ISSN 0819-2642
ISBN 0 7340 1720 0

THE UNIVERSITY OF MELBOURNE

THE UNIVERSITY OF MELBOURNE

DEPARTMENT OF ECONOMICS

RESEARCH PAPER NUMBER 779

MARCH 2001

SIMULATING COHORT DEMOGRAPHIC CHARACTERISTICS FOR AUSTRALIA

by

Justin van de Ven

Department of Economics
The University of Melbourne
Melbourne Victoria 3010
Australia.

Simulating Cohort Demographic Characteristics for Australia

J. van de Ven[1]
University of Melbourne
e-mail: justin.vandeven@sant.ox.ac.uk

March 1, 2001

[1]I should like to thank John Creedy and John Muellbauer for their extensive support and advice throughout the construction of the model. My thanks are also extended to the Henderson Foundation for its financial support, and to the HRD for making available the data sources used. Any omissions or errors are my own.

Abstract

This paper describes a dynamic microsimulation model of cohort demographics that has been developed to consider the working lifetime in Australia. The model provides a valuable resource for studying dynamics in Australia since, in addition to the usual advantages of microsimulation analysis, Australian panel data is scarce. The model possess a highly flexible structure. Specifically, the modular framework of the model permits the addition or omission of cohort characteristics as desired, and the use of transition probability functions facilitates transparent sensitivity analysis.

1 Introduction

This paper describes a dynamic microsimulation model of cohort demographics developed to consider the working-lifetime in Australia. Microsimulation models were first used for economic analysis by Orcutt (1957), and are now commonly employed to undertake policy analyses in many countries around the world. The feature that distinguishes microsimulation models from their macro based counterparts is that each micro-unit (also referred to as agent) from a given population is individually represented.[1] This feature makes microsimulation models useful for undertaking distributional analyses.

Microsimulation models are classified as either dynamic or static, depending upon how (and whether) the population is aged. Unlike static microsimulation models, dynamic microsimulation is designed specifically to consider the effects of counterfactual conditions on a population of agents through time. The ability to consider couterfactual experiments means that dynamic microsimulation models are capable of providing insights that survey data cannot. The limitations of survey data are compounded in Australia, where the few panel data sets are small, both in terms of duration and breadth, compared to those of many other countries. This makes the current model particularly useful for analyses that focus on intertemporal change in Australia, or where the desired period of measurement is in excess of a single year.

Most microsimulation models that are currently in use are static. Prominent examples of these include, STINMOD (Australia; refer to NATSEM, Australia), POLIMOD (UK; see Redmond *et al.*, 1998), TRIM2 (US; see Giannarelli, 1992), SPSP (Canada; refer to Statistics Canada), GMOD (Germany), SWITCH (Ireland), LOTTE (Norway), FASIT (Sweden), and CSO (Hungary).[2] However, advances in computing power, and the availability of increasingly detailed survey data have led to an increase in both the number, and sophistication, of dynamic microsimulation models. Some of the dynamic models in use include ASPEN (US; see Basu *et al.*, 1998), CORSIM (US; see Caldwell, 1996), HARDING (Australia; see Harding, 1993), and SESIM (Sweden), while many more are currently being developed.

The model described in this paper is comprised of three components that generate marital status, age of spouse, and number of dependants for a co-

[1]For macro-based models that study the impact of policy changes, see Dervis *et al.*(1982), Taylor (1990), and De Janvry *et al.* (1991). These are examples of Computable General Equilibrium models. Most micro-based models are constructed using a partial equilibrium framework. For examples of micro-based models that use a general equilibrium framework, see Meagher (1993), and Congneau (2000).

[2]For useful surveys, refer to Sutherland (1995), and Merz (1991).

hort of individuals aged 20 in 1970. Individual characteristics are generated at annual intervals for every cohort age between 20 and 55, thereby capturing the working-lifetime.[3] Unlike the architecture of most microsimulation models the current model has been developed to facilitate transparent sensitivity analysis. This objective has led to the adoption of a highly parsimonious structure.

Specifically, each of the three model components use transition probability functions to generate change wherever practicable, which replace the transition matrices that are typical of microsimulation models.[4] Transition probability functions enable sensitivity analysis to be undertaken by varying a few well defined parameters, as opposed to the relatively opaque element-by-element adjustment required for transition matrices. Similarly the effects of observed temporal trends are incorporated into the model using simple functions to facilitate model transparency.

Most microsimulation models generate a large number of characteristics for each individual to make possible a broad range of analyses. Given the relatively few characteristics generated by the current model, the ability to include additional characteristics as required is a fundamental feature of the modular structure adopted. Following the addition of cohort earnings, for example, the model is capable of analysing the redistributive effect of income taxes and a range of transfer schemes that comprise approximately 70 per cent of Australian social security expenditure, excluding pensions for the retired.[5]

Dynamic microsimulation models can be distinguished by the extent to which they incorporate agent specific behavioural responses. Given the aging populations observed in many countries, attention has been focused in recent years on the responsiveness of labour supply, savings, and fertility to alternative tax and transfer systems.[6] However, unlike the models used to examine these issues, the simulation procedure described in this paper makes no adjustment for behavioural responses. This property may be thought to call into question the extent to which analyses based upon the current model are of practical use. Specifically, given that many tax and transfer schemes

[3]Following the age of 55, retirement has a dominant effect upon annual measures of income inequality. See, for example, Figure 2 of Creedy and van de Ven (1999), which uses an earlier version of the present model.

[4]See Harding (1993).

[5]The model described here forms part of a larger microsimulation model developed to examine the redistributive effects of taxes in Australia. See van de Ven (2001) for a detailed description of the income components of the larger simulation model.

[6]see Macunovich (1998), and Hotz *et al.* (1997) for surveys of the fertility literature, and Auerback (1997) on savings.

are designed to affect agent behaviour, the predicted impact of such schemes derived from simulations that omit behavioural responses must be fundamentally inaccurate.

In response to this criticism, not all fiscal reforms cause the behaviour of agents to change. Furthermore, given that no model, however complex, can possibly capture the full extent of real-world diversity, any prediction derived from simulation methods must be treated with a degree of caution. With regard to microsimulation models that project labour supply responses, for example, it is possible to find reports of wage elasticities that range from small negative values to measures just over one. Although it is reasonable to suspect that a small positive wage elasticity is likely to reflect most applied cases, the accompanying uncertainty regarding the 'true' value cannot be ignored. In this sense, the first order effects generated by microsimulation models that omit behavioural responses provide a means of making unambiguous statements of the kind, "if behaviour remained unchanged..." Accompanying sensitivity analysis associated with any expected changes in behaviour can, of course, be subsequently undertaken. The extent to which the first order effects of policy change are of practical interest is evidenced by the continued use of, and focus on, static microsimulation models, which include no behavioural responses.

Section 2 provides an overview of the simulation procedure. A detailed description of how cohort marital status, spouse age, and number of dependants are simulated is provided in sections 3 to 5. Concluding comments are made in section 6.

2 The Simulation Procedure

Given the lack of suitable panel data for calibration, the model generates annual characteristics for a cohort of individuals and, where relevant, their nuclear families so that the demographic profiles at any given cohort age reflect data derived from Australian cross-sectional surveys. Heterogeneity between individuals is restricted to the following four characteristics:

1. sex
2. marital status
3. relative age of spouse (where relevant)
4. number of dependants (children aged 16 and under)

The sex of each individual is allocated such that 50.92 % of the simulated cohort is identified as male, consistent with census estimates of the 20 year old Australian population in 1970.[7] In any given year the remaining characteristics are generated for each individual using a 'linear' procedure. This procedure is depicted by the flow chart of Figure 1.

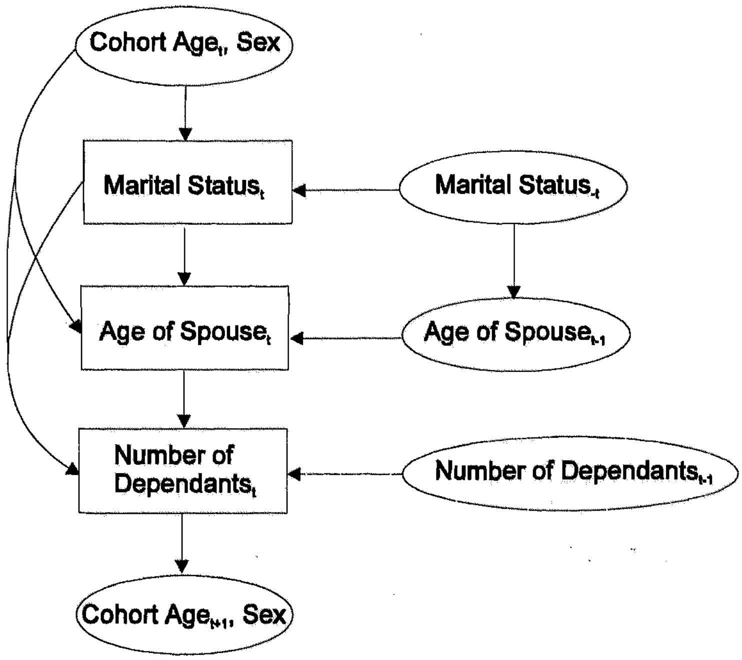


Figure 1: Stylised Demographic Simulation Procedure

In Figure 1, t subscripts refer to the reference period, and $-t$ subscripts refer to the entire simulated history up to period t. Characteristics included in elliptical frames are endogenous inputs used to generate period t characteristics, and arrows denote links. Hence, the age of a cohort member's spouse is generated with reference to the marital status, age, and sex of the cohort member in the generated period, and the age of the spouse (if they existed) in the previous period.

The following three sections describe the procedures used to simulate marital status, age of spouse, and the births of children. A detailed description of the methods used to calibrate each of the model components is also provided.

[7] See Commonwealth Bureau of Census and Statistics Publication Reference Number 4.15, Table 2.

3 Marital Status

Marital status is the first characteristic that is determined by the simulation procedure (after the trivial generation of age and sex) as illustrated by Figure 1. The focus here is on 'registered marital status' for which a marital certificate is held in accordance with the Australian Marriage Act 1961 (Commonwealth). This may be contrasted with 'social marital status', which, in addition to couples holding registered marital status, includes de facto relationships as defined under common law. During recent decades there has been a significant trend away from registered marital status. Associated with this shift is an increase in the importance of de facto relationships for any analysis based on the (nuclear) family. There are, however, important reasons why registered and social marital status should not be treated equivalently. Akerlof (1998 p. 290), for example, states that "on getting married men tend to change their behaviour." This leads him to describe the ceremony associated with registered marital status as a rite-of-passage such that, "after the wedding the life of the bride and of the groom will be changed. This can be modelled as a change in utility: with marriage the bride and the groom will have increased commitment toward each other, toward their offspring and, if religious, toward the Lord Himself." (p. 291).

In this framework, two extremes may be defined; one in which de facto couples act as though they are unattached in any formal respect, and another where de facto couples act as though they are 'registered married' in all respects. The focus on registered marital status taken here is equivalent to the first of these two extremes. Sensitivity analysis with respect to the other extreme may be undertaken by adjusting the marital rates used to include all 'socially married' couples.

Marital status is randomly allocated in the first simulated year of the cohort, based on sex, and the associated proportion of the Australian population married at 20 years of age in the year 1970.[8] Transition probabilities are used to generate changes in individual marital status for all subsequent years. Specifically, for an individual who was identified as married in the previous year, the probability that they cease to be married in the current year is equal to their risk of divorce, which is dependent on their age, sex, and the calendar year, plus the likelihood that their partner dies, which is dependent on their partner's age, sex, and the calendar year.[9] Alternatively,

[8] 1.429 % of males and 8.834 % of females aged 20 in 1970 are assumed to be married - estimates obtained from table 4.3 of ABS publication Cat. No. 3310.0, 1998.

[9] Risk of divorce is modelled independently from years of marriage to simplify the modelling procedure used. The model does, however, partially capture a relationship between years of marriage and probability of divorce insofar as the raw data used for

for an individual who was identified as unmarried in the previous year, the probability of marriage is dependent not only on their age, sex, and calendar year, but also on whether they had been identified as married in some previous year. It is assumed that no widows or divorcees are included in the first year of the simulated cohort.

Individual marital status is identified by generating a uniform random number between zero and one and comparing the value obtained with the associated transition probability, a method that is commonly referred to as a Monte Carlo procedure. If, for example, the value of the random variable is less than the associated probability of entering into marriage for a previously unmarried individual, then they are identified as married for the given year. To identify a marriage dissolution, on the other hand, the random variable must have a value less than or equal to the sum of the probabilities of divorce and partner death. Where a marriage dissolution is identified, divorce and death are distinguished by recognising divorce if the random variable is less than or equal to the probability of divorce, and death otherwise. If the random variable takes the value 0.11, for example, where the probability of divorce is equal to 0.10 and the probability of spouse death is equal to 0.04, then a marriage dissolution resulting from spouse death would be identified.

This allocating procedure is depicted by Figure 2, where; n defines the uniform random number, (t-1) denotes individual status in the previous period, (-t) denotes individual status for the entire simulated history, (t) denotes allocated state in current period, r_{DIV} is the probability of divorce, r_{DEATH} is the probability that spouse dies, r_M is the probability of remarriage, and r_{NM} is the probability of first marriage.

The probabilities referred to above are based on raw data derived from statistical tables published by the Australian Bureau of Statistics (ABS), which are listed in Appendix A. These tables provide Australian marital, divorce, and death rates by sex and calendar year for five year age groups between 15 and 59 years of age. Marital rates are divided into 'First marriage' and 'Remarriage' groups, which were calculated respectively by dividing the total number of individuals marrying, and remarrying, by the number of persons never married, and widowed or divorced. The tables listed in Appendix A provide marital rates for 11 years between 1966 and 1998. Similarly, the divorce rates and death rates were calculated as proportions of the married, and total, Australian populations by age group and sex, for, respectively, 15 and 16 years between 1966 and 1998.

Probability of Marriage. The simulation process requires the first marriage and remarriage rates of males and females. Each of these four rates are

calibration exhibit reduced probability of divorce with age.

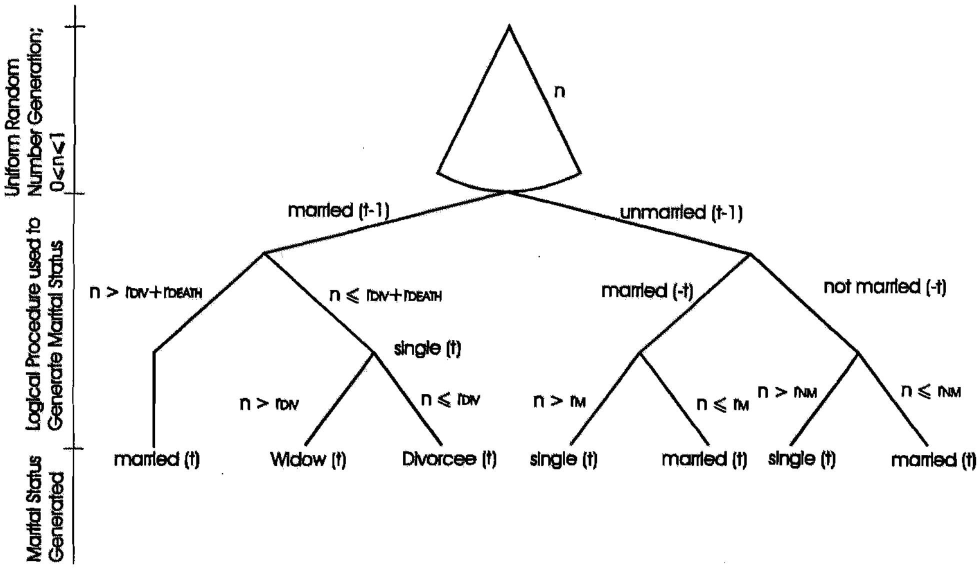


Figure 2: Marital Status Allocation Procedure

stored using the same underlying functional forms, and so specific reference is made to the rates applicable for males at first marriage for ease of exposition, where associated tables and figures for the remaining three rates are provided in Appendix B.

In the absence of any other distributional information, the rates provided by the statistical tables for the five year age groups were allocated to the mid-ages of the respective groups to enable regression estimates to be calculated for the transition probability functions. Cross-sectional regressions of the probability of marriage with respect to age were undertaken first, and the coefficients obtained were subsequently regressed against time to capture observed trends. As can be seen from Figure 3, the distribution of raw cross-sectional rates with respect to age suggest the use of a spline comprised of a linear segment (for the first two age groups) and an exponential (for all subsequent age groups). Estimates were consequently calculated for the following two models:

$$\begin{aligned} r_t &= \alpha + \beta t & t &= 17, 22 \\ &= \exp\left[\gamma + \delta t + \phi t^2\right] & t &> 22 \end{aligned} \tag{1}$$

where r_t is the relevant marriage rate for age t. A summary of the associated regression output is provided in Table 1.[10]

[10]In all cases, standard errors are provided in parentheses. Standard errors and R

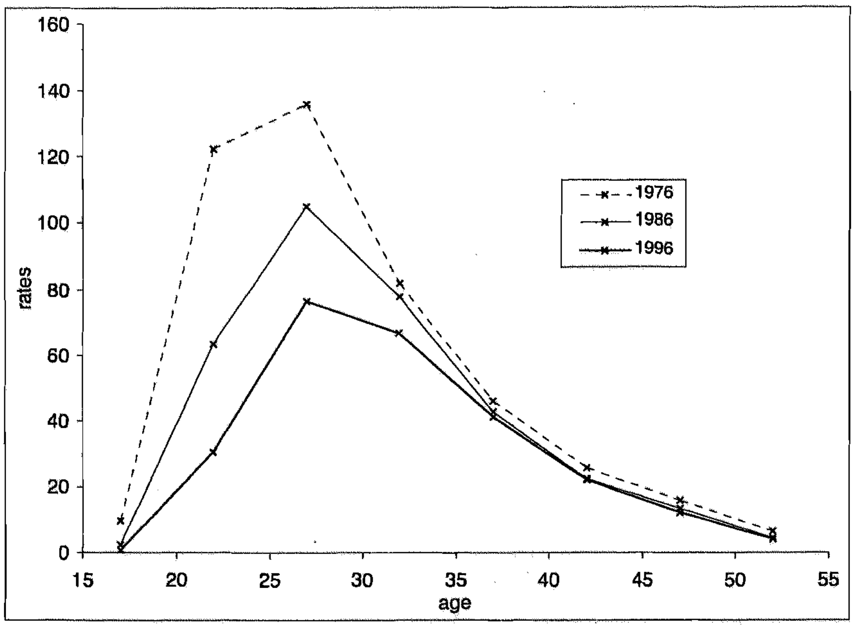


Figure 3: Male First Marriage Rates per 1000 Never Married Males

As can be seen from Table 1, the magnitude of both the linear segment coefficients, α and β, decrease systematically with time, albeit at a decreasing rate. With regard to the exponential segment, all of the regressions undertaken explain the observed variation of data well, where the lowest R squared is equal to 0.9935. Temporal trends are also evident for the second equation, where the values of γ and ϕ systematically fall with time, while the value of δ increases. These effects reflect the fact that the function $(\gamma + \delta t + \phi t^2)$, is approximately linear in 1966, and exhibits progressively more curvature with time, where the magnitude of the slope and intercept systematically decrease.

Linear functions with respect to time are adopted to capture the observed time trends for each of γ, δ, and ϕ, while exponential functions are used to capture the decreasing rate of the observed time trend for the linear model coefficients, α, and β. These simple functional forms are consistent with the unidirectional slope of the time trends discussed above and will aid analysis associated with variation of the time trend. The regression results for each of the associated equations are provided in Table 2.

squared statistics are not provided for the first equation because it possess' no degrees of freedom

Table 1: Cross-sectional Male First Marriage Rate Regression Output

Year	α	β	γ	δ	ϕ	R Square
	c	age	c	age	age^2	(quadratic)
1966	27.58	-453.96	7.8502	-0.0769	-7.419E-04	0.9943
			(1.2289)	(0.0643)	(8.10E-04)	
1971	32.72	-539.94	7.5191	-0.0623	-8.876E-04	0.9963
			(0.9647)	(0.0505)	(6.36E-04)	
1976	22.54	-373.28	6.5216	-0.0299	-1.127E-03	0.9961
			(0.8902)	(0.0466)	(5.87E-04)	
1981	16.68	-278.06	5.6481	0.0143	-1.719E-03	0.9990
			(0.4588)	(0.0240)	(3.03E-04)	
1983	14.72	-246.24	4.9898	0.0499	-2.200E-03	0.9969
			(0.8297)	(0.0434)	(5.47E-04)	
1986	12.20	-205.00	4.5172	0.0722	-2.474E-03	0.9944
			(1.1236)	(0.0588)	(7.41E-04)	
1988	10.72	-179.84	3.5804	0.1247	-3.180E-03	0.9954
			(1.0503)	(0.0550)	(6.93E-04)	
1993	7.86	-132.22	3.1854	0.1383	-3.300E-03	0.9965
			(0.8839)	(0.0462)	(5.83E-04)	
1996	5.88	-98.96	2.2684	0.1777	-3.721E-03	0.9962
			(0.8835)	(0.0462)	(5.83E-04)	
1997	5.52	-92.84	1.6631	0.2099	-4.131E-03	0.9958
			(0.9343)	(0.0489)	(6.16E-04)	
1998	5.36	-90.12	1.8186	0.2009	-3.987E-03	0.9935
			(1.1447)	(0.0599)	(7.55E-04)	

As can be seen from Table 2, all of the functions capture the observed variation well, with high R squares and low relative standard errors. Given the simplicity of the functional forms adopted, an obvious concern is the ability of the associated model to reflect the data upon which it is based. Figure 4 depicts the model data derived from the functional forms incorporating the regression coefficients of Table 2, and the raw data upon which the model is based for three years, which span the entire period for which raw statistical data were used. As can be seen from Figure 4, data derived from the regression coefficients closely reflect the associated raw data, adequately capturing both the observed trend and the cross-sectional distribution of first time marital rates for males.

Probability of Divorce. As for marital rates, divorce rates are summarised

Table 2: Intertemporal Male First Marriage Rate Regression Output

	c	year	R Square
Exponentional cross-sectional segment; Linear time trend			
	c	year	R Square
γ	406.94	-0.2027	0.9826
	(17.853)	(8.99E-03)	
δ	-18.80	0.00950655	0.9726
	(1.056)	(5.32E-04)	
ϕ	0.2236	-1.139E-04	0.9609
	(1.52E-02)	(7.65E-06)	
Linear cross-sectional segment; Exponential time trend			
	c	year	R Square
$\ln(-\alpha)$	137.062	-0.06635	0.9975
	(2.3393)	(1.18E-03)	
$\ln(\beta)$	135.656	-0.06706	0.9978
	(2.2206)	(1.12E-03)	

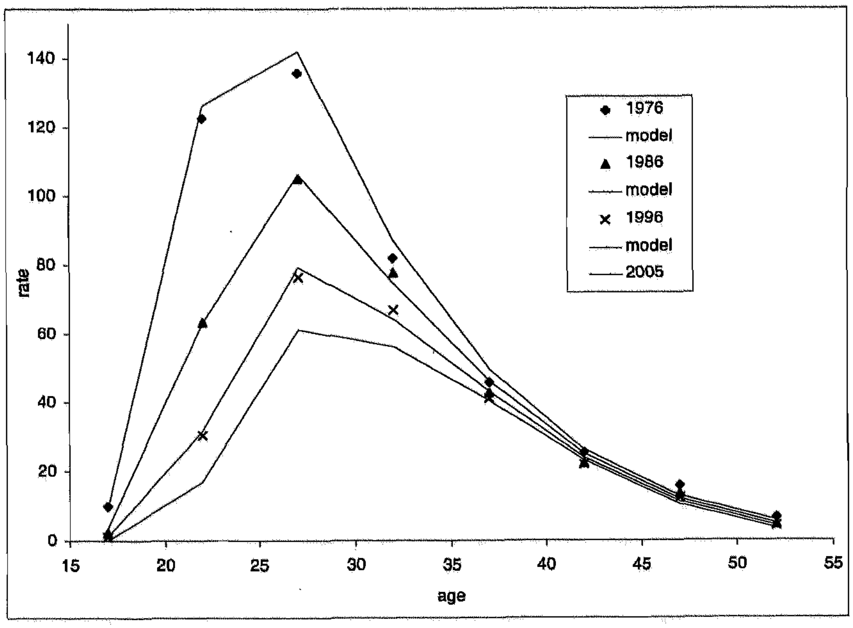


Figure 4: Male First Marriage Rates per 1000 Never Married Males - Model versus Raw Data

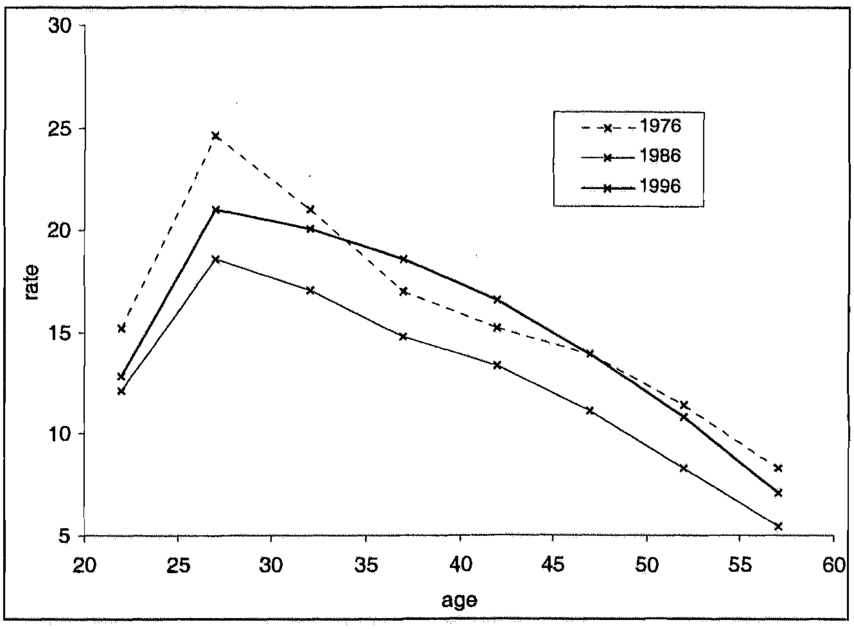


Figure 5: Male Divorce per 1000 Married Males by Age.

by first obtaining cross-sectional estimates and then regressing the coefficients obtained against time to capture temporal trend. Figure 5 depicts the raw cross-sectional divorce rates of males by age for the three years, 1976, 1986, and 1996. These divorce rates, which are reflective of the associated female rates and those for the other years used to calibrate the model, indicate a general downward trend with respect to age following the 20 to 24 year old age group, which deviates from the trend. Furthermore, the downward trend exhibits some curvature, where the 1976 and 1996 rates are, respectively, convex and concave to the age axis. These observations suggest the use of a quadratic to model all rates for ages following the 20 to 24 year old group, where 20 to 24 year olds are modelled separately. The results of regressing the divorce rate against a quadratic function of age are provided for males in Table 3.

As can be seen from Table 3, the constant and age square coefficients tend to decrease with time, while the age coefficient tends to increase. The age squared (age) coefficient passes from the positive to the negative (negative to the positive) domain between 1981 and 1982 (1984 and 1993), and is consequently not found to be significantly different from zero about these years. This variation reflects the observed trend where the associated cross-sectional relationships tend from convex to concave with respect to the age axis. The

Table 3: Cross-sectional Male Divorce Rate Regression Output

Year	c	age	age^2	R Square
1976	45.9905	-0.9486	5.238E-03	0.9843
	(6.3560)	(0.3147)	(3.73E-03)	
1981	39.8389	-0.7016	1.762E-03	0.9990
	(1.7663)	(0.0875)	(1.04E-03)	
1982	37.1339	-0.4934	-1.048E-03	0.9978
	(2.6788)	(0.1326)	(1.57E-03)	
1983	28.8740	-0.1419	-4.714E-03	0.9975
	(2.6429)	(0.1309)	(1.55E-03)	
1984	27.6727	-0.1124	-4.810E-03	0.9962
	(3.1694)	(0.1569)	(1.86E-03)	
1985	22.5022	0.0247	-5.762E-03	0.9943
	(3.1694)	(0.1569)	(1.86E-03)	
1986	22.2902	0.0010	-5.190E-03	0.9980
	(1.9167)	(0.0949)	(1.12E-03)	
1987	23.8592	-0.0483	-4.952E-03	0.9966
	(2.6825)	(0.1328)	(1.57E-03)	
1988	24.7258	-0.0527	-5.095E-03	0.9967
	(2.7481)	(0.1361)	(1.61E-03)	
1993	15.2380	0.5041	-1.157E-02	0.9936
	(3.7666)	(0.1865)	(2.21E-03)	
1994	14.0953	0.5401	-1.176E-02	0.9957
	(2.9562)	(0.1464)	(1.73E-03)	
1995	13.0183	0.6089	-1.257E-02	0.9897
	(4.6249)	(0.2290)	(2.71E-03)	
1996	18.4841	0.4076	-1.052E-02	0.9977
	(2.3100)	(0.1144)	(1.35E-03)	
1997	20.5770	0.2941	-9.190E-03	0.9912
	(4.5006)	(0.2228)	(2.64E-03)	
1998	18.0458	0.4273	-1.081E-02	0.9902
	(4.7985)	(0.2376)	(2.81E-03)	

Table 4: Intertemporal Male Divorce Rate Regression Output

Males aged 25 years and over				
coefficient	c	year	year^2	R Square
c	3.382E+05	−338.962	0.08493	0.8981
	(8.68E+04)	(87.327)	(0.02196)	
age	−11287.972	11.2918	-2.824E-03	0.9128
	(3851.191)	(3.8742)	(9.74E-04)	
age^2	124.9398	−0.12498	3.125E-05	0.9177
	(41.8674)	(0.04212)	(1.06E-05)	

Males aged between 20 and 24 years			
c	year	year^2	R Square
91219.63	−91.6954	0.02305	0.6885
(19059.4)	(19.1731)	(4.82E-03)	

fact that the quadratic function of age captures most of the observed variation of divorce rates, is evidenced by the high R squares obtained.

The decreasing rate of change observed for the time trends of the three coefficients imply the use of a quadratic by year. On the other hand, the divorce rates of the 20 to 24 year old age group start at relatively high values and fall to a minimum in 1990 before increasing in later years in a relationship with time that can also be modelled using a quadratic. Associated regression results are provided in Table 4, while Figure 6 enables a comparison between the raw cross-sectional divorce rates and the rates derived using the coefficients displayed in Table 4.

Probability of Spouse Death. Figure 7 depicts the raw cross-sectional death rates of males for a number of years between 1973 and 1998. The regular shape of the cross-sectional distributions depicted suggest the use of a cubic function with respect to age, which is supported by the high R squared values of the associated regression results as depicted in Table 5.

One of the most striking features displayed by the raw cross-sectional death rates in Figure 7, is the decrease in curvature that occurred between 1973 and 1998. For age groups below 35 years, the temporal trend observed is relatively small, where each of the six cross-sectional series exhibit approximately the same inflection point (around 30 years of age), and a slight reduction is observed for death rates following the local maximum at 22 years. The decreased curvature observed is driven by substantial reductions for the death rates of age groups older than 35 years, where the death rate of the 55 to 59 year old age group in 1998 is less than half of the associated rate in

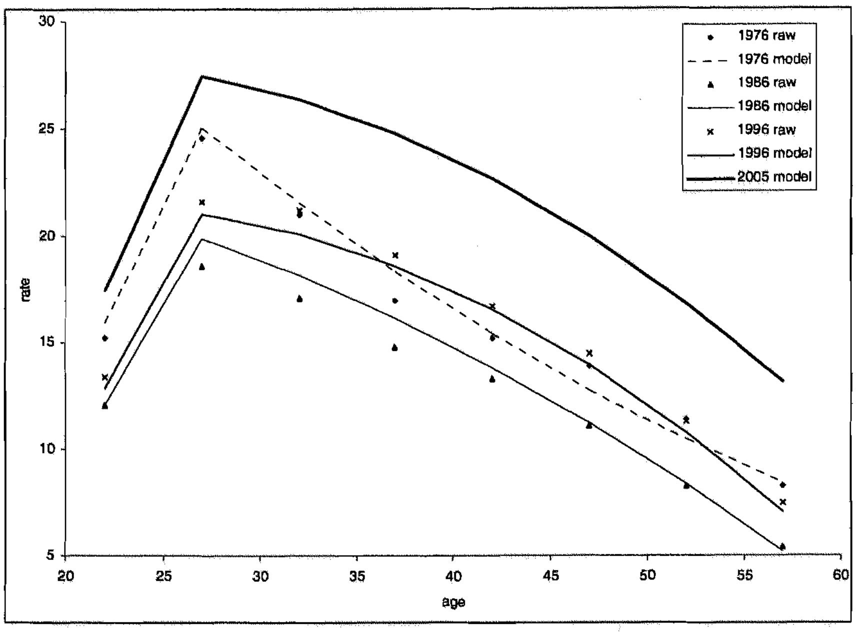


Figure 6: Male Divorce Rates per 1000 Married Males - Model versus Raw Data

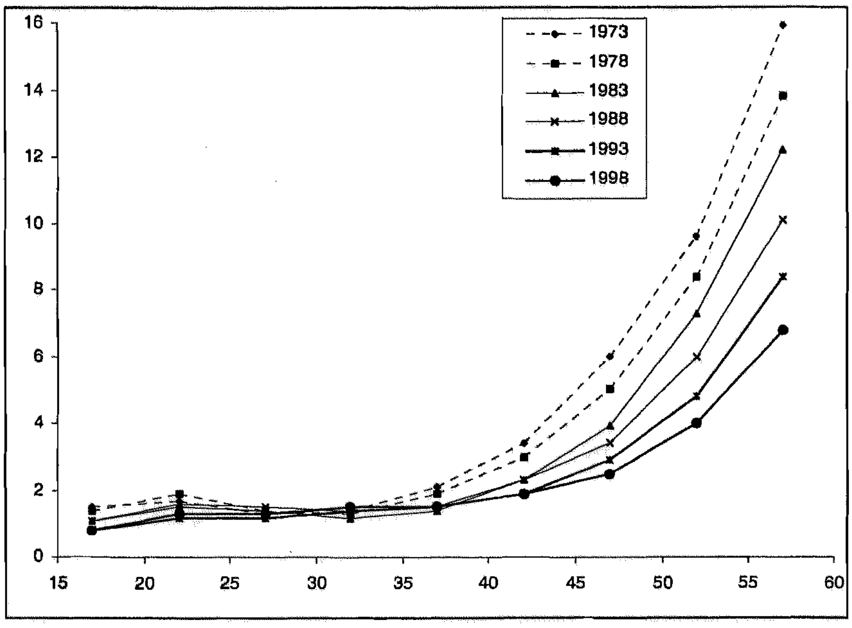


Figure 7: Male Cross-sectional Death Rates per 1000 Population, by Age.

Table 5: Cross-sectional Male Death Rate Regression Output

Year	c	age	age^2	age^3	R Square
1968	-7.6279	1.0947	-0.04235	5.394E-04	0.9992
	(1.760)	(0.1631)	(4.67E-03)	(4.19E-05)	
1973	-4.6463	0.8029	-0.03359	4.525E-04	0.9991
	(1.687)	(0.1562)	(4.47E-03)	(4.01E-05)	
1978	-5.5908	0.8798	-0.03481	4.444E-04	0.9989
	(1.585)	(0.1468)	(4.20E-03)	(3.77E-05)	
1982	-6.6458	0.9615	-0.03673	4.525E-04	0.9996
	(0.905)	(0.0838)	(2.40E-03)	(2.15E-05)	
1983	-8.3313	1.1107	-0.04095	4.875E-04	0.9997
	(0.724)	(0.0670)	(1.92E-03)	(1.72E-05)	
1984	-8.5729	1.1103	-0.04026	4.741E-04	0.9987
	(1.504)	(0.1393)	(3.99E-03)	(3.57E-05)	
1985	-8.6379	1.1323	-0.04106	4.801E-04	0.9992
	(1.090)	(0.1009)	(2.89E-03)	(2.59E-05)	
1986	-7.6931	1.0201	-0.03702	4.350E-04	0.9988
	(1.281)	(0.1187)	(3.40E-03)	(3.05E-05)	
1987	9.4831	1.1914	-0.04211	4.815E-04	0.9987
	(1.347)	(0.1247)	(3.57E-03)	(3.20E-05)	
1988	-7.9186	1.0296	-0.03648	4.202E-04	0.9991
	(1.008)	(0.0934)	(2.67E-03)	(2.40E-05)	
1993	-8.0900	0.9731	-0.03299	3.677E-04	0.9974
	(1.408)	(0.1304)	(3.73E-03)	(3.35E-05)	
1994	-8.2636	0.9739	-0.03231	3.542E-04	0.9949
	(1.858)	(0.1721)	(4.93E-03)	(4.42E-05)	
1995	-7.3278	0.8857	-0.02970	3.286E-04	0.9971
	(1.340)	(0.1241)	(3.55E-03)	(3.18E-05)	
1996	-6.9809	0.8472	-0.02852	3.178E-04	0.9971
	(1.341)	(0.1242)	(3.56E-03)	(3.19E-05)	
1997	-5.6848	0.7241	-0.02471	2.788E-04	0.9985
	(0.860)	(0.0796)	(2.28E-03)	(2.04E-05)	
1998	-7.3321	0.8727	-0.02861	3.091E-04	0.9967
	(1.226)	(0.1135)	(3.25E-03)	(2.91E-05)	

Table 6: Time series Male Death Rate Regression Output

Coefficient	c	year	R Square
constant	49.6527	-2.8731E-02	0.03887
	(75.8611)	(3.818E-02)	
age	10.8967	-4.9938E-03	0.10793
	(7.6231)	(3.837E-03)	
age^2	-0.87478	4.2263E-04	0.46066
	(0.2428)	(1.222E-04)	
age^3	1.5636E-02	-7.6621E-06	0.74930
	(2.35E-03)	(1.185E-06)	

1973. The regression results displayed in Table 5 reflect these observations, where the strength of the temporal trend increases with the power on age.

Each of the coefficients depicted in Table 5 were regressed against linear functions with respect to year of observation. The associated regression results are presented in Table 6.

The results displayed in Table 6 support the observations made above, where the linear functions with respect to time explain an increasing proportion of the coefficient variation as the power on age increases. Despite the poor regression results associated with the constant term as presented in Table 6, a linear function with respect to time was found to reflect the raw data more accurately than imposing the sample mean, and is consequently adopted. Figure 8 enables a comparison of the rates derived from the coefficients displayed in Table 6, with the raw data upon which the model is based.

4 Age of Spouse

Where an individual is identified as newly married, the age of their spouse is generated by taking a random selection from an underlying distribution, which is conditional on the individual's age, sex, and calendar year.[11] The simulation process consequently requires the distribution of bride age by groom age and calendar year, and the distribution of groom age by bride age and calendar year.[12] These conditional distributions were obtained using

[11]This is subject to an imposed minimum of 17 years.

[12]The distribution of bride age by groom age and calendar year is required to generate the spouses of males in the simulation cohort, and vice versa for females. These distributions are evidently related. They are, however, modelled separately to simplify the simulation procedure. This artificial separation is valid given that the simulated cohort is

data derived from statistical tables listed in Appendix A, which detail the number of marriages by age of bride and groom that took place during the odd numbered years between 1973 and 1995 inclusive.

Figure 9 displays the distribution of bride ages observed for 20, 30, 40, and 50 year old grooms in 1995. This figure indicates that the distribution of bride ages is skewed, where the mean, and variance both increase, and the direction of skew tends from left to right as the groom age increases. The systematic variation observed for the bride age distribution with respect to groom age implies, and is implied by, a similar systematic variation of the groom age distribution with respect to bride age. The distribution of bride age for a given groom age and marriage year is consequently modelled by a parallel procedure to the one used to model the distribution of groom age for a given bride age and marriage year. To avoid unnecessary repetition, the following exposition makes specific reference to the calibration undertaken for the bride age distribution, where associated tables and figures can be found in Appendix C for the groom age distribution.

A log-normal was selected to characterise the required distributions because it is flexible, exhibits skew, and is fully defined by two meaningful parameters; the mean and variance of logs. The tables listed in Appendix A partially aggregate data into age groups. Specifically, for 1995 and 1993, marriages are aggregated for grooms or brides under the age of 18, over the age of 59, and in the five year age group between 55 and 59. In addition, marriages are aggregated for brides in the five year age group between 50 and 54. Data for the other years examined are subject to a higher degree of aggregation where, for all years between 1991 and 1973, five year age groups are imposed for marriages of brides or grooms aged between 30 and 59. Where aggregated five year age group data were used, the respective number of marriages was divided by five and associated with the mid-age of the group.[13] Marriages recorded in the lowest age group were divided by three as an approximation and allocated to the upper age of the group, whereas the upper age group was omitted from the analysis. The distributions subsequently obtained from the statistical tables are consequently incomplete, which prevents direct calculation of the required mean and variance of logs from the raw data. Hence, as a first step in the calibration procedure, the means and variances were estimated by eye, where evaluation was based on the ability of the associated

assumed to draw spouses from an infinitely large underlying population.

[13]If 520 marriages were identified between brides aged 27 and grooms aged between 30 and 34, for example, 104 marriages would be identified between brides aged 27 and grooms aged 32. Alternatively, if 1000 marriages were identified between brides aged between 30 and 34 and grooms aged between 35 and 39, 250 marriages would be identified between brides aged 32 and grooms aged 37.

distributions to reflect the raw data. It was found that, between the male ages of 20 and 42, bride age distributions are positively skewed, while negative skew was observed for male age groups between 43 and 55. Given that the log-normal distribution is only able to describe positively skewed data, it was necessary to re-format bride ages for the groom age groups above 42. This involved a simple transformation where the age of a given bride in the group was subtracted from the maximum bride age plus one.[14] The 'rough' estimates obtained from this manual estimation procedure were subsequently adjusted to smooth associated temporal and cross-sectional trends.

The adjusted estimates of means and variances derived for each of the male age groups are stored by the simulation model in summary equations. Specifically, linear and cubic equations are used respectively to characterise the cross-sectional means and variances obtained for the positively skewed age groups. Figures 10 and 11 depict these equations for three years, which are indicative of the observations obtained for the entire sample period. Associated regression results are displayed respectively in Tables 7 and 8.

Both Figures 10 and 11, and Tables 7 and 8 indicate that the equations reflect the raw data well, with high R squares, relatively low standard errors, and a close reflection of observed variation. The figures indicate the presence of strong temporal trends, which are also born out by the regression results. The linear equations associated with the mean of log bride ages, tend to shift up and decrease slightly in slope with time. This implies an increase in the average age of brides, where the effect is reduced at higher groom ages, a result that is consistent with the increase observed in section 3 for mean marriage age. Similarly, the cubic equations associated with the variance of log bride ages tend to exhibit more curvature resulting in an overall decrease in variance at higher groom ages.

Due to the level of aggregation imposed by the tables that provide data for the years prior to 1993, only two data points for each year could be obtained for the mean and variance of the distributions for groom ages between 43 and 55. The implied estimation problems associated with the cross-sectional equations are, however, mitigated by two factors. First, as indicated by section 3, relatively few marriages take place after the age of 42. Second, data are available for single year age groups for 1993 and 1995. The 1995 data, which reflect the data for 1993, are presented in Figures 12 and 13 for, respectively, mean and variance of log bride ages.

Figures 12 and 13 indicate that the associated relationships of mean and variance with age may be adequately summarised by linear equations. It is

[14]Maximum bride age used for the analysis was 57, hence bride ages were subtracted from 58 for the negatively skewed data.

Table 7: Cross-section Regression - Mean log Bride Age

Year	const.	age(groom)	R square
1995	2.5607	0.02433	0.9877
	(2.30E-02)	(8.18E-04)	
1993	2.5500	0.02439	0.9904
	(2.03E-02)	(7.25E-04)	
1991	2.5392	0.02445	0.9927
	(1.78E-02)	(6.34E-04)	
1989	2.5284	0.02451	0.9945
	(1.54E-02)	(5.48E-04)	
1987	2.5176	0.02457	0.9960
	(1.31E-02)	(4.67E-04)	
1985	2.5068	0.02463	0.9972
	(1.11E-02)	(3.96E-04)	
1983	2.4960	0.02470	0.9979
	(9.57E-03)	(3.41E-04)	
1981	2.4852	0.02476	0.9983
	(8.70E-03)	(3.10E-04)	
1979	2.4745	0.02482	0.9983
	(8.72E-03)	(3.11E-04)	
1977	2.4637	0.02488	0.9979
	(9.63E-03)	(3.43E-04)	
1975	2.4529	0.02494	0.9972
	(1.12E-02)	(3.99E-04)	
1973	2.4421	0.02501	0.9961
	(1.32E-02)	(4.71E-04)	

Table 8: Cross-section Regression - Variance log Bride Age

Year	const.	age	age^2	age^3	R square
1995	2.083E-01	-2.314E-02	8.324E-04	-9.000E-06	0.9994
	(9.85E-03)	(1.02E-03)	(3.44E-05)	(3.76E-07)	
1993	1.908E-01	-2.142E-02	7.783E-04	-8.423E-06	0.9995
	(9.86E-03)	(1.02E-03)	(3.45E-05)	(3.76E-07)	
1991	1.732E-01	-1.970E-02	7.241E-04	-7.845E-06	0.9993
	(1.18E-02)	(1.23E-03)	(4.13E-05)	(4.50E-07)	
1989	1.557E-01	-1.798E-02	6.700E-04	-7.267E-06	0.9990
	(1.49E-02)	(1.55E-03)	(5.22E-05)	(5.70E-07)	
1987	1.382E-01	-1.626E-02	6.159E-04	-6.689E-06	0.9985
	(1.87E-02)	(1.94E-03)	(6.53E-05)	(7.12E-07)	
1985	1.206E-01	-1.454E-02	5.618E-04	-6.111E-06	0.9980
	(2.27E-02)	(2.36E-03)	(7.94E-05)	(8.66E-07)	
1983	1.031E-01	-1.282E-02	5.077E-04	-5.533E-06	0.9974
	(2.69E-02)	(2.80E-03)	(9.42E-05)	(1.03E-06)	
1981	8.553E-02	-1.110E-02	4.536E-04	-4.955E-06	0.9967
	(3.12E-02)	(3.24E-03)	(1.09E-04)	(1.19E-06)	
1979	6.798E-02	-9.379E-03	3.994E-04	-4.378E-06	0.9960
	(3.56E-02)	(3.70E-03)	(1.25E-04)	(1.36E-06)	
1977	5.044E-02	-7.659E-03	3.453E-04	-3.800E-06	0.9953
	(4.00E-02)	(4.16E-03)	(1.40E-04)	(1.53E-06)	
1975	3.290E-02	-5.939E-03	2.912E-04	-3.222E-06	0.9946
	(4.45E-02)	(4.62E-03)	(1.56E-04)	(1.70E-06)	
1973	1.536E-02	-4.219E-03	2.371E-04	-2.644E-06	0.9939
	(4.89E-02)	(5.08E-03)	(1.71E-04)	(1.87E-06)	

consequently reasonable to assume, on the basis of the available evidence, that linear equations can be used to approximate the variation observed for the years prior to 1993. This is fortunate given that linear equations can be (exactly) estimated by the available data. Table 9 provides estimates of linear equations derived from cross-sectional regressions for the negatively skewed distributions of bride ages.

The coefficients displayed in Table 9 exhibit temporal trends that are consistent with those of the positively skewed age of bride distributions. Specifically, in general, the decreased constant and increased slope of the cross-sectional equations characterising mean log (adjusted) bride age imply that the associated mean is reduced at a decreased rate with groom age as the year is increased. Given the adjustment made to bride ages for the negatively

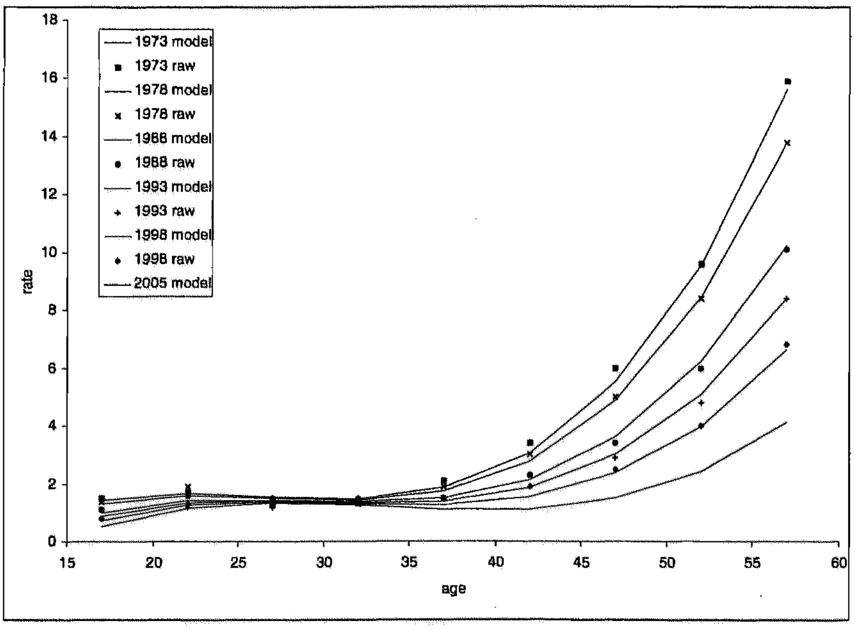


Figure 8: Male Death Rates per 1000 Male Population - Model versus Raw Data

Table 9: Cross-section Regression for Negatively Skewed Distributions- Mean and Variance log Bride Age

	Mean log Bride Age		Variance log Bride Age	
Year	const.	age	const.	age
1995	5.7623	-0.06530	-0.18053	8.690E-03
1993	5.6339	-0.06249	-0.18205	8.782E-03
1991	5.9006	-0.06732	-0.18358	8.874E-03
1989	6.0296	-0.06980	-0.18510	8.966E-03
1987	6.2178	-0.07342	-0.18663	9.058E-03
1985	6.3520	-0.07600	-0.18815	9.150E-03
1983	6.4997	-0.07884	-0.18967	9.242E-03
1981	6.6141	-0.08104	-0.19120	9.334E-03
1979	6.7503	-0.08366	-0.19272	9.426E-03
1977	6.8283	-0.08516	-0.19425	9.518E-03
1975	6.9843	-0.08816	-0.19577	9.610E-03
1973	7.1008	-0.09040	-0.19729	9.702E-03

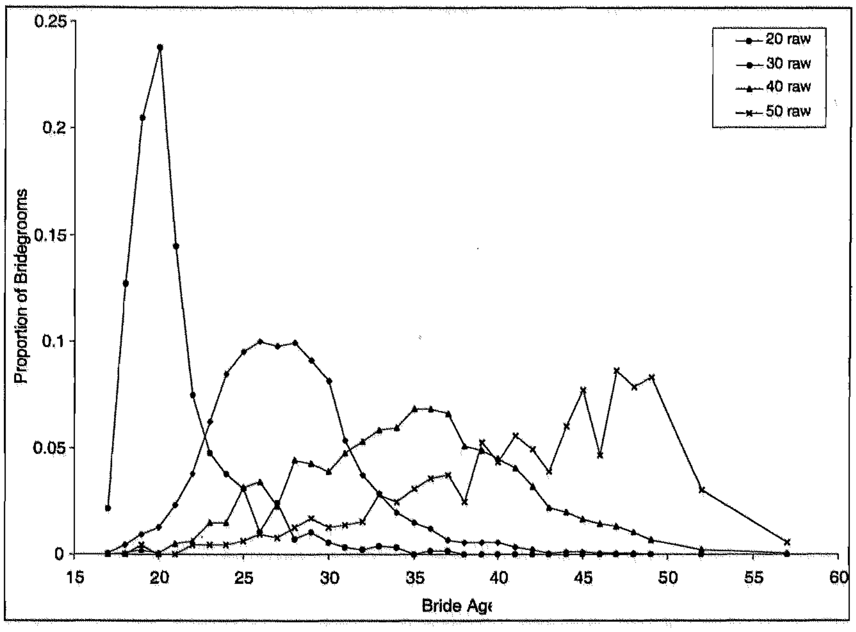


Figure 9: Observed Distribution of Age of Bride for Bridegrooms Aged 20, 30, 40, and 50 Years in 1995

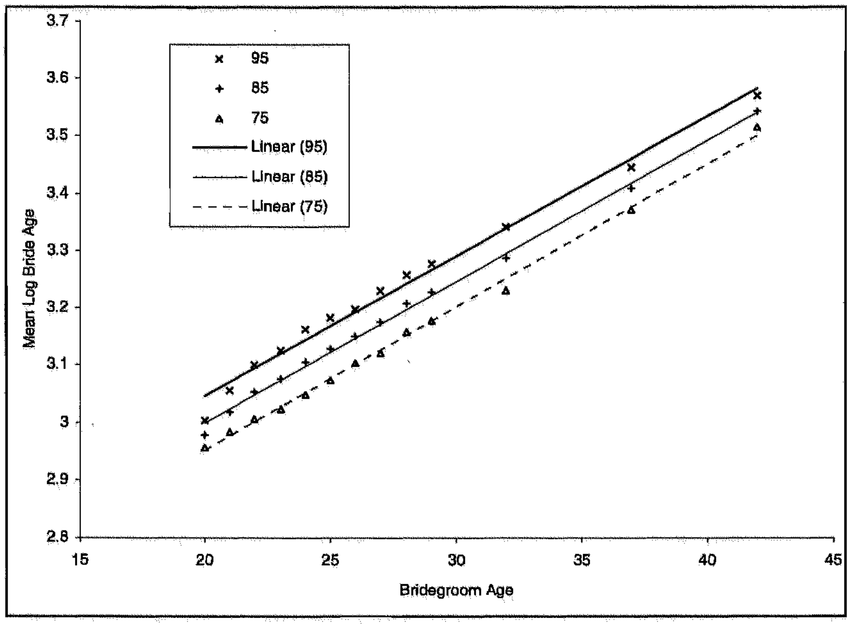


Figure 10: Mean log Bride Age by Groom Age - Cross-sectional observations.

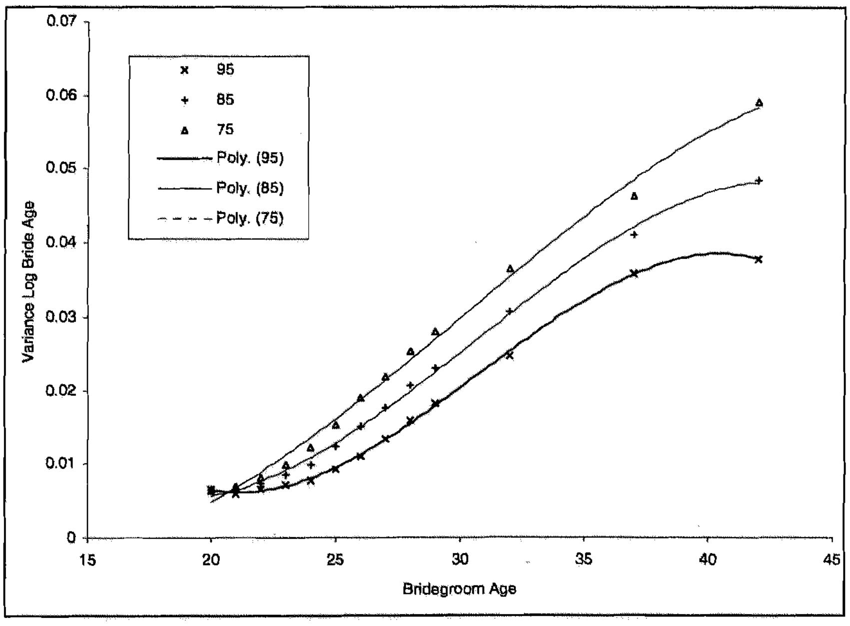


Figure 11: Variance log Bride Age by Groom Age - Cross-sectional observations.

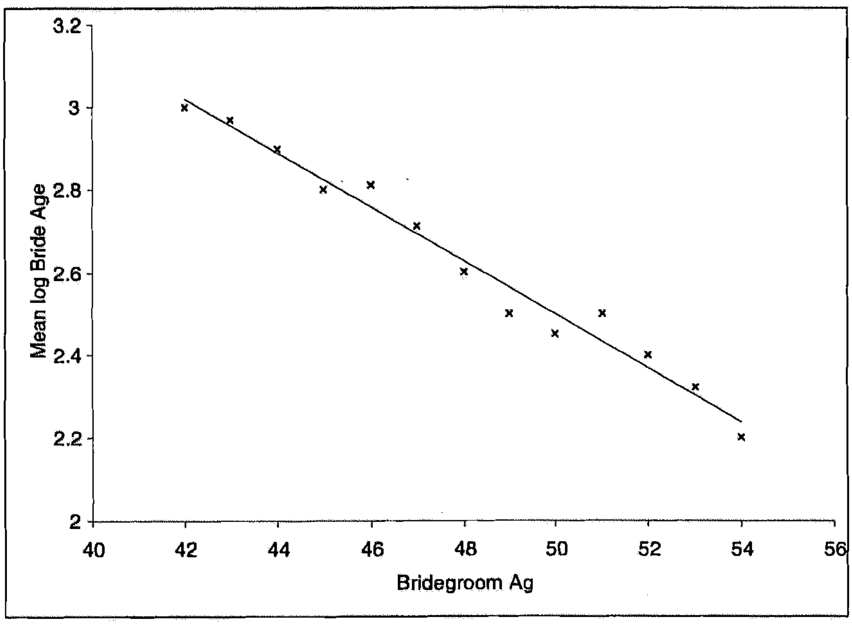


Figure 12: Mean log (Adjusted) Bride Ages by Groom Age

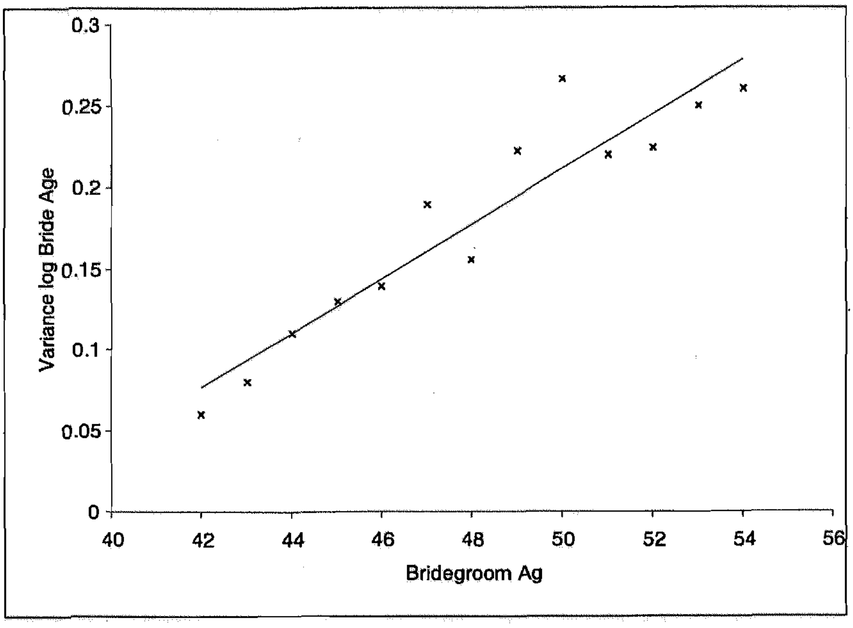


Figure 13: Variance log (Adjusted) Bride Ages by Groom Age

skewed distribution, this implies that bride age, on average, increases with time, as was observed for the positively skewed data. Similarly, the increase in the constant associated with the variance of log bride age is dominated by the decrease in observed slope such that the variance of log bride ages decreases with time, consistent with the trend observed for the positively skewed data.

To account for the observed temporal trends, the coefficients presented in Tables 7, 8, and 9 were regressed against linear equations with respect to calendar year. Due to the paucity of the available data, linear equations were used to 'smooth' the rough estimates of mean and variance obtained from the raw data by visual comparison. The coefficients displayed in the three tables consequently exhibit precise linear variation with time. The associated temporal estimates are provided in Table 10.

The most important test of the model is it's ability to reflect the raw data upon which it is based. To this end Figures 14, 15, and 16, which span the entire period of years examined, have been produced. These figures are typical of other combinations of groom age and calendar year and indicate that, despite it's obvious simplicity, the model closely reflects the raw data upon which it is based.

Table 10: Temporal Variation of Regression Coefficients

coefficient	const.	year
Mean log bride age - Groom age ≤ 42		
const.	-8.19786	5.393E-03
age	8.604E-02	-3.093E-05
Variance log bride age - Groom age ≤ 42		
const.	-17.29023	8.771E-03
age	1.69243	-8.599E-04
age^2	-5.315E-02	2.706E-05
age^3	5.674E-04	-2.889E-07
Mean log bride age - Groom age > 42		
const.	139.27376	-6.698E-02
age	-2.59842	1.271E-03
Variance log bride age - Groom age > 42		
const.	-1.70072	7.620E-04
age	0.10046	-4.600E-05

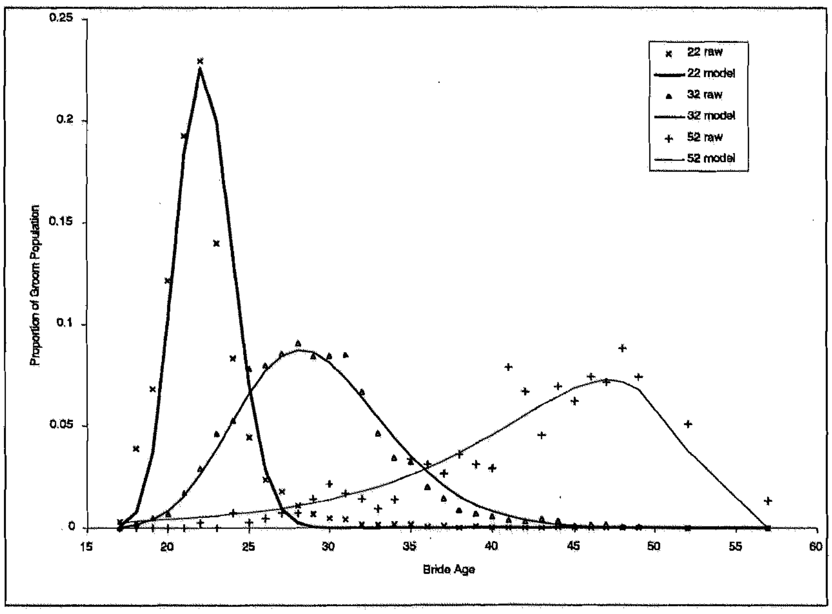


Figure 14: Model versus Raw Data - 1995

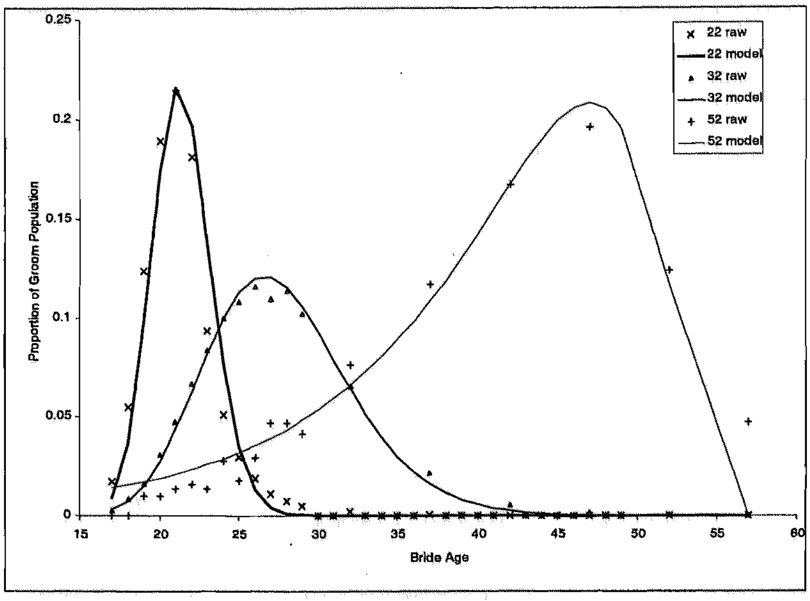


Figure 15: Model versus Raw Data - 1985

5 Number of Dependants

The number of dependant children for any individual in each year is equal to the number in the previous year, plus the number of dependants gained, less the number of dependants lost. Gains and losses are generated by the model subject to the following assumptions:

1. No individual can gain an additional dependant out of wedlock. Given that data are not available for the fathers of births out of wedlock, this assumption has been imposed to maintain some consistency with respect to the modelling routine of males and females. The effects of relaxing this assumption may be examined as part of the sensitivity analysis associated with marital status as referred to previously, where the definition of marital status may be expanded to encapsulate a broader set of personal relationships.

2. No individual marrying into the simulation cohort has a dependant prior to marriage.

3. No more than one child per year can be born to a couple.

4. Once born, each child remains a dependant for taxation and welfare purposes for 16 years. This is satisfied by assuming that each child in

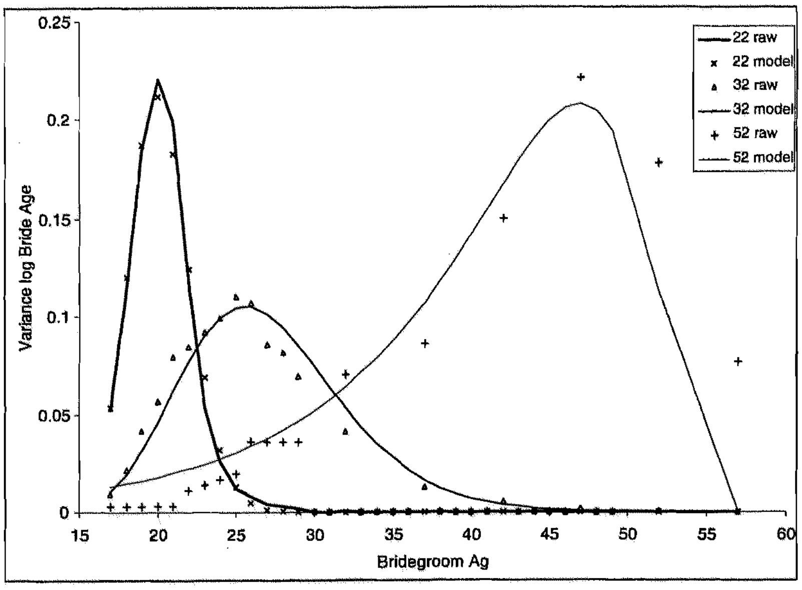


Figure 16: Model versus Raw Data - 1973

the model cohort remains a full-time student up to the age of 16 with an annual income less than \$6919.00 as at 20th of March 1998.

5. No female under the age of 18 can give birth.[15]

6. No female over the age of 44 can give birth.

7. The probability of giving birth to each additional child after the fifth is equal to the probability of giving birth to the fifth child.

8. In the case of divorce, the probability that any existing dependants remain with the male cohort members is exogenously set to θ_m, and θ_f for female cohort members. As a base condition, $\theta_f = 0.8 = 1 - \theta_m$.[16]

Given the assumptions described above, the loss of a dependant occurs either when they reach 17 years of age, or through divorce, in which case a Monte Carlo procedure is used to identify continued guardianship. Alternatively, the 'birth' of an additional dependant is generated for married couples

[15] This assumption is partially relaxed by the way in which the probability of giving birth is calibrated.

[16] θ_m and θ_f are specified separately so that the extreme conditions, where cohort members are always assumed to keep ($\theta_m = \theta_f = 1$), and always assumed to loose ($\theta_m = \theta_f = 0$) their dependants in the case of divorce, may be considered.

using a Monte Carlo procedure based on the age of the wife, and the previous issue.[17]

Due to a lack of suitable raw data, the probabilities required to generate child births were calculated using a recursive method based on a pseudo panel data set. Specifically, the tables listed in Appendix A provide estimates of the total married female population by age for each year between 1976 and 1996. In addition, the tables either provide the proportion of married women giving birth by age and previous issue, or enable the proportion to be determined for eight years between 1976 and 1995. Estimates fitting the respective time series and cross-sectional trends were used to expand these data sets to 27 consecutive years between 1971 and 1997.

Given the assumption that no wife under the age of 18 can give birth, the probability of an 18 year-old wife giving birth to her first child is equated to the proportion of 18 year-old wives who gave birth to their first child in 1971. The probability of a 19 year-old wife giving birth to her first child was determined by first multiplying the proportion of 19 year-old wives giving birth to their first child in 1972 by the respective married population of 19 year-old wives, providing the total number of 19 year-olds giving birth to their first child in 1972. This figure was divided by the number of 19 year-old wives who had yet to give birth, which was derived by subtracting the number of 18 year-old wives giving birth to their first child in 1971 from the total number of 19 year-old wives in 1972. The probabilities associated with consecutively higher aged wives giving birth to their first child (up to the age of 44) were determined using a similar process. This is generalised formally by:

$$p_{i,y} = \frac{\delta_{i,y} N_{i,y}}{N_{i,y} - \sum_{t=18}^{i-1} \delta_{t,(y-i+t)} N_{t,(y-i+t)}} \tag{2}$$

where $p_{i,y} =$ Probability that a wife of age i gives birth to her first child in year y

$\delta_{i,y} =$ Proportion of wives of age i giving birth to their first child in year y

$N_{i,y} =$ Number of wives of age i in year y

It is evident that $p_{i,y}$ provides only a rough approximation of the actual probability that a wife of age i gives birth to her first child in year y. Specifically, aside from the inaccuracy of the raw data used, mothers entering into $N_{i,y}$ via marriage or immigration bias $p_{i,y}$ downward while married mothers from previous years who exit the population through death, divorce, or im-

[17] 'Previous issue' is the number of children to whom the female has previously given birth.

migration bias $p_{i,y}$ upward. Hence the values of $p_{i,y}$ obtained from equation (2) provide a first estimate for calibrating the model.

The probabilities associated with each of the succeeding issues were obtained via a process similar to that of the first child as described above, by drawing upon the population defined by the preceding issue. The population defined by the preceding issue was obtained using a simple inflow-outflow equation (2), whereby:

$$N^c_{i,y} = N^c_{i-1,y-1} + \delta^c_{i-1,y-1} N_{i-1,y-1} - \delta^{c+1}_{i-1,y-1} N_{i-1,y-1} \tag{3}$$

where $\delta^c_{i,y} =$ Proportion of wives of age i giving birth to their cth child in year y.

$N^c_{i,y} =$ Number of wives of age i who had given birth to c children by the start of year y.

from which the associated probabilities, $p^c_{i,y}$, were calculated using:

$$p^c_{i,y} = \frac{\delta^c_{i,y} N_{i,y}}{N^c_{i,y}} \tag{4}$$

The crude probabilities obtained from this procedure were adjusted to ensure that the births generated by the model reflect the raw data upon which it is based. The resultant probabilities, which are stored by the model in a transition matrix, are provided in Table 11.

Figure 17 provides a comparison between the characteristics of a generated population of 50,000 individuals, and the raw data upon which the generating procedure is based. As can be seen, the generated data provide a reasonably close reflection of the raw data used.

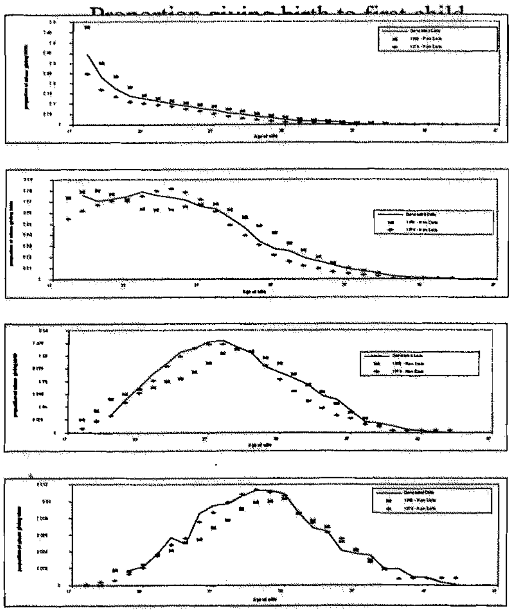
Due to a lack of raw data, no estimation of the time trend for births could be obtained. Specifically, the method used to generate dependants assumes that the year of observation is irrelevant with respect to the probability of a wife giving birth. That is, a wife aged 20 in 1970 has the same probability of giving birth as a wife aged 20 in 1990. This problem is partially mitigated by the method used to calculate the required probabilities, which incorporates the time trend associated with the cohort of females aged 18 in 1971.[18] The method used therefore partially incorporates the observed time trend, a trend that can be varied by adjustment of the individual probabilities in the transition matrix in order to undertake sensitivity analysis.

[18] This is reflected in Figure 17, where the simulated data are, in general, closer to the 1976 raw data for early ages, and closer to the 1990 raw data for later ages.

Table 11: Probability of Wife Giving Birth, by Age and Previous Issue

	Number of Previous Issue				
Age of wife	0	1	2	3	4 and over
18	3.50E-01	0.00E+00	0.00E+00	0.00E+00	0.00E+00
19	2.50E-01	4.10E-01	0.00E+00	0.00E+00	0.00E+00
20	2.00E-01	3.84E-01	1.91E-01	0.00E+00	0.00E+00
21	1.80E-01	4.30E-01	1.95E-01	2.79E-01	0.00E+00
22	1.68E-01	4.50E-01	1.92E-01	1.98E-01	3.47E-01
23	1.55E-01	4.50E-01	2.11E-01	1.80E-01	2.59E-01
24	1.50E-01	4.43E-01	1.79E-01	1.48E-01	1.97E-01
25	1.45E-01	4.40E-01	1.93E-01	9.80E-02	1.75E-01
26	1.30E-01	4.30E-01	1.84E-01	1.29E-01	1.28E-01
27	1.20E-01	4.35E-01	1.73E-01	1.17E-01	1.21E-01
28	1.10E-01	4.25E-01	1.52E-01	9.00E-02	1.11E-01
29	1.00E-01	3.80E-01	1.38E-01	9.25E-02	9.75E-02
30	8.00E-02	3.20E-01	1.15E-01	7.52E-02	8.04E-02
31	7.00E-02	2.50E-01	9.65E-02	6.77E-02	7.34E-02
32	5.00E-02	2.46E-01	8.77E-02	6.69E-02	5.79E-02
33	4.00E-02	2.18E-01	7.87E-02	4.15E-02	4.03E-02
34	4.00E-02	1.85E-01	6.17E-02	3.73E-02	3.55E-02
35	4.00E-02	1.64E-01	5.23E-02	2.75E-02	3.33E-02
36	2.00E-02	1.25E-01	4.43E-02	2.11E-02	2.94E-02
37	1.00E-02	9.94E-02	1.81E-02	1.68E-02	2.30E-02
38	5.00E-03	7.07E-02	1.79E-02	1.23E-02	2.08E-02
39	9.00E-04	4.93E-02	8.23E-03	8.52E-03	8.80E-03
40	5.00E-04	3.09E-02	5.29E-03	5.47E-03	7.83E-03
41	3.00E-04	1.86E-02	3.24E-03	3.21E-03	4.97E-03
42	2.00E-04	1.08E-02	1.81E-03	2.02E-03	3.54E-03
43	1.00E-04	5.21E-03	9.87E-04	1.09E-03	1.45E-03
44	1.00E-04	1.61E-03	2.45E-04	3.32E-04	0.00E+00

Proportion giving birth to first child



Proportion giving birth third child.

Proportion giving birth to fourth child.

Figure 17: Proportion of Wives Giving Birth by Age and Previous Issue: Generated versus Raw Data

6 Conclusion

This paper has outlined a microsimulation model of cohort demographics that is based upon the Australian population. The model has been created to enable analyses to be undertaken for the working-lifetime and is of particular value given the scarcity of Australian panel data. Care has been taken to structure the model to facilitate transparent sensitivity analysis, subject to the limitations imposed by the data used for calibration.

Marital status and the number of dependant children are generated as a series of binomial events based upon transition probabilities. Although the transition probabilities used to generate the births of children are stored in a transition matrix, transition probability functions are used to generate changing marital status. The use of functions to summarise the required transition probabilities is a departure from the standard microsimulation model architecture, which facilitates transparent sensitivity analysis. Functions are also used to summarise the data used to generate the age of spouses, which are randomly drawn from conditional distributions.

Although the functions used are statistical in nature and not based on theoretical maximising behaviour, they are shown to describe the observed data well. Specifically, the regular nature of the observed demographic trends imply that little detail is lost due to the restrictions imposed by the functional forms adopted. In addition, the modular nature of the model allows theoretical foundations of the functions used to be added, should they be desired, without affecting the remainder of the model.

7 References

Akerlof, G. A. (1998), "Men without children", *The Economic Journal*, 108, pp. 287-309.

Auerbach, A. J. (1997), Fiscal Policy: Lessons from Economic Research. London: MIT Press.

Basu, N., Pryor, R. & Quint, T. (1998), ASPEN: A microsimulation model of the economy", Computational Economics, 12, pp. 223-41.

Caldwell, S. (1996), Health, wealth, pensions, and life paths: The CORSIM dynamic microsimulation model. In Microsimulation and Public Policy: Selected Papers from the IARIW Special Conference on Microsimulation and Public Policy. Edited by A. Harding. Oxford: Elsevier, North-Holland.

Cogneau, D. & Robillard, A.-S. (2000), Growth, distribution and poverty in Madagascar: Learning from a microsimulation model in a general equilibrium framework", Trade and Macroeconomics Division, Internatinal Food Policy Research Institute, Discussion Paper, 61, pp. .

Creedy, J. & van de Ven, J. (1999), "The effects of selected Australian taxes and transfers on annual and lifetime inequality", *The Australian Journal of Labour Economics*, 3, pp. 1-22.

DeJanvry, A., Sadoulet, E. & Fargeix, A. (1991), Politically feasible and equitable adjustment: Some alternatives for Ecuador", World Development, 19, pp. 1577-94.

Dervis, K., DeMelo, J. & Robinson, S. (1982), General Equilibrium Models forDevelopment Policy. Cambridge: Cambridge University Press.

Giannarelli, L. (1992), An Analysts Guide to TRIM2, the Transfer Income Model, Version 2. Washington DC: Urban Institute Press.

Harding, A. (1993), Lifetime Income Distribution and Redistribution: Applications of a Microsimulation Model. London: North-Holland.

Hotz, V. J., Klerman, J. A. & Willis, R. J. (1997), The economics of fertility in developed countries. In Handbook of Population and Family Economics. Edited by M. R. Rosenzweig & O. Stark. New York: Elsevier Science.

Macunovich, D. (1998), Fertility and the Easterlin hypothesis: An assessment of the literature", Journal of Population Economics, 11, pp. 53-111.

Merz, J. (1991), Microsimulation - A survey of principles, developments, and applications", International Journal of Forecasting, 7, pp. 77-104.

Orcutt, G. (1957), A new type of socio-economic system", Review of Economics and Statistics, 58, pp. 773-97.

Redmond, G., Sutherland, H. & Wilson, M. (1998), The Arithmetic of Tax and Social Security Reform: A Users Guide to Microsimulation Methods and Analysis. Cambridge: Cambridge University Press.

Sutherland, H. (1995), Static microsimulation models in Europe: A survey", University of Cambridge, Department of Applied Economics Working Paper, Amalgamated Series: 9523, , pp. .

Taylor, L. (1990), Socially Relevant Policy Analysis. Structural Computable General Equilibrium Models for the Developing World. Cambridge: MIT Press

van de Ven, J. (2001), "Simulating Cohort Earnings for Australia", *University of Melbourne Department of Economics Research Paper.*

A Statistical Tables

Data sources are as follows:

1. Marital Status

    - Marital Rates

      ABS Cat No. 3310.0 (1998) Table 2.8
      ABS Cat No. 3306.0 (1988) Table 6
      ABS Cat No. 3306.0 (1985) Table 6

    - Divorce Rates

      ABS Cat No. 3310.0 (1998) Table 3.5
      ABS Cat No. 3307.0 (1987) Table 3

    - Death Rates

      ABS Cat No. 3302.0 (1998) Table 2.3
      ABS Cat No. 3302.0 (1987) Table 4

2. Relative Age of Spouse

    - Age of Bridegroom and Bride

      ABS Cat No. 3310.0 (1995) Table 10
      ABS Cat No. 3306.0 (1993, 1991) Table 10
      ABS Cat No. 3320.0 (1989, 1987, 1985) Table 9
      ABS Cat No. 3320.0 (1983, 1981, 1979, 1977, 1975, 1973) Table 7

3. Number of Dependants

    - Married Female Population

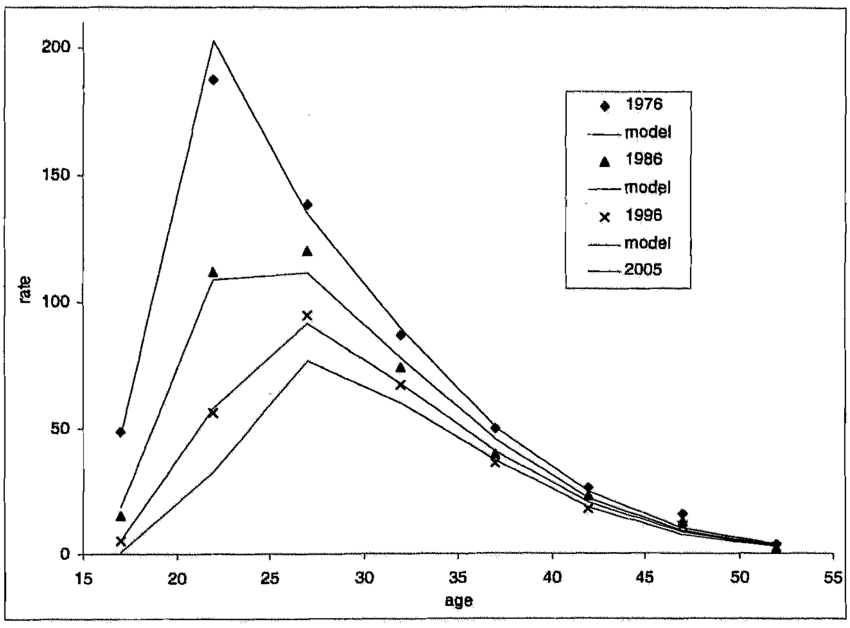
      ABS Cat No. 3101.0 (1995) Table 13

ABS Cat No. 3310.0 (1995) Table 20
ABS Cat No. 3310.0 (1994) Table 27
ABS Cat No. 3306.0 (1993) Table 19
ABS Cat No. 3220.0 (1990) Table 1
ABS Cat No. 3220.0 (1988) Table 1-8

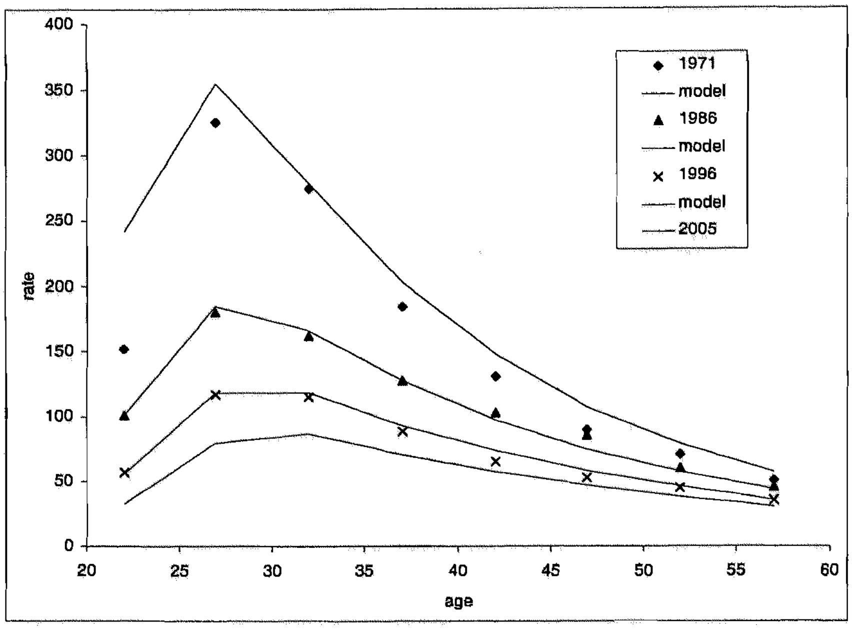
- Number of Births by Previous Issue

ABS Cat No. 3301.0 (1995, 1994) Table 22
ABS Cat No. 3301.0 (1993, 1990) Table 15
ABS Cat No. 3301.0 (1985, 1983, 1981) Table 13
ABS Cat No. 3301.0 (1976) Table 10

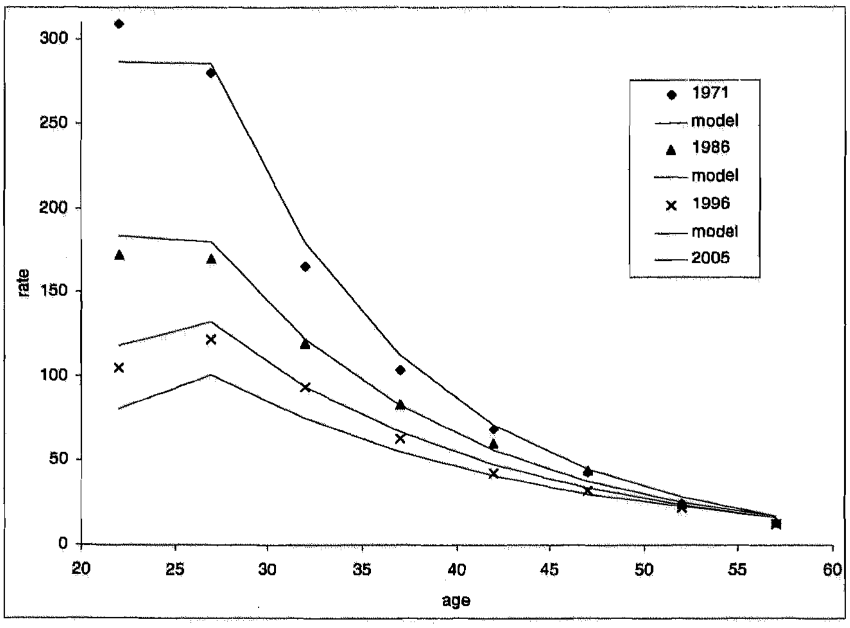
B Supplementary Figures - marital status



Female First Marriage Rates per 1000 Never Married Females - Model versus Raw Data



Male Remarriage Rates per 1000 Widowed or Divorced Males - Model versus Raw Data



Female Remarriage Rates per 1000 Widowed or Divorced Females - Model versus Raw Data

Table 12: Intertemporal Female First Marriage Rate Regression Output

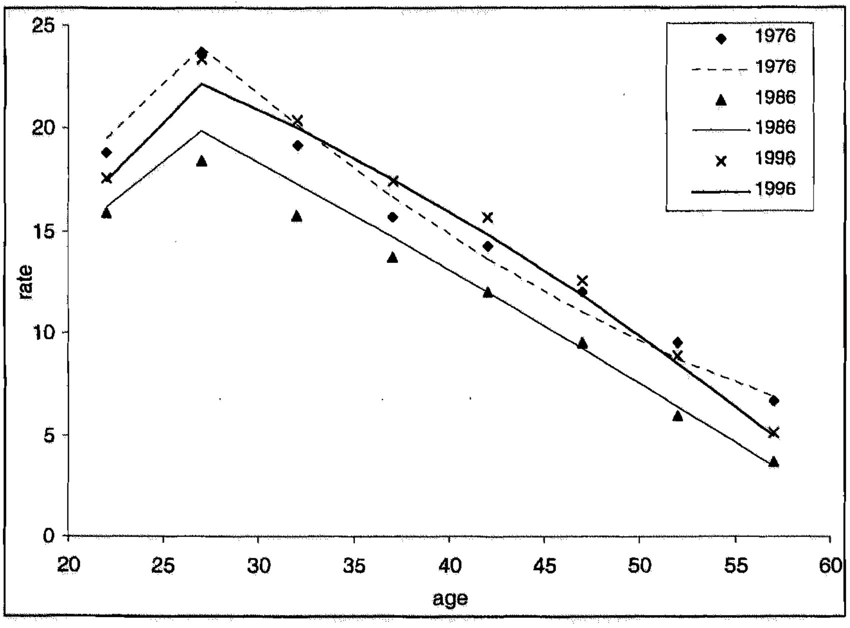
Exponentional cross-sectional segment linear time trend				
		c	year	R Square
γ	c	177.898	-0.08777	0.6310
		(44.414)	(0.0224)	
δ	age	-7.2771	3.733E-3	0.5209
		(2.369)	(1.19E-3)	
ϕ	age^2	0.08504	-4.460E-5	0.4779
		(0.0309)	(1.553E-5)	
Linear cross-sectional segment exponential time trend				
		c	year	R Square
$\ln(-\alpha)$		107.549	-0.05130	0.9732
		(5.976)	(0.00301)	
$\ln(\beta)$		110.910	-0.05439	0.9813
		(5.268)	(0.00265)	

Table 13: Intertemporal Male Remarriage Rate Regression Output

Exponentional cross-sectional segment linear time trend				
		c	year	R Square
γ	c	117.145	-0.0555	0.9597
		(7.5346)	(3.796E-3)	
δ	age	-1.3598	6.576E-4	0.8421
		(0.1884)	(9.489E-5)	
20-24 and 25-29 year old exponential time trends				
		c	year	R Square
$\ln(r_{22})$		121.271	-0.05874	0.7633
		(22.978)	(0.0116)	
$\ln(r_{27})$		92.356	-0.04388	0.9515
		(6.9575)	(0.00350)	

Table 14: Intertemporal Female Remarriage Rate Regression Output

Exponentional cross-sectional segment linear time trend				
		c	year	R Square
γ	c	120.610	-0.05705	0.9221
		(10.972)	(5.527E-3)	
δ	age	-2.0113	9.732E-4	0.8032
		(0.3187)	(1.606E-4)	
Linear cross-sectional segment exponential time trend				
		c	year	R Square
$\ln(-300-\alpha)$		222.404	-0.10971	0.6039
		(62.424)	(0.03142)	
$\ln(-15+\beta)$		146.906	-0.07276	0.7889
		(26.456)	(0.01332)	

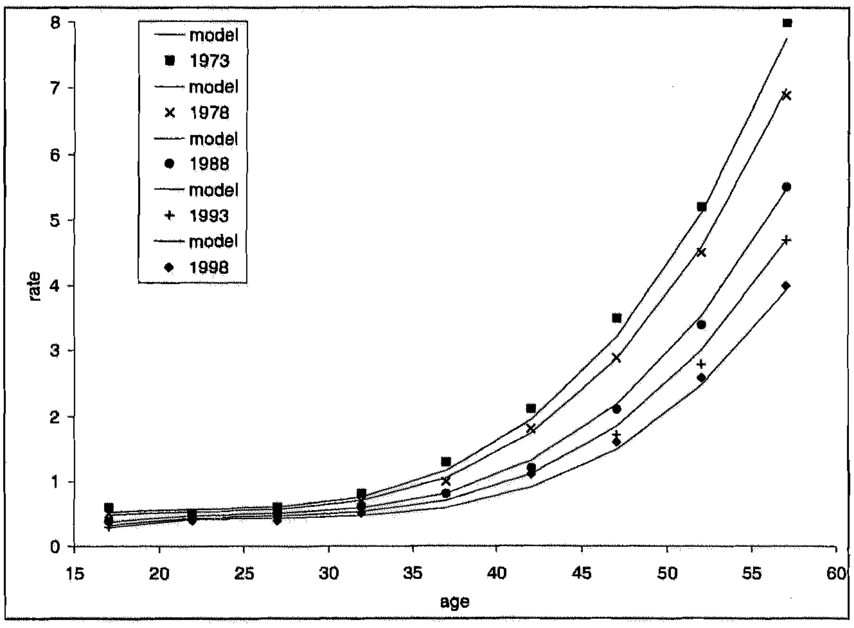


Female Divorce Rates per 1000 Married Females - Model versus Raw Data

Table 15: Intertemporal Female Divorce Rate Regression Output

Exponentional cross-sectional segment linear time trend				
coefficient	c	year	$year^2$	R Square
c	315197.4	−316.276	0.079347	0.7071
	(110335)	(110.99)	(0.02791)	
age	−9939.71	9.9535	-2.491E-03	0.7865
	(4694)	(4.7221)	(1.19E-03)	
age^2	107.363	−0.10745	2.688E-05	0.8313
	(49.05)	(0.04934)	(1.24E-05)	

Quadratic function for 20 to 24 year old age group			
c	year	$year^2$	R Square
91460.9	−91.9865	0.02313	0.649827505
(19381)	(19.497)	(4.90E-03)	

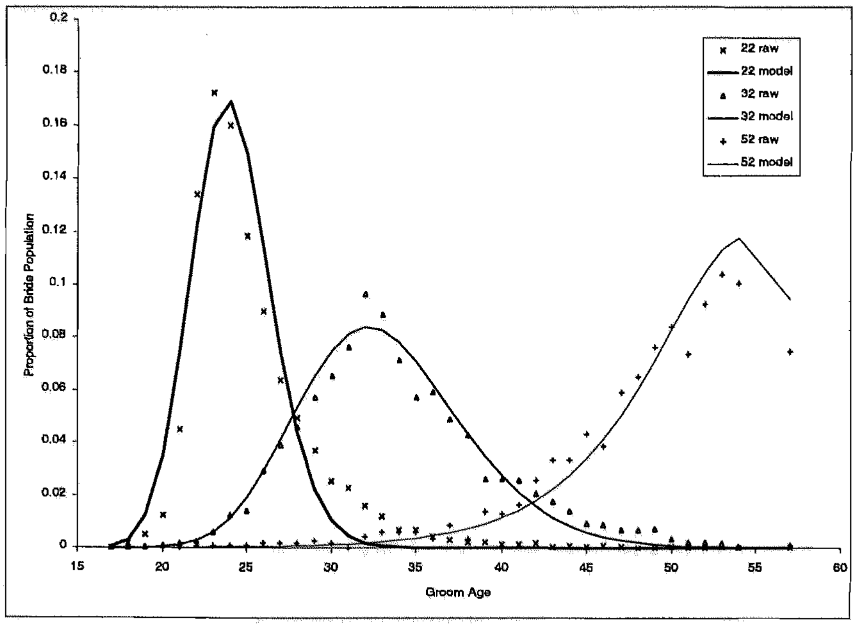


Female Death Rates per 1000 Female Population - Model versus Raw Data

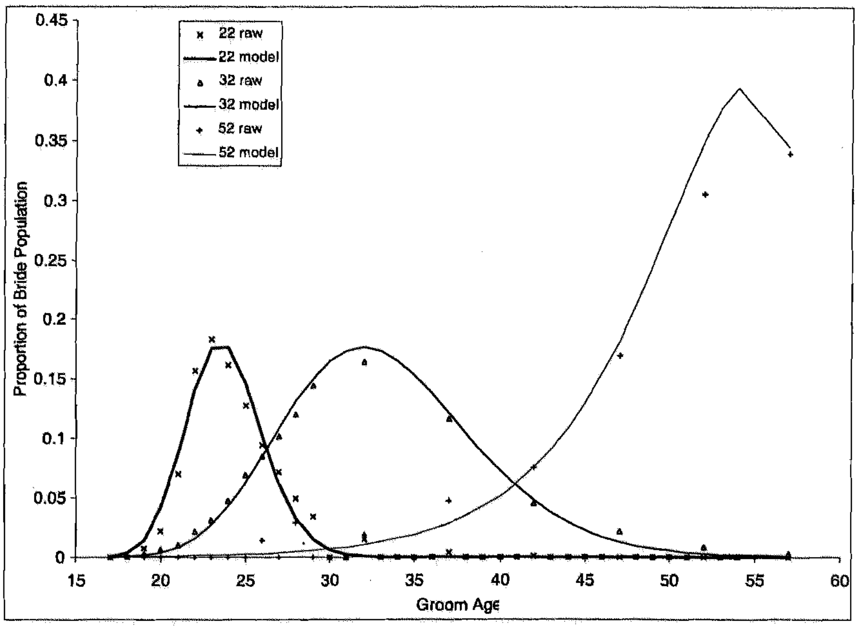
Table 16: Time series Female Death Rate Regression Output

Coefficient	c	year	R Square
constant	74.3756	-3.8333E-02	0.21923
	(38.412)	(1.93E-02)	
age	-3.2733	1.7823E-03	0.05614
	(3.881)	(1.95E-03)	
age^2	-4.5623E-02	1.7579E-05	0.00586
	(0.12157)	(6.12E-05)	
age^3	3.0780E-03	-1.4761E-06	0.32188
	(1.13E-03)	(5.73E-07)	

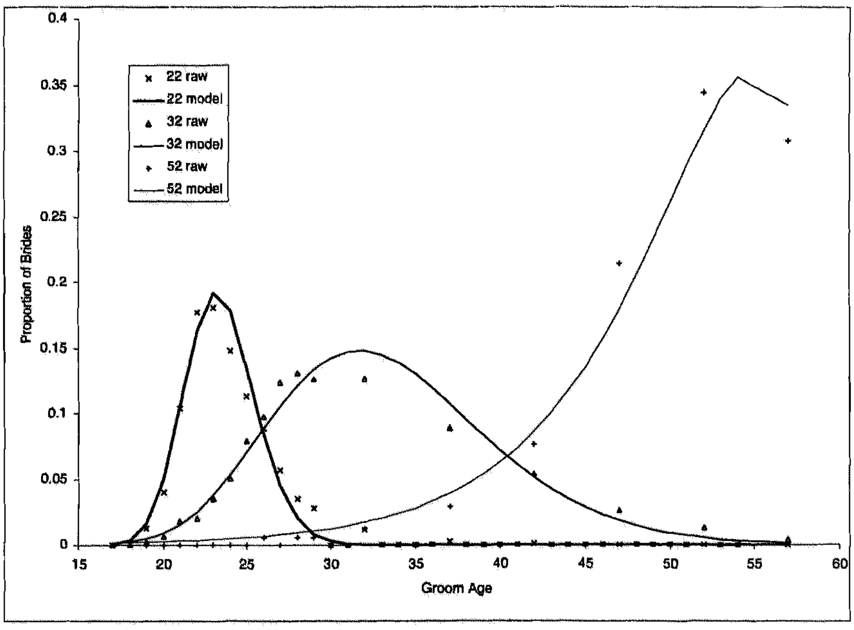
C Supplementary Figures - spouse age



Distribution of Groom Age Given Bride Age - Model versus Raw Data 1995.



Distribution of Groom Age Given Bride Age - Model versus Raw Data 1985.



Distribution of Groom Age Given Bride Age - Model versus Raw Data 1973.

Table 17: Temporal Variation of Regression Coefficients - distribution of Groom Age

coefficient	const.	year
Mean log groom age - Bride age < 40		
const.	-2.3220	2.421E-03
age	1.380E-01	-5.411E-05
Variance log groom age - Bride age < 40		
const.	3.2236	-1.475E-03
age	-6.687E-01	3.207E-04
age^2	3.278E-02	-1.594E-05
age^3	-4.243E-04	2.076E-07
Mean log groom age - Bride age $\geq$ 40		
const.	-3.0666	4.408E-03
age	5.235E-03	-3.337E-05
Variance log groom age - Bride age $\geq$ 42		
const.	-9.9733	4.913E-03
age	3.205E-01	-1.578E-04

NO.	AUTHOR/S	TITLE	DATE	INTERNAT. WORKING PAPER NO.	ISBN NO.	TOTAL NO. OF PAGES
771	Carol Johnston & Nilss Olekalns	Enriching the Learning Experience: A CALM Approach	December 2000	IWP 708	0 7340 1712 X	29
772	Robert Dixon	Australian Labour Force Data: How Representative is the 'Population Represented by the Matched Sample'?	January 2001	IWP 709	0 7340 1713 8	33
773	Lisa Cameron & Deborah Cobb-Clark	Old-Age Support in Developing Countries: Labor Supply, Intergenerational Transfers and Living Arrangements	January 2001	IWP 710	0 7340 1714 6	37
774	John Creedy	Quadratic Utility, Labour Supply and The Welfare Effects of Tax Changes	January 2001	IWP 711	0 7340 1715 4	16
775	John Creedy & Norman Gemmell	Public Finance and Public Education in a General Equilibrium Endogenous Growth Model	January 2001	IWP 712	0 7340 1716 2	39
776	Justin van de Ven	Distributional Limits and the Gini Coefficient	January 2001	IWP 713	0 7340 1717 0	20
777	John Stachurski	Stochastic Optimal Growth with Unbounded Shock	February 2001	IWP 714	0 7340 1718 9	31
778	Joseph G. Hirschberg, Esfandiar Maasoumi & Daniel J. Slottje	Clusters of Attributes and Well-Being in the US	February 2001	IWP 715	0 7340 1719 7	26
779	Justin van de Ven	Simulating Cohort Demographic Characteristics for Australia	March 2001	IWP 716	0 7340 1720 0	42
780	Justin van de Ven	Simulating Cohort Earnings for Australia	March 2001	IWP 717	0 7340 1721 9	27
781	Justin van de Ven, John Creedy & Peter J. Lambert	Close Equals and Calculation of the Vertical, Horizontal and Reranking Effects of Taxation	March 2001	IWP 718	0 7340 1722 7	15

RESEARCH PAPER SERIES - RECENT PUBLICATIONS IN THE DEPARTMENT OF ECONOMICS

NO.	AUTHOR/S	TITLE	DATE	INTERNAT. WORKING PAPER NO.	ISBN NO.	TOTAL NO. OF PAGES
782	John Creedy & Justin van de Ven	Taxation, Reranking and Equivalence Scales	March 2001	IWP 719	0 7340 1723 5	32