# The Impact of Duration on Disability Transitions: Evidence from the Cardiovascular Health Study

Liming Cai[1]
Nathaniel Schenker[1]
James Lubitz[1]
Paula Diehr[2]
Alice Arnold[2]
Linda P. Fried[3]

1. National Center for Health Statistics, CDC
2. University of Washington
3. Johns Hopkins University

## 1. Introduction

The multi-state life table (MSLT) model is a widely used statistical tool to analyze sequence of multiple and recurrent events in a target population (Schoen, 1987). Researchers in social science have used the MSLT model to study changes in health, marriage, residence, labor force participation and etc. The appeal of MSLT to a large extent lies in it intuitively simple assumption of a first-order Markov chain, where the transition probabilities are conditional on the current status only. Recently, however, there is increasing evidence that changes in the health status of the elderly may also depend on the length of duration in current states (Crimmins and Saito, 1993; Maddox, Clark and Steinhauser, 1994; Hardy and Gill, 2005). These observations suggest the need to fit a duration model that formally incorporates the duration effect on transition probabilities.

Due to the problem of left censoring, fitting a duration model such as the semi-Markov process (SMP) model presents problems. Left censoring occurs when the length of time already spent in the baseline health states when subjects come under observation, $R$, is unknown (Mayer and Tuma, 1990). This is true for most of the longitudinal health surveys on the US elderly (e.g., the Medicare Current Beneficiary Survey, the Health and Retirement Survey, the National Long-Term Care Survey and the Longitudinal Studies of Aging I and II). In order to apply the SMP model to left-censored data, analysts often exclude these baseline spells (Allison, 1984), where a spell is defined as a subject's sojourn in a particular state, or make unrealistic assumptions such as $R$=0.

Excluding the baseline spells typically results in substantial loss of information. The $R$=0 assumption is generally incorrect and may lead to severe bias in parameter estimates unless the hazard rates are constant (Heckman and Singer, 1986), in which case there is no duration effect and thus no need for an SMP model.

Recently a more general approach to fitting the SMP model to left-censored data by using the stochastic EM (expectation-maximization) algorithm, denoted as SMP-EM, has been developed (Cai, Schenker and Lubitz, 2006, manuscript under review). This new approach is based on the use of an analog to the stochastic EM algorithm (Celeux and Diebolt, 1985), which is a variant of the popular EM algorithm (Dempster, Laird and Rubin, 1977). They argued that the unknown $R$ could be treated as a form of missing data and iteratively imputed so that the SMP model could be estimated based on all of the observed data using a conditional likelihood approach developed by Lancaster (1979). The stochastic EM algorithm yielded robust and quick convergence to common stationary distributions of coefficient estimates under different initial assumptions about the unobserved durations $R$. They compared the SMP-EM and MSLT estimates of disability incidence and recovery rates and found substantial differences in these estimates.

Cai, Schenker and Lubitz (2006) used the Medicare Current Beneficiary Survey (MCBS) to develop the approach and to illustrate its applications. The MCBS has a relatively short follow up period – respondents remain in the MCBS for a maximum of 3.5 years. Due to this limitation, it is difficult to conclude whether the new SMP-EM approach is valid, or how close the SMP-EM estimates are to the empirical rates. Data with a longer observation period therefore are needed to better evaluate the new approach.

In comparison, the Cardiovascular Health Study (CHS) is a better source of data for both fitting the SMP model using the stochastic EM algorithm and for evaluating its performance. The CHS has two unique advantages to suit our study: a long observation period of up to 16 years and persons who have IADL and ADL limitations at baseline were asked about how long they have had them. That is, for those who are IADL or ADL disabled at baseline, their $R$ is known. CHS data therefore will help us to address three important questions here. First, it will allow us to examine if the SMP-EM estimates are similar to the estimates based on fitting the SMP model using the known values of $R$. If the estimates are similar, then SMP-EM can be considered a valid approach to analyzing left-censored longitudinal data. Second, it will allow detailed assessment of the dynamic process to see if changes in functional status over time are indeed a duration-dependent process and if so, whether the duration effect is more accurately estimated by the SMP model. Finally, we will compare estimates from both SMP-EM and MSLT with the CHS data to help assess the adequacy of the current specification and implementation of SMP-EM, and to help future development of more realistic duration models. Comparisons include disability incidence, recovery and death, as well as summary measures such as life expectancy. In the following sections, we will describe the data and methods in detail, and conclude the study by summarizing the results and their implications.

## 2. Data

The Cardiovascular Health Study (CHS) is a population-based longitudinal study of coronary heart disease and stroke in adults aged 65 years and older (Fried et al., 1991). CHS has been used in longitudinal studies on changes in health status over time (Diehr et al., 1998; 2005). When the study started in 1989-1990, just over 5200 men and women (N=5201) were recruited from four communities in the US (Forsyth County, NC; Sacramento County, CA; Washington County, MD; and Pittsburgh, PA). An additional 687 African Americans were added in 1992-1993, bringing the total sample size of CHS to 5,888 people 65 years and older. Eligible participants were sampled from Medicare enrollment files. Those eligible included all persons living in the household of each individual sampled from the Medicare sampling frame, who were 65 years or older at the time of examination, were non-institutionalized, were expected to remain in the area for the next three years, and were able to give informed consent and did not require a proxy respondent at baseline. Potentially eligible individuals who were wheelchair-bound in the home at baseline or were receiving hospice treatment, radiation therapy or chemotherapy for cancer were excluded.

In this study functional disability is defined on the basis of IADL and ADL limitations. IADL limitation is defined as having difficulty performing any of six IADL activities because of health problems: using the telephone, doing light housework, doing heavy housework, preparing one's own meals, shopping for personal items, and managing money. ADL limitation is defined as having difficulty performing any of six ADL activities because of health problems: bathing/showering, dressing, eating, getting in or out of bed or chairs, walking, and using the toilet. We classify a living person into two mutually exclusive functional states: active health and disability. Disability includes one or more functional limitations in either IADL or ADL; active health is free of any functional limitation in both IADL and ADL. Death is the third and absorbing health state.

Between baseline (1989-1990) and 1998-1999, the study conducted annual clinic visits and asked detailed questions on functional status limitation on every visit. Since 1998-1999 the study continued with phone interviews every six months through the present time. Information on deaths of sampled persons was only available through 2003, however.

The annual clinic visit was supplemented by telephone follow-ups every six months after the clinic visit. These interviews asked the sampled person whether their overall functional status had changed and, if so, in what direction. We use these telephone interviews to impute missing clinic visit data wherever appropriate. The 6-month telephone interviews after 1998-1999 (when annual clinic visits were discontinued) were more detailed and we used them to derive annual observations of functional status through 2002-2003. The final analysis sample contains annual observations of functional status of up to 14 years for

5,840 persons aged 65 years and older after excluding 48 persons whose sample records are insufficient for the present analysis.

## 3. Methods

The application of the stochastic EM algorithm to fitting an SMP model involves two steps. The E-step imputes the unknown *R*-values for the left-censored spells based on the current parameter estimates for the SMP model, and the M-step fits the SMP model to the "pseudo-complete" data resulting from imputation. The algorithm iterates back and forth between the E- and M-steps, successively updating the imputed values of *R* and then the estimated parameters of the SMP model, until the distributions of the SMP model parameter estimates stabilize.

### *3.1 The E-step: Imputing R-values for left-censored spells*

The stochastic E-step performs a random draw from the approximate conditional distribution of the missing data. Thus, for our application, we needed to first approximate the conditional distribution of unknown *R*-values, given the observed data and the estimates from the previous iteration. This was accomplished by simulating a cohort from which imputed values of *R* were drawn. Since the youngest persons in our sample came under observation at age 65, the simulated cohort must be younger than 65 in order for the *R*s to be imputed. We decided to simulate a cohort of age 55 so that the estimated transition rates are likely to remain valid for the old-age functional status transitions under our study.

To illustrate the steps involved in simulating the 55-year old cohort, Figure 0 displays a schematic diagram for a hypothetical 55-year old man. First, an initial functional status and an initial elapsed duration in that state were assigned to the man. Transition probabilities based on the most recently fitted SMP model were then used to simulate the man's transitions from one year to the next, until his death. Further details of these steps are as follows:

- **Initial functional status**

  The distribution of functional status for 55-year olds is not directly observed in the CHS; instead it was estimated from a multinomial logistic regression fitted to the observed data at ages 65 and over by gender. A random draw from the estimated distribution was used as the initial functional status for the 55-year old man.

- **Initial elapsed duration**

  The initial elapsed duration was drawn from the distribution of imputed $R$s at age 65 from the previous stochastic E-step, under the assumption that the distributions of elapsed durations are similar at ages 55 and 65. This assumption could be replaced by another one, if desired.

- **Transition probabilities**

  Given the man's functional status and elapsed duration at age 55, along with the most recent estimates of the SMP model parameters, we computed the probabilities of belonging to each possible state at age 56, and then used these transition probabilities to randomly generate his status at age 56. If he died, then the simulation stopped and we returned to simulate a new person. If he survived to age 56, then a new set of transition probabilities was computed, corresponding to his status and duration at age 56, and his status at age 57 was generated. Such yearly transitions were generated until his death.

In the present study, we simulated a cohort of 5,000 persons for each E-step. The simulated spells for these 5,000 people were large enough to allow all of left-censored spells to be imputed. These imputed left-censored spells were joined with the other spells with observed beginnings to form an updated set of "pseudo–complete" data for the next M-step.

### *3.2 The M-step: Fitting the SMP model to the pseudo-complete data via the conditional likelihood approach*

Given a set of pseudo-complete data from the E-step, the M-step fits the SMP model using the conditional likelihood approach (Lancaster, 1979). The CHS data, like many other longitudinal studies, schedules interviews at discrete time intervals. Events are only known to have occurred somewhere between annual observations. The imprecise measurement of event timing suggests that a discrete-time hazard model is appropriate (Allison, 1982). However, the proposed algorithm, including both the E-step and the M-step, can also be implemented to fit a continuous-time SMP model.

Using PROC LOGISTIC in SAS, we fit a discrete-time hazard model using the pseudo-complete data with age at the beginning of the spell, duration of current status and its natural log as covariates. The dependent variable is functional status at the end of each age interval. The fitted coefficients are used to estimate age-duration-state-specific transition probabilities for use in the next iteration of E-step. The transition probability estimates are extrapolated to ages under 65 to facilitate simulations of the 55-year old cohort.

### *3.3 Initialization and convergence of the stochastic EM algorithm*

In this study, we selected three different sets of $R$ values to initialize the algorithm: 1) $R$=0, 2) $R$=2$T$, and 3) $R$=4$T$. $T$ is the observed portion of the total duration $W$. The assumption that $R$=0 implies that $W$ is equal to the observed duration $T$ for each left-censored spell, and it represents a rather extreme assumption about the hidden durations. The other two assumptions represent naïve assumptions that the unobserved durations are proportional to $T$.

The convergence behavior of the stochastic EM algorithm is presented in Figure 1, where we show the sequence of coefficient estimates for both age and duration effects from the first 75 iterations of fitting SMP-EM to the sub-sample of respondents who are disabled at baseline. The use of this subsample is explained in the following sections. Each panel in the figures includes three unique series that correspond to the different sets of initial values for $R$.

It appears that the distributions of estimates for both the duration and the age effect have converged quickly. There are substantial differences in the initial estimates. These initial differences are reduced quickly, and the algorithm appears converging to common stationary distributions after just a few iterations. Had there not been iterations, the estimates of the parameters would have been dependent on the assumed values for *R*.

## 4. Analysis Plan

We will address three questions in this study. The first question asks whether the SMP-EM estimates are similar to the estimates based on fitting the SMP model using known values of *R*. Since CHS does not ask those who are in active health at baseline how long they have been in active health, the recalled duration *R*s are known only for those who are disabled at baseline. Therefore we perform two separate tests using different subsets of the CHS sample.

The first subsample includes only those persons who are disabled at baseline. This results in a sample of 2,327 persons with 5,834 spells. We use SMP-RD to denote fitting the SMP model using the recalled duration in the disabled sample. The second is a sub-sample of CHS split at the end of wave 7 (the 6th year). This subset only includes those spells that span across wave 7 for those people who have stayed in the study longer than 6 years. In this subset we use the observed duration between baseline and wave 7 as the new elapsed duration. The rationale for creating such a subset was two-fold. First, the observed duration of functional status is likely more reliable than the recalled duration. Second, this sub-sample included those who were healthy at baseline. The limitation of this split sample is the loss of information on those individuals who have stayed in CHS for less than 6 years (1,086 persons) and the earlier spells that have ended before wave 7 for the remaining persons (5,387 spells). This sub-sample included 4,754 persons with 7,669 spells. We use SMP-OB to denote fitting the SMP model using the observed durations in this sample.

To answer the first question, we compared the distributions of imputed $R^*$s with that of the known durations, estimates of life expectancy at age 65, the pattern of disability incidence, recovery and mortality rates as a function of duration, based on SMP-RD or SMP-OB, with estimates from SMP-EM. The estimates based on SMP-RD and SMP-OB are considered the true estimates of the SMP model. They were estimated in similar manner as the M-step in the SMP-EM approach using Lancaster's conditional likelihood approach (1979). For both SMP-RD and SMP-OB estimates , they were summarized from a 100,000 65-year old cohort simulated using the estimated transition probabilities. The SMP-EM estimates were averaged over 10 successive iterations (from iterations 26-35) of the $R$=0 series. At each of the 10 iterations, estimates of interest were obtained by summarizing the characteristics of a simulated cohort of 100,000 65-year olds using the transition probabilities estimated in the current M-step. The distribution of functional status at 65 for the simulated cohort in SMP-RD, SMP-OB and SMP-EM were equal to the sample proportions.

To facilitate assessment of the differences in life expectancy estimates, we estimated their standard errors using bootstrap method (Lohr, 1999). We sampled $n_h$-1 respondents with replacement from each of the four communities in the CHS sample, where $n_h$ is the number of respondents in county $h$. We fit a separate discrete-time hazard model to each bootstrapped sample and simulate a cohort of 50,000 65-year olds to estimate life expectancy at 65. We repeated the bootstrap procedure 100 times. The standard deviations of the 100 estimates of life expectancy are the standard errors of the original point estimates.

The second question evaluates the existence of a duration effect on transition probabilities in CHS data and the accuracy of the SMP-EM and MSLT estimates. To answer this question, we examined the patterns of empirical disability incidence, recovery and mortality rates in relation to duration, and the prediction errors for the SMP-EM and MSLT estimates. The transition probabilities for the MSLT model was estimated by fitting a discrete-time hazard model with functional status at the end of each age interval as the dependent variable and current age as the only covariate. To derive estimates of incidence, recovery and mortality rates, we used average estimates from 50 sets of prediction-validation samples. For each set, the 5,840-person CHS sample was evenly split into a prediction sample, where SMP-EM and MSLT were fit, and a validation sample, where predictions over a respondent's lifetime were made using his or her baseline characteristics (age, functional status and recalled duration).[1] For each validation sample, the SMP-EM estimates are averages from iterations 41-50 from the $R$=0 series. For both SMP-EM and MSLT, predicted lifetime data is right censored at the end of the actual observation for the respondent.

[1] Elapsed durations for those active at baseline are assumed to be the maximum of imputed $R^*$ using the prediction sample.

Prediction error is measured by mean absolute percentage error (MAPE). It is calculated as the average percent of prediction errors between predicted (SMP-EM) and empirical rates over the range of durations (1, 2,..., 10). MAPE is a simple and useful measure of prediction accuracy, and was used in Porell, Tompkins and Turner (1990) to assess the accuracy of projected Medicare reimbursements.

Finally, we compared the SMP-EM estimates of age-specific transition rates with those from MSLT and CHS data, and life expectancy estimates from SMP-EM and MSLT. The purpose of this comparison is to evaluate the current specification of SMP-EM, and to highlight areas of potential misspecification to help future development of more realistic duration models. We examine age-specific rates because age-specific MSLT rates are theoretically weighted averages of the SMP rates. Given current age $x$ and status $i$, the conditional probability of entering state $y_{x+1}=j$ in a SMP model is given by $P_S = \Pr(y_{x+1} = j \mid y_x = i, d = l)$, where $d = l$ indicates duration of $l$ years in current state $i$. On the other hand, the conditional probability at current age $x$ under MSLT is given by $P_M = \Pr(y_{x+1} = j \mid y_x = i)$ . It is straightforward to see that the $P_M$ is a weighted average of $P_S$ ( $P_M = \sum_l \Pr(d = l \mid y_x = i) * P_S$ ), with the weight, $\Pr(d = l \mid y_x = i)$, being the probability of duration $l$ conditional on current status. In the SMP-EM approach, $P_S$ is estimated by the hazard model in the M-step, and the duration distributions are imputed in the E-step. The hypothesis that age-specific rates from both MSLT and SMP-EM should be similar thus depends on whether *both* components, the weights and $P_S$ , are correctly estimated.

Estimates of life expectancy at 65 (total, active and disabled) are of unique interest to demographic research in general, and to demographic research of healthy aging in particular. Given the nature of these summary measures, comparison of LE estimates is a more comprehensive assessment of the differences between the two approaches to analyzing the dynamics of functional status changes. Unlike the previous examination of transition rates, however, comparison of LE estimates is complicated by the fact that there is not a *true* life expectancy value to compare with. We only have estimates that at best are approximate: the MSLT estimates ignore the duration effect altogether; while the SMP-EM estimates are affected by both the misspecification of duration effect and the imputation of duration distribution. We therefore cannot provide any measure of prediction accuracy for this comparison.

The procedure for comparing SMP-EM and MSLT estimates with estimates derived directly from the CHS data is similar to the previous one, where 50 sets of prediction-validation samples are used. Life expectancy estimates from SMP-EM and MSLT are produced using the full CHS sample. Given dependence on duration and current status in a SMP model, status-based SMP-EM estimates of LE at 65 are weighted averages of LE estimates for individual members of the cohort in current states, with the weights being the conditional distribution of durations. Population-based estimates are weighted averages of the status-based estimates, with the weights being the distribution of states at 65. Bootstrap standard errors are also provided to assess the statistical significance of the differences.

## 5. Results

### *5.1*

Figures 2 & 3 compare the histograms of imputed $R^*$s with the $R$s that are known. The histogram of recalled $R$s in the disabled sample show a large concentration around 2 years of length, for which the distribution of imputed $R^*$s has failed to reproduce accurately (Figure 2). The distributions of imputed and recalled durations at higher values are more similar. Using the split sample, the histogram of observed $R$s (censored at the end of the $6^{th}$ year) indicates a bi-modal shape at both ends (Figure 3). To make fair comparisons here, we censored the imputed $R^*$s at the end of the $6^{th}$ year too. But the distribution of imputed $R^*$s cannot reproduce the bi-modal shape; it shows a significant concentration at the upper end.

Table 1 presents estimates of total, active and disabled life expectancy (TLE, ALE and DLE) from the three SMP models and their standard errors. Panel A compares the estimates from SMP-EM and SMP-RD using the disabled sample; none of the estimates are significantly different. The estimates in Panel B are based on fitting SMP-EM and SMP-OB to the split sample. For all 65-year olds and their subgroups, the SMP-EM estimates of TLE and DLE are significantly lower than SMP-OB estimates. Estimates of ALE are not significantly different.

Figure 4 compares the recovery and mortality rates as a function of duration using the disabled sample.[2] Recovery rate is defined the probability of experiencing recovery from disability within the next year for those who are currently disabled. Mortality rates are estimated for persons in both current states (active and disabled). The SMP-EM and SMP-RD estimates of both recovery (Figure 4A) and mortality rates (Figure 4B) indicate similar patterns, with a cross-over at the 5$^{th}$ year of duration. The SMP-RD estimates of mortality rates seem to level off after six years in disability, while the SMP-EM estimates continue to rise with duration. Figure 5 compares the estimated incidence, recovery and mortality rates using the split sample. In all three panels of Figure 5, the values and patterns of SMP-EM and SMP-OB estimates are similar.

In summary, the comparisons suggested that fitting the SMP model by using the stochastic EM algorithm may be a valid approach to the left censoring problem. The SMP-EM estimates of incidence, recovery and mortality are reasonably close to the SMP estimates using the known values of *R*s. The shapes of the histograms did not match exactly, especially in Figure 3 where the true distribution is bi-modal. But the mismatch did not affect the estimates of

---

[2] Incidence rates are identical under SMP-EM and SMP-RD since spells that end in disability incidence are completely observed after the study began and do not need imputation from the stochastic EM algorithm.

transition rates since the relationship of incidence and recovery to duration should be independent of the distribution of $R$s.

***5.2***

Figure 6 shows the SMP-EM estimates of disability incidence, recovery and mortality rates as a function of duration, and compared them with the empirical rates, using 50 sets of prediction-validation samples. Two observations can be made from the three panels. First, the CHS empirical rates exhibit substantial duration effects on all three types of transition rates. Incidence and recovery rates fall with duration (Figures 6A and 6B); while mortality rates rise with duration (Figure 6C). Second, SMP-EM estimates exhibit similar patterns of duration dependence as the CHS empirical rates, and the prediction errors are smaller than those for the MSLT estimates. The MAPE for recovery and mortality rates estimated by SMP-EM are 9% and 11%, while for MSLT estimates they are 116% and 48%, respectively. The MAPE for SMP-EM incidence rate is relatively large (63%), although still less than that for MSLT estimates (81%). The larger error in predicting incidence rates is likely due to the greater variability at longer durations (Figure 6A). Overall, the above comparisons suggest that the SMP model is likely to be a more appropriate model for analyzing such duration-dependent processes of functional status changes over time, and the form of duration dependence is relatively well estimated in the current specification of SMP-EM.

***5.3***

The three panels in Figure 7 indicate that the MSLT estimates of disability incidence and recovery by age are closer to empirical rates than SMP-EM estimates, while the prediction errors are similar for mortality rates. The MAPE for MSLT estimates of incidence, recovery and mortality rates are 7%, 11% and 10% respectively, while the MAPE for SMP-EM estimates are 26%, 20% and 11%. The small prediction errors for the MSLT estimates are not surprising: both the MSLT and empirical rates are conditional on age only; MSLT only uses a logistic regression to smooth the estimates.

Table 2 indicates that the SMP-EM and the MSLT estimates of LE for the entire 65-year old population and its distribution in active and disabled states are not statistically different. Status-based LE estimates show some significant differences, however. For active elderly, SMP-EM estimate of ALE is significantly higher. For disabled elderly, SMP-EM estimates of TLE and ALE are significantly shorter, while the estimate of DLE is significantly higher. SMP-EM thus implies that disabled 65-year olds would live shorter lives and spend more time in disability than MSLT estimates imply.

## 6. Conclusion

Our study has two important findings. First, the approach of fitting SMP model to left-censored data using the stochastic EM algorithm appears to perform as well as semi-Markov models that have the advantage of actual duration data. The SMP-EM estimates of disability incidence, recovery and mortality rates are similar to the true SMP estimates that use the known values of $R$. This should encourage further tests of the SMP model with left-censored longitudinal data.

Second, thanks to the long observation period of CHS, we were able to show clearly that the process of functional status transitions is indeed duration-dependent, and that the SMP-EM estimates of transition rates are closer to the empirical values than the MSLT estimates.

In some cases the performance of the SMP-EM model was less than ideal. The distribution of imputed $R^*$ did not quite match the true bi-modal distribution in the split sample, although the mismatch, as other results indicated, did not affect duration-dependent estimates of transition rates since the form of duration dependence is independent of the shape of the duration distribution. On the other hand, the shape of the distribution of imputed $R^*$ matters significantly to the comparisons of age-specific transition rates. As we saw earlier, age-specific rates are weighted averages of the duration-dependent rates, with the weights given by the distribution of duration. Discrepancies in the distribution of $R^*$ and $R$ therefore contributed to the worse results for SMP-EM in the comparisons of age-specific rates.

However, this should *not* be interpreted as evidence that the MSLT is a superior model. We need to carefully distinguish the *presentation* of the data from the *generation* of the data. The MSLT model ignores the duration effect completely, thus its characterization of the dynamics that generated the individual's event history is flawed. The SMP model is a better model because it realistically takes into account the duration effect on the individual's changes in health status over time. On the other hand, any data, no matter how it is generated, can be presented in age-specific format. The MSLT estimates performed well because the generation of MSLT estimates coincides with the presentation of CHS estimates. The SMP-EM estimates, on the other hand, are affected by estimates of both the duration distribution and the duration effect. Although the application of SMP model is an improvement in describing the dynamics, in its current form of specification it is not yet good enough to fully replicate the CHS data generation process, resulting in its poor performance in the comparisons of age-specific rates.

The current SMP-EM specification can be improved in at least two ways. First, the functional form of duration dependence is probably too simple in the present analysis. Our discrete-time hazard model only included the linear effects of duration and its natural logarithm in the logistic regression. The true, and likely more complex, form of duration dependence needs to be determined. Second, the uniqueness of some of the first spells is insufficiently dealt with in the current SMP-EM. For some sampled people, their first entry into disability occurred during the study period. Since initial onset is likely different from subsequent spells, a different set of transition probabilities may be necessary for these first spells (see discussions of a similar problem in Moffit and Rendall (1995)). Imposing restrictions on these prior durations, such as assuming no disability from birth to the beginning of the observation period, is possible within the current framework of the stochastic EM algorithm. However, it is difficult to identify the people for whom such restrictions should be imposed without additional information.

The use of SMP model potentially allows a more flexible and realistic description of the trajectories of changes in functional status by conditioning the likelihood of an event on not only the current status, but also its duration. Future research on duration models should also consider including other characteristics such as the presence and timing of past events to improve the understanding of the disablement and recovery processes.

## Reference


Allison, P.D. 1982. "Discrete-time methods for the analysis of event histories." In *Sociological Methodology* 1982, edited by S. Leinhardt, 61-98. San Fransisco: Jossey-Bass.

Allison, P.D. 1984. *Event History Analysis, Regression for Longitudinal Event Data.* Beverly Hills, California: Sage.

Celeux, G. and J. Diebolt. 1985. "The SEM Algorithm: A Probabilistic Teacher Algorithm Derived From the EM Algorithm for the Mixture Problem." *Comp. Statis. Quart.* 2: 73-82.

Cai, L., N. Schenker and J. Lubitz. 2006. "Analysis of Functional Status Transitions Using a Semi-Markov Process Model in the Presence of Left-Censored Spells." *Manuscript under review.*

Crimmins, E.M. and Y. Saito. 1993. "Getting better and getting worse." *Journal of Aging and Health* 5 (1): 3-36.

Dempster, A.P., N.M. Laird and D.B. Rubin. 1977. "Maximum Likelihood Estimation from Incomplete Data via the EM Algorithm (With Discussion)." *Journal of Royal Statistical Society B* 39: 1-38.

Diehr, P., D.L. Patrick, D.E. Bild, G.L. Burke and J.D. Williamson. 1998. "Predicting future years of healthy life for older adults." *Journal of Clinical Epidemiology* 51 (4): 343-353.

Diehr, P., L.L. Johnson, D.L. Patrick, D.E. B. Psaty. 2005. “Methods for Incorporating Death into Health-Related Variables in Longitudinal Studies.” *Journal of Clinical Epidemiology* 58: 1115-1124.

Fried, L.P., N.O. Borhani, P. Enright, C.D. Furberg; J.M. Gardin, R.A, Kronmal, L.H. Kuller, T.A. Manolio, M.B. Mittelmark, A. Newman, D. O'Leary, B. Psaty, P. Rautaharju and R. Tracy. 1991. “The Cardiovascular Health Study: design and rationale.” *Ann Epidemiol*. 3(1); 263-276.

Hardy, S.E. and T.M. Gill. 2005. “Factors Associated With Recovery of Independence Among Newly Disabled Older Persons.” *Archives of Internal Medicine* 165: 106-112.

Heckman, J.J. and Singer, B. (1986) Econometric analysis of longitudinal data. In Handbook of Econometrics, Vol. III, edited by Griliches, Z. and Intriligator, M.D. North-Holldand, Amsterdam, 1689-1763.

Lancaster, T. (1979) Econometric methods for the duration of unemployment. *Econometrica*, 47, 939-956.

Lohr, S. (1999) *Sampling: Design and Analysis*. Pacific Grove, CA: Duxbury Press.

Maddox, G.L., D.O. Clark and K. Steinhauser. 1994. “Dynamics of functional impairment in late adulthood.” *Social Science and Medicine* 38 (7): 925-36.

Mayer, K.U. and Tuma N.B. (1990) *Event History in Life Course Research*, University of Wisconsin Press, Madison.

Porell, F.W., C.P. Tompkins and W.M. Turner. 1990. "Alternative Geographic Configurations for Medicare Payments to Health Maintenance Organizations." *Health Care Financing Review* 11 (3): 17-30.

Schoen, R. (1987) *Modeling Multigroup Populations.* New York: Plenum Press.

**Figure 1. Schematic diagram of simulation of a 55-year old man in the stochastic E-step**

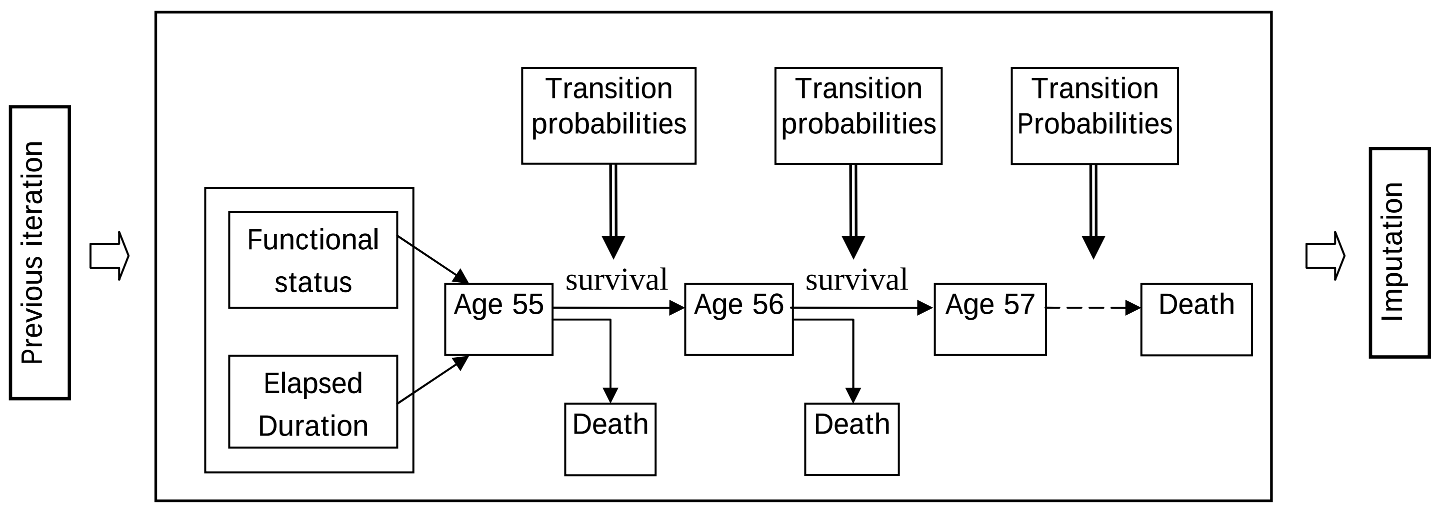
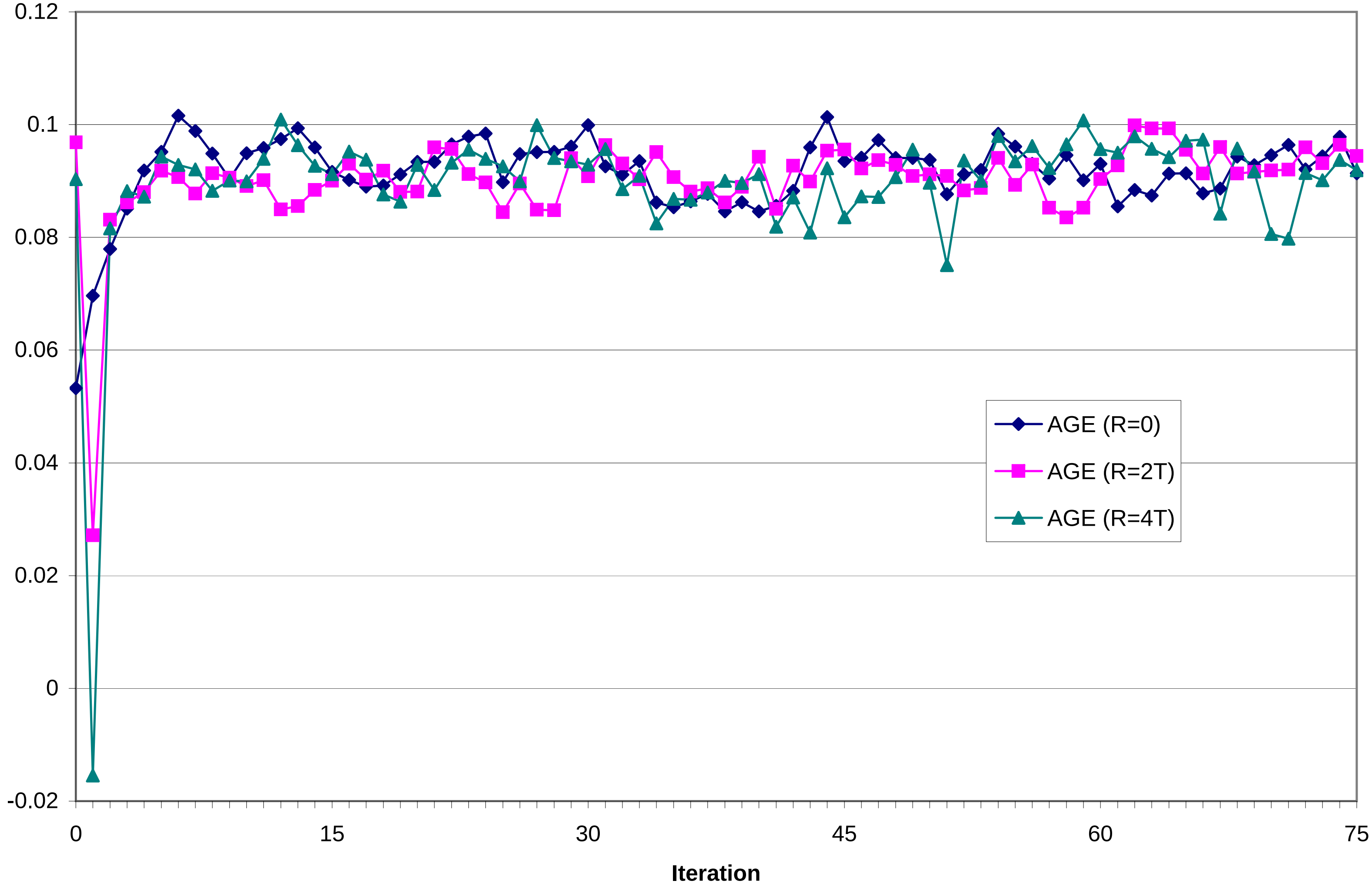


**Fig. 2 Convergence pattern of coefficient estimates from iterations of SMP-EM**

**A. Estimates for age**



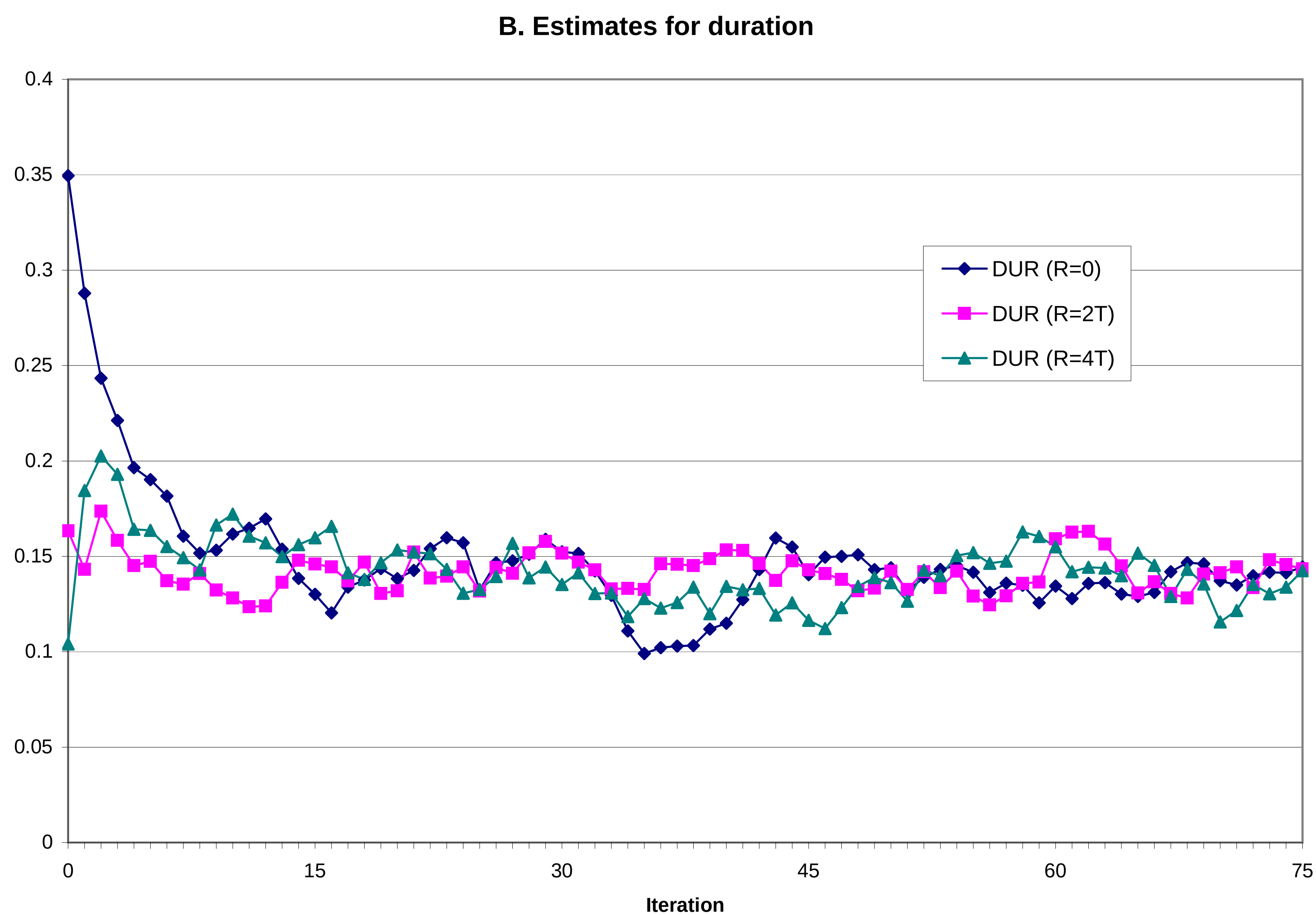

B. Estimates for duration
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
15
30
45
60
75
Iteration
DUR (R=0)
DUR (R=2T)
DUR (R=4T)

Fig. 3 Histogram of imputed and recalled R based on the disabled sample

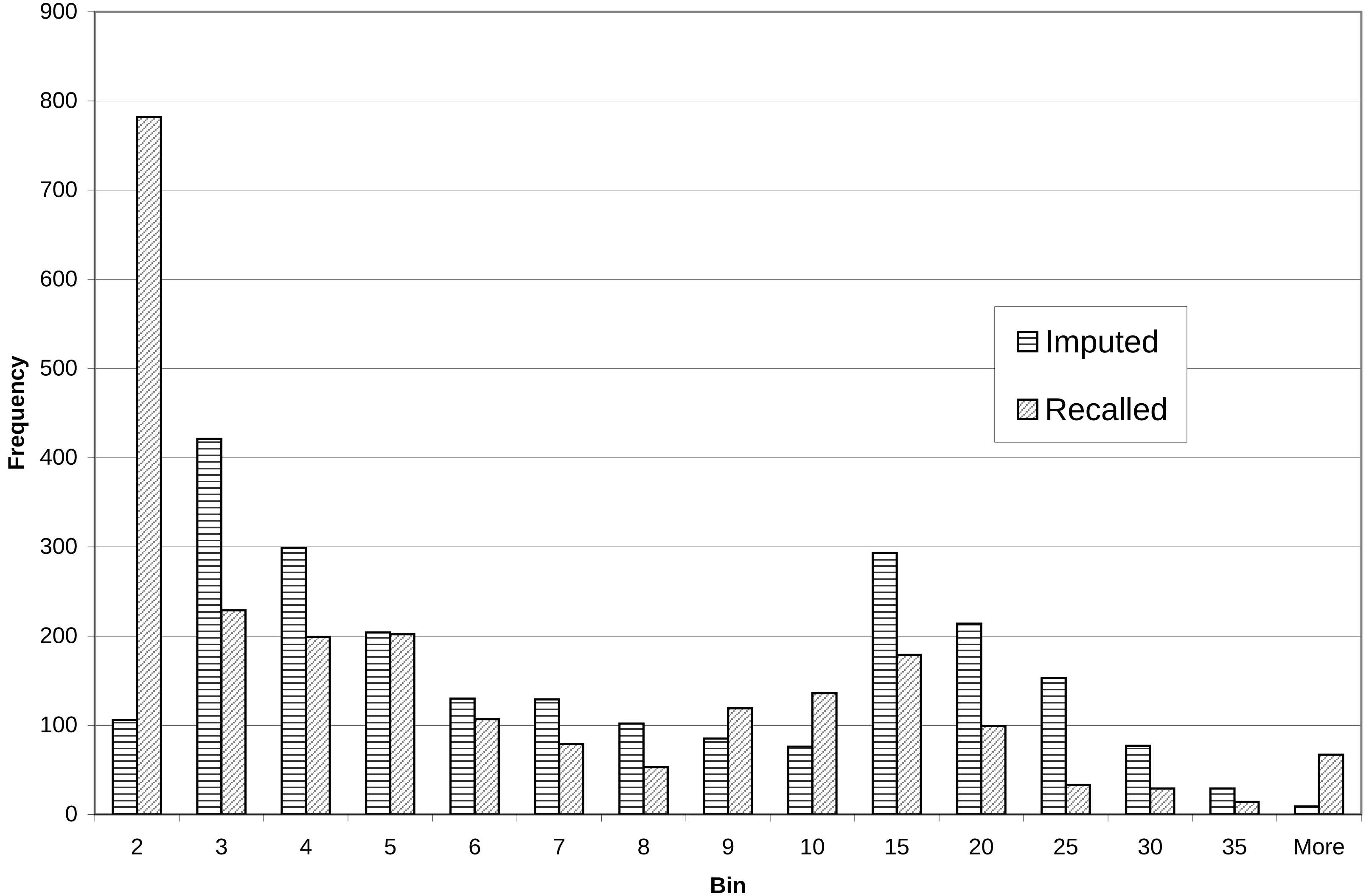


Fig. 4 Histogram of imputed and observed R based on the split sample

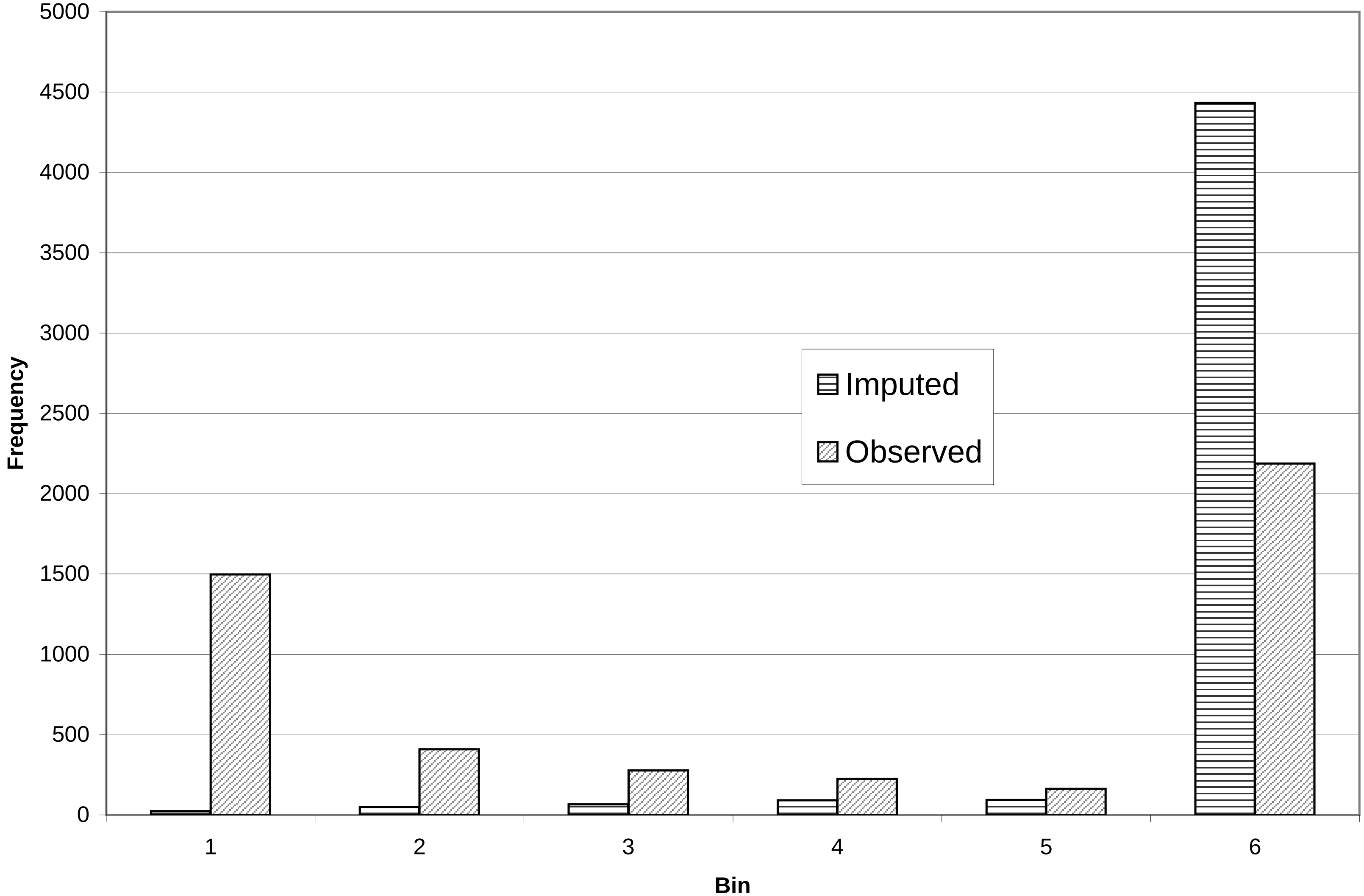


Table 1 Comparison of LE Estimates at Age 65 from SMP models

| | | Status @ 65 | TLE | ALE | DLE |
|---|---|---|---|---|---|
| A. Disabled sample, both sexes | SMP-EM | Disabled | 20.9 (1.6) | 8.5 (1.7) | 12.5 (0.3) |
| | SMP-RD | Disabled | 22.6 (1.6) | 9.7 (1.6) | 12.9 (0.4) |
| B. Full sample vertically split at the end of year 6, both sexes | SMP-EM | All | 23.7 (0.3) | 14.9 (0.8) | 8.8 (0.6) |
| | | Active | 24.7 (0.2) | 17.5 (0.3) | 7.2 (0.2) |
| | | Disabled | 19.5 (0.4) | 3.7 (0.5) | 15.8 (0.5) |
| | SMP-OB | All | 25.0 (0.3) | 14.4 (0.3) | 10.6 (0.3) |
| | | Active | 25.7 (0.4) | 16.9 (0.4) | 8.9 (0.4) |
| | | Disabled | 22.3 (0.6) | 4.9 (0.5) | 17.4 (0.5) |

*Note:* SMP-EM refers to fitting the SMP model using the EM algorithm, SMP-RD refers to fitting the SMP model using recalled duration and SMP-OB refers to fitting the SMP model using observed duration between baseline and split. TLE is total life expectancy. ALE is active life expectancy, the expected time to spend in active health during one's remaining life. DLE is disabled life expectancy, the expected time to spend in functional disability during one's remaining life. TLE is the sum of ALE and DLE.

Fig. 5A Disability Recovery Rates Predicted by SMP-EM and SMP-RD (recalled) Based on Disabled Sample

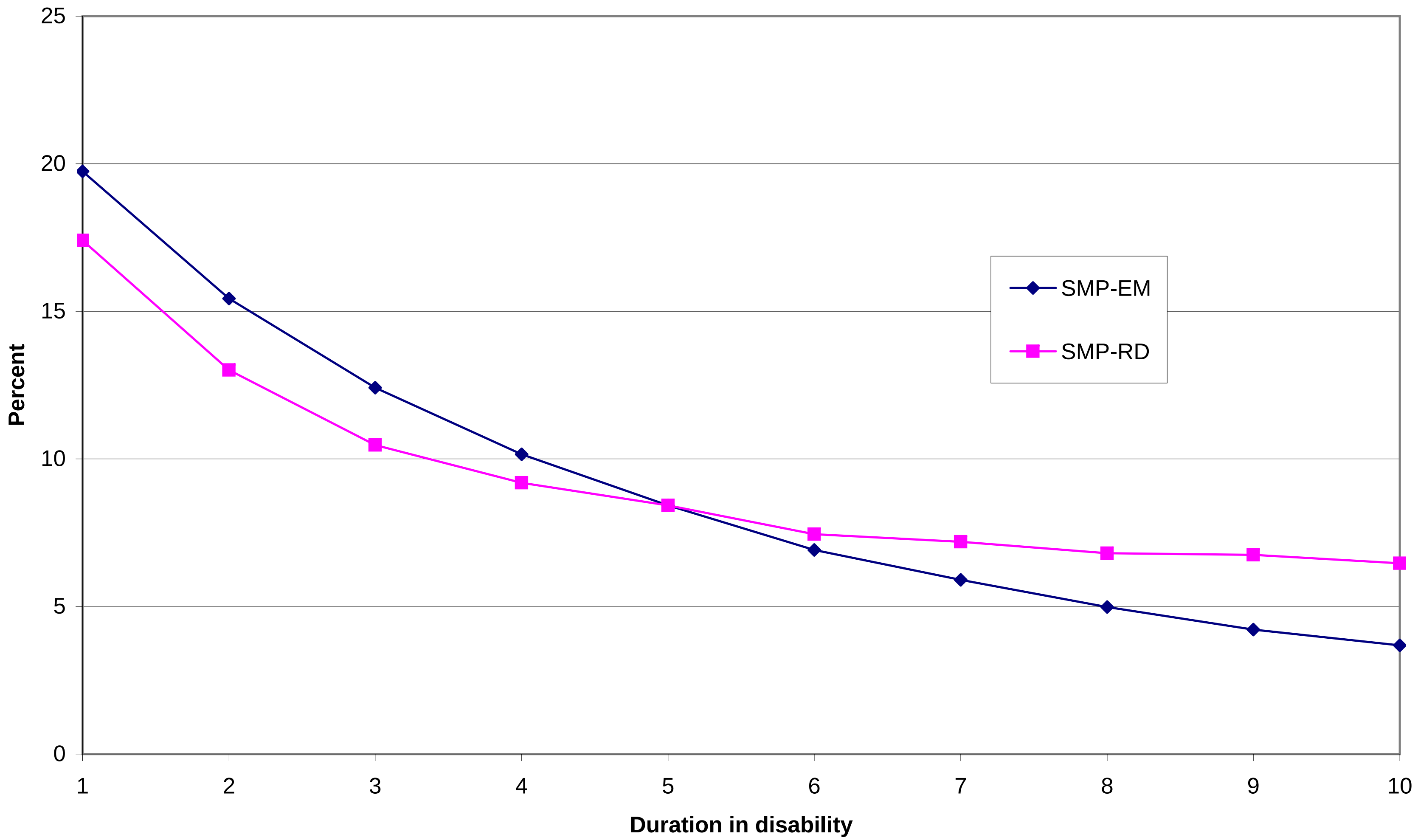


**Fig. 5B Mortality Rates Predicted by SMP-EM and SMP-RD (recalled) For Disabled Elderly Based on Disabled Sample**

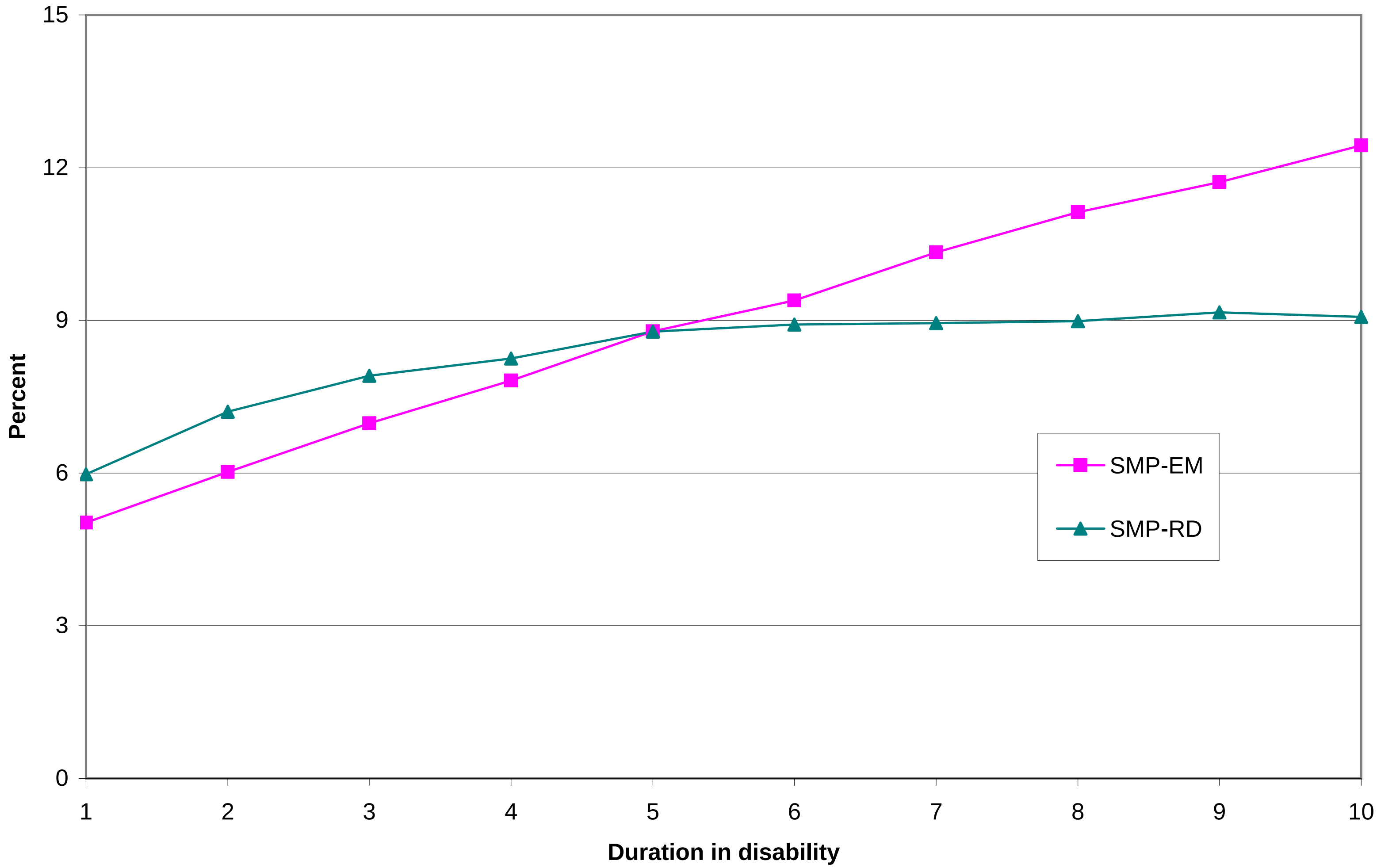


Fig. 6A Disability Incidence Rates Predicted by SMP-EM and SMP-OB (observed) Based on CHS Sample Split at the End of Year 6

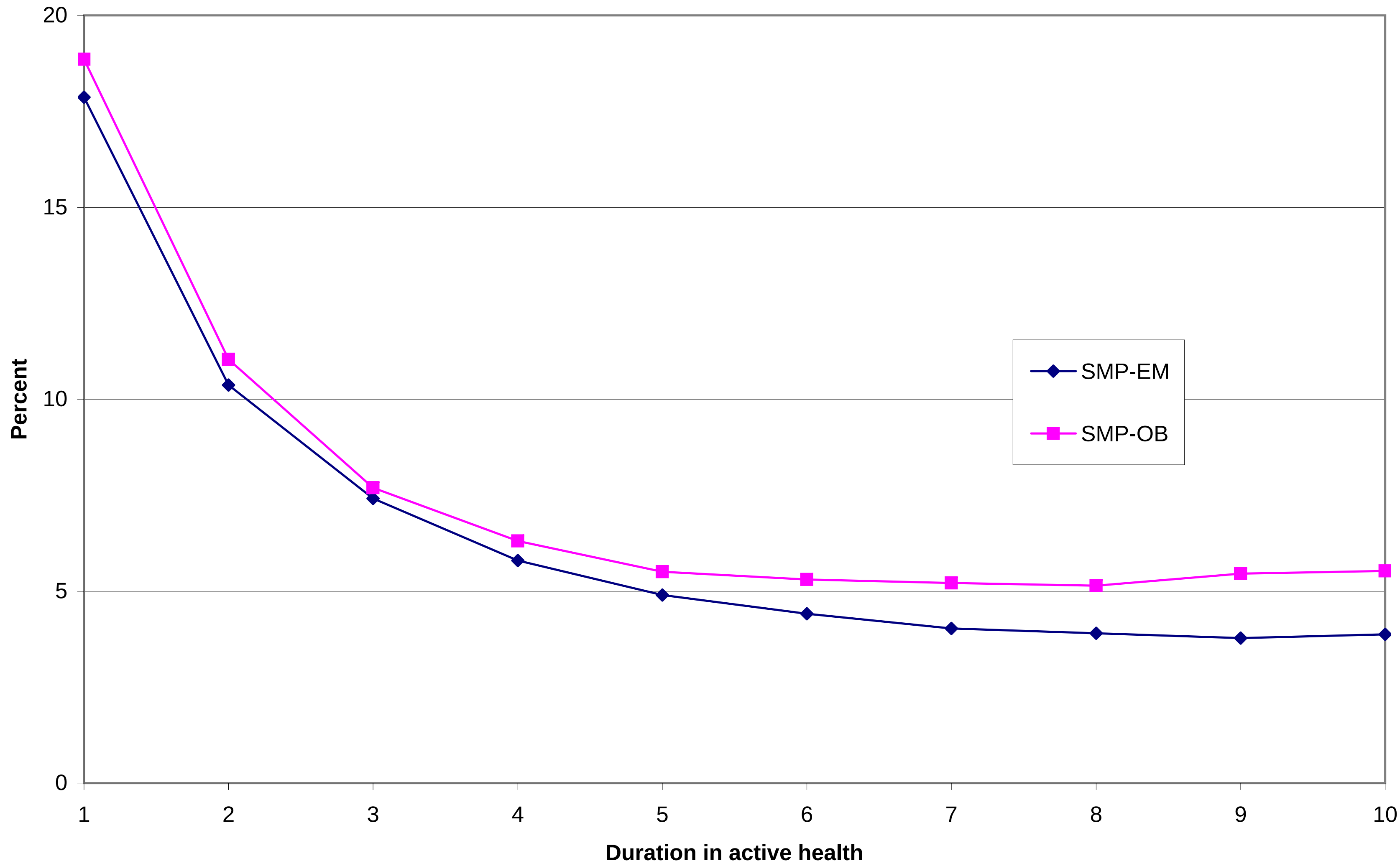


**Fig. 6B Disability Recovery Rates Predicted by SMP-EM and SMP-OB (observed) Based on CHS Sample Split at the End of Year 6**

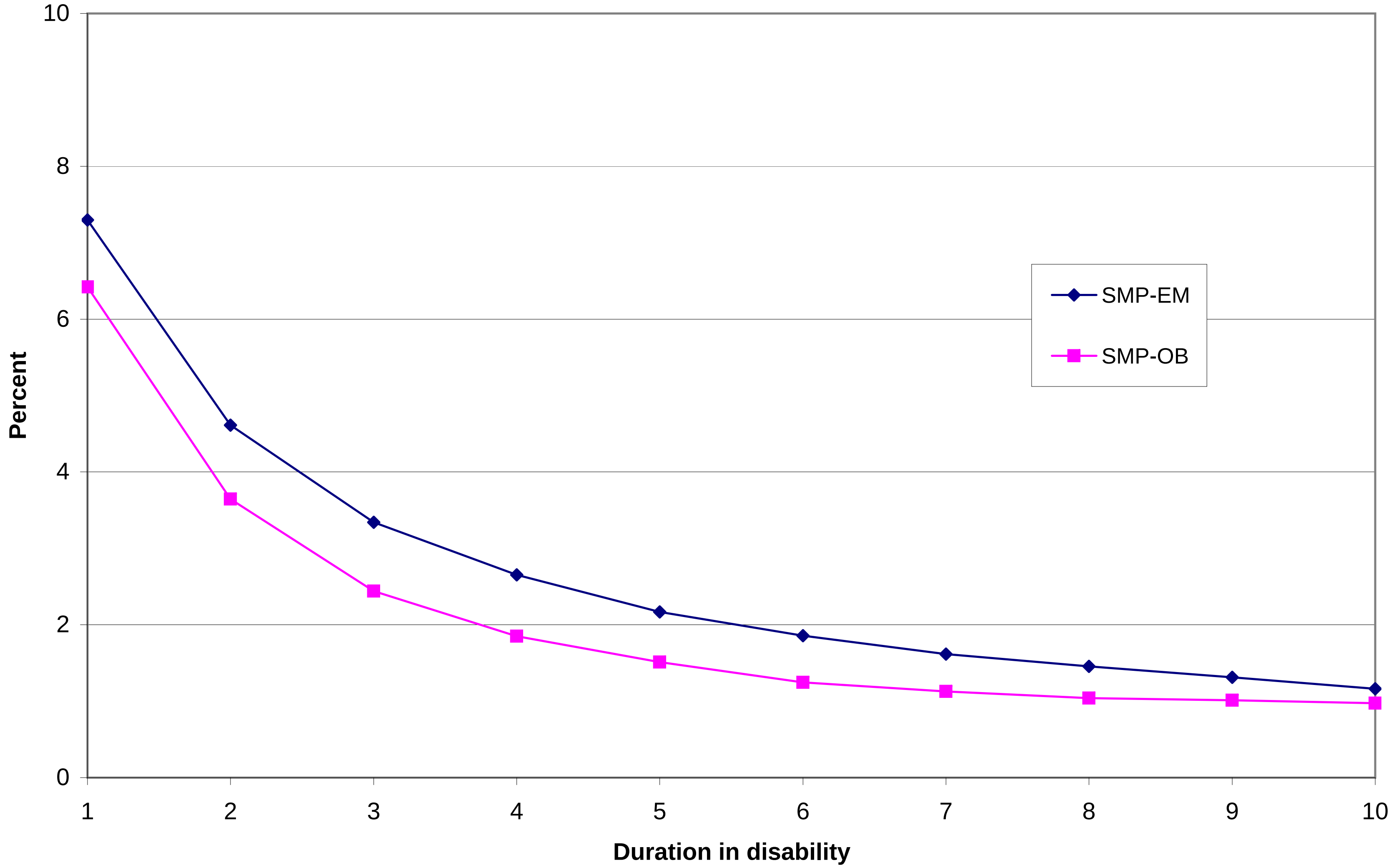


**Fig. 6C Mortality Rates Predicted by SMP-EM and SMP-OB (observed) For All Elderly Based on CHS Sample Split at the End of Year 6**

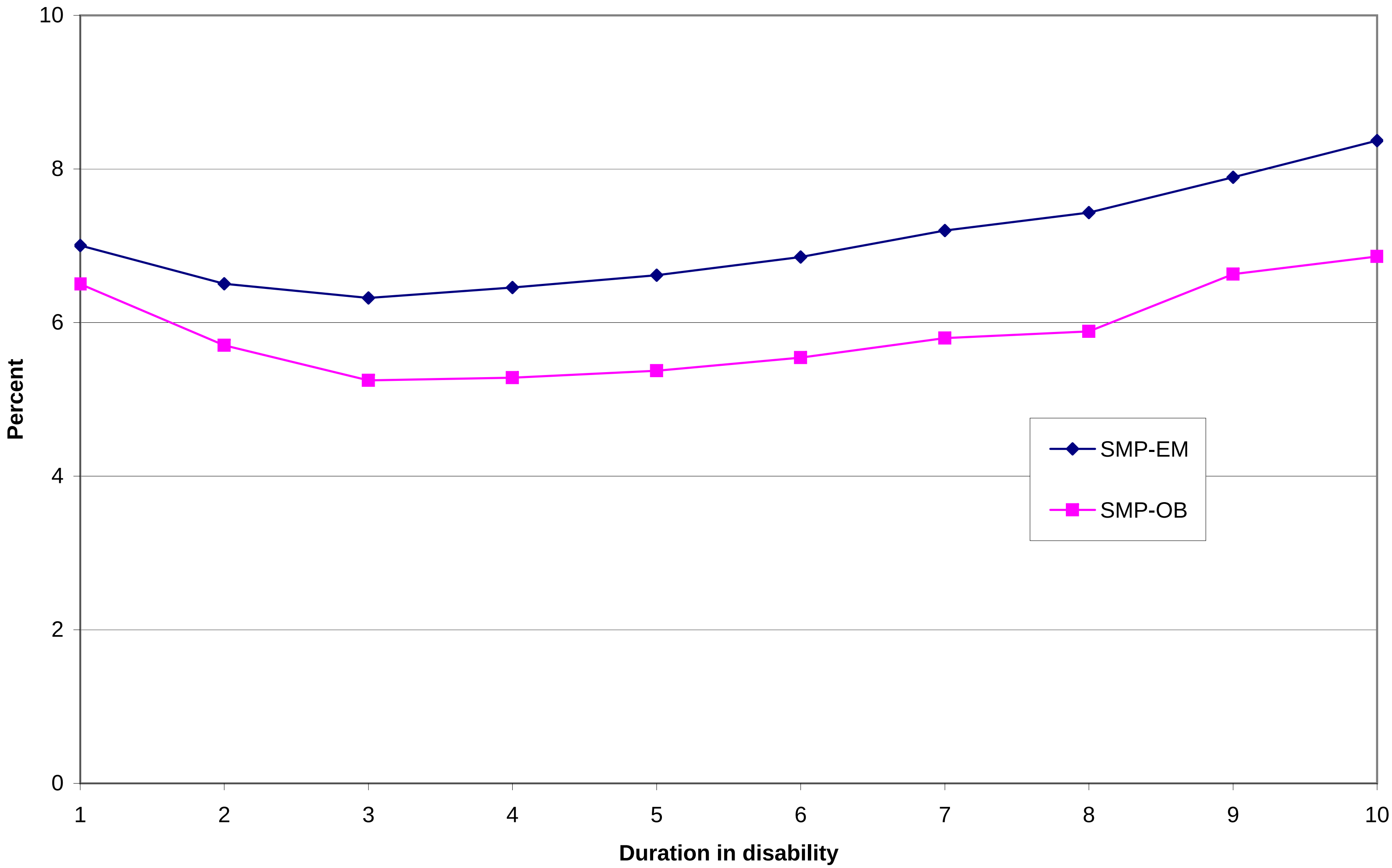


Fig. 7A Empirical and Predicted Disability Incidence Rates From SMP-EM and MSLT Based on CHS Validation Samples

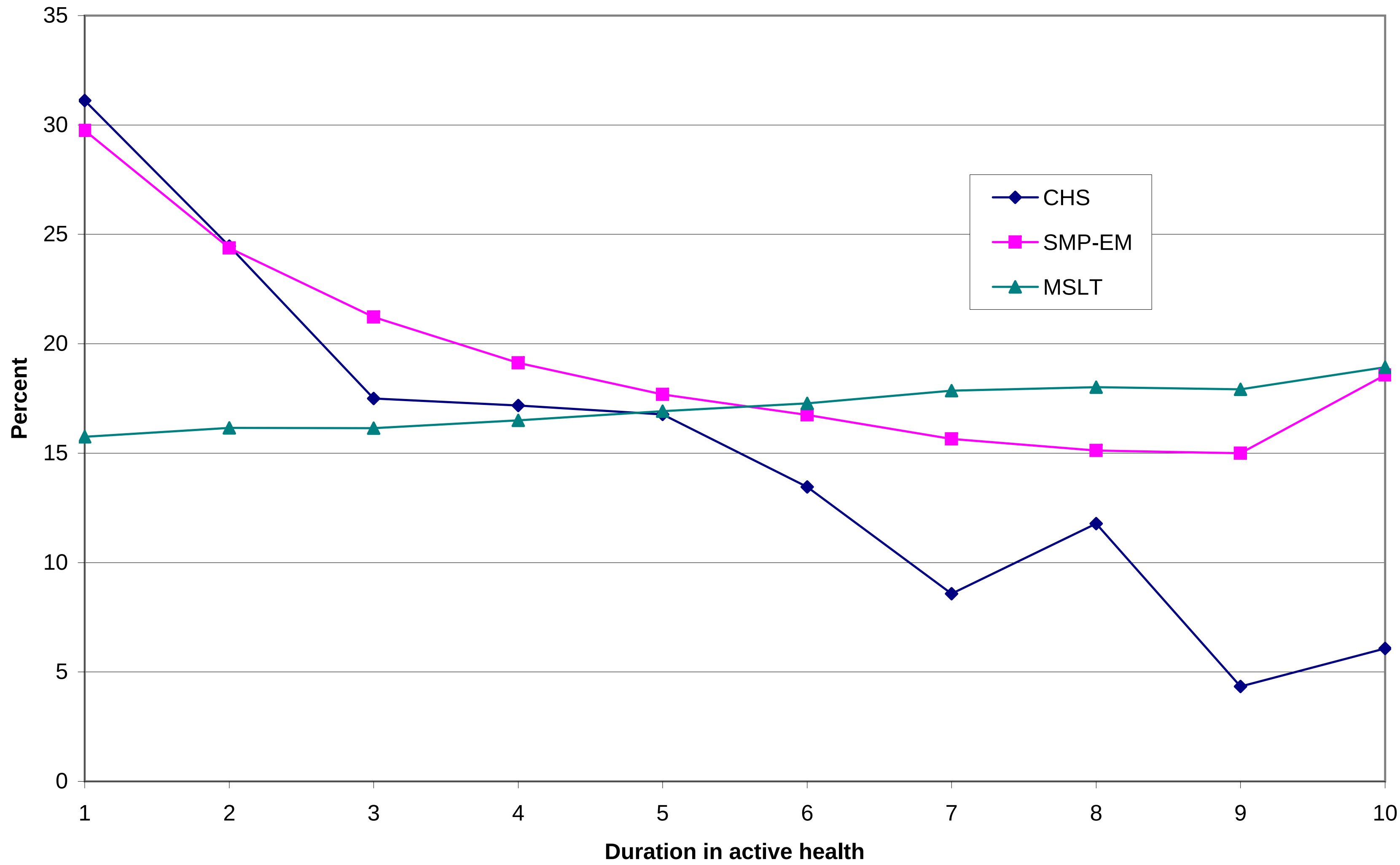


**Fig. 7B Empirical and Predicted Disability Recovery Rates From SMP-EM and MSLT Based on CHS Validation Sample**

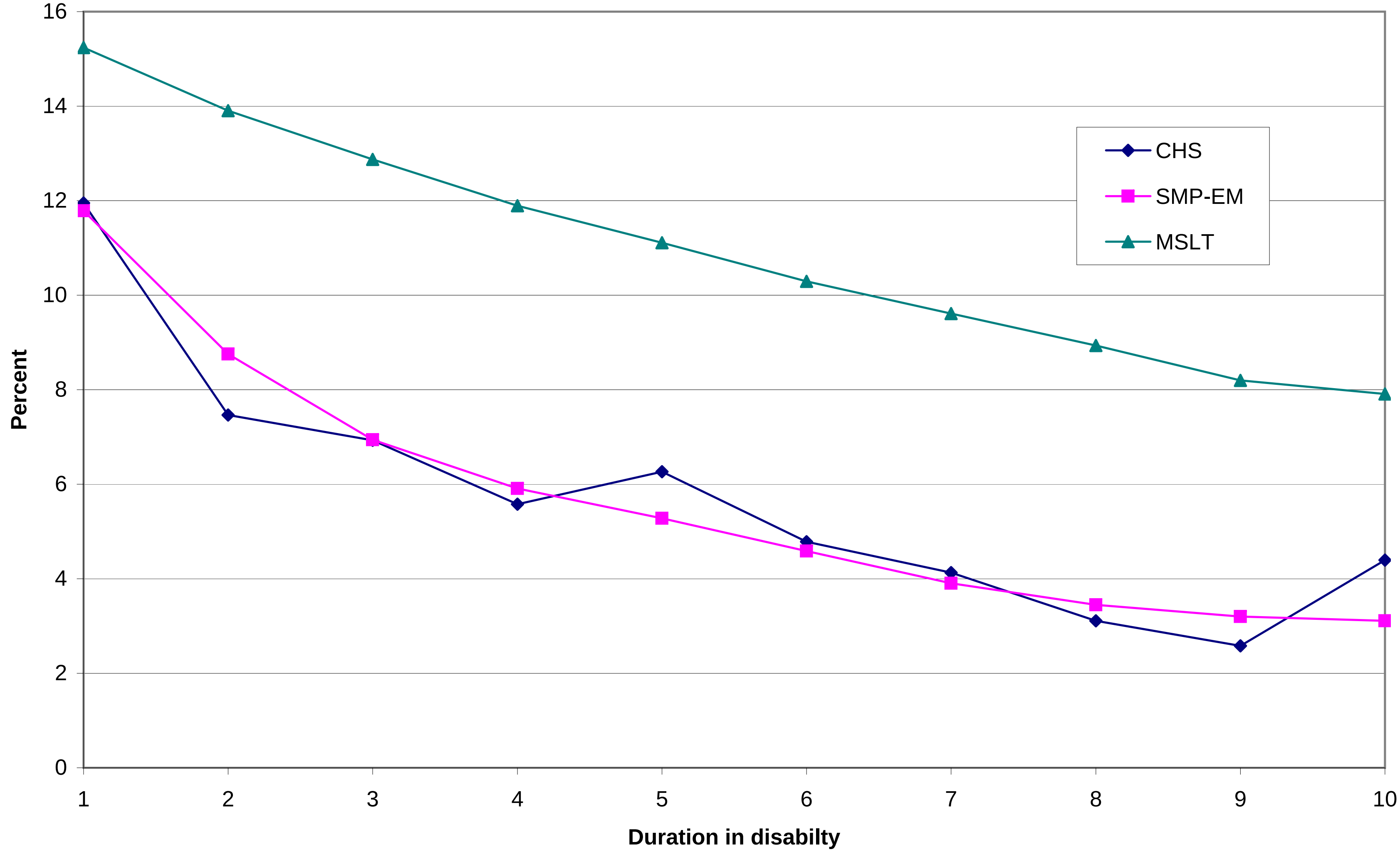
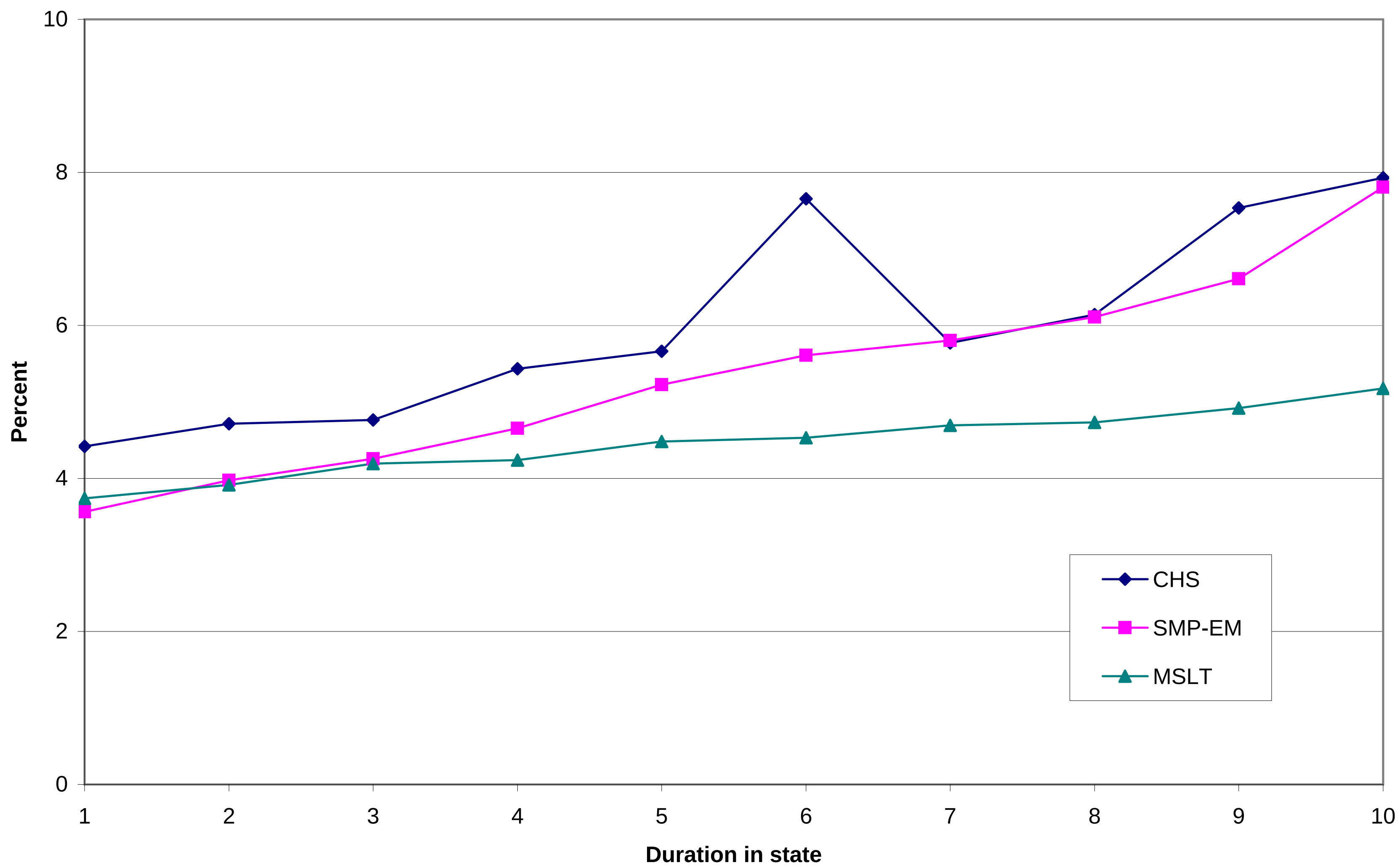


**Fig. 7C Empirical and Predicted Mortality Rates For All Elderly From SMP-EM and MSLT Based on CHS Validation Sample**



## Fig. 8A Age-Specific Empirical and Predicted Disability Incidence Rates From SMP-EM and MSLT Based on CHS Validation Sample

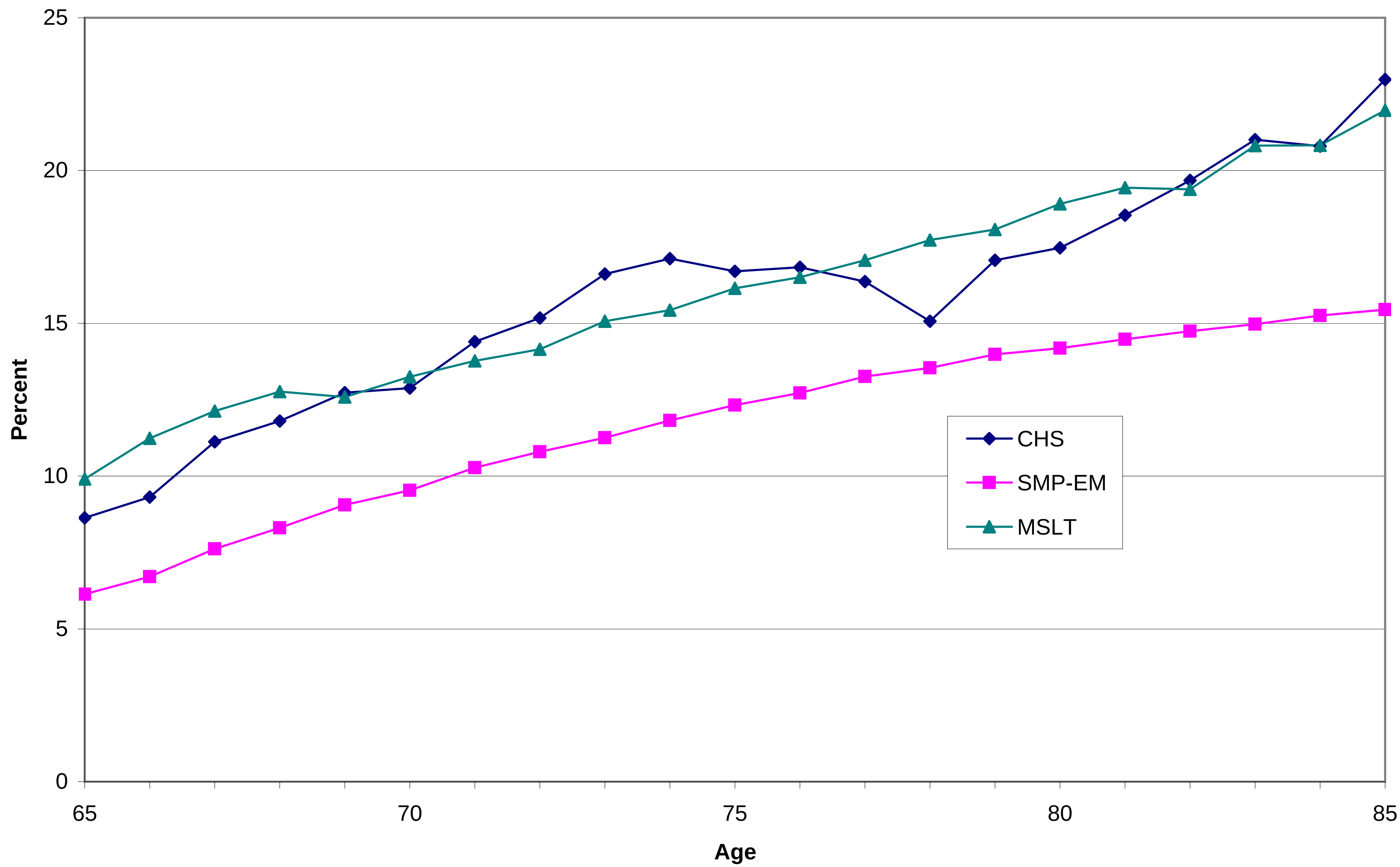
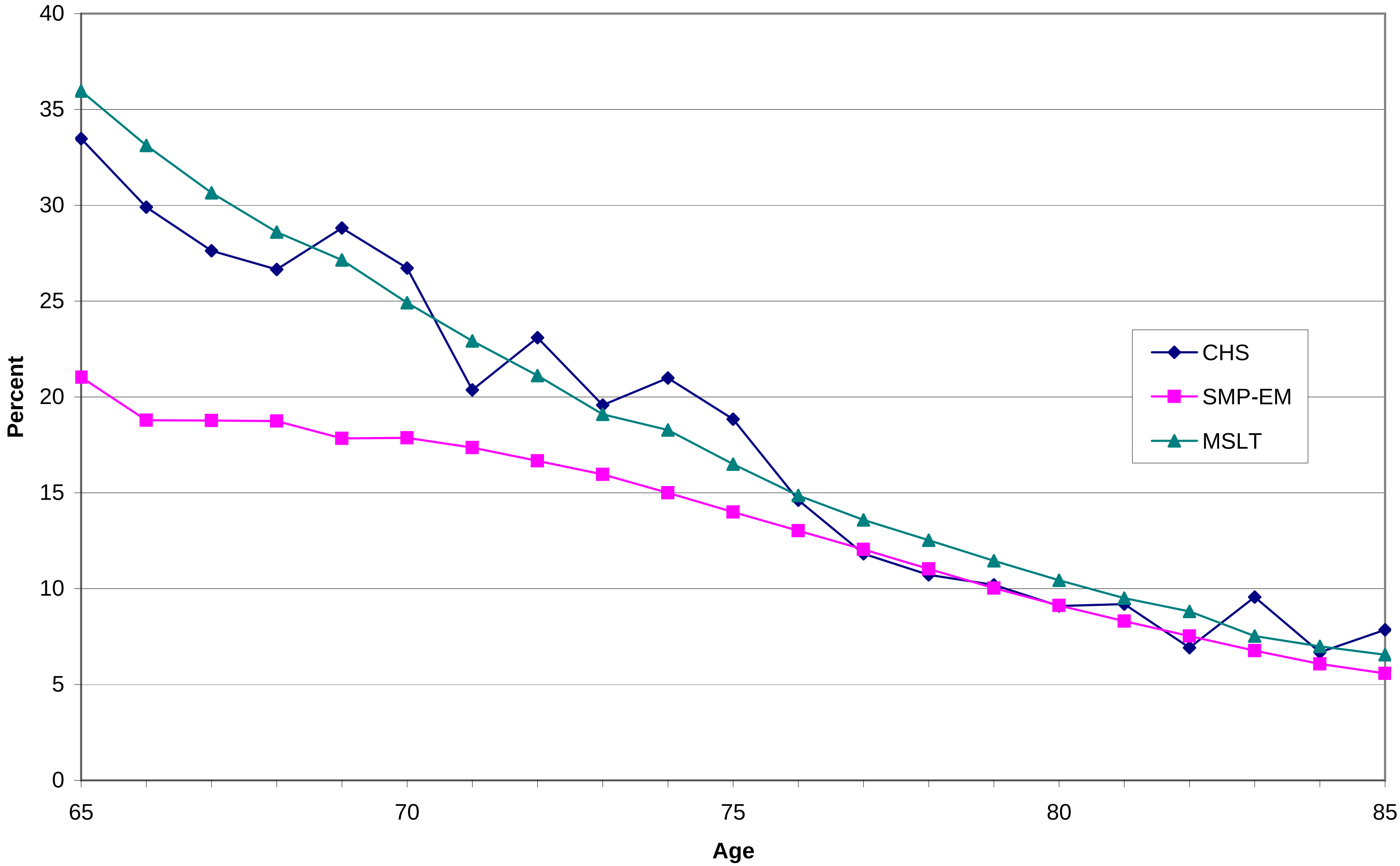


**Fig. 8B Age-Specific Empirical and Predicted Disability Recovery Rates From SMP-EM and MSLT Based on CHS Validation Sample**



## Fig. 8C Age-Specific Empirical and Predicted Mortality Rates For All Elderly From SMP-EM and MSLT Based on CHS Validation Sample

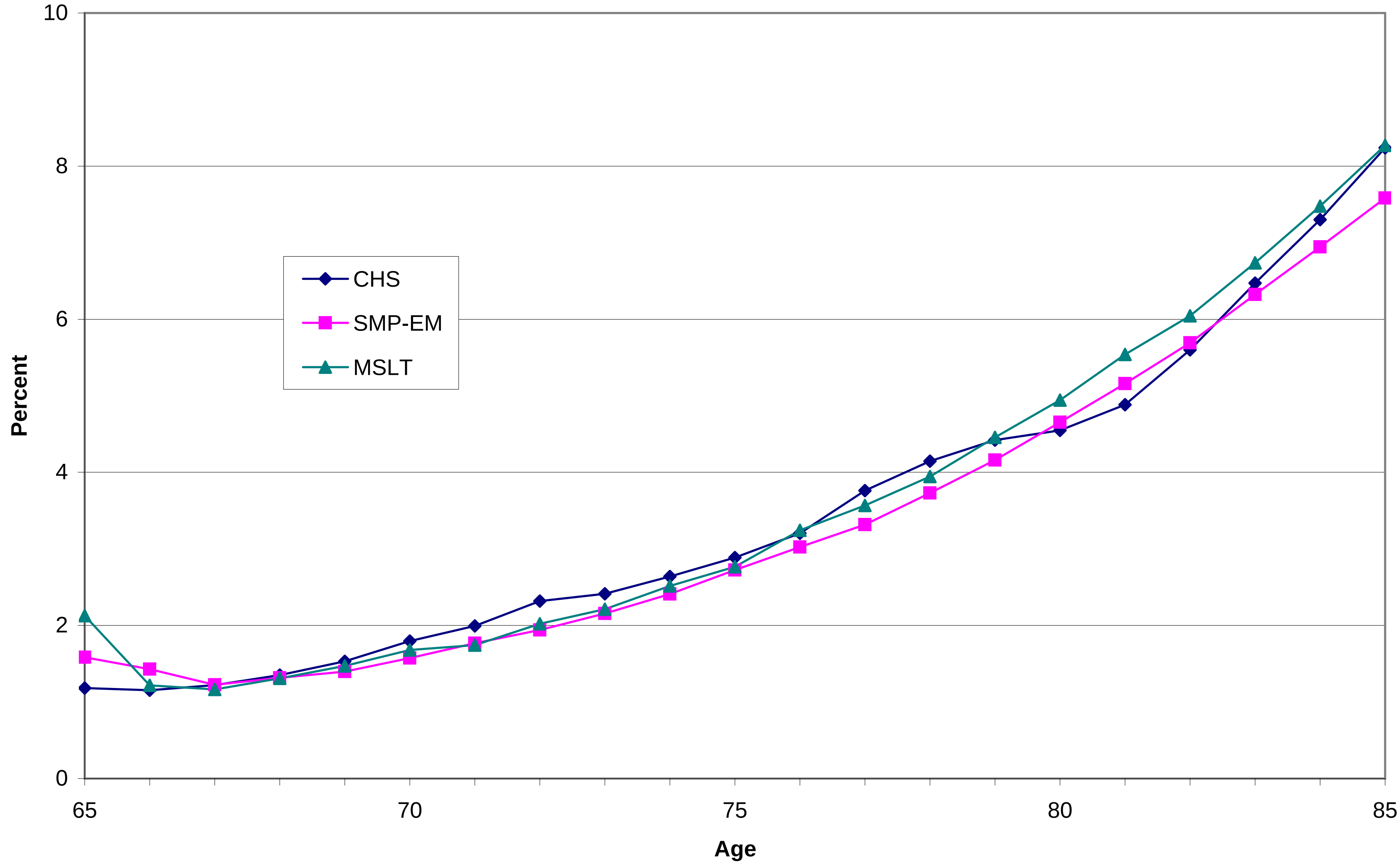


Table 2 LE Estimates at Age 65 Using SMP-EM and MSLT Based on Full CHS Sample

| | Status @ 65 | TLE | ALE | DLE |
|---|---|---|---|---|
| SMP-EM | All | 20.0 (0.14) | 10.4 (0.15) | 9.6 (0.13) |
| | Active | 20.9 (0.17) | 12.4 (0.17) | 8.4 (0.15) |
| | Disabled | 18.6 (0.20) | 7.0 (0.22) | 11.6 (0.19) |
| MSLT | All | 20.0 (0.14) | 10.7 (0.08) | 9.3 (0.12) |
| | Active | 20.2 (0.15) | 11.4 (0.09) | 8.9 (0.12) |
| | Disabled | 19.7 (0.16) | 9.4 (0.11) | 10.3 (0.14) |

*Note:* SMP-EM refers to fitting the SMP model using the stochastic EM algorithm, and MSLT refers to the multi-state life table model.