MODELING AND FORECASTING FIRST MARRIAGE: A LATENT FUNCTION APPROACH

Nan Li[1]
United Nations
Population Division
New York, NY 10017
li32@un.org

Zheng Wu
Department of Sociology
University of Victoria
Victoria, BC V8W 3P5 Canada
zhengwu@uvic.ca

April 11, 2008

[1] The views and opinions expressed in this paper are those of the authors and do not necessarily reflect those of the United Nations.

ABSTRACT

Hernes (1972) proposed a deductive model for diffusion processes and applied it to first marriage. Using a latent function, we develop a generalized diffusion model that includes the Hernes model as a special case. The Hernes model has a linear latent function, and hence is the simplest among the generalized Hernes models, for the latent function can be nonlinear to describe other diffusion processes. The linear latent function gives two advantages to the Hernes model. First, forecasts are based on linear extensions of historical trends, and the confidence intervals of forecasts are obtained analytically. Second, the starting and closing ages of the model (of first marriage) can be chosen less arbitrarily. The optimal starting and closing ages are determined by examining empirical data and, e.g., by identifying and removing outliers of a linear trend, which are difficult to accomplish in nonlinear models. We use data on first marriage from the U.S. and Canada to demonstrate these advantages, and forecast declines and delays in first marriage for Canadian cohorts.

Key words: Diffusion models, cohort forecasts, first marriage

MODELING AND FORECASTING FIRST MARRIAGE: A LATENT FUNCTION APPROACH

The question of declining marriage occupies a major place in demography, because this transnational trend signals a historic demographic transition. Although marriage rates have been declining in most advanced industrial countries for several decades, especially when compared with the exceptionally high marriage rates during the 1950s, individuals are not willing to abandon marriage (e.g., Edin, 2000; Waller, 2001). Even though nonmarital unions have indeed become increasingly salient, a growing number of individuals are delaying marriage until older ages. This socio-demographic phenomenon associates with a number of structural and individual-level factors, such as the socioeconomic shifts that have encroached upon men's capacities to earn a family wage and increased women's employment opportunities (Becker, 1981; Oppenheimer, 1994), and ideational change concerning the meaning of marriage (Lesthaeghe, 1983). The focus on declining marriage is important because this change reverberates through societies, affecting population replacement, support ratios, and economic production, among other things, and therefore represents a potential catalyst for other major socio-demographic shifts. Hence, understanding the levels of declining marriage, and being able to forecast future marriage rates offers a vital perspective on the demographic change looming upon the horizon.

Forecasting cohort marriage rates requires either parametric or non-parametric models. Non-parametric models, e.g., Li and Wu (2003), are based upon weak distributional assumptions and require data from a large number of successive cohorts. Although parametric models are based on strong distributional assumptions, they operate on data from single cohorts, which explains why they have yielded more applications than non-parametric models.

Using Swedish data, Coale and McNeil (1972) proposed an empirical parametric model for first marriage, which has been widely used, e.g., in survey data applications (Rodriguez and Trussell, 1980), identifying marriage patterns (Bloom and Bennett, 1990), and forecasting the level of first marriage (Goldstein and Kenney, 2001). An obvious problem of the Coale-McNeil model is that it is empirical, namely depending entirely on Swedish data (Cherlin, 1990).

Viewing the entry into first marriage as a diffusion process, during which unmarried individuals are pressed to marry by sensing how many people are already married and, in the meanwhile, their own declining "marriagebility", Hernes (1972) proposed a deductive model of marriage that does not depend on empirical data. The Hernes model was used for forecasting only recently, and was found performing better than the Coale-McNeil model when applied to U.S. data (Goldstein and Kenney, 2001).

In this paper, we use a latent function to develop a model of the diffusion process, which generalizes the Hernes nonlinear model and bases its distributional assumption on empirical observation. When the latent function changes linearly, the model reduces to the Hernes model. A linear latent function offers a number of advantages. In this paper, we focus on two such advantages. First, forecasts are based on linear extensions of historical (linear) trends and the confidence intervals of forecasts can be obtained analytically. Second, the starting and closing ages of the model can be chosen less arbitrarily. Optimal starting and closing ages are determined by empirical observation and, e.g., by identifying and removing observations that do not fit the linear trend, which can hardly be achieved in non-linear models. The paper first reviews the Hernes model. It then introduces a latent function to model marriageability and generalizes the Hernes model. Use of this model in stochastic forecasts is discussed next and examples using Hernes' data and that from Canadian census data are presented.

The Hernes Model

Our objective is to develop a formal diffusion model based upon Hernes' (1972) seminal research. Diffusion refers to the spread of an innovation within a social system. Diffusion is not about the increasing prevalence of an innovation *per se*, but embodies a causal process influencing how innovations spread (Strang and Soule, 1998).

The Hernes model conceptualizes first marriage formation as a diffusion process within a cohort, and posits that social pressure to get married and marriageability (or eligibility) affect the extent and the rate of diffusion (Burch, 1993). The social pressure to marry increases as being married becomes more prevalent within cohorts, for single people experience more social alienation and stigma when being in a marriage is the dominant status among their fellow cohort members (Hernes 1972). Marriageability depends an individual's age and marriage pool, with marriage eligibility decreasing over time as an individual's age increases and the number of available partners decreases. According to Nazio and Blossfeld (2003), social learning refers to the "knowledge awareness" gleaned from previous cohorts. Younger cohorts are more likely to adopt an innovation practiced within an older, adjacent cohort, and cross-cohort communication channels are a factor driving change within social systems. The experiences of older cohorts offer younger cohorts information about the potential costs and benefits of an innovation, and reduce the "risk" of adopting an innovation by increasing its social visibility and acceptability. Peer pressure influences the diffusion of an innovation *within* a cohort because potential adopters can learn from observing and interacting with innovators among their peer group. Lastly, eligibility is a function of age combined with the number partners available for forming a union.

For a cohort aged x, Hernes (1972) denotes marriageability by $f(x)$, and uses the proportion of ever married, $p(x)$ to describe the social pressure to marry. Assuming that the rate of change in the probability of getting married is the same for unmarried individuals at the same age, Hernes proposes that

$$\frac{dp(x)}{dx} = f(x)p(x)[1 - p(x)]. \tag{1}$$

The next step is selecting a parametric model for $f(x)$ and empirically estimating its parameters. Although it is reasonable to argue that $f(x)$ should decline with age, Hernes assumes an exponential function for $f(x)$

$$f(x) = ab^x, \tag{2}$$

where a is the average initial marriageability, and $b < 1$. Equation (2) expresses a constant decline (depreciation) in marriageability. Under this assumption, the Hernes model is expressed as

$$p(x) = \frac{1}{1 + \dfrac{1 - p(s)}{p(s)} \cdot \dfrac{\exp[a/\log(b)]}{\exp\{[a/\log(b)]\,b^{(x-s)}\}}}, \quad s \le x \le t, \tag{3}$$

where s and t are the starting and closing ages of the model, respectively. Using equation (3), parameters a and b can be estimated using the observed values of $p(x)$, and the model values of $p(x)$ are then expressed and compared with the observed values. For a cohort that has not completed first marriage experience, parameters a and b can also be estimated on the basis of the observed values of $p(x)$, and future levels of first marriage can then be forecast using equation (3).

Although equation (3) provides a close fit to the first marriage experience of American cohorts born before 1925 (Hernes, 1972) and after 1945 (Goldstein and Kenney, 2001), the distributional assumption about the function form of *f(x)* , i.e., equation (2), is rather arbitrary. This problem can be resolved because we can indirectly measure *f(x)* using equation (1). Our strategy is to use empirical observation as the basis of making the assumption about the parametric form of *f(x)*.

Modeling Marriageability with a Latent function

In fact, equation (1) is solvable without assuming a parametric form for *f(x)*. By solving equation (1), the general form of the Hernes model can be expressed as

$$p(x) = \frac{1}{1 + \frac{1-p(s)}{p(s)} \exp[-\int_s^x f(y)dy]} = \frac{1}{1 + \frac{1-p(s)}{p(s)} \exp\{-\int_s^x \exp[g(y)]dy\}}, \qquad (4)$$

which can utilize any form of *f(y)* or *g(y) = log[f(y)]*. When *f(y)* is exponential, *g(y)* is linear, (4) reduces to (3).

Defining *g(x)* as a latent function, we obtain the values of *g(x)*, using the discrete form, expressed in single years, $\frac{dp(x)}{dx} \approx \frac{p(x+1) - p(x-1)}{2}$ and equation (1), as

$$g(x) = \log\{\frac{p(x+1) - p(x-1)}{2p(x)[1-p(x)]}\}, \quad x = s+1, \ s+2, \ ..., \ t-1. \qquad (5)$$

Applying equation (5) to the data for the U.S. presented in Hernes' paper, we obtain the values of *g(x)* in Figures 1 and 2.

< Figures 1 and 2 About Here >

Among the parametric models, especially the parsimonious ones, linear functions seem to be the promising choice to fit the values of *g(x)* in Figures 1 and 2. We therefore conclude that the Hernes' assumption of linear *g(x)* is appropriate on the basis of empirical observation, not for general *g(x)* but for the *g(x)* shown in the figures. From this perspective, the Hernes model has a linear latent function *g(y)*, and hence is the simplest among the generalized models in (4) that may have latent functions of more complex forms.

Since empirical data are not perfect, even if the relationship were perfect linear, we would expect to see deviations from linearity in observed data. To take account of error in the data, let:

$$g(x) = \alpha + \beta \cdot x + \sigma \cdot e(x), \tag{6}$$

where *e(x)* is a random disturbance on the underlying linear pattern. Assuming that *e(x)*'s are independent each other and standard normal, parameters *α, β*, and *σ* can be estimated by the ordinary least squares (OLS) method,

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} t-s-2 & \sum_{x=s+1}^{t-1} x \\ \sum_{x=s+1}^{t-1} x & \sum_{x=s+1}^{t-1} x^2 \end{pmatrix} \begin{pmatrix} \sum_{x=s+1}^{t-1} g(x) \\ \sum_{x=s+1}^{t-1} xg(x) \end{pmatrix},$$
$$\sigma = \sqrt{\frac{\sum_{x=s=1}^{t-1} [g(x) - \alpha - \beta \cdot x]^2}{t-s-2}}. \tag{7}$$

Since the formal statistical test could only reject the null hypothesis that *g(x)* does not change over *x* and cannot specify whether or not the change is linear, we test the linearity assumption empirically using the R-square value

$$R^2 = 1 - \frac{\sum_{x=s+1}^{t-1} [g(x) - \alpha - \beta x]^2}{\sum_{x=s+1}^{t-1} [g(x) - \frac{\sum_{x=s+1}^{t-1} g(x)}{t-s-2}]^2} \tag{8}$$

that indicates the fraction of *g(x)*'s variance that is explained by the linear model.

Stochastic Forecasts

For cohorts at age t, say, age 30 years or younger, the processes of entering first marriage are incomplete, but their future marriage levels can be forecast on the basis of their past experience. Because point forecasts can hardly be accurate, we provide a stochastic forecast that includes not only point estimates but also their standard errors. We first forecast the latent function $g(x)$, in which forecast errors come from estimating parameters and future random disturbances. The standard errors in estimating α and β are given by,

$$
\begin{aligned}
se(\alpha) &= \sigma\sqrt{\frac{1}{t-s-2}}, \\
se(\beta) &= \sigma\sqrt{\frac{1}{\sum_{x=s=1}^{t-1} x^2}}.
\end{aligned}
\tag{9}
$$

Applying equation (6) to ages $y \geq t$, and noting that estimated errors are caused by $e(x)$ at ages $x < t$ and therefore are independent from the random disturbance $e(y)$, the $g(y)$ is forecast as

$$
\begin{aligned}
g(y) &= g(t-1) + [\beta + se(\beta)e(y)][y-(t-1)] + \sigma\sum_{z=t}^{y} e(z), \\
&= g(t-1) + [y-(t-1)]\beta + [y-(t-1)]se(\beta)e(y) + \sigma\sqrt{y-t}\cdot e(y), \\
&= g(t-1) + [y-(t-1)]\beta + \sqrt{\{[y-(t-1)]se(\beta)\}^2 + (y-t)\sigma}\cdot e(y).
\end{aligned}
\tag{10}
$$

Specifically, the mean ($g_m(y)$), the upper ($g_u(y)$) and lower ($g_l(y)$) bounds of the 95% confidence interval are given by

$$
\begin{aligned}
g_m(y) &= g(t-1) + [y-(t-1)]\beta, \\
g_u(y) &= g(t-1) + [y-(t-1)]\beta + 1.96\sqrt{\{[y-(t-1)]se(\beta)\}^2 + (y-t)\sigma}, \\
g_l(y) &= g(t-1) + [y-(t-1)]\beta - 1.96\sqrt{\{[y-(t-1)]se(\beta)\}^2 + (y-t)\sigma}.
\end{aligned}
\tag{11}
$$

Using equation (4), the median[2], the upper and lower bounds of the 95% confidence interval of the ever-married proportion are forecasted as

---

[2] The forecast of $p(x)$ has a central value that is equal to the median, but not the mean, because the forecasts are initially made on $g(x)$, which are then non-linearly transformed to $p(x)$.

$$p_m(y) = \frac{1}{1 + \frac{1-p(s)}{p(s)} \exp\{-\int_s^y \exp[g_m(z)]\, dz\}},$$
$$p_u(y) = \frac{1}{1 + \frac{1-p(s)}{p(s)} \exp\{-\int_s^y \exp[g_u(z)]\, dz\}}, \qquad (12)$$
$$p_l(y) = \frac{1}{1 + \frac{1-p(s)}{p(s)} \exp\{-\int_s^y \exp[g_l(z)]\, dz\}}.$$

Without the linear latent function, it would be impossible to express the uncertainty in forecasting analytically.

An Example and An Application

Using the U.S. census data presented in Hernes' paper, we show the improvement achieved by using the linear latent function approach. We present R-square values in Table 1, which are used to test the linearity assumption. It is clear that the linearity assumption works better for U.S. White population and younger cohorts. Errors in the data may be responsible for the large distortions from the linearity in the non-White and in older cohorts.

< Table 1 About Here >

Fitting the model to data from 2001 Canadian census (Statistics Canada, 2007), the linearity assumption holds better for women and for older cohorts than men and younger cohorts (see Table 2). Overall, for both Canadian and U.S. data, the differences in R-square values are minimal, suggesting that assumption of a linear form for *g(x)* is acceptable.

< Table 2 About Here >

In order to explore the consequences of major distortions from the linearity, let us look at the U.S. non-White female cohort born in 1885 (which has the lowest R-square value, see Table 1). Figure 3 illustrates the deviations between the observed values (circles) and the fitted values (lines). The top-left panel of Figure 3 shows that the main reason for the poor linear fit is that the observed $p(x)$ values at the starting and ending ages deviate markedly from the linear trend.

< Figure 3 About Here >

Choosing the starting age without arbitrariness is a long standing problem in modeling diffusion processes (see, Billari, 2001). This problem is especially important for the Hernes model because the model is sensitive to the starting value $p(s)$, as can be seen in equations (3) and (4). Moreover, selecting a closing age in marriage models is equally problematic, because above age 45 few people remain single, which makes the increase of $p(x)$ trivial. Furthermore, mortality can offset the slight increase of $p(x)$ expected from marriage alone, a limitation that most marriage models do not take into account. In the Hernes model, for example, $p(x)$ is an increasing function since its derivative is positive (see equation (1)). Using a linear latent function helps identify outliers, the observed data that deviate markedly from the linearity, and offers a possible strategy for adjustment, i.e., removing the outliers. In the case of the non-White female cohort born in 1885, removing the outliers raises the R-square value from 0.8612 to 0.9436 (not shown in the table), which significantly improves the fit of the model (see the lower two panels of Figure 3). Without using a linear latent function, outliers would be less traceable.

We now turn to forecast using the U.S. White female cohort born in 1925 as an example. The values of $p(x)$ presented by Hernes (1972) end at age 34, and we forecast them to age 45. Using equation (11), we first forecast $g(x)$. The forecast values are shown in the top-left panel of Figure 4, as a linear extrapolation of linear fit to the historical data. A linear extension of a linear trend could provide perhaps the most convincing forecast. Indeed, one reason that the Lee-Carter method (Lee and Carter, 1992) gains popularity in mortality forecasts is that the forecast is based on a *linear* index *k(t)*.

< Figure 4 About Here >

Using equation (12) and the extrapolation of *g(x)*, the values of *p(x)* can also be extrapolated as shown in the top-right panel of Figure 4. The proportion ever-married is forecast to rise slightly within narrow 95% confidence intervals. The forecast pattern of first marriage is fairly consistent with the observed trend as presented, for instance, by Goldstein and Kenney (2001).

Now suppose that the proportions ever-married were unknown at ages older than 24, and they need to be estimated. These estimates, based on 10 years of data, are presented in the lower-left panels of Figure 4. Confidence intervals are clearly wider than those using 20 years of data (see the top right panel), and widen further when only 7 years of data are used as the basis for forecasts (see the lower right panel).

This exercise demonstrates that this approach cannot produce reliable forecast when too few observations are available because there is too much uncertainty in forecasting the proportion ever-married at age 45, for example, from data that stop at age 21 (lower right panel in Figure 4). As equation (9) indicates, with fewer data the errors in estimates increase. Yet, the top left panel of Figure 4 shows that with data observed for ages 15 – 25, the linear *g(x)* can capture much of overall linear trend, but that observed from ages 15 – 21 cannot.

Turning to the case of Canadian data, Figures 5 and 6 show the forecasts of proportions ever-married for women and men, respectively. The forecasts are based on the 2001 census data from Canadian cohorts of 1960 – 74. Overall, the model fits the data well and shows the declining trend in the proportion ever-married (see Bloom and Bennett, 1990) or the delay of marriage documented by Goldstein and Kenney (2001). For example, for Canadian women, the *p(45)* and the mean age at first marriage are forecast to be 0.79 and 25.1 years for the oldest cohort (1960 – 64), but 0.71 and 26.4 years for the youngest cohort (1970 – 74).

< Figures 5 and 6 About Here >

Summary

In this paper, building on the seminal work of Hernes (1972), we used a linear latent function to develop a generalized diffusion model that includes the Hernes model as a special case. Among the generalized models, the original Hernes model is the simplest one, which has a linear latent function. The linear latent function gives two important advantages to the Hernes model. First, forecasts are based on linear extensions of historical trends, and the confidence intervals of forecasts are obtained analytically. Second, the starting and ending ages of the model could be chosen less arbitrarily by examining observed data and identifying outliers of a linear trend, which can hardly be accomplished in nonlinear models. Using data from the U.S. and Canada, we demonstrated that the latent function changes approximately linearly for the diffusion process of first marriage. Our forecasts of young Canadian cohorts suggest both decline and delay in first marriage. In our future work, we consider applying the Hernes model with a latent function to other datasets and in other countries. We also wish to apply the generalized Hernes model to cover demographic events other than first marriage, for which the latent function may be nonlinear.

References

Becker, G. (1981). *A Treatise on the Family*. Cambridge, MA: Harvard University Press.

Billari, Francesco. 2001. A log-logistic regression model for a transition rate with a starting threshold. *Population Studies, 55*: 15 – 24.

Bloom, David E. and Neil G. Bennett. 1990. Modeling American marriage patterns. *Journal of the American Statistical Association, 85*: 1009 – 1017.

Burch, Thomas K. 1993. Theory, computers, and the parameterization of demographic behaviour. *International Population Conference*, IUSSP, Montreal, Quebec, Vol. 3, 377 – 88.

Cherlin, Andrew J. 1990. The strange career of the Harvard-Yale study. *Public Opinion Quarterly, 54*: 117 – 124. .

Coale, Ansley J. and Donald R. McNeil. 1972. The distribution of age at first marriage in a female cohort. *Journal of the American Statistical Association, 67*: 743 – 749.

Edin, K. 2000. What do lo-income single mothers say about marriage? Social Problems 47: 112-133.

Goldstein, Joshua R. and Catherine T. Kenney. 2001. Marriage delayed or marriage foregone? New cohort forecasts of first marriage for U.S. women. *American Sociological Review, 66*: 506 – 519.

Hernes, Gugmund. 1972. The process of entry into first marriage. *American Sociological Review, 37*: 173 – 182.

Lee, R. D. and L. Carter. 1992. Modeling and forecasting the time series of U.S. mortality. *Journal of the American Statistical Association* 87: 659 – 71.

Lesthaeghe, R. (1983). A century of demographic change in Western Europe: An exploration of underlying dimensions. *Population and Development Review*, 9, 411 – 435.

Li, Nan and Zheng Wu. 2003. Forecasting cohort incomplete fertility: A method and an application. *Population Studies*, 57: 303 – 320.

Nazio, Tiziana & Hans-Peter Blossfeld. 2003. The diffusion of cohabitation among young women in West Germany, East Germany, and Italy. *European Journal of Population*, 19, 47 – 82.

Oppenheimer, Valerie K. 1994. Women's rising employment and the future of the family in industrial societies. *Population and Development Review*, 20, 293 – 342.

Rodriguez, German and James Trussell.1980. Maximum likelihood estimation of the parameters of Coale's model nuptiality schedule from survey data. *Technical Bulletin 7. World Fertility Survey, London, England.*

Statistics Canada. 2007. Canadian Census. Retrieved Online: http://www12.statcan.ca/english/census01/home/Index.cfm. Ottawa: Statistics Canada.

Strang, David & Sarah A. Soule. 1998. Diffusion in organizations and social movements: From hybrid corns to poison pills. *Annual Review of Sociology*, 24, 265 – 90.

U.S. Bureau of the Census. 1999. Table MS-2. Estimated median age at first marriage, by sex, 1890 to the present. Retrieved Online: http://www.census.gov/population/socdemo/ms-la/tabms-2.txt.

Waller, M. R. 2001. High hopes: Unwed parents' expectations about marriage. Children and Youth Services Review, 23: 457-484.

Table 1. R-square values for American cohorts by sex

Cohort	Male	Female
White American cohort born in 1885	0.9568	0.9916
Non-white American cohort born in 1885	0.9206	0.8612
White American cohort born in 1925	0.9886	0.9856
Non-white American cohort born in 1925	0.9751	0.9613

Data Source : Hernes, 1972 (Tables 2 - 3).

Table 2. R-square values for Canadian cohorts by sex

Cohort	Age	Male	Age	Female
Cohort born in 1960 - 64	18 - 41	0.9873	17 - 37	0.9885
Cohort born in 1965 - 69	18 - 36	0.9744	15 - 35	0.9798
Cohort born in 1970 - 74	19 - 30	0.8963	17 - 30	0.9590

Note: R-square values after removing outliers in the models.

Data Source: The 2001 Canadian census.

Figure 1. The latent functions of the U.S. cohorts born in 1925

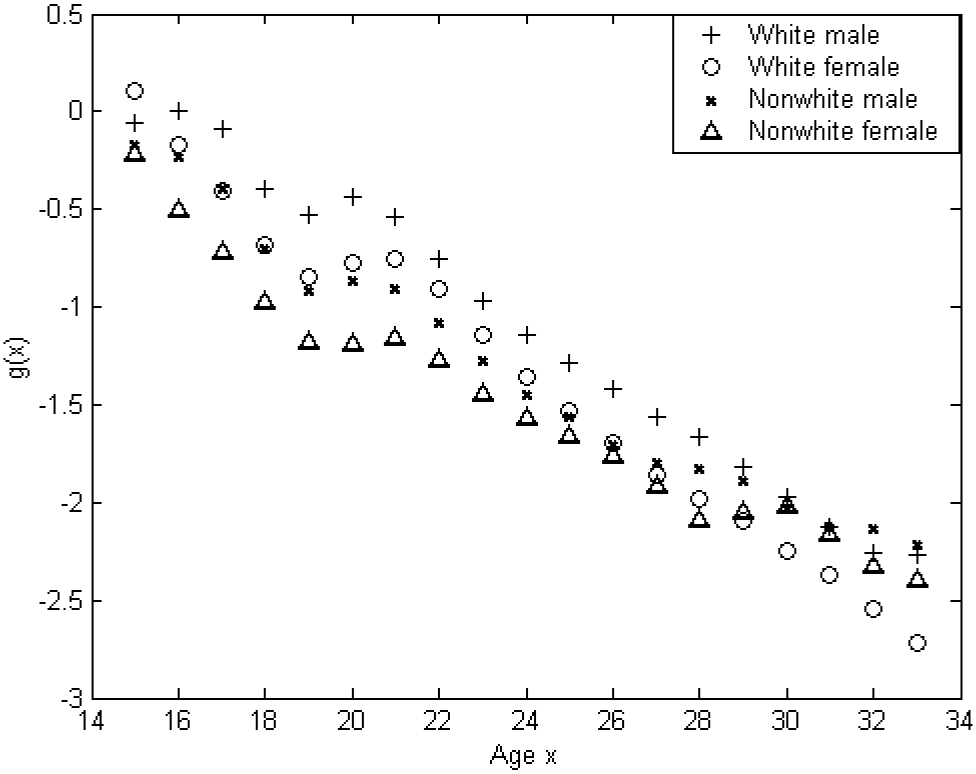

Figure 2. The latent functions of the U.S. cohorts born in 1885

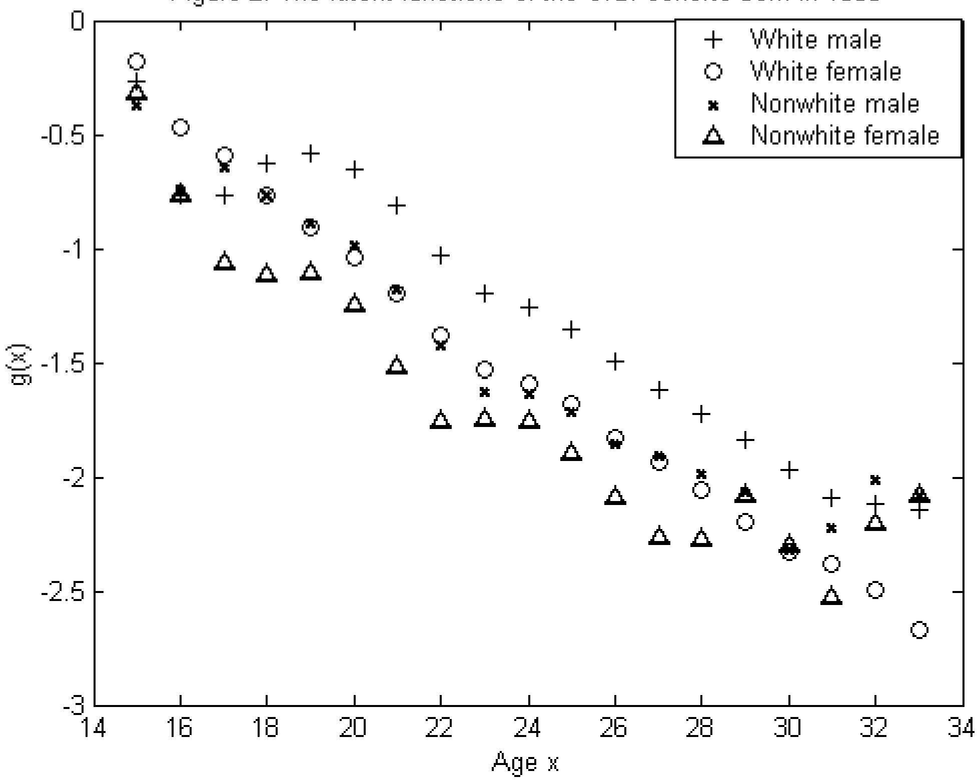

Figure 3. The latent function g(x) and proportion ever-married p(x)
for the U.S. Nonwhite female cohort born in 1885

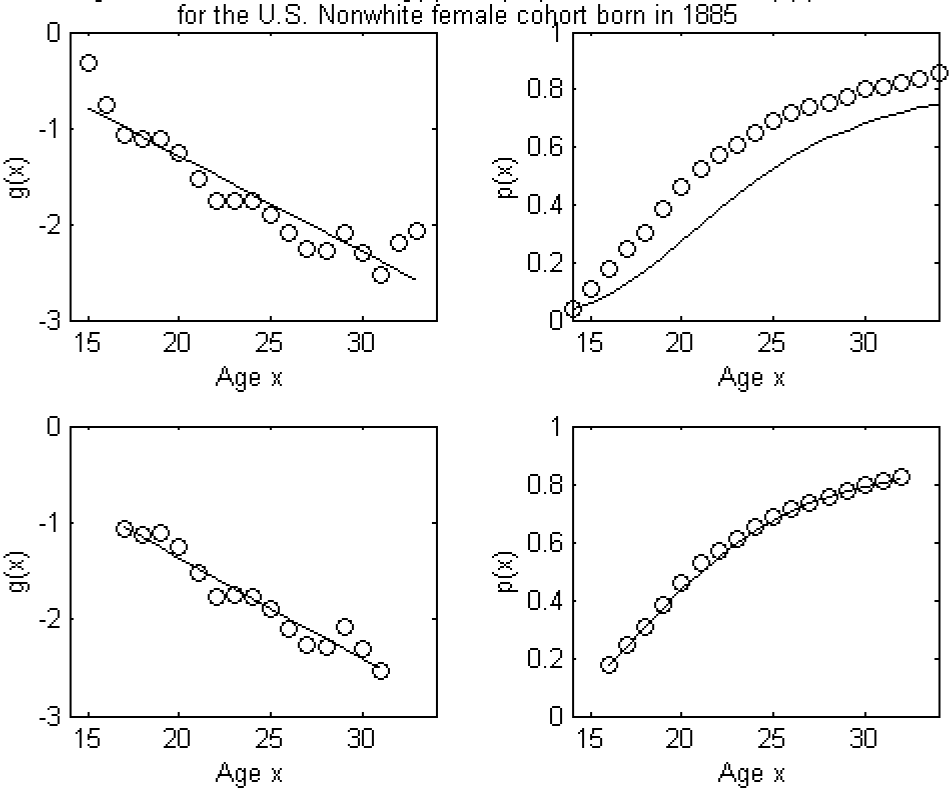

Figure 4. Forecasts of g(x) and p(x) for the U.S. White female cohort born in 1925, using data in different age ranges

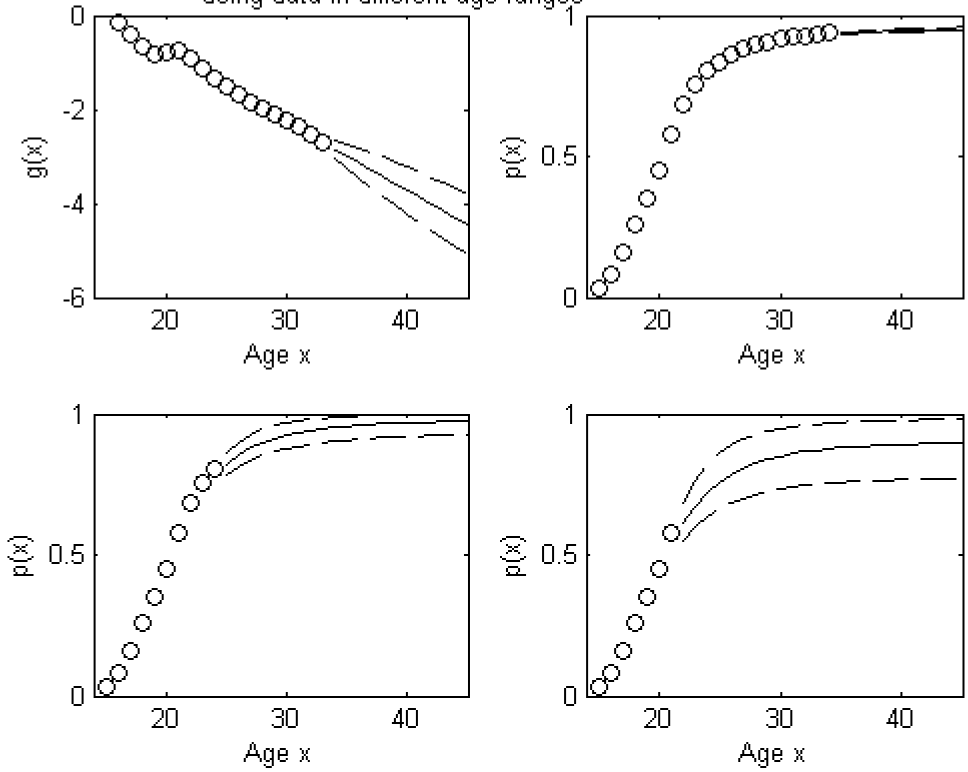

Figure 5. Observations and forecasts for Canadian female cohorts

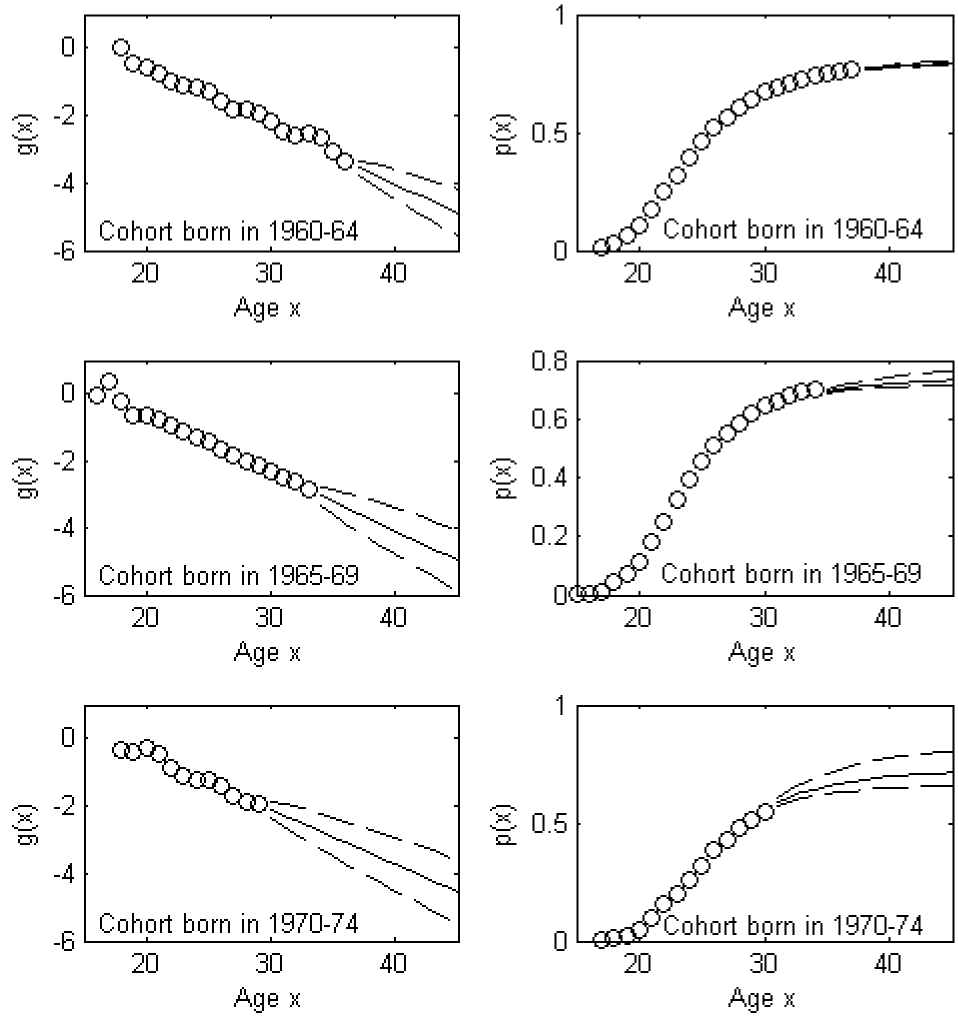

Figure 6. Observations and forecasts for Canadian male cohorts