WP 36
31 May 2010

UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

STATISTICAL OFFICE OF THE EUROPEAN UNION (EUROSTAT)

CONFERENCE OF EUROPEAN STATISTICIANS

Joint Eurostat/UNECE Work Session on Demographic Projections
(28-30 April 2010, Lisbon, Portugal)

Item 12 – Stochastic techniques for demographic projections

A Mate-Matching Algorithm for Continuous-Time Microsimulation Models

Sabine Zinn, Max Planck Institute for Demographic Research, Germany

1. INTRODUCTION

In the social sciences, including demography, microsimulation is an approach that models dynamics of a system, the population, society or economy by modelling the behavior of its micro-units (typically an individual or a household). Microsimulations go back to the work of Orcutt (Orcutt 1957, Orcutt et al. 1961) and are techniques to produce population projections (Imhoff and Post 1998, Wolf 2001). The central unit of a demographic microsimulation is an individual's life-course, which is characterized by a sequence of demographic events such as birth, marriage, childbirth, divorce, retirement and finally death. In dynamic microsimulation models, aging of the micro units leads to a time-varying population structure. Both age and calendar time have to enter a realistic demographic microsimulation model, and both can be treated as discrete (usually in units of years or months) or as continuous (Galler 1997). Advantages and disadvantages of discrete- and continuous-time models are discussed at length in Satyabudhi and Onggo (2008), Willekens (2009), and Galler (1997).

Independently of how time is treated, in a demographic microsimulation model relationships between individuals should be regarded. Neglecting kinship and partnership relations during simulation is a source of biased outcomes. Murphy and Wang (2001), for example, argue that in the U.S., Italy, Norway and Poland "the relationship between fertility of successive generations is becoming stronger with time" (Murphy and Wang 2001, p. 1). Hence, in a related microsimulation model the ignorance of mother - daughter relationships would lead to a distorted fertility pattern. Moreover, a population's fertility and mortality pattern strongly depends on marriage and partnership processes. A woman, for instance, who lives in a partnership, has a much higher probability of childbirth than a single woman. As another example, mortality of married men is lower than of unmarried men. Ignoring the impact that a spouse has on an individual's life course is, therefore, tenuous.

In modelling demographic kinship different problems have to be addressed.

1. How can we model and simulate the onset of relationships?
2. How can we model and simulate correlation between life-courses?
3. How can we model and simulate dissolution events?

In this article we focus on the first problem, with particular emphasis on mate-matching in continuous-time microsimulations.

We structure our work as follows. In a first step we review mate-matching algorithms of existing microsimulations. We continue, in section 3, by describing a mate-matching algorithm that works in continuous time. In section 4, we illustrate the capabilities of the present approach. Running simulations, we forecast a synthetic population based on the population of the Netherlands. We conclude in section 5 by validating the new procedure, and we give an outlook of how we can improve our work.

2. A REVIEW OF MATING MODELS

When developing mating models, several difficulties have to be borne in mind and tackled. The first one concerns the high data demands that are necessary to map mate-matching mechanisms properly. In order to determine synthetic couples, data from actual couples are necessary. Retrospective surveys that record the attributes of spouses before and after the onset of partnerships would be most appropriate. However, in the majority of cases such surveys only gather very limited information on partnership relations (Huinink and Feldhaus 2009). Problems also arise because actual mating processes comprise mechanisms that are not observable, e.g., the formation rules. The modeler is forced to set assumptions and hypothetical rules about mating behavior. Only the effect of these can be simulated and hypothetical outcomes can be contrasted with those actually observed (de Vos and Palloni 1989). A further difficulty emerges due to the fact that in microsimulation models the temporal progression is modelled in reverse: First an event that leads to a partnership onset is simulated and then the question of defining a proper spouse is addressed. Consequently, if a transition implies the onset of cohabitation or a marriage, an individual has to become part of a couple. Otherwise, the outcome of the underlying stochastic model is disregarded.

In the field of demographic microsimulations three different approaches exist to model the onset of partnerships, and marriages in particular: (1) ignoring mating processes, (2) open mating models, and (3) closed mating models. We describe these approaches in detail below, paying special attention to the above listed problems. We do not provide a categorization of existing microsimulations according to the type of mating model used. For an overview of an extensive range of microsimulation models, see O'Donoghue (1999), Zaidi and Rake (2001) and Spielauer (2002).

2.1 No mating model

Ignoring mating processes means that individual life-courses are modelled independently of each other. If an individual experiences a partnership event, no synthetic couple is created. As a result, it is impossible to map inter-dependencies between life-course events of spouses.

Ignoring mating is reasonable in the context of one-sex models. An example of this type is the family microsimulation FAMSIM that specifies partnership forms as attributes of women (Spielauer and Neuwirth 2001).

In general, ignoring mating processes has some advantages: Neither data for actual couples is required, nor hypothetical rules that describe mating processes. However, skipping mating processes has two substantial defects. Their effects on the population composition are neglected. For example, a restricted number of available spouses might affect the characteristics of formed partnerships. This, in turn, might have an impact on the number of partnership dissolutions. Furthermore, ignoring relationships between individuals prevents the modelling of demographic kinship. As a result, it is, for example, not possible to account for intergenerational inter-dependencies, such as the correlation between the fertility behavior of mothers and daughters.

2.2 Open mating model

In an open model, appropriate spouses are created 'ex nihilo'. Their attributes are generated in such a manner that the characteristics of the newly created couples resemble actual ones. Age and educational attainment are the attributes that are usually regarded as being essential in this context.

As it is not necessary to identify a proper spouse within the model population, the implementation of open models is straightforward. A further advantage of this approach is that simulations for individuals (and their immediate families) can be run independently of other individuals (O'Donoghue 1999). However, open models reveal three major problems. First of all, it is not assured that the newly created individuals are representative of the target population. This might be the case in one-sex models, but the situation differs if whole populations are mapped. Another problem is the interpretation of an open model. The purpose of a microsimulation is to model population dynamics realistically. However, it is not realistic to pull an appropriate spouse out of the hat when needed. A problem that is related to this point is the missing retrospective life-course of a newly created spouse. Smith (1987) describes a method that tackles this problem by creating spouses whose only characteristic is that they are at the right age (relying on age differences that are reported for real couples). A more viable solution is proposed by Holmer et al. (2009). They suggest sampling complete retrospective life courses of spouses from retrospective surveys. Nevertheless, this presupposes the availability of huge sets of event history data.

The discrete-time microsimulation CAMSIM (Smith 1987), and the continuous-time microsimulations PENSIM (Holmer et al. 2009) and LifePaths (Statistics Canada 2004) are examples of the use of open mating models.

2.3 Closed mating model

In a closed model, marriage and consensual union partners have to be found among existing individuals. To do so, several problems have to be addressed:

1. How can we determine who is searching?
2. Who matches whom?
3. When are couples formed?
4. What is the data base needed in order to form synthetic couples that resemble actual ones?

To the knowledge of the author, closed mating models have so far only been realized for discrete-time microsimulations.

In a discrete microsimulation model, time changes in discrete steps. After each step all individuals of the model population are inspected whether they experience an event in the next interval and, if yes, which event. This simulation procedure implicitly determines who is searching and also when couples are formed. Searching individuals are collected in a partnership market. A 'partnership market' is a construct that is used to pool all those individuals who look for a spouse. The notion 'market' might be confusing at this point as it suggests that searching individuals are trading with restricted sets of goods. However, in the literature a 'partnership market' is described as a pool of prospective spouses. From the technical point of view, the partnership market can be regarded as a sorted list of individuals. A sorting criterion is, e.g., age.

It has to be ensured that in the matching process those individuals who look for a marriage partner are not paired with persons that search for a common-law spouse, and vice versa. Consequently, the partnership market contains two separated sets of individuals: 'cohabitation-willing' and 'marriage-willing' individuals. The latter ones are assumed not to live in cohabitation, when they enter the market. Individuals, who live in cohabitation, already have spouses. If they experience marriage events, their cohabitations are simply converted into marriages. A few microsimulations exist where the partnership market is only accessible for marriage-willing individuals and not for those who look for a common-law spouse. This constraint has been identified as a major inconsistency (Leblanc et al. 2009). Annual or monthly partnership markets are an obvious choice in discrete models. After one year or one month, the partnership market is depleted by pairing individuals.

So far we have addressed the questions of how and when. Two questions remain: Who mates whom, and what data is needed to construct couples that resemble actual/ observed ones? Both questions concern the mating rules that are applied to match individuals. Generally, microsimulation models employ two types of mating rules: stable and stochastic ones (Perese 2002). In order to determine the quality of potential pairings, mating rules make use of a compatibility measure. This measure quantifies the quality of a respective pairing dependent on the attributes of the potential spouses.

2.3.1 Compatibility Measure

The compatibility measure is a function that associates female and male attributes with a positive real number. This number indicates how compatible a woman and a man are. A large value is a sign for high compatibility. Likewise, a small value points to incompatibility.

We introduce some notation.

- At a fixed point in time the partnership market comprises m women and n men.
- We denote female attributes by w_i, $j=1,\ldots, m$ and the male attributes by m_j, $j=1,\ldots, n$.
- The set F comprises all w_i and the set M all m_j.

The compatibility measure C is defined as follows C: $F \times M \rightarrow R_0^+$. Often, the domain of C comprises only the age and educational attainment of males and females. However, the domain of C differs between microsimulation models. Bacon and Pennec (2007) provide an extensive review of attributes that have been employed in mating models. Two different specifications of C are used: distance functions, and logit models. Distance functions are employed for minimization of the discrepancy in the attributes of spouses. Examples are the French microsimulation DESTINIE, and DYNASIM. DESTINIE employs a sum of squared differences (Duee 2005), and DYNASIM embodies an exponential distance function (Perese 2002).

In logit models, the idea is to quantify compatibility by the likelihood of a union between potential pairs (Perese 2002, Bouffard et al. 2001). In order to account for different types of partnerships (cohabitations/ marriages) and to differentiate between first and higher order partnerships, typically more than one logit model is applied. Data on observed couples are used to estimate the coefficients of these logit models. Empirical findings show that partners tend to have similar ages and education. Therefore, ideally, the estimated coefficients are in accordance with the theory of assortative mating (Bouffard et al. 2001, Leblanc et al. 2009).[1]

There exists one problem that most probably emerges in a partnership market: What to do if there are more men than women, or vice versa? Three alternative strategies have been proposed in the literature.

1. In the current period, the model treats remainder individuals as though they did not enter the partnership pool. They are simply left unmarried.
    a. In the next period, they are at risk to experience a partnership event (Perese 2002, Leblanc et al. 2009).
    b. In the next period, they are automatically members of the pool of potential spouses (Hammel et al. 1990).
2. Individuals are added or removed as needed (Leblanc et al. 2009).
3. A totally different approach is simulating partnership events only for females. Appropriate spouses are taken from the pool of unpaired eligible men (Hammel et al. 1976, King et al. 1999).

The first two strategies imply disregarding the outcome of the stochastic model of a dynamic microsimulation. For each individual in the marriage market a partnership event has been simulated. Therefore, removing excess from the marriage pool means to ignore a simulated event. Likewise, adding individuals to the pool means that individuals who are not simulated to experience a marriage event get married. The last strategy implies that female partnership behavior completely determines the dynamics in male partnership behavior. Bacon and Pennec (2007) suggest to use the second strategy and to embed an alignment procedure into the matching process. Their idea is to add and remove individuals until a pre-defined number of couples has been created.

2.3.2 Stable Mating Rules

In order to match individuals both stable and stochastic rules can be employed. The problem of finding stable mating rules is equivalent to the stable marriage problem. Gale and Shapley (1962, p. 11) describe the problem as follows: "A certain community consists of n men and n women. Each person ranks those of the opposite-sex in accordance with his or her preferences for a marriage partner. We seek a satisfactory way of marrying off all members of the community. [...] we call a set of marriages unstable [...] if under it there are a man and a woman who are not married to each other but prefer each other to their actual mates." They prove that, for any number of men and women, it is always possible to solve the problem and make all marriages stable.

The stable marriage approach requires that, in the partnership market, each woman expresses her preference regarding each man and vice versa. Gale and Shapley (1962) developed an algorithm that produces a set of stable marriages. It is based on a sequence of proposals from men to women. Each man proposes, in descending order, to the women according to his preferences. A man is pausing when a woman agrees to consider his proposal. He is continuing if a proposal is immediately or subsequently rejected. When a woman receives a proposal, she rejects if she already holds a better proposal (relying on her preferences). Otherwise, she agrees to hold the proposal for consideration. In doing so, she rejects any poorer proposal that she may hold. This procedure assures that no man can have a better partner than he gets in this matching and no woman can have a worse one. Consequently, the Gale-Shapley algorithm produces marriages that greatly favor the men's preferences. Since 1962 several studies have developed the topic and improved the algorithm. Researchers mainly work on two defects of the Gale-Shapley algorithm. The gender that is allowed to give the first bed is favored. Furthermore, the classical Gale-Shapley algorithm does not result in a unique set of stable marriages.

Relying on the theory of assortative mating, the degree of compatibility between a woman and a man is a proper measure of their preferences for each other. From now on, we quantify this degree using the compatibility measure *C* that has been introduced in paragraph 2.3.1. Its usage for constructing synthetic couples allows a simplified version of the stable marriage algorithm (Bouffard et al. 2001). The preference of a woman for a man might differ from the man's preference for the woman. The compatibility measure, however, is symmetric in its treatment of men and women.

Using the notation given in the previous paragraph, we define (Perese 2002): A stable set of pairings is reached, if for all couples (w_i, m_i) and (w_j, m_j), $i \neq j$, the following condition holds true:

$C(w_i, m_i) > C(w_i, m_j)$ *or* $C(w_j, m_j) > C(w_i, m_j)$. Only if both inequalities fail, would a set of pairings be unstable.

[1] Naturally, persons that are closely related show strong assortative characteristics. Therefore, in order to avoid incestuous pairings.

Applying the following algorithm, a stable set marriages/ consensual unions can be determined:

1. Men and women are separated into two sets.
2. A compatibility measure is computed for all potential pairs.
3. All pairings are ordered according to their compatibility measure (in descending order).
4. The best match is paired. (Those two individuals are matched that have the highest degree of compatibility.)
5. All pairings that include one of the spouses of the newly formed couple are removed from the list of potential couples.
6. The compatibility of the remaining individuals is re-ranked and the next most compatible couple is paired.

This procedure is repeated until all matches have been made.

A selection rule has to be applied if there are ties. The algorithm has a time complexity of $O(n^2)$.

The stable mating approach has three desirable features (Bouffard et al. 2001):

- It is based on extensive research.
- It is easy to understand and implement.
- Due to the usage of the compatibility measure, both sexes are equally treated. Neither sex is favored.

Nevertheless, the stable mating algorithm suffers from a remarkable defect. In the beginning it produces couples that have high compatibility measures. Finally, only individuals who do not match well remain in the pool. Therefore, matches are created that would have a very low probability in reality. Bouffard et al. (2001) study the effects of this imbalance. In order to analyze the suitability of the outcomes of the stable mating approach, they used the Canadian 1981 census. A logit model was employed to measure compatibility. In their examination, they found that the stable mating algorithm produced too many "extreme" pairings (e.g. age differences of spouses greater than 20 years). Furthermore, they point to a misuse of the information that the compatibility measure comprises: A logit model maps the likelihood of a potential pairing, not one of an optimal match.[2] Bouffard et al. (2001) also note on some attempts to overcome the problems of a stable mating algorithm. However, none of the proposed modification leads to a significant improvement in the results.

Recently, Leblanc et al. (2009) described an algorithm (ODD, order of decreasing difficulty) that first finds good matches for those individuals who show undesirable characteristics. It processes to construct pairs along the order of decreasing difficulty.[3] However, this algorithm is also not capable of reproducing actual data

Randomly reducing the pool of prospective spouses is another approach to solve the problem of "bad" matches (Cumpston 2009, Leblanc et al. 2009).[4] The corresponding procedure is as follows:

1. An individual i is randomly drawn from the pool of prospective spouses.
2. A certain number p of opposite-sex individuals is also randomly selected from the pool.
3. Of these p potential partners, that one is chosen that holds the highest compatibility with i. This procedure is repeated until the marriage market is depleted.

Leblanc et al. (2009) reveal that this modified approach still suffers from the problems of the original stable mating algorithm. They find that "the algorithm generates far too many marriages with extreme age differences" (Leblanc et al. 2009, p. 18). Moreover, they figure out that the constructed matchings show a distribution of compatibility measures that diverges significantly from the one estimated from real data.

In conclusion, the main inconsistency of the stable mating algorithm is the incongruity of the concept of the compatibility measure and the matching algorithm itself: For "any arbitrary pairing, the measure's value should be proportional to the probability that those persons end up marrying. The stable marriage algorithm, however, actively departs from this property because, in its quest for stability, it disproportionately favors for pairings with high compatibility values" (Bouffard et al. 2001, p. 15). Stochastic mating rules are an option to overcome that problem.

---

[2] Regarding this point Bouffard et al. (2001) refer to conclusions drawn by Easther, R. and J. Vink (2000): "A Stochastic Marriage Market for CORSIM", Technical paper. However, despite her greatest efforts the author could not access this article.

[3] Concerning the development of the ODD algorithm Leblanc et al. (2009) refer to the work of Redway, H. They cite a presentation of Redway, H. (July 2005): "Data Fusion by Statistical Matching", Model Development Unit Presentation.

[4] Leblanc et al. (2009) call their technique 'tournament' algorithm and Cumpston (2009) denotes his approach 'best of n' matching. Notwithstanding, regarding the concept both methods are equivalent.

The Swedish spatial microsimulation model SVERIGE (Holm et al. 2002) and the U.S. policy microsimulation POLISIM (Caldwell et al. 1999, Caldwell 1996, Bouffard et al. 2001, O'Harra and Sabelhaus 2002) are examples of the employment of a stable mating algorithm.

2.3.3 Stochastic Mating Rules

In a stochastic mating model, the outcome of a stochastic experiment decides whether a match between two potential spouses occurs. The compatibility measure between a woman and a man indicates the probability of a respective match. A stochastic matching procedure ensures that individuals with a low compatibility also have a chance to get matched. With regard to their compatibility, constructed couples are thus not necessarily optimal ones. As a result, the occurrence of "extreme" matchings is less likely. The latter is a big advantage over the stable mating algorithm.

In microsimulation models, basically, three variants of stochastic mating are applied:

1. Male-dominant mate-matching,
2. female-dominant mate-matching, and
3. mixed-dominant mate-matching.

While describing the related algorithms, we make use of the notation that has been introduced in paragraph 2.3.1.

Variant 1: Male-dominant

In an algorithm for male-dominant mate-matching, men choose their spouses from a list of eligible women. The DYNASIM team has developed an efficient algorithm that produces acceptable results in linear time (Perese 2002). The algorithm is described in Figure 1.

Perese (2002) deems the exponential distance function that DYNASIM includes, to be inappropriate for mapping compatibility. Therefore, he replaces it by logit models. He finds that, due to this replacement, many potential couples have very low compatibility measures. As a result, the probability of producing matches declines, and more iterations are needed to find a proper spouse. A significant increase in the algorithm's run time is the consequence. By applying normalized compatibility measures in DYNASIM's mate-matching algorithm, Perese tackles this problem. Before matching starts, for each man, the highest compatibility value that he can achieve with a woman is determined. Subsequently, all compatibility measures that a man exhibits with potential spouses are divided by this highest value. The normalization ensures that at least with one woman a man has a compatibility value of 1. Until a match is made, a searching man scans a random number of women. Perese (2002) argues that "this technique creates a more randomized process than the one employed in DYNASIM, which arbitrarily limits the search to *10* women for each man before a match is made with certainty" (Perese 2002, p. 17). However, the modification causes an increase of the theoretical run time of the original algorithm. The modified algorithm has time complexity $O(n^2)$. The processing of the revised algorithm is given in Figure 2. In order to test the algorithm, Perese estimated logit models using SIPP survey data.[5] Simulation runs showed that the algorithm closely replicated actual data.

[5] SIPP is short for Survey of Income and Program Participation. It is a U.S. survey that records in one of its modules marriages that started in 1994, 1995, or 1996.

1. All women $w_1, \cdots, w_m$ are put in random order into a list L_1. Likewise, all men $m_1, \cdots, m_n$ are inserted in random order in a second list L_2.

2. Initialize a counter variable for men: $j = 1$.

3. Initialize a counter variable for women: $i = 1$.

4. Take the jth man m_j of L_2.

5. In order to record already computed compatibility measures with man m_j, initialize a K-dimensional vector: $V = (V_i)_{i=1,\cdots,K} = (0, \cdots, 0)$. $K = 10$ for males younger than 35, or $K = 20$ otherwise.

6. Take the ith woman w_i of L_1.

7. Compute the corresponding compatibility measure $C(w_i, m_j)$. Set $V_i = C(w_i, m_j)$.

8. Draw random number R uniformly distributed between 0 and 1.

9. If $R > C(w_i, m_j)$ the match between m_j and w_i is made.

   9.1 Remove w_i from list L_1.

   9.2 Otherwise, $i = i + 1$.

10. Repeat steps 6 to 9 until either

    - a match has been made, or
    - $i = K$.

    In the latter case m_j is matched with the woman with whom he possesses the highest compatibility $\max\{V_i, i = 1, \cdots, K\}$.

11. Set $i = 1$ and increment j by 1.

Repeat steps 3 to 11 until either L_1 or L_2 is empty.

Figure 1 - Male-dominant mate-matching algorithm

Variant 2: Female-dominant

An algorithm for female-dominant mate-matching is basically equivalent to one for male-dominant mate-matching. Only, the roles of women and men are reversed: a searching woman is allowed to choose among men. Examples of the employment of female-dominant mate-matching procedures are: the first version of SOCSIM (Hammel et al. 1976, Chapter 9), and the DYNAMOD (Kelly 2003, King et al. 1999) microsimulation.

1. Put all women $w_1, \cdots, w_m$ in random order into a list L_1. Likewise, insert all men $m_1, \cdots, m_n$ in random order in a second list L_2.

2 Initialize a counter variable for men: $j = 1$.

3. Take the jth men m_j of the second list.

4. For all members w_i of L_1 compute $C(w_i, m_j)$. Assign $v = \max\{C(w_i, m_j), i = 1, \cdots, m\}$.

5. Initialize a counter variable for women: $i = 1$.

6. Take the ith woman w_i of L_1.

7. Compute the corresponding compatibility measure $C(w_i, m_j)$. Normalize the measure: $\tilde{C}(w_i, m_j) = C(w_i, m_j)/v$.

8. Draw random number R uniformly distributed between 0 and 1.

9. If $R < \tilde{C}(w_i, m_j)$ the match between m_j and w_i is made.

    9.1 Remove w_i from list L_1.

    9.2 Otherwise, $i = i + 1$.

10. Repeat steps 6 to 9 until a match has been made.

11. Set $i = 0$ and increment j by 1.

Repeat steps 3 to 11 until the list of bachelors is empty.

Figure 2 - Male-dominant mate-matching algorithm with normalized compatibility measures

Variant 3: Mixed-dominant

In a mixed-dominant mate-matching procedure both genders are equally treated. We differentiate in this article between two variants. We call the first ‘the sequential approach’ and the second one ‘the concurrent approach’. The sequential approach has been realized in the SOCSIM microsimulation (Hammel et al. 1990) and the concurrent approach in the U.S. CORSIM microsimulation and the Canadian DYNASIM microsimulation.

In the sequential approach individuals ‘seek a [...] partner in random order from among the members of the opposite sex in accordance with their criteria of preference” (Wachter 1995, p. 7). (In our terminology, the latter are outcomes of a compatibility measure.) The actual spouse is chosen at random from the opposite-sex candidates with the highest compatibility.

The corresponding processing is presented in Figure 3. Unsuccessful suitors (if in step 5, $z=0$) remain unpaired. In step 5.3, an arbitrary number between 1 and z can be set for M. If $M=1$, the sequential algorithm and the revised stable mating algorithm in paragraph 2.3.2 are equivalent. In its worst case, i.e. $M=max(n,m)$, the sequential approach has time complexity of $O(n^2)$. It has linear complexity if the minimal criteria (step 3) are rather restrictive and L_2 has only few elements. Wachter (1995) and Wachter et al. (1998) find that the algorithm produces feasible results.

The concurrent approach of stochastic mate-matching goes back to the work of Vink and Easther (Bouffard et al. 2001). It had been developed originally for the CORSIM microsimulation model. Its functionality and the suitability of the outcome are extensively discussed by Bouffard et al. (2001). They deem the approach to be suitable for reproducing actual data. The algorithm comprises the steps described in Figure 4. As compatibility values have always to be computed for all potential pairings, the algorithm has time complexity of $O(n^2)$.

1. Sort in random order all searching individuals $w_1, \cdots, w_m$ and $m_1, \cdots, m_n$ into a list L_1.
2. Initialize a counter $k = 1$.
3. Draw at random an individual i from L_1.
4. Insert all opposite-sex individuals that meet some minimal criteria into a second list L_2. The number ofi ndividuals in L_2 is z.
5. If $z > 0$:
   - 5.1 Compute compatibility measures c_p between i and all individuals in L_2, $p = 1, \cdots, z$.
   - 5.2 Sort all individuals in L_2 in decreasing order according to their compatibility measures.
   - 5.3 Take the first M members of the list. Their respective compatibility measures are $\tilde{c}_1, \cdots, \tilde{c}_M$.
   - 5.4 Draw random number R uniformly distributed between 0 and $\sum_{p=1}^{M} \tilde{c}_p$.
   - 5.5 Match i and that individual h for which $\sum_{p=1}^{h-1} \tilde{c}_p \leq R < \sum_{p=1}^{h} \tilde{c}_p$.
6. Increment k by 1.

Repeat steps 3 to 6, until $k = \min(n, m)$.

Figure 3 - Mixed-dominant sequential mate-matching algorithm

In conclusion, we draw the following lessons from our literature review:

- Closed models are easier to interpret than open models. They enable us to regard the effects of mating processes on the population composition.
- In order to measure the compatibility between two persons, logit models are more appropriate than distance functions.
- Each strategy that has been proposed to obtain the same numbers of women and men in the partnership market shows defects.
- Stochastic mate-matching procedures resemble actual data better than stable mating procedures. The outcome of a stochastic mate-matching algorithm is not significantly affected by the chosen variant (male-, female-, or mix-dominant).
- In the context of stochastic mate-matching, a 'sequential approach' is on average more efficient than a 'concurrent approach'.

The aim of this article is to construct a mate-matching algorithm that works in continuous-time. Relying on the results of our literature review, we opt for a closed model that embodies a 'sequential' stochastic mate-matching procedure.

1. Compute the compatibility measures $C(w_i, m_j)$ for all possible combinations of women $w_1, \cdots, w_m$ and men $m_1, \cdots, m_n$ in the partnership market. Put them into a $m \times n$ matrix $D = d_{i,j}$. The rows of D refer to the women in the market, and the columns to the men.
2. Initialize two variables in order to map the effective dimension of D: $p = m, q = n$. The number of rows is given by p and the number of columns by q.
3. Calculate the total sum over all compatibility measures: $C_s = \sum_{i=1}^{p} \sum_{j=1}^{q} d_{i,j}$.
4. Build a matrix $C = c_{i,j}$ containing the row-wise cumulated compatibility measures:

$$\begin{pmatrix} c_{1,1} & \cdots & c_{1,q} \\ c_{2,1} & \cdots & c_{2,q} \\ \vdots & \ddots & \vdots \\ c_{p,1} & \cdots & c_{p,q} \end{pmatrix} = \begin{pmatrix} C(w_1, m_1) & \cdots & \sum_{j=1}^{q} C(w_1, m_j) \\ \sum_{j=1}^{q} C(w_1, m_j) + C(w_2, m_1) & \cdots & \sum_{i=1}^{2} \sum_{j=1}^{q} C(w_i, m_j) \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{p-1} \sum_{j=1}^{q} C(w_i, m_j) + C(w_p, m_1) & \cdots & C_s \end{pmatrix}$$

5. Draw a random number R uniformly distributed between 0 and C_s.
6. Match (w_1, m_1), if $0 < R \leq c_{1,1}$
   (w_i, m_1), if $c_{i-1,q} \leq R < c_{i,1}$
   (w_i, m_j), if $j > 1, c_{i,j-1} \leq R < c_{i,j}$.
7. Remove every row and every column from D that refers to either the woman or the man of the newly created couple. Consequently, the number of both rows and columns of D is decremented by 1: $p = p - 1, q = q - 1$.

Repeat steps 3 to 7 until $\min(n, m)$ couples are formed.

Figure 4 - Mixed-dominant concurrent mate-matching algorithm

3. MATE-MATCHING IN CONTINUOUS-TIME

In the previous section we have extensively discussed advantages and disadvantages of existing mating models. We deem a closed mating model to be preferable to an open model. As mentioned above, closed models have only been embedded into discrete-time microsimulation models. In a continuous-time microsimulation, individual life-courses are described as sequences of events (Gampe and Zinn 2007, Zinn et al. 2009). The occurrence of events is not determined at fixed points in time, but continuously. Consequently, the probability of concurrent events is practically zero. Individuals will never experience partnership events at the same time. Therefore, annual or monthly partnership markets are not suitable.

In a closed model, couples are constructed between existing individuals. For consistency reasons a partnership between individuals has to have a clearly defined formation time. We illustrate the problem using an example. A woman I_1 experiences the onset of a partnership at time t_1 and a man I_2 at time t_2. Without loss of generality, we assume $t_2 < t_1$. An intuitive way to compute a formation time $\tilde{t}$ of a partnership between I_1 and I_2 is $\tilde{t} = t_2 + c\,(t_1 - t_2)$, $c \in [0,1]$. Then, instead of t_1 and t_2, for both I_1 and I_2 the adjusted $\tilde{t}$ is used as starting time of a partnership. Changing simulated event times in this way, affects the outcome of the microsimulation model. In order to avoid significantly biased outcome, we have to assure that $t_1 - t_2$ is small. We define accordingly that I_1 and I_2 can only be regarded as potential spouses, if $t_1 \sim t_2$, where

$$t_1 \sim t_2 : \ t_1 \in \Gamma_2 \text{ and } t_2 \in \Gamma_1,$$
$$\Gamma_1 = [\, min(t_S, t_1 - B),\ max(t_1 + B, t_E)\,] \quad and \quad \Gamma_2 = [\, min(t_S, t_2 - B),\ max(t_2 + B, t_E)\,]$$

t_S is the simulation start time, and t_E the simulation stop time. B is an arbitrary time period that is shorter than one year. We call Γ_i the 'searching period' of I_i, $i = 1, 2$. With respect to this definition, only individuals can date if their searching periods overlap. In the subsequent, we use $B = 0.5$ *years* and $c = 0.5$. The latter results in $\tilde{t} = t_2 + 0.5(t_1 - t_2)$.

Even if the searching periods of mating willing individuals overlap, their characteristics might not match. A bachelor could be more than twenty years older than a bachelorette. Therefore, besides event times, also individual characteristics have to be checked for conformance. For this purpose, we use a compatibility measure like it has been introduced in section 2.3.1. We employ logit models to evaluate how well the characteristics of potential spouses fit together.

As the microsimulation model that we use is a generic model, the state space is not fixed. Only individual age and gender are mandatory. Depending on the problem to be studied, other relevant demographic states are considered. We can only include covariates in the logit models that the state space comprises. For example, the state space contains the state variables educational attainment, marital status, and fertility. Consequently, we can only include these variables into the logit model. An additional inclusion of, for example, ethnicity would be meaningless.

In order to simulate life-course events, we generate waiting times to next events (Gampe and Zinn 2007, Zinn et al. 2009). For example, at simulation start time t_S a woman I_1 is a_0 years old. She has never been married and is childless at this time. Conditioned on her current state, her age and the current calendar time, we simulate a waiting time of $w = 3.6$ *years* to a marriage event. Due to this simulation technique, we already know in advance when individuals will experience partnership events. We employ a partnership market to collect 'mating-minded' individuals. As soon as a partnership event has been simulated, an individual becomes member in this market. He or she leaves the partnership market either after he/ she has found a proper spouse or his/ her searching period is expired. In contrast to partnership markets of discrete-time models, individuals can enter and leave the market over the complete simulation time range.

We implement the partnership market using a so called marriage queue M. The marriage queue consists of all unpaired individuals who want to date someone. Each individual in the queue is equipped with a stamp that indicates the time of the upcoming partnership event. The woman I_1 of the example enters the market at time t_S and her waiting time to marriage is *3.6 years*. Her searching period, therefore, is $\Gamma_1 = [t_S + 3.1 \text{ years}, t_S + 4.1 \text{ years}]$. An appropriate spouse for I_1 has to exhibit a searching period that overlaps Γ_1. Furthermore, the joint characteristics of I_1 and of a potential spouse have to resemble joint characteristics of actual couples.

In our mate-matching procedure we take into account that individuals have cognitive constraints regarding their social network size. Because of the size of their neocortex, humans are restricted to social networks with approximately *150* members (Hill and Dunbar 2003). Considering this fact, for each 'mating-minded' individual we restrict the number of potential spouses. We set an upper bound that follows a normal distribution with expectation $\mu = 120$ and standard deviation $\sigma = 30$. Furthermore, we assign to each individual a random value that captures his/ her aspiration level regarding a partner. If the compatibility measure between an individual and a potential spouse exceeds the aspiration level, he/ she accepts the pairing. We state that the aspiration level follows a beta distribution. Relying on the theory of initial parental investment, women are choosier than men concerning their partners (Trivers 1972). We parameterize the beta distribution for females and males accordingly.

In order to construct synthetic couples in continuous-time, we use a modified version of the sequential stochastic mate-matching procedure that we have introduced in paragraph 2.3.3. If an individual I_i experiences the onset of a partnership, the processing described in Figure 5 is triggered.

1. Determine the searching period Γ_i of I_i.

2. Generate I_i's level of aspiration: a_i.

3. If the marriage queue M is empty, insert I_i.

Otherwise:

- 3.1 Draw random number N, normally distributed with expectation $\mu = 120$ and standard deviation $\sigma = 30$.
- 3.2 If N is greater than the number $\bar{N}$ of individuals in the marriage queue, assign $N = \bar{N}$.
- 3.3 Of M, take randomly N individuals whose searching periods overlap Γ_i. Insert them into the so called working marriage queue W.
- 3.4 Remove those from W who are of the same sex like I_i.
- 3.5 Remove those from W who do not meet some minimal criteria.
- 3.6 If W is empty, insert I_i into M.

  Otherwise:

  - (i) Initialize $j = 1$.
  - (ii) Take the j th individual I_j of W. I_j has aspiration level a_j.
  - (iii) Compute the compatibility measure $c_{ij} = C(w_i, m_j)$ or $c_{ij} = C(w_j, m_i)$, respectively, between I_i and I_j.
  - (iv) If $a_i < c_{ij}$ and $a_j < c_{ij}$, the individuals I_i and I_j get paired. Remove I_j from M.
  - (v) Otherwise, reduce the aspiration level of I_j: $a_j = \max(0, a_j - 0.1)$, and increment j by 1.

  Repeat steps (ii) to (v) until either I_i is paired or all individuals of W have been inspected.

If no appropriate spouse can be found for I_i, he/ she is enqueued into M.

Figure 5 - Mate-matching algorithm for continuous-time microsimulation models

In the description of the mate-matching algorithm we apply SOCSIM terminology. Both terms 'marriage queue' and 'working marriage queue' have been introduced by Hammel et al. (1990). In step 3.5, the minimal criteria are: no incest, no remarriage of previously divorced couples, and no extreme age differences between the spouses.

The presented mate-matching algorithm does not assure that each searching individual will be paired. Mate-matching malfunctions, if a seeker is not able to find within his/ her searching period a spouse with compatible characteristics. In order to be successful, each seeker has to have access to a pool of potential spouses. This can only be assured, if the model population maps a large proportion of an actual population.

Notwithstanding, if the searching period of a 'mating-minded' individual is expired, three options exist:

A. Extend the searching period. The individual remains in the marriage queue.

B. Send the individual unpaired back to model population. The individual is removed from the marriage queue. He/ she is again at risk to experience a partnership event.

C. Let a proper spouse immigrate or let the individual emigrate. The individual is removed from the marriage queue.

The last idea is borrowed from open models. An appropriate spouse is created 'ex nihilo'. Each of these options entails a major difficulty. Extending the searching period means shifting the time of the partnership event. Rejecting a seeker implies ignoring an already scheduled event. Allowing too many immigrated spouses, spoils the representativeness of the model population. Consequently, in order to assure plausible outcomes, searching periods that have expired without success have to be an exception.

4. MATE-MATCHING IN PRACTICE

The developed mate-matching algorithm has been implemented in the MicMac microsimulation tool (Zinn et al. 2009, http://www.nidi.knaw.nl/en/micmac/). In order to illustrate its capability we run simulations to forecast a synthetic population based on the population of the Netherlands. The state space that we employ for this purpose consists of the following state variables (corresponding values are given after the colons):

- gender: female, male
- marital status: living at parental home, first married never lived in a union before, first married cohabiting before, remarried, living alone never lived in a union before, living alone cohabiting before, living alone married before, first cohabitation never lived in a union before, higher order cohabitation never married before, cohabitation married before
- fertility: childless, one child, two children, three and more children
- educational attainment: primary, lower secondary and upper secondary plus tertiary education

We run simulations over 17 years, starting from January 1, 2004 up to December 31, 2020. During simulation, we only focus on individuals aged between 0 and 63. The initial population consists of 80,459 males and 82,121 females (which corresponds to 1% of the real Dutch population at January 1, 2004). During simulation individuals can experience the following events: giving birth, leaving parental home, launching a cohabitation, marrying, getting divorced and separated, raising their educational level, and dying. In order to assure that each 'mating-minded' individual is matched, we apply option A that is described above.

The initial population and transition rates have been estimated using different European data sources. We have applied the EUROPOP 2004 projections for the Netherlands (baseline scenario)[6] provided by EuroStat. This data set comprises for the years 2004 to 2050 information on Dutch mortality and fertility. We further have used the Fertility and Family Survey for the Netherlands (FFS_NL) conducted between February and May 2003. This survey contains micro information on fertility behavior and changes in the marital status. Data on educational attainment has been taken from Goujon (2008).

We have constructed the initial population using the method of iterative proportional fitting (ITF). In order to estimate fertility rates and transition rates regarding marital status we have employed a slightly modified version of MAPLE (Impicciatore and Billari 2007). Mortality rates have been taken from the EUROPOP 2004 projections.

The proposed mate-matching procedure requires the computation of compatibility measures between potential spouses. For this purpose we use logit models. We have estimated the effects of the age of spouses, their educational attainment, and their marital history before pairing, as well as the number of children with former partners. For estimating the models, we have used the first wave of the Netherlands Kinship Panel Study that has been conducted in the period from 2002 to 2004 (Dykstra et al. 2005). Only partnerships have been included that started in the years from 1990 to 2002.

We perform 10 simulation runs. During simulation all demographic events are tracked. Because of unsuccessful search about 10 percent of partnership onsets have to be shifted. In order to evaluate the suitability of the proposed mate-matching algorithm we have analyzed the distribution of joint-spouses characteristics, with special emphasis on differences in educational attainment and age. We compare the differences in the educational level of synthetic couples to those of couples given in the NKPS (in the range from 1990 to 2002).

Figure 6 contrasts simulated and actual data concerning the educational level of the spouses of married males with lower secondary education (graph on left hand side) and cohabitating males with upper secondary and tertiary education (graph on right hand side). The graphs depict the distribution of the educational level of the female spouses. The simulation produces on average 11 (24, 65) percent marriages in which the male has a lower secondary education and the female a primary (lower secondary, upper secondary/ tertiary) education. This is contrasted by 15 (25, 60) percent of comparable marriages in the data. In 8 (22, 70) percent of the synthetic cohabitations highly educated males are paired with females who hold a primary (lower secondary, upper secondary/ tertiary) education. By contrast, the NKSP comprises 8 (7, 85) percent of such couples. Consequently, compared to the observed numbers, in the simulation highly educated males are more often cohabitating with females who hold a middle education; and less often with females who hold a similar educational level. This discrepancy is caused by the fact that in the simulation the individual aspiration level decreased with age: As a consequence the older an individual is at his/her partnership onset the higher is the probability of a non-assortative match. In general, the simulation satisfactorily captures the overall pattern of differences in the educational level of spouses.

[6] cp. http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/, accessed March 2010

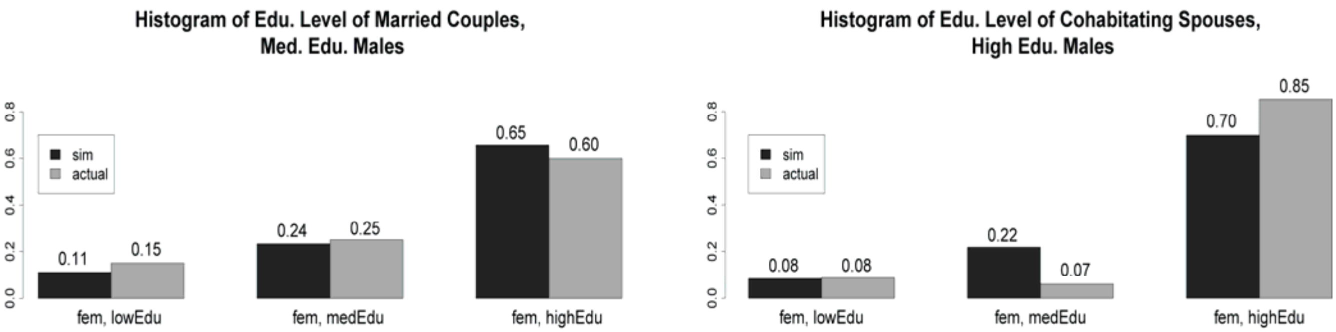

Figure 6 - Differences in the educational level of spouses in observed and simulated couples (low Edu.: primary education, med. Edu.: lower secondary, high Edu.: upper secondary and tertiary education)

Figure 7 depicts the distribution of age differences of cohabitating and married spouses (age of male minus age of female). The shapes of the respective simulated and actual frequency distributions are nearly identical. Discrete mate-matching algorithms generally produce age difference distributions that are too flat (Leblanc et al. 2009). They are not capable to reproduce the observed peak at differences of [-1,1]. The proposed mate-matching algorithm is able to overcome this problem. It only produces slightly more cohabiting couples in the case of cohabitations where the female is much older than the male. A reason for this discrepancy could be the small number of corresponding cases in both the actual and the simulated data.

5. DISCUSSION AND OUTLOOK

After an extensive literature review, we have proposed a stochastic mate-matching algorithm for continuous-time microsimulations. We have demonstrated its capability using data on fertility and marriage behavior of the contemporary Netherlands. Population forecasts have been conducted over 17 years, from 2004 to 2020. We found that the proposed algorithm produces acceptable result. It reproduces the observed peak of the frequency distribution of age differences of spouses. A problem arises due to the shifts of event times that are performed if mate seekers are not successful in time. Such shifts provoke distorted simulation outcome. A solution to the problem might be the usage of a combination of the options A, B, and C listed in section 3. Moreover, preferences concerning the characteristics of a partner can be subject to secular trends, and such changes over time could be included in the mate-matching process. Currently, we only use compatibility measures that are based on actual preference pattern. Generally, pairing individuals raises problems concerning the modelling and simulation of linked lives. For example, if an individual experiences the onset of a dissolution event, what happens to his/her spouse? A general model to establish and simulate linked lives in continuous-time is currently under examination.

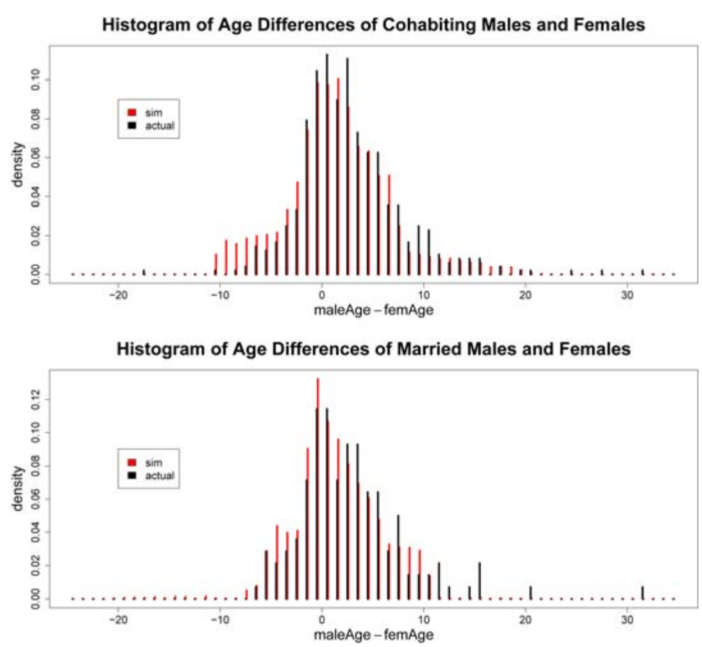

Figure 7 - Age differences of spouses in observed and simulated couples

Acknowledgements

This work is entirely part of the MicMac project. The Centraal Bureau voor Statistiek in the Netherlands has provided the data of the Fertility and Family Survey for the Netherlands (FFS NL). The Netherland Kinship Panel Study is funded by grant 480-10-009 from Major Investments Fund of the Netherlands Organization for Scientific Research (NWO), and by the Netherlands Interdisciplinary Demographic Institute (NIDI), Utrecht University, the University of Amsterdam and Tilburg University. This paper would never have been written without the help of a couple of persons. Particular thanks go to Jutta Gampe for her fruitful supervision. Ekaterina Ogurtsova assisted in preparating the input data files for the microsimulation example. Jim Oeppen helped with general advices. Basic ideas of mate-matching in continuous time go back to discussions with Jan Himmelspach and Frans Willekens

REFERENCES

Bacon, B., and S. Pennec. 2007, November. APPSIM - modelling family formation and dissolution. Working Paper 4, University of Canberra, National Centre for Social and Economic Modelling.

Bouffard, N., R. Easther, T. Johnson, R. Morrison, and J. Vink. 2001. Matchmaker, matchmaker, make me a match. Brazilian Electronic Journal of Economics 4 (2).

Caldwell, S. B. 1996. Corsim 3.0: Technical and user documentation. Unpublished manuscript, Cornell University, Department of Sociology.

Caldwell, S. B., M. Favreault, A. Gantman, J. Gokhale, T. Johnson, and L. J. Kotlikoff. 1999. Social security's treatment of postwar Americans. Tax Policy and the Economy 13:109–148.

Cumpston, R. 2009. Acceleration, alignment and matching in multi-purpose household microsimulations. Paper presented at the 2nd general conference of the International Microsimulation Association, Ottawa, Australian National University, South Melbourne.

de Vos, S., and A. Palloni. 1989. Formal models and methods for the analysis of kinship and household organisation. Population Index 55 (2): 174–198.

Duee, M. 2005. La modélisation des comportements démographiques dans le modèle de microsimulation DESTINIE. Technical report, INSEE, Institute National de la Statistique et des Etudes Economiques, Paris.

Dykstra, P. A., M. Kalmijn, T. C. Knijn, A. C. L. Aafke E. Komter, and C. H. Mulder. 2005. Codebook of the Netherlands kinship panel study, a multi-actor, multi-method panel study on solidarity in family relationships, Wave 1. NKPS Working Paper 4, NIDI, The Hague.

Gale, D., and L. S. Shapley. 1962. College admissions and the stability of marriage. The American Mathematical Monthly 69(1): 9–15.

Galler, H. 1997. Discrete-time and continuous-time approaches to dynamic microsimulation (reconsidered). Technical report, NATSEM - National Centre for Social and Economic Modelling, Faculty of Management, University of Canberra.

Gampe, J., and S. Zinn. 2007. Description of the microsimulation model. Technical report, MicMac project, MPIDR, Rostock.

Goujon, A. 2008. Report on projections by the level of education (Future human capital: Estimates and projections of education transition probabilities). Technical report, MicMac project, Vienna Institute of Demography, Austrian Academy of Sciences.

Hammel, E., D. Hutchinson, K. Wachter, R. Lundy, and R. Deutel. 1976. The SOCSIM demographic-sociological microsimulation program: Operating manual. Institute of International Studies, University of California, Berkeley. Research Series, NO. 27.

Hammel, E., C. Mason, and C. Wachter. 1990. SOCSIM II: A sociodemographic microsimulation program rev. 1.0. operating manual. Regents of the University of California.

Hill, R. A., and R. I. M. Dunbar. 2003. Social network size in humans. Human Nature 14 (1): 53–72.

Holm, E., K. Holme, K. Mkil, M. Mattsson-Kauppi, and G. Mrtvik. 2002. The SVERIGE spatial microsimulation model: content, validation, and example applications. Work report GERUM 2002:4, Kulturgeografiska institutionen/ SMC.

Holmer, M., A. Janney, and B. Cohen. 2009. PENSIM overview. Policy Simulation Group and the U.S. Department of Labour.

Huinink, J., and M. Feldhaus. 2009. Family research from the life course perspective. International Sociology 24 (299): 299–324.

Imhoff, E., and W. Post. 1998. Microsimulation methods for population projection. Population: An English Selection, New Methodological Approaches in Social Sciences 10 (1): 97–138.

Impicciatore, R., and F. C. Billari. 2007. Age profiles estimation for family and fertility events based on micro data: The MAPLE (method for age profile longitudinal estimation). In Work Sessions on Demographic Projections.

Kelly, S. 2003, December. Australia's microsimulation model - DYNAMOD. Technical report, National Centre for Social and Economic Modelling, University of Canberra.

King, A., H. Baekgaard, and M. Robinson. 1999. Dynamod-2: An overview. Technical report, NATSEM, National Center for Social and Economic Modelling, University of Canberra.

Leblanc, N., R. Morrison, and H. Redway. 2009. A match made in silicon: Marriage matching algorithms for dynamic microsimulation. Paper presented at the 2nd general conference of the International Microsimulation Association, DYNACAN project, Ottawa.

Murphy, M., and D. Wang. 2001. Family-level continuities in childbearing in low-fertility societies. European Journal of Population 17:75–96.

O'Donoghue, C. 1999. Dynamic microsimulation: A methodological survey. Brazilian Electronic Journal of Economics 4(2).

O'Harra, J., and J. Sabelhaus. 2002, October. Projecting longitudinal marriage patterns for long-run policy analysis. Technical Paper 2002-2, Congressional Budget Office Washington, DC.

Orcutt, G. 1957. A new type of socio-economic system. Review of Economics and Statistics 39 (2): 116–123.

Orcutt, G., M. Greenberger, J. Korbel, and A. Rivlin. 1961. Microanalysis of socioeconomic systems: A simulation study. New York: Harper & Row.

Perese, K. 2002, November. Mate matching for microsimulation models. Technical Paper 3, Long-Term Modeling Group, Congressional Budget Office, Washington, DC.

Satyabudhi, B., and S. Onggo. 2008. Parallel discrete-event simulation of population dynamics. In Proceedings of the 2008 Winter Simulation Conference, ed. S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, and J. W. Fowler, 1047–1054: Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Smith, J. 1987. Family demography: Methods and their application, Chapter The computer simulation of kin sets and kin counts, 249–266. Clarendon Press, Oxford.

Spielauer, M. 2002. The potential dynamic of microsimulation in family studies: A review and some lessons for FAMSIM+. Working Paper 18, Austrian Institute for Family Studies, Austria.

Spielauer, M., and N. Neuwirth. 2001. Family microsimulation. Brazilian Electronic Journal of Economics 4 (2). Statistics Canada 2004. The lifepath microsimulation model: An overview.

Trivers, R. 1972. Parental investment and sexual selection, Chapter Sexual selection and the descent of man, 1871-1971, 136–179. Aldine, Chicago.

Wachter, K. 1995, April. 2030's seniors: Kin and step-kin. Technical report, University of California, Berkeley.

Wachter, K., D. Blackwell, and E. Hammel. 1998, February. Testing the validity of kinship microsimulation: An update. Technical report, University of Carlifornia, Berkeley.

Willekens, F. 2009. New frontiers in microsimulation modelling, Chapter Continuous-time microsimulation in longitudinal analysis, 353–376. Ashgate, Farnham, Surrey, UK. (Book released during 2nd International Microsimulation Conference, Ottawa).

Wolf, D. 2001. The role of microsimulation in longitudinal data analysis. Special Issue on Longitudinal Methodology, Canadian Studies in Population 28 (2): 313–339.

Zaidi, A., and K. Rake. 2001. Dynamic microsimulation models: A review and some lessons for SAGE.

Zinn, S., J. Himmelspach, J. Gampe, and A. M. Uhrmacher. 2009. Mic-core: A tool for microsimulation. In Proceedings of the 2009 Winter Simulation Conference, ed. M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls. Piscataway, NJ: IEEE.