs. Smoothed Estimate of the Latent State Process in a Simple Linear State Space Model

In our case the time horizon, denoted by T , is 7, whereas in Koopman et al. (2008), $T > 14,000$ which also limits the benefits from the conditional simulation of $\psi(t)$ in the model estimation.

The approximating Gaussian state space model is however used to retrieve the unobserved frailty process based on the parameter estimates and observations which we discuss in the next section.

2.3 Recovery of the Frailty Process

The main idea of the state space model is to use a linear function to link the observation (whether a transition occurs, denoted by $Y_{sk}(t)$) and the corresponding log transition rate, which is dependent on the state v_t . Intuitively, the larger the magnitude of the transition rate, the more likely a transition will occur.

In order to estimate the frailty factor in the frailty model with trend specified in Equation (4) the following state space model representation is used:

$$\begin{cases} v_t = F_t v_{t-1} + \tilde{R}_t \eta_t, & \eta_t \sim NIID(\mu, \sigma^2), \\ Y_{sk}(t) = Z_{skt} v_t + \xi_{skt}, & \xi_{skt} \sim N(c_{skt}, C_{skt}), \end{cases} \quad (11)$$

where the “state” v_t is a vector containing the constant coefficients, coefficients of the observable variables, and the “frailty” $\psi(t)$. The system matrix F_t and the selection matrix \tilde{R}_t are defined according to our specification of $\psi(t)$ in Equation (5). The vector Z_{skt} contains the observable information and the parameter α_s . We have

$$v_t = \left(\beta_1, \dots, \beta_S, \gamma_1^{age}, \dots, \gamma_S^{age}, \gamma_1^{female}, \dots, \gamma_S^{female}, \phi_1, \dots, \phi_S, \psi(t) \right)', \quad (12)$$

and

$$Z_{skt} = \{e'_s, e'_s \otimes x, e'_s \otimes F, e'_s \otimes t, \alpha_s\}, \quad (13)$$

where e_s is the s th column of an $S \times S$ identity matrix I_S . With this model representation we see that $\ln\{\lambda_{skx}(t)\} = Z_{skt} v_t$ if the person is exposed to the risk, which is equivalent to Equation (4).

The matrix F_t and the vector \tilde{R}_t are defined to be consistent with the form of the latent process $\psi(t)$. For the frailty model specified in Equations (4) and (5), this gives

$$F_t = I_p, \quad \tilde{R}_t = \begin{bmatrix} O_{p-1} \\ 1 \end{bmatrix}, \quad (14)$$

where I_p is a $p \times p$ identity matrix and O_{p-1} is a vector of zeros of length $p - 1$, with p being the total number of elements in the vector v_t . For instance, in the frailty model, $p = 17$.

The state space model specified in Equation (11) is then used to estimate the path of the systematic latent factor given the observations and the estimates of the parameters:

$$\left\{ \beta_1, \dots, \beta_S, \gamma_1^{age}, \dots, \gamma_S^{age}, \gamma_1^{female}, \dots, \gamma_S^{female}, \phi_1, \dots, \phi_S, \psi(t), \alpha_1, \dots, \alpha_S \right\}.$$

In order to make the model parsimonious we assume $C_{skt} = \kappa_t^2$ and $c_{skt} = \zeta_t$. Following Durbin and Koopman (1997), we choose κ_t^2 and ζ_t such that the non-Gaussian density and the approximating Gaussian density are as close as possible in the neighbourhood of the last element in v_t , i.e. $\psi(t)$. This requires

$$\frac{\partial l_t(v)}{\partial \psi(t)} = 0, \quad (15)$$

and

$$\frac{\partial^2 l_t(v)}{\partial \psi(t)^2} = 0, \quad (16)$$

where $l_t(v) = \ln p(z|v, \mathcal{F}_T)_t - \ln g(z|v, \mathcal{F}_T)_t$, $p(z|v, \mathcal{F}_T)$ denotes the non-Gaussian density, and $g(z|v, \mathcal{F}_T)$ denotes the approximating Gaussian density. Solving Equations (15) and (16) simultaneously gives the estimates for κ_t^2 and ζ_t .

To determine the estimate of the frailty, an initial guess for ψ , denoted by $\hat{\psi}^{[0]}$, is used with the Kalman filter and smoother by computing κ_t^2 and ζ_t iteratively, based on estimates of ψ denoted by $\hat{\psi}^{[j]}$ for $j = 1, 2, \dots$. This continues until convergence of parameter estimates is achieved.

3 Health and Retirement Study Data

We use the Health and Retirement Study (HRS) data from the University of Michigan to estimate the health transition models. The HRS data is a comprehensive and ongoing U.S. national longitudinal household survey of people aged 50 and above starting from 1992.

The cohorts are interviewed every two years, covering information on basic demographics, income, assets, health status, health care expenditures, job history, family structure, etc. We use data from wave 1998 onward as there were inconsistencies in the survey questions before wave 1998 (Shao et al., 2015). The latest available wave at the time of writing is in 2012, which contributes an additional wave of data to that included in Fong et al. (2015) and Shao et al. (2015).

Health states and transitions are determined using the HRS data on self-reported difficulties in six Activities of Daily Livings (ADLs) which allow assessment of incidence of disability and recovery. The six ADLs are dressing, walking, bathing, eating, transferring (i.e. getting in and out of bed or up from chair), and toileting. Two or more difficulties in any of the six ADLs is categorised as the disabled state, otherwise individuals are in the healthy state if alive. The death, and the date of death, of an individual are also reported in the data allowing identification of transitions to death. Since there are approximately two years between waves, we assume transitions to disability or recovery occur at the mid-point of the period. This mid-point assumption is also used for deaths if the exact death date is not given in the data. Withdrawals are treated as non-informative right censoring.

Crude age- and sex-specific transition rates for each sex are calculated as:

$$\begin{aligned}
 m_s(x) &= \frac{Y_s(x)}{R_s(x)} \\
 &= \frac{\text{number of transitions of type } s \text{ for those aged } x \text{ last birthday}}{\text{exposure years to transition type } s \text{ for those aged } x \text{ last birthday}}, \quad (17)
 \end{aligned}$$

where x is the integer age. The log crude rates are shown in Figure 3 and these are consistent with the results in Fong et al. (2015).

The logarithms of the transition intensities show an almost linear pattern with respect to age for both genders, which supports the proportional hazard specification in Equation (1). The gender difference in the crude transition rates across ages is close to parallel or insignificant.

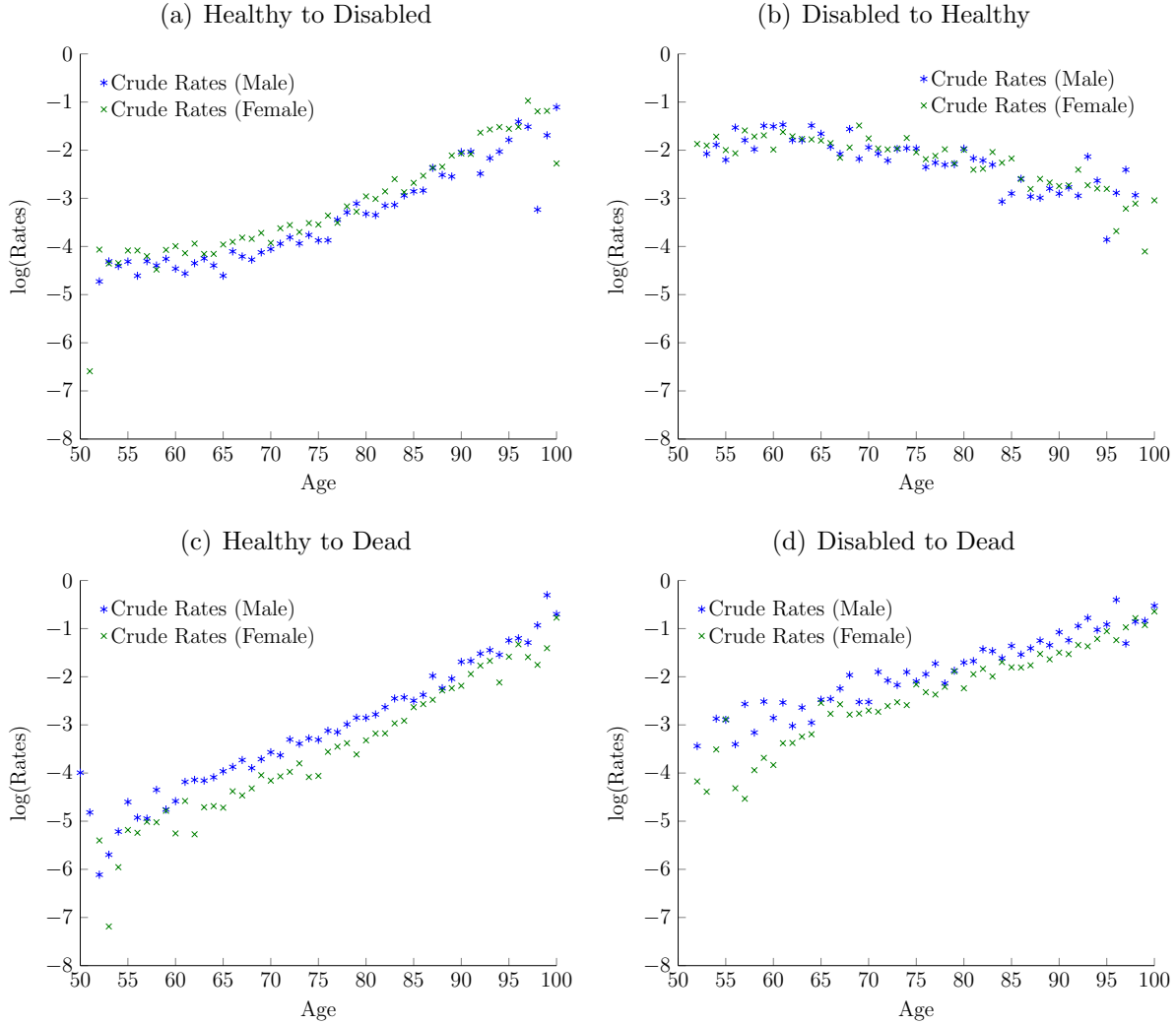


Figure 3. Log transformation of crude transition rates.

This is supported by the results in Xu et al. (2015) where it is shown that the male and female sub-populations in the HRS data do not have significantly different deviations from aggregate systematic mortality improvement. Slight curvature is observed in the plots for disability inception rates (healthy to disabled). Females have higher incidence of disability than males. Females generally have lower mortality rates than males at all ages regardless of health states. Recovery rates are more noisy and decline with age at older ages for both males and females.

4 Estimation of Model Parameters and Frailty Factor

Table 1. Parameter estimates (Monte Carlo MLE)

	Transition Type $s =$	H - M 1	M - H 2	H - D 3	M - D 4
No Frailty	$\hat{\beta}_s$	-7.9488*** (0.1080)	0.9150*** (0.1263)	-10.0296*** (0.1155)	-6.2067*** (0.1467)
	$\hat{\gamma}_s^{age}$	0.0678*** (0.0014)	-0.0320*** (0.0017)	0.1001*** (0.0014)	0.0648*** (0.0017)
	$\hat{\gamma}_s^{female}$	0.2894*** (0.0284)	0.0501 (0.0400)	-0.4558*** (0.0271)	-0.3775*** (0.0357)
	Log Likelihood	-51,772			
No Frailty w/ Time	$\hat{\beta}_s$	-7.9224*** (0.1085)	0.9319*** (0.1267)	-9.8901*** (0.1156)	-6.1982*** (0.1485)
	$\hat{\gamma}_s^{age}$	0.0681*** (0.0014)	-0.0321*** (0.0017)	0.1016*** (0.0014)	0.0648*** (0.0018)
	$\hat{\gamma}_s^{female}$	0.2900*** (0.0284)	0.0493 (0.0400)	-0.4567*** (0.0271)	-0.3777*** (0.0357)
	$\hat{\phi}_s$	-0.0143** (0.0069)	-0.0025 (0.0096)	-0.0674*** (0.0069)	-0.0029 (0.0086)
	Log Likelihood	-51,722			
Frailty	$\hat{\beta}_s$	-7.9237*** (0.1085)	0.9163*** (0.1268)	-9.8886*** (0.1157)	-6.1963*** (0.1484)
	$\hat{\gamma}_s^{age}$	0.0682*** (0.0014)	-0.0318*** (0.0017)	0.1015*** (0.0014)	0.0648*** (0.0018)
	$\hat{\gamma}_s^{female}$	0.2896*** (0.0284)	0.0481 (0.0400)	-0.4568*** (0.0271)	-0.3780*** (0.0357)
	$\hat{\phi}_s$	-0.0117 (0.0075)	0.0243** (0.0102)	-0.0777*** (0.0076)	0.0036 (0.0094)
	$\hat{\alpha}_s$	0.0173 (0.0202)	0.2065*** (0.0304)	-0.0664*** (0.0204)	0.0449 (0.0253)
	Log Likelihood	-51,697			

H = “healthy/non-disabled”, M = “morbid/disabled”, and D = “dead”

* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

$\lambda_{skx}(t)$ calculated from figures above are bi-annual rates, and for the frailty model $N = 1,000$

Table 1 gives the parameter estimates for the models. The no frailty models and the frailty model give consistent estimation results for age and gender effects.

- **Age:** Transition rates are all strongly age-dependent and consistent with the results in Fong et al. (2015). Disability and mortality rates (H - M, H - D, and M - D) increase with age and recovery rates from disability (M - H) decrease. Exits from disability can

be either by recovery or death. The estimated $\hat{\gamma}_2^{age}$ and $\hat{\gamma}_4^{age}$, as well as the $\hat{\beta}$'s, show that before the age of 75 approximately, exit from disability for males is mostly due to recovery, consistent with the crude data in Figure 3. The mortality rate for a healthy individual rises with age more quickly than the disability rate ($\hat{\gamma}_3^{age}$ vs. $\hat{\gamma}_1^{age}$). Thus exit from the healthy state is mostly by death rather than disability at older ages. As we show later, this pattern of rates means that for a healthy population aged 50, the number of disabled will increase initially and then start declining around the age of 75.

- **Gender:** As expected, gender plays a significant role in most health state transitions, although there is no significant gender difference in recovery rates. Females have higher risks of becoming disabled than males, in line with the results in Fong et al. (2015). For the frailty model, $\hat{\gamma}_1^{female} = 0.2896$ whereas $\hat{\gamma}_1^{age} = 0.0682$, which implies that, all else being equal, a healthy female aged 60 has the same risk of disability as a healthy male aged 64 (approximately). Females have significantly lower mortality rates than males, regardless of health status ($\hat{\gamma}_3^{female} < 0$ and $\hat{\gamma}_4^{female} < 0$). A female aged 60 has an equivalent risk of dying as her 55.5-year-old male counterpart if she is non-disabled or to her 54-year-old male counterpart if she is disabled. As we show later, this pattern of rates results in a more rectangular survival curve and longer life expectancy for females. Since women have higher disability rates than men, women spend more time in disability than men even though they live longer.
- **Time trend:** There has been a significant mortality improvement trend for the healthy population in both the no frailty model and the frailty model. In the frailty model, the estimated $\hat{\phi}_3$ was consistently significant in the Monte Carlo MLE for differing values of N ranging from 100 to 1,000, with value around -0.07 . Since the length between jumps in time is 2 years (survey cycle of the HRS), this value of $\hat{\phi}_3$ implies that, all else being equal, a healthy 67-year-old in 2012 has roughly the same mortality rate as a healthy 65-year-old in 2006. There has also been improvement in healthy-disabled

transition rates, but this improvement is not significant in the frailty model.

- **Systematic uncertainty** (“frailty”): For the frailty $\psi(t)$, recovery rates and healthy lives’ mortality rates are significantly impacted by the stochastic frailty process. Thus, M - H and H - D transitions need to take into account systematic uncertainties. The signs of $\hat{\alpha}_2$ and $\hat{\alpha}_3$ are different, so that the frailty factor has opposite effects on recovery rates and healthy mortality rates.

Using these estimated parameters, the posterior mean of the frailty process is estimated as described in Section 2. The results are shown in Figure 4, where each year on the horizontal axis represents a wave, e.g. $\phi(2010)$ denotes the frailty factor for the period between 2010 and 2012. The posterior mean of the frailty has been slightly declining from 1998 to 2010 and the sign of the frailty is negative in the later years.

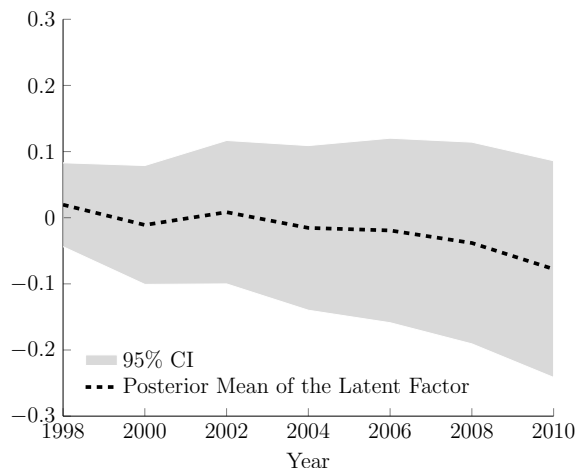


Figure 4. Posterior mean of the latent frailty factor $\psi(t)$.

Kass and Raftery (1995) use the Bayes factor for model comparison, given by $B_{10} = \frac{\mathbb{P}(\text{Data}|H_1)}{\mathbb{P}(\text{Data}|H_0)} = \frac{\hat{L}_{H_1}(\hat{\theta}_{H_1}|\mathcal{F}_T)}{\hat{L}_{H_0}(\hat{\theta}_{H_0}|\mathcal{F}_T)}$. Table 2 shows the results for model comparison for each pair of models denoted by H_0 and H_1 respectively.

We see that the inclusion of the time trend and the frailty factor improves model fit. We will use all these models to project simulated lives for healthy individuals in order to assess the

Table 2. Model comparison based on test statistic $2 \ln(B_{10})$.

H_1	H_0	
	No Frailty	No Frailty w/Time
No Frailty w/Time	100	-
Frailty w/Time	150	50

impact of systematic trend and uncertainty on expected future lifetimes as well as expected healthy life times.

5 Survival Curves

We use simulation to quantify future lifetime and future healthy lifetime along with the uncertainty around the expected values implied by the estimated models. Survival curves are initially produced from the simulations before considering the proportion of individuals in the disabled state and the uncertainty in these proportions. We then consider life expectancies. We do not include parameter uncertainty in the simulations. The standard errors of parameter estimates are relatively small as shown in Table 1 and the conclusions that we draw based on the systematic uncertainty are not changed if this was included.

We use the parameter estimates for the three models to simulate the life path of a healthy individual at differing ages for a large number of times. The maximum attainable age is 120. The simulation is performed for ages $x = 50, 55, 60, 65, 70$ and 75 , and for both genders. For the no-frailty models, with and without time trend, 10,000 homogeneous lives are simulated for each starting age and gender. For the frailty model, 10,000 paths of the latent frailty factor are simulated and 10,000 homogeneous lives are simulated for each path. The initial value for the simulated frailty process is set to be the posterior mean of ψ in 2010. The simulation results in this study provide predicted life trajectories for an individual aged x in 2010.

Figure 5 shows the transition rates, along with uncertainties for the frailty model given

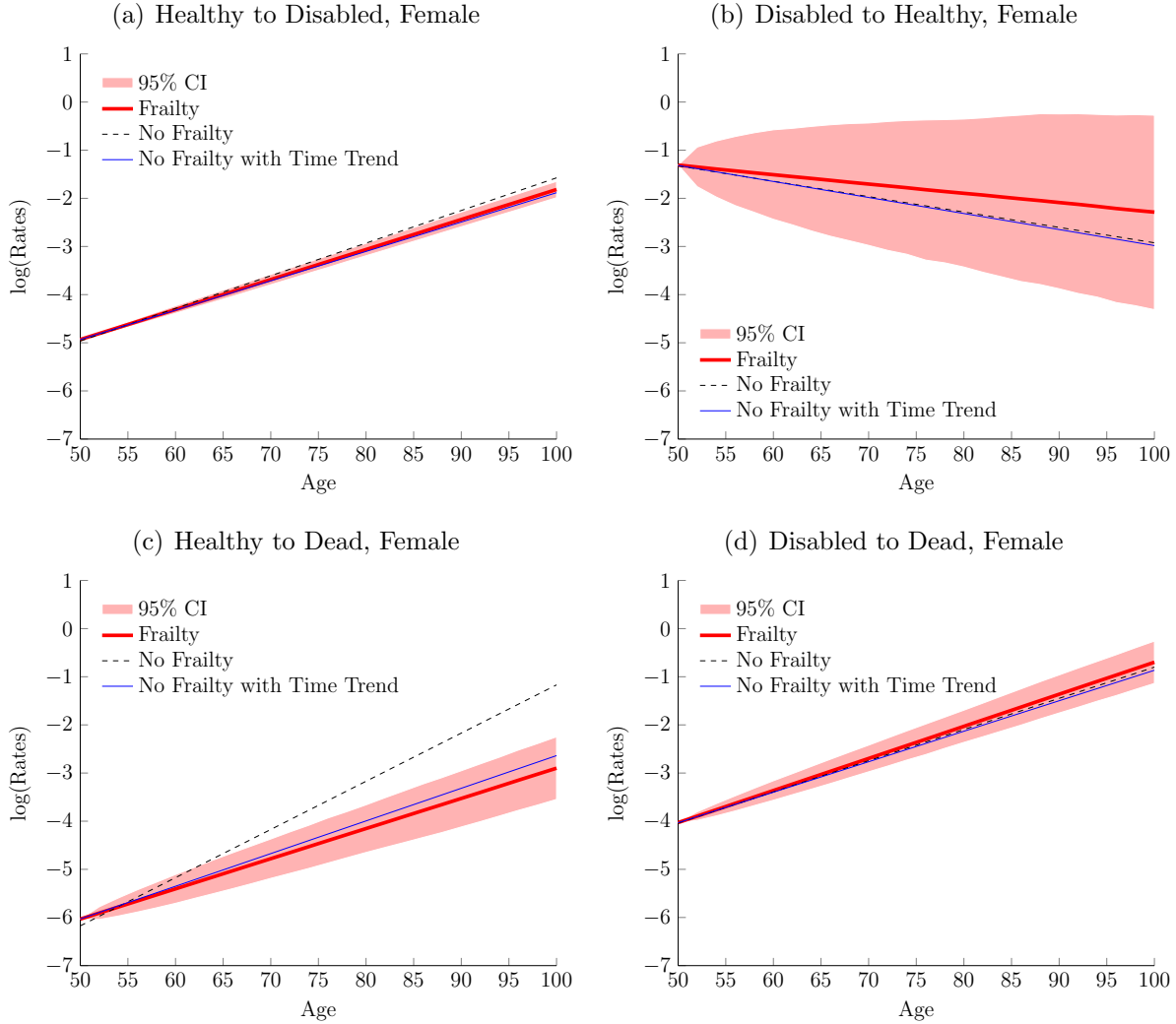


Figure 5. Simulated transition rates for females (annualised).

by 95% confidence intervals, for females aged 50 in 2010 until the age of 100, based on the parameter estimates. The disability rates show small reductions by age when a time trend is included and there is little uncertainty in these rates arising from the frailty factor. Systematic mortality improvement for the healthy population over time is significant, shown by the much flatter slope of the healthy to dead transition rates with respect to age for the models with time trend. The frailty produces greater uncertainties in recovery rates and healthy mortality rates than in the other two transition rates. In particular, recovery rates show a high level of uncertainty arising from the frailty factor, such that it is not possible to observe any statistically significant impact of a time trend. Similar comments apply to

the disabled mortality rates where the uncertainty arising from the frailty factor is lower but there is no significant effect of the time trend. The confidence intervals for both healthy to disabled and healthy to dead transitions show that systematic improvement over time has been significant. Similar results hold for males except that males have lower disability rates and higher mortality rates than their female counterparts.

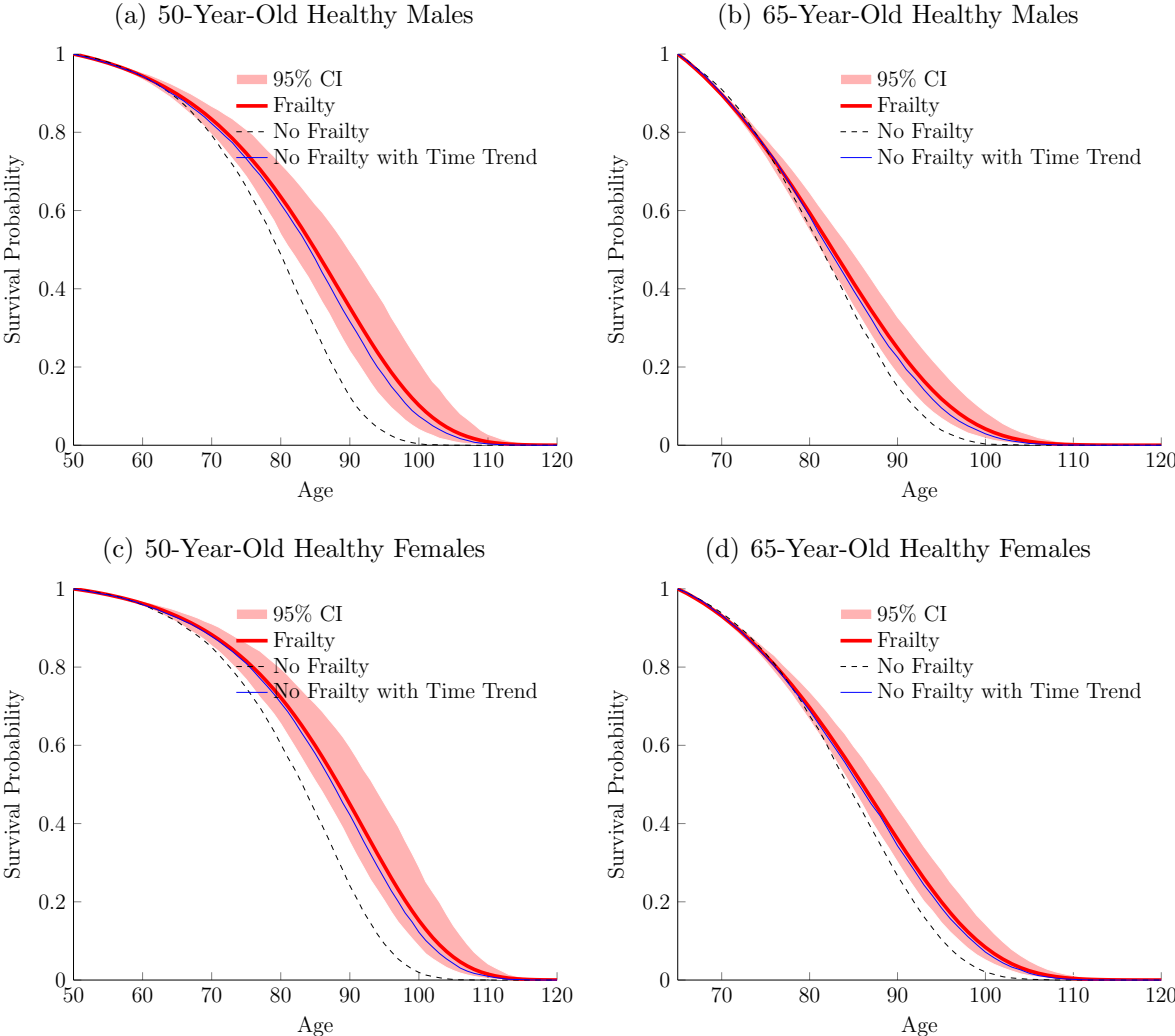


Figure 6. Simulated survival curves for for $x = 50, 65$.

Figure 6 shows the simulated survival probabilities for ages 50 and 65. The frailty model survival probabilities are significantly higher than those for the no-frailty model without the time trend, as expected. For the no frailty model with a linear time trend, there is less difference between the survival curves. The mortality improvement in terms of older-age

survival probabilities for a 50 year-old healthy individual is more substantial than those for a 65 year-old healthy individual. This reflects the fact that the younger cohort will experience future systematic improvement for a longer period than the older generations. Females have more rectangular survival curves than their male counterparts, reflecting the fact that females have significantly lower mortality rates than males regardless of health status.

6 Distribution of Disability

Although survival curves capture the impact of the systematic improvement factor on mortality, they do not allow us to quantify the impact on disability. To do this we simulate a cohort of healthy individuals and estimate the proportion that are disabled under the different models, along with the uncertainty in these proportions from the frailty factor. Figure 7 shows the expected percentage of an initial 10,000 healthy individuals aged x in 2010 that will be in the disabled state in the future, for initial ages $x = 50$ and 65. The results reflect the interaction between mortality and disability. Since these percentages are based on the initial 10,000 lives, they also reflect the relative numbers disabled at future ages.

In Figure 7 the proportion of individuals in disability initially rises with age and declines around the age of 80. Since females have higher disability rates and lower mortality rates, the peaks of the curves for females reach higher values and take place at slightly older ages than males. Initially, higher disability rates lead to a higher proportion of disabled individuals and as age increases, more healthy females than healthy males survive due to lower healthy mortality rates. This results in a higher exposure to becoming disabled for the female population. As noted earlier, at older ages, when the exit from disability is mainly driven by death rather than recovery, disabled females also have lower mortality risks than their male counterparts, leading to a delayed decline of the proportion disabled.

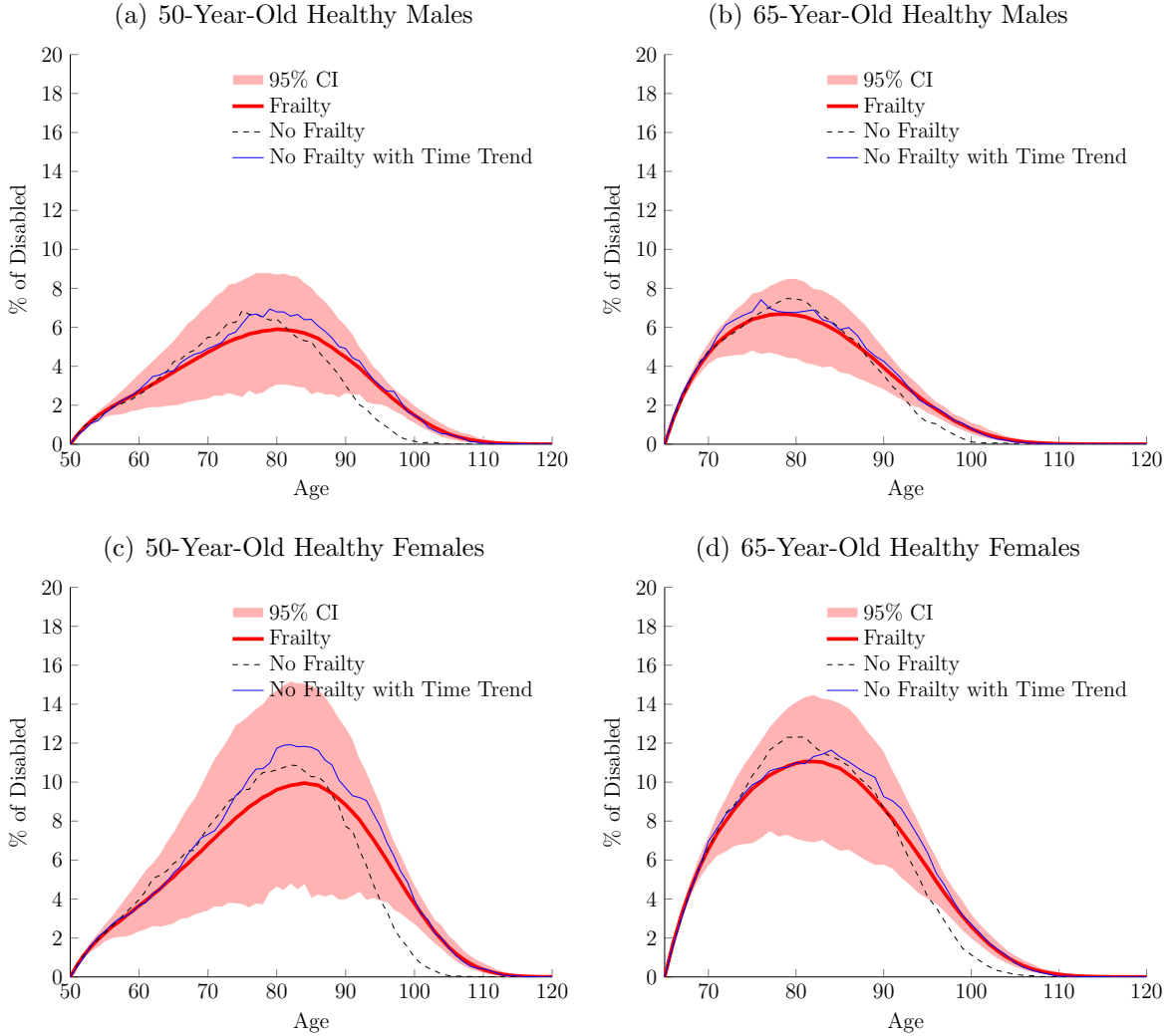


Figure 7. Simulated proportion of disabled individuals for $x = 50, 65$.

The frailty factor generates significant variations in the expected percentage of disabled individuals after around 5 years from the initial age x . This reflects the large uncertainties in the mortality and recovery rates. Nonetheless, the uncertainty reduces at very old ages as the total population alive reduces. Over the first 5 years of the simulations the frailty model produces almost the same expected number of disabled individuals as the no-frailty models. The transition rates in Figure 5 show that the improvement in the disability rates is not as profound, despite the moderate statistical significance of $\hat{\phi}_1$. A postponement of the disability inception age is expected for the population as more individuals live longer and become disabled at older ages which will arise due to the higher exposures to becoming

disabled at old ages (more old-age survivors) rather than a significant reduction in disability transition rates. These results are consistent with Crimmins and Beltrán-Sánchez (2011) who claim there has been no elimination or delay of diseases in the U.S.

At the older ages the impact of the time trend from the systematic improvement factor becomes clear with more people being in disability than is expected from the no-frailty model without time trend. Under the models with time trend there are more healthy survivors at older ages, arising from the improvement in healthy mortality rates, leading to higher levels of disability transitions at older ages. The impacts of systematic improvements on disability prevalence are not likely to be observed until the older ages.

Cross sectional data and analysis of recent disability data will not detect these longer run changes, which requires the modeling and simulation approach used here.

7 Expected Future Lifetimes

Systematic improvement will lead to higher life expectancies. The frailty model with time trend allows us to determine the impact of uncertainty on both expected future life times and the confidence interval for this life expectancy. Table 3 shows simulation results for remaining life expectancy for a healthy individual aged $x = 50, 55, 60, 65, 70,$ and 75 . Comparing the models with time trend to the no-frailty model without time trend we see that the remaining life expectancy is significantly extended due to the systematic mortality improvement. For example, a 65-year-old healthy male, on average, will live to age 81.8 if projected without time trend and to 82.9 with time trend. Using the frailty model this same 65 year old male would on average live to 83.4, with a 95% confidence interval between 82 and 85.1. Similar results apply for a 65-year-old healthy woman who on average have 2.3 to 3 additional years of remaining life expectancy than their male counterparts.

At the younger ages, 65 and below, the lower confidence interval in the frailty model is higher

Table 3. Simulated expected lifetime

	Age	Number of Years		
		No Frailty	No Frailty w/Time	Frailty
Males	50	29	33.2	34.2 (30.7, 37.5)
	55	24.9	27.7	28.5 (26.0, 31.1)
	60	20.6	22.4	23.1 (21.1, 25.1)
	65	16.8	17.9	18.4 (17.0, 20.1)
	70	13.4	13.9	14.1 (13.0, 15.4)
	75	10.3	10.4	10.6 (10.0, 11.4)
Females	50	32.2	36.5	37.2 (34.2, 40.5)
	55	27.8	30.8	31.7 (29.5, 34.2)
	60	23.5	25.7	26.2 (24.5, 28.5)
	65	19.6	20.9	21.3 (19.9, 23.1)
	70	15.9	16.5	16.8 (15.8, 18.2)
	75	12.6	12.8	12.9 (12.3, 13.7)

Note: maximum attainable age is assumed to be 120.

than the expected future life time for the no frailty model without time trend. Projections of future expected life times for these ages that do not include time trend will significantly understate future survival prospects. The difference in life expectancy between the frailty model with time trend and the no frailty models is much reduced at the older ages, as is the uncertainty in the expected future lifetime.

To consider the impact of systematic improvement allowing for disability, we estimate the expected future life time in the healthy state. Table 4 shows the simulation results for remaining healthy life expectancy for a healthy individual aged $x = 50, 55, 60, 65, 70,$ and 75 . Reflecting the results for total remaining life expectancy, there is an increase in remaining healthy life expectancy, which decreases with the initial age x . For example, a 65-year-old healthy male and a 65-year-old healthy female are predicted to spend approximately 1.5 and 1.7 more years, respectively, in the healthy state if systematic improvement in transition rates is not included.

There is an increase, or expansion, in healthy life expectancy. Since total life expectancy also increases, the expected time in the disabled state can be determined by differencing the total life expectancy and the healthy life expectancy. In absolute terms the expected time

Table 4. Simulated expected lifetime in healthy state

		Number of Years		
	Age	No Frailty	No Frailty w/Time	Frailty
Males	50	27.2	31.1	32.2 (28.5, 36.9)
	55	23.1	25.7	26.7 (23.9, 29.8)
	60	18.9	20.5	21.4 (19.3, 24.1)
	65	15.3	16.3	16.8 (15.4, 18.8)
	70	12.0	12.5	12.7 (11.6, 14.4)
	75	9.2	9.2	9.5 (8.7, 10.4)
Females	50	29.1	32.7	33.9 (30.1, 39.2)
	55	24.7	27.2	28.4 (25.3, 32.1)
	60	20.5	22.4	23.1 (20.9, 26.3)
	65	16.8	18.0	18.5 (16.8, 20.9)
	70	13.4	13.9	14.2 (12.8, 16.2)
	75	10.3	10.5	10.7 (9.8, 11.8)

Note: maximum attainable age is assumed to be 120.

spent in disability, taking into account systematic trend, increases by a small amount for most ages.

To better understand the impact of systematic improvement estimated from HRS data on morbidity compression or expansion we consider the proportion of life expectancy that is healthy life expectancy and compare the results for the no frailty model with the frailty model with time trend. The ratio is calculated as:

$$\frac{\text{HLE}}{\text{TLE}} = \frac{\text{healthy life expectancy}}{\text{total life expectancy}}. \quad (18)$$

Table 5 shows the means and quantiles of the HLE/TLE ratios derived from the simulations.

We see that, for both genders, the HLE/TLE ratio declines with age so that more time is spent on average in disability for older aged individuals. The ratio for females declines with a steeper slope than for males, so females spend more time on average in disability as a proportion of remaining lifetime than males.

Table 5. HLE/TLE

	Age	No Frailty	No Frailty w/Time	Frailty
Males	50	0.938	0.936	0.941 (0.917, 0.967)
	55	0.929	0.928	0.934 (0.910, 0.955)
	60	0.920	0.916	0.925 (0.903, 0.948)
	65	0.911	0.909	0.915 (0.896, 0.938)
	70	0.900	0.897	0.903 (0.884, 0.928)
	75	0.889	0.891	0.892 (0.874, 0.913)
Females	50	0.901	0.897	0.908 (0.866, 0.951)
	55	0.887	0.884	0.897 (0.859, 0.934)
	60	0.874	0.873	0.882 (0.848, 0.922)
	65	0.859	0.859	0.867 (0.834, 0.907)
	70	0.841	0.840	0.847 (0.809, 0.888)
	75	0.822	0.825	0.828 (0.796, 0.861)

Note: maximum attainable age is assumed to be 120.

Although the average ratio is slightly lower for most ages when a time trend is included in the no-frailty model, suggesting some expansion in time spent in disability, this is the reverse for the frailty model with time trend. The mean of the HLE/TLE ratio for the frailty model is slightly higher than the no-frailty models. An examination of Figure 8 highlights the substantial uncertainties around the expected values. Allowing for this, based on the HLE/TLE ratios, there is no strong evidence for either morbidity compression or expansion based on the HRS data for the United States. Variations around the mean of the HLE/TLE ratio for females are larger than those for males.

8 Conclusions

Systematic improvement in mortality has attracted significant attention in recent years. Despite this, stochastic health transitions in LTC have received little attention. Transition rates are typically modeled without a stochastic component for systematic risk, and estimated mostly using aggregate data, rather than individual level data. In this paper we introduce the multi-state latent factor intensity (MLFI) model, which has been used in modeling credit rating transitions, into LTC modeling and estimate it using individual panel data from the HRS. The MLFI model allows us to incorporate a time trend and frailty capturing systematic

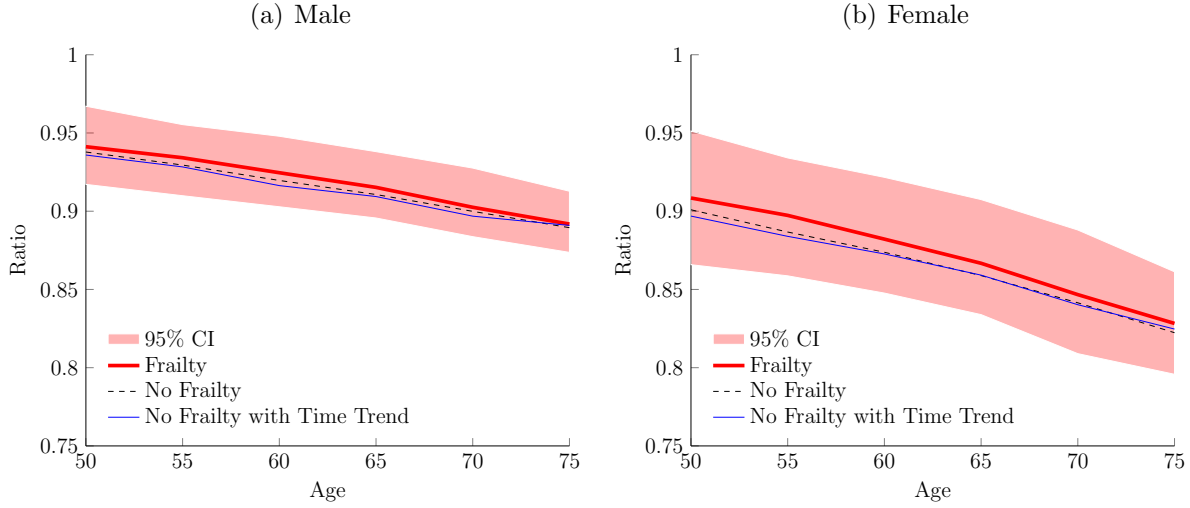


Figure 8. Ratio of healthy life expectancy (HLE) to total life expectancy (TLE).

trend and uncertainty in a parsimonious way, compared to the Lee-Carter specification used in Majer et al. (2013). The model is flexible and can easily incorporate covariates.

Fitting the model to HRS individual level data shows a significant improvement in mortality rates for both genders and allows us to quantify the impact of the systematic factor, or frailty, on disability rates and recovery rates. Mortality rates for disabled lives show less improvement than for healthy lives. Recovery rates from disability are found to be highly uncertain arising from the inclusion of the stochastic frailty factor. Survival probabilities improve and more healthy individuals are expected to survive to old ages resulting in higher exposures to disability. This results in an expected postponement of the disability inception age.

Healthy life expectancy, for both genders, is expected to increase although the expected time spent in disability is expected to increase much less in absolute terms. Uncertainty in the ratio of healthy life expectancy to total life expectancy shows that there is a high level of uncertainty as to the extent of morbidity compression or expansion for the U.S. population aged 50 and above. While total life expectancy has expanded and the number of old-aged disabled individuals has increased, the proportion of expected time spent in disability reflects

the significant variability in the number of disabled individuals in our simulations.

These results demonstrate how studies that do not include systematic trend and uncertainty may well come to differing conclusions around whether or not there has been expansion or contraction in disability. Our results suggest that there may, on average, be some disability compression, but the uncertainty in future mortality and disability rates may well lead to either compression or expansion.

Importantly, our methodology can be applied to a range of significant research issues around modeling and projecting disability including estimation of long term care costs, budgetary implications for government of improving longevity and the risks and uncertainties for pricing and risk management of LTC insurance products.

Acknowledgement

The authors acknowledge the financial support of the Australian Research Council Centre of Excellence in Population Ageing Research (project number CE110001029). Li also acknowledges the financial support of UNSW Business School. Opinions and errors are solely those of the authors and not of the institutions providing funding for this study or with which the authors are affiliated.

References

- Biessy, G. (2015). Long-term care insurance: A multi-state semi-markov model to describe the dependency process for elderly people. SCOR Global Life working paper.
- Biffis, E. (2005). Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics*, 37(3):443–468.
- Brown, J. and Warshawsky, M. (2013). The life care annuity: A new empirical examination of an insurance innovation that addresses problems in the markets for life annuities and long-term care insurance. *Journal of Risk and Insurance*, 80(3):677–704.
- Cairns, A. J., Blake, D., and Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance*, 73(4):687–718.
- Cairns, A. J., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., and Khalaf-Allah, M. (2011). Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, 48(3):355–367.
- Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., and Balevich, I. (2009). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *North American Actuarial Journal*, 13(1):1–35.
- Crimmins, E. M. and Beltrán-Sánchez, H. (2011). Mortality and morbidity trends: Is there compression of morbidity? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 66(1):75–86.
- Dahl, M. (2004). Stochastic mortality in life insurance: Market reserves and mortality-linked insurance contracts. *Insurance: Mathematics and Economics*, 35(1):113–136.

- Durbin, J. and Koopman, S. J. (1997). Monte Carlo maximum likelihood estimation for non-Gaussian state space models. *Biometrika*, 84(3):669–684.
- Fong, J. H., Shao, A. W., and Sherris, M. (2015). Multi-state actuarial models of functional disability. *North American Actuarial Journal*, 19(1):41–59.
- Fries, J. F. (1980). Aging, natural death, and the compression of morbidity. *The New England Journal of Medicine*, 303(3):130–135.
- Gruenberg, E. M. (1977). The failures of success. *The Milbank Memorial Fund Quarterly. Health and Society*, 55(1):3–24.
- Hardy, S. E. and Gill, T. M. (2004). Recovery from disability among community-dwelling older persons. *Journal of the American Medical Association*, 291(13):1596–1602.
- Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.
- Koopman, S. J., Lucas, A., and Monteiro, A. (2008). The multi-state latent factor intensity model for credit rating transitions. *Journal of Econometrics*, 142(1):399–424.
- Kwon, H.-S. and Jones, B. L. (2008). Applications of a multi-state risk factor/mortality model in life insurance. *Insurance: Mathematics and Economics*, 43(3):394 – 402.
- Lee, R. (2000). The Lee-Carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal*, 4(1):80–91.
- Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting US mortality. *Journal of the American Statistical Association*, 87(419):659–671.
- Lepez, V. (2006). Trajectoires en dépendance des personnes âgées: Modélisation, estimation et application en assurance vie. *Mémoire d'Actuariat, Centre d'Etudes Actuarielles*.

- Leung, E. (2006). A multiple state model for pricing and reserving private long term care insurance contracts in Australia. *Australian Actuarial Journal*, 12(2):187–247.
- Levantesi, S. and Menzietti, M. (2012). Managing longevity and disability risks in life annuities with long term care. *Insurance: Mathematics and Economics*, 50(3):391–401.
- Liu, X. and Lin, X. S. (2012). A subordinated markov model for stochastic mortality. *European Actuarial Journal*, 2(1):105–127.
- Majer, I. M., Stevens, R., Nusselder, W. J., Mackenbach, J. P., and van Baal, P. H. (2013). Modeling and forecasting health expectancy: Theoretical framework and application. *Demography*, 50(2):673–697.
- Manton, K. G. (1982). Changing concepts of morbidity and mortality in the elderly population. *The Milbank Memorial Fund Quarterly. Health and Society*, 60(2):183–244.
- Olivieri, A. and Pitacco, E. (2001). Facing LTC risks. In *Proceedings of the XXXII International ASTIN Colloquium*. Washington.
- Pritchard, D. (2006). Modeling disability in long-term care insurance. *North American Actuarial Journal*, 10(4):48–75.
- Rickayzen, B. D. and Walsh, D. E. (2002). A multi-state model of disability for the United Kingdom: Implications for future need for long-term care for the elderly. *British Actuarial Journal*, 8(02):341–393.
- Schrager, D. F. (2006). Affine stochastic mortality. *Insurance: Mathematics and Economics*, 38(1):81–97.
- Shao, A. W., Sherris, M., and Fong, J. H. (2015). Product pricing and solvency capital requirements for long-term care insurance. *Scandinavian Actuarial Journal*, doi:10.1080/03461238.2015.1095793.

- Stallard, E. (2007). Trajectories of morbidity, disability, and mortality among the U.S. elderly population: Evidence from the 1984 - 1999 NLTCs. *North American Actuarial Journal*, 11(3):16–53.
- Stallard, E. (2011). Estimates of the incidence, prevalence, duration, intensity, and cost of chronic disability among the U.S. elderly. *North American Actuarial Journal*, 15(1):32–58.
- Tomas, J. and Planchet, F. (2013). Multidimensional smoothing by adaptive local kernel-weighted log-likelihood: Application to long-term care insurance. *Insurance: Mathematics and Economics*, 52(3):573–589.
- Treasury of the Commonwealth of Australia (2015). 2015 Intergenerational report, Australia in 2055. Parkes, A.C.T: Treasury.
- Xu, M., Sherris, M., and Meyricke, R. (2015). Mortality heterogeneity and systematic mortality improvement. UNSW Business School Research Paper No. 2015ACTL23.