

Assortative Mating and the Reversal of Gender Inequality in Education in Europe – An Agent-Based Model*

André Grow

Jan Van Bavel

Yolien De Hauw

University of Leuven, Centre for Sociological Research

Prepared for “European Population Conference 2014“, 25-28 June 2014, Budapest

Abstract

In Europe, higher education was mostly a male domain until the 1970s. Since then, the gender gap has decreased and by now, women largely excel men in terms of educational attainment. Education is known to affect various forms of reproductive behaviour, such as union formation and union dissolution. However, to date we know little about how the reversal of gender inequality in education (from here on RGE) might have affected these behaviours. With this paper, we provide one of the first steps towards filling this lacuna. We develop an agent-based computational model that enables us to study the mechanisms that link RGE to patterns of assortative mating across European countries.

Our model builds on the notion that mate search is an adaptive process. In this view, individuals have aspirations for partners with certain characteristics. These aspirations develop and change in response to experiences on the marriage market. For instance, individuals who fail to find a partner of similar age with a desired tertiary degree might lower their aspirations and might become willing to accept partners with lower educational degrees or partners who are considerably older. We argue that this process, in combination with individuals’ preferences for the education, age, and earnings potential of prospective partners, has created complex dynamics that link RGE with patterns of assortative mating on each of the three dimensions. Our model enables us to assess this proposition with analytical rigor. It also enables us to assess the relative impact that each preference might have had on observed patterns of assortative mating.

Keywords

Assortative Mating, Gender Inequality, Agent-Based Modelling

*Please direct correspondence to: André Grow, Centre for Sociological Research / Family & Population Studies (FaPOS), Faculty of Social Sciences, University of Leuven, Parkstraat 45 bus 3601, 3000 Leuven, Belgium

e-mail: Andre.Grow@soc.kuleuven.be

phone: +32 16 323256

The research leading to these results has received funding from the European Research Council under the ERC Starting Grant “Implications of the Shifting Gender Balance in Education for Reproductive Behaviour in Europe” (Principal investigator: Jan Van Bavel)

ASSORTATIVE MATING AND THE REVERSAL OF GENDER INEQUALITY IN EDUCATION IN EUROPE – AN AGENT BASED MODEL

Over the last decades, Europe has witnessed dramatic changes in the relative educational attainment of men and women. Until the 1970s, higher education was mostly a male domain. Since then, the gender gap has decreased and by now, women largely excel men in terms of participation and success in higher education (Schofer & Meyer, 2005). Education is known to affect various forms of reproductive behaviour, such as union formation and union dissolution (Lutz, Cuaresma, & Sanderson, 2008). However, to date we know little about how the reversal of gender inequality in education (from here on RGE) might have affected these behaviours (Van Bavel, 2012). In this paper, we provide one of the first steps towards filling this lacuna. We develop an agent-based computational model that enables us to study the mechanisms that link RGE to patterns of assortative mating across European countries.

Assortative mating is the “nonrandom matching of individuals into relationships” (Schwartz, 2013: 452) and can be based on various individual characteristics. Esteve et al. (2012) recently provided first insights into the effects that RGE might have had on *educational assortative mating*. Education is a particularly relevant characteristic, because it strongly affects individuals’ access to valuable resources. A non-random matching of partners based on education therefore can lead to an unequal distribution of resources across couples and thereby can increase social inequality (Blossfeld, 2009). In 20th-century Europe, the dominant pattern of assortative mating has been educational homogamy combined with female hypergamy and male hypogamy. That is, women have tended to marry men who were at least as highly educated as themselves, whereas men have tended to marry women who were at most as highly educated as themselves (Van Bavel, 2012). Esteve et al.’s results suggest that RGE might have changed this pattern by shifting the relative prevalence of

female hypergamy and hypogamy. That is, Esteve et al. showed that in countries in which the gender gap has become lower or has changed to the advantage of women, female hypogamy has increased relatively to female hypergamy.

The findings of Esteve et al. (2012) are instructive and illustrate the important implications that RGE might have had for patterns of assortative mating in Europe. However, the correlational nature of the evidence leaves open what causal mechanisms might underlie the observed association. To gain insights into these mechanisms, we follow the lead of earlier research on human mating behaviour and make use of agent-based computational (ABC) modelling (Billari, Ongaro, & Prskawetz, 2003; Todd, Hills, & Hendrickson, 2013). ABC modelling allows us to test conceptual frameworks of how population-level mating patterns might have emerged from the behaviour of the individuals that make up these populations. In particular, ABC modelling enables us to ‘turn on and off’ certain aspects of human mate search and thereby allows us to assess how certain preference structures relate to macro level patterns of assortative mating (cf. Epstein & Axtell, 1996).

In developing our model, we draw on recent work by Van Bavel (2012), who suggested that patterns of assortative mating might be the result of individual partner preference structures that encompass multiple characteristics, such as education and age. In this view, the interplay between these preferences might affect multiple dimensions of assortative mating. To illustrate such interplay, consider the following thought experiment. Due to RGE, highly educated women born in recent cohorts experience a lack of similarly educated men within their own age groups. If women would prefer a similarly educated partner but were indifferent about his age, they might try to satisfy this preference by looking for a mate in older cohorts, in which highly educated men are relatively abundant. If a substantive number of women were able to find a partner in these cohorts, RIGE might have little effect on educational assortative mating, but might lead to an increase in age heterogamy to the

disadvantage of women. However, it is known that often there are limits to the age difference that women are willing to accept in their partners (e.g., England & McClintock, 2009). This can create conflicting pressures for women, given that they somehow need to trade-off preferences for education and age during partner search. Contingent on how women deal with this trade-off, RGE might have affected both patterns of educational assortative mating and patterns of age homogamy.

In line with Van Bavel's (2012) reasoning, our model centres on the notion that mate search is an adaptive process (Simão & Todd, 2002). In this view, individuals have aspirations for partners with certain characteristics. These aspirations develop and change in response to experiences during the search process. In developing the model, we partly draw on existing models that also centre on the notions that human mate selection is an adaptive process (Simão & Todd, 2002, 2003; Todd, Billari, & Simão, 2005; Todd & Billari, 2003; Todd & Miller, 1999) and that humans select their partners based on multiple characteristics (French & Kus, 2008). Our focus is on the characteristics education, age, and earnings prospects. We selected these dimensions because they have been shown to be important selection criteria during partner search (Boxer, Noonan, & Whelan, 2013; e.g., Buss & Barnes, 1986; Buss, Shackelford, Kirkpatrick, & Larsen, 2001; Skopek, Schulz, & Blossfeld, 2010), and because assortative mating on these dimensions has received much attention in the sociological literature (e.g., Kalmijn & Flap, 2001; Kalmijn, 1998; Mare, 1991; Schwartz, 2010).

In what follows, we first discuss the theoretical background of our model and describe its main elements. Next we present the outcomes of systematic computational experiments. We close with a discussion and outlook for future research.

BACKGROUND

Our model is based both on insights from empirical research and on insights from earlier modelling work. First, a large number of studies suggest that individuals tend to evaluate potential partners on multiple dimensions and the differences in the import that men and women tend to attach to some of these dimensions is comparatively stable across time and across cultural contexts (e.g., Buss et al., 1990, 2001). Given a set of preferences, both men and women often require prospective partners to meet minimal standards on the characteristics they perceive as particularly important. If prospective partners fail to meet some of these aspirations, even outstanding qualities on other characteristics often cannot compensate for these shortcomings. Women, for example, have been found to attach import to the earnings prospects of their partners and require that potential suitors meet some minimal standard on this characteristic; only when this standard is met, they also consider dimensions for which they have no such minimal standards (cf., Li & Kenrick, 2006; Li, 2007; Li et al., 2013; Li, Bailey, Kenrick, & Linsenmeier, 2002). Based on these insights, we assume that agents have preferences for partners with certain qualities on several important characteristics (i.e. education, age, and earnings prospects), and that outstanding qualities of a prospective partner on some of these characteristics cannot make up for shortcomings on other characteristics.

Second, our selection of the characteristics education, age, and earnings prospects is based on the insight that these characteristics typically rank high in importance both in the responses to survey items by individuals and in the results of studies that try to uncover preferences from observed mating behaviour (e.g., Boxer et al., 2013; Skopek et al., 2010); furthermore, they have been theorized to play an important role in the changes that RGE might induce (Van Bavel, 2012). Based on earlier research, we assume that both male and female agents prefer partners with similar educational backgrounds and prefer partners with

high earnings prospect (e.g., Buss et al., 1990, 2001; Skopek et al., 2010). By contrast, we assume that male agents prefer young partners (regardless of their own age), whereas female agents prefer partners who are somewhat older than themselves (e.g., England & McClintock, 2009). We do not model other characteristics that might affect partner selection; for simplicity, we assume that these characteristics are not systematically associated with the characteristics that we model and thus can be treated as noise in the modeling process.

Third, earlier modelling work has identified several behavioural principles that together can generate key characteristics of macro level mating patterns. Our model incorporates some of these principles to increase the match between the dynamics that our model generates and the dynamics that occur in the real world. More specifically, we assume that agents engage in courtship periods in which they date potential partners while looking for better alternatives (e.g., Simão & Todd, 2003). During this period, they can get to know each other better, while minimizing the opportunity costs from bonding permanently with somebody who might not be a perfect fit. If no better alternative comes along, they form a permanent bond (i.e. they marry) with their current partner. Furthermore, we assume that agents experience subjective pressure to find a permanent partner, and this pressure increases over their life course (cf. French & Kus, 2008). More specifically, pressure increases with age and this increase is stronger for female than for male agents. The stronger pressure becomes, the less selective individuals become in choosing a partner and the more willing they become to accept partners who do not meet their all their aspirations.

The outcome that we focus on is assortative mating in terms of education, age, and earnings prospects in married couples in which the female partner is married for the first time. We aim at studying the link between this outcome and the relative educational attainment of men and women, as well as the structures of earnings potentials among men and women over time.

MODEL

Agents and Their Characteristics

Our model simulates the process of union formation in discrete time steps and all time related elements are expressed in these steps; ten time steps represent one year. The model starts with a population of M male and F female individuals, which are represented by agents i . Each agent can be described by its gender g_i (male or female), its age a_i (measured in time steps), the highest educational level that it will ever attain s_i ($s_i \in \{0, 1, 2, 3\}$), its earnings prospects y_i ($y_i \in \{1, 2, \dots, 10\}$), its relationship status r_i (single, dating, or married), and the number of time steps it is already in a relation with its current partner c_i ($c_i > 0$ if i is dating or married; $c_i = 0$ if i is single).¹ The implementation of agents' highest educational level and earnings prospects is aligned with the data that we use to initialize agent cohorts (see the section 'Creating Agent Cohorts' for details).

Agents enter the model with $a_i = 0$, but only start dating (i.e. enter the marriage market) after 160 time steps ($A_{\text{dating}} = 160$). This corresponds to an age of 16 years in real life. Until this age, agents are considered adolescents and are not viewed as potential partners by others. They are thus excluded from the dating process described below. From the moment they reach the age of $t_i = 160$, they are considered adults and take part in the dating processes.

Partner Preferences

Agents try to find permanent partners who have a high subjective mate value v_{ij} , which is determined by preferences for the education, age, and earnings prospects of partners. We assume that both male and female agents prefer partners who have a similar educational

¹ We use the letter s for representing education (derived from the word 'schooling') instead of e to avoid confusion with earnings potential, and with the exponential function that in math is represented by this letter and to avoid confusion with the characteristic 'earnings prospects'. For the same reason, we use the letter Y for representing agent's earnings potentials.

background and high earnings prospects. By contrast, male agents prefer young partners regardless of their own age, whereas female agents prefer partners who are somewhat older than themselves. We implement these gender specific preferences by

$$v_{ij} = \begin{cases} \left(\frac{S_{\max} - |s_i - s_j|}{S_{\max}} \right)^{w_s^m} \left(\frac{y_i}{Y_{\max}} \right)^{w_y^m} \left(\frac{A_{\max} - a_j}{A_{\max} - A_{\text{dating}}} \right)^{w_a^m} & \text{if } g_i = \text{male} \\ \left(\frac{S_{\max} - |s_i - s_j|}{S_{\max}} \right)^{w_s^f} \left(\frac{y_i}{Y_{\max}} \right)^{w_y^f} \left(\frac{A_{\max} - |a_i - a_j + \sigma|}{A_{\max}} \right)^{w_a^f} & \text{if } g_i = \text{female} \end{cases} . \quad (1)$$

In Eq. (1), S_{\max} , Y_{\max} , and A_{\max} are the maximal education, earnings prospects, and age that agents can attain. The parameter σ governs for female agents how much older an ideal partner should be. The parameters w_s , w_y , and w_a govern the importance that male and female agents (indicated by superscript m and f) attach to each of these characteristics. For both male and female agents, when $w_s > 0$, the more similar a potential partner j is in terms of educational background, the higher i perceives j 's mate value. Furthermore, when $w_y > 0$, the higher the earnings prospects of a potential partner, the higher is this agent's mate value in the eyes of i . Finally, for male agents, when $w_a^m > 0$, the mate value of potential partner decreases with its age; for female agents, when $w_a^f > 0$, the mate value of potential partners is maximal when they are σ time steps older and decreases when they are younger or older than this ideal age.

Eq. (1) only contains preferences for characteristics that are central to our research interest. We assume for simplicity that other mating relevant characteristics are uncorrelated with these characteristics and can be treated as noise. We implement this assumption by modelling dating and mating decisions stochastically, as discussed in the following sections.

Meeting, Dating, and Mating

In each time step, adult agents can meet other adult agents (somebody new, or somebody they have already met in the past). The probability that this happens depends on

how long they are already in a relation with their current partner, if they have one. More formally, the probability that agent i meets another agent in the current time step ($P(i \text{ meet})$) is determined by

$$P(i \text{ meet}) = e^{-(c_i\beta)} . \quad (2)$$

In Eq. (2), β is an exogenous factor that determines the effect that the time that agent i is already in a relation with its current partner (c_i) has on the probability that i will meet somebody. From here on, we refer to β also as the ‘intimacy factor’. When $\beta > 0$, the longer a given agent is already in its current relationship (i.e. the larger c_i becomes), the less likely the agent becomes to meet others. This decrease accelerates with increasing values of β . Note that for single agents, c_i is always 0; the probability that they will meet somebody is therefore always 1.

If it has been determined that agent i will meet somebody in the current time step, one member j of the adult agent population of the opposite gender is randomly selected with a probability proportional to $P(j \text{ meet})$ over this set.²

Whenever two adult agents meet, both decide whether they want to start dating the respective other. Dating means here that they begin a serious relationship that might lead to marriage. Single agents consider any other agent as a potential partner. By contrast, agents who are currently in a relationship only consider those as potential partners whose subjective mate value is higher than the subjective mate value of their current partner (i.e. when $v_{ij}^{\text{alternative}} > v_{ij}^{\text{partner}}$). When both *are willing* to date, they become partners and enter a relationship (and leave possible current relationships). However, when at least one of them is *not willing* to date, they will not date and both remain single (or remain with their current

² In practice, the interaction partner was not selected from the entire adult population of the opposite gender. Instead, we drew a random sample of 50 members of the opposite sex, and then select the interaction partner proportional to $P(j \text{ meet})$ across this smaller set. This approach greatly reduced computation time.

partner if they have one). The probability that a given agent is willing to date a given potential partner j ($P(i \text{ date } j)$) is determined by

$$P(i \text{ date } j) = \left(1 - e^{-(a_i v_{ij} \kappa)}\right) e^{-(c_i \beta)} . \quad (3)$$

In Eq. (3), κ is an exogenous factor that governs the effect that the age of agent i (a_i) has on its willingness to start a relationship with somebody else. From here on, we refer to κ also as the ‘age pressure factor’. In general, the probability that agents are willing to date somebody else increases with their age (a_i) at a given value of κ (when $\kappa > 0$) and this effect is stronger for potential partners with higher subjective mate value (v_{ij}). Note again that for single agents, c_i is always equal to zero. For such agents, the second term of Eq. (3) is therefore always equal to 1. As a consequence, all that matters for their willingness to start dating somebody is this their own age (a_i) and the subjective mate value of the potential partner (v_{ij}), in combination with the age pressure factor (κ). By contrast, for agents who currently have a partner, the value of $P(i \text{ date } j)$ is attenuated by the time they are already in the relationship (c_i), in combination with the intimacy factor (β). This applies to both agents who are only dating their partner and agents who are married to their partner. This implies that divorces are possible and that agents can remarry.

The longer a given agent is already dating, the more willing it becomes to marry and thus to propose marriage to its current partner. From the moment in which agent i (or j) proposes marriage to its partner j (i), the proposal remains intact until j (i) also proposes to i (j), or until of them terminates the relation (e.g., due to death). If both are willing to marry each other (i.e. if both propose), they get married. We model the probability that agent i becomes willing to marry its current partner j ($P(i \text{ marry } j)$) by

$$P(i \text{ marry } j) = \left(1 - e^{-(a_i v_{ij} \kappa)}\right) \left(1 - e^{-(c_i \beta)}\right) . \quad (4)$$

Eq. (4) holds that older agents are generally more willing to marry their partners (i.e. they are more likely to propose marriage), and this willingness increases with the mate value of their partner (v_{ij}), with the length of their relationship (c_i), with the age pressure (κ), and with the intimacy factor (β).

Death and Reproduction

Agents leave the population either when they die or when they reach the maximal age of $A_{max} = 600$, which corresponds to the age of 60 years in real life. We remove agents at the age of 60, given that in reality only a very small share of individuals experience their first marriage above this age; they are also very unlikely to be the first marriage partners of others. Before this age, there is a certain probability that a given agent will die in the current time step, and this probability increases convexly with its age. More specifically, the probability of death in a given time step is determined by

$$P(i \text{ death}) = d \left(\frac{A_{max} - |a_i - A_{max}|}{A_{max}} \right)^{w_d}, \quad (5)$$

where d is a factor that determines the maximum probability of death at the age of 60 and w_d governs the shape of the function.

For each agent who dies or leaves the population because it has reached A_{max} , one new agent of the same gender is created that enters the simulation in the next simulation step with $a_i = 0$.

With this approach for modelling death and reproduction, we make the simplifying assumptions that population size remains stable over time, that there are no differences in mortality between men and women, and that the relative numbers of men and women at birth are balanced and do not become systematically imbalanced over time.

Creating Agent Cohorts

We initialize the distribution of educational degrees (s_i) and earnings prospects (y_i) of agent populations based on empirical data. For initializing agent cohorts in terms of s_i , we used data provided by the International Institute for Applied Systems Analysis/Vienna Institute of Demography (IIASA/VID; KC et al., 2010; Lutz, Goujon, KC, & Sanderson, 2007). The IIASA/VID provide reconstructions (from 1970 until 1995) and projections (from 2000 until 2050) of the distribution of educational attainment for a large number of countries.³ Educational attainment is recorded based on the international standard for coding educational degrees (ISCED), grouped into four categories.⁴ We used this information to directly initialize agent cohorts.

An example helps illustrating this procedure. Assume that the simulation process starts with a population born in the year 1941. According to the IIASA/VID data shown in Table 1, in 1975, of those men born between 1941 and 1945 in Belgium (i.e. of those who were 30-34 years old in 1975), about 22% had attained first or second stage tertiary education (ISCED codes 5 and 6). Five years later, this value had increased to about 23%. Based on this information, male agents who enter the simulation between 1941 and 1945 are assigned a s_i value of 3 with probability .22. Male agents who enter the simulation about 5 years later (i.e. about 50 simulation steps later) are assigned a s_i value of 3 with probability .23.

–Table 1 here–

³ We use the global education trend (GET) scenario for projections.

⁴ Coded as 0 = no education, 1 = ‘pre-primary education’ / ‘primary education or first stage of basic education’ (ISCED codes 0 and 1), 2 = ‘lower secondary’ / ‘second stage of basic education’ / ‘post-secondary non-tertiary education’ (ISCED codes 2, 3, and 4), 3 = ‘first stage of tertiary education’ / ‘second stage of tertiary education’ (ISCED codes 5 and 6).

For initializing agent cohorts in terms of y_i , we used data from the cross-sectional version of the European Union Statistics on Income and Living Conditions (EU-SILC).⁵ The EU-SILC is particularly attractive for our purposes, given that it is a representative household survey that provides information about respondents' (and their partners') education, annual income, and recent labour history. This enabled us to reconstruct differences in earnings prospects for those European countries that are included in the survey.

Based on earlier research on earnings trajectories and life-time earnings potentials (e.g., Björklund, 1993; Blomquist, 1981; Haider & Solon, 2006; Klevmarken, 1982; Rosen & Taubman, 1982), we focused on the annual net income in the year prior to the year of the survey of men and women in the age range 35 to 50 years who were not retired or self-employed in this period; the income of individuals who did not work was coded as 0. Our focus on the age range 35 to 50 years is based on the insight that income during this period tends to be a good predictor of individuals' life-time income. Based on this selection criterion, Table 2 provides an overview of the number of observations broken down by country, five-year cohort, gender, and educational level. We used two (i.e. tertiary vs. lower than tertiary) instead of four educational levels, to avoid that cell counts become too low. To increase reliability and prevent biases due to random fluctuations, in our analysis we excluded cells that had less than 100 observations and imputed the information as described below.

–Table 2 here–

We operationalized respondents' earnings prospects as their income relatively to the income of the top earnings category (i.e. the lower boundary of the top income decile in the observed data) of members of their own five-year birth cohort, expressed in ten categories (1 = earns up to 10% of the income of the top earnings category, 2 = earns 10% to 20%, ..., 10

⁵ Survey years 2004 to 2010

= earns more than 90%). Focusing on relative income has the advantage that we do not need to explicitly model changes in nominal income over time (Bosworth, Burtless, & Steuerle, 2000). Furthermore, we categorize respondents' income relatively to members of their own five-year birth cohort, rather than to all 35 to 50 year old respondents, given that current income is a good estimator of individuals' earnings trajectory relatively to individuals who are in the same phase of their life, but not relatively to older or younger individuals. The reason is that older individuals tend to earn more than younger individuals. Thus, even if two individuals are on exactly the same trajectory, the fact that their income is located at different stages of this trajectory might lead to the incorrect conclusion that they are on different trajectories.

Based on this operationalization, we first estimated separately by country, gender, and educational level the likelihood that respondents fell into one of the 10 earning prospects categories. We did this by means of a multinomial logistic regression model in which we included respondents' five-year birth cohort as a continuous variable to be able to capture trends over; we also included its squared value to capture potentially non-linear trends.⁶ Next, based on the estimates obtained in the multinomial logistic regression, we constructed hypothetical distributions of earnings prospects for cohorts that are not included in the EU-SILC data, but which we model in our simulations. We also used these estimates to construct distributions of earnings prospects for those cells in Table 1 that we had removed due to lower numbers of observations.⁷

Figure 1 provides an example of the outcome that the foregoing procedure generated. The figure suggests that during the period for which there is observed data, in Belgium there was a trend among highly educated women towards higher relative income (which we

⁶ In these variables, each cohort was represented by the middle of the respective five-year interval.

⁷ We applied weights both when calculating the top income decile and when estimating the regression model.

assume to indicate increasing earning potentials in this demographic groups over time), whereas there was a trend towards lower relative income among highly educated men. The multinomial regression model projects this trend into the future, for cohorts for which no observational data is available yet.

–Figure 1 here–

We used the information illustrated in Figure 1 for initializing agent cohorts similarly to the data on education. For example, assume that there is a female agent who enters the population in simulation year 1956 and who has been assigned an educational level of $s_i = 3$ (i.e. a tertiary degree). According to the data shown in Figure 1, in Belgium, about 9% of women with a tertiary degree born between 1956 and 1960 fell into the earnings prospect category 10. Based on this, there is a probability of .09 that the agent will be assigned the value $y_i = 10$ at the day of its birth. Note that the reduction to two educational categories in the estimation process means that all agents with $s_i < 3$ have the same probability to be assigned one of the ten categories of y_i

Scheduling of Simulation Steps

Each time step/simulation step consists of the following sub-steps:

- 1) Increase agents' age (a_i), and the relationship times (c_i) of those agents who currently are in a relation (either dating or married), by 1
- 2) Select agents randomly one at a time (without replacement) and
 1. Determine whether and whom they meet in the current time step⁸
 2. If a given agent meets another agent, determine for each of them whether they want to start dating the respective other

⁸ Given that this is determined for each agent successively, agents can meet more than one other agent in each time step.

3. If both agents want to date, remove possible relations with current partners and create a new relation between them
- 3) For all agents who currently are dating somebody, determine whether they want to propose marriage to their current partner.
- 4) Wed agents who both proposed marriage to each other
- 5) Determine for each agent whether it dies at the end of the current simulation step or leaves the population because it has reached the maximal age
- 6) Replace each agent who leaves the population by one agent of the same gender

When a given simulation run is initialized, there cannot be any relations among the first set of agents that have developed through mating behaviour. To produce more realistic starting conditions, we initialized each simulation run with a burn-in phase of 120 years (i.e. 1200 simulation steps, which represents two full agent life-cycles). This phase begins with creating M male F female agents, for which there is a 40% probability that they are in the adolescent phase (i.e. the value of t_i is randomly drawn from the range $0 \leq a_i < A_{dating}$) and a 60% probability that they are in the adult phase (i.e. the value of t_i is randomly drawn from the range $A_{dating} \leq a_i \leq A_{max}$). This creates an age-structure that is roughly similar to the contemporary age structure in Europe. As we discuss below, we simulated mating decisions between the years 1941 and 2010. The education and earnings prospects of initial agents, and of all other agents born during the burn-in phase, is based on the distributions for the five-year cohort born in (1935-1940]. Once the burn-in phase is over, the simulation proceeds as described above, starting in the year 1941.

COMPUTATIONAL EXPERIMENTS AND RESULTS

Experimental Design and Outcome Measures

ABC modelling allows us to ‘turn on and off’ certain aspects of human mate search and this enables us to assess how a given preference relates to macro level patterns of assortative

mating. We made use of this in the following way. We created a benchmark model in which we implemented preference structures that we had selected based on both insights from earlier research and on model exploration (see details below). In line with the elements of Eq. (1), we refer to this model as the *SYA-model*, indicating that agents select partners based on preferences for education (*S*), earnings prospects (*Y*), and age (*A*). Using this as a starting point, we created six different versions of the *SYA-model* in which we ‘turned on’ only a selection of these preferences. We refer to these models by omitting the letter relating to the preferences that are turned off (i.e. *S-model*, *A-model*, *Y-model*, *SA-model*, *SY-model*, *AY-model*). Additionally, we created a model in which we turned off all preferences. In this model, mating occurs purely at random and this makes this model a useful baseline against which to assess how much a given preference as contributed to patterns of assortative mating, over and above the effects of mere chance processes. We refer to this mode as the *R-model*.

The following parameterization was the same in each model. There were $M = F = 400$ agents that started dating at the age of $A_{dating} = 160$ and left the population at the age of $A_{max} = 600$. Based on insights from earlier research and based on model exploration, we used the gendered specific age-pressure factors $\kappa^{male} = .00125$ and $\kappa^{female} = .25$. This means that female agents earlier feel an increasing urge to find a partner. We used a gender neutral intimacy factor of $\beta = .015$. This means that both female and male agents become similarly attached to their partners over time. We assumed that the probability that a given agents dies in a given simulation year increases convexly (i.e. $w_d = 1.5$) up to 2.5% ($d = .0025$, which amounts to 2.5% when it is applied in each of the ten time steps of a simulation year) at the maximal age (i.e. at $A_{max} = 600$) and we assumed that women prefer partners who are $\sigma = 25$ time steps (i.e. 2.5 years) older.

Between the different versions of the model, we varied parameters as shown in Table 3. We selected parameters as to reflect known differences in the relative importance than men

and women tend to attach to these characteristics. That is, women tend to attach more import to both the education and their earnings prospects than men (Buss & Barnes, 1986; Buss et al., 1990, 2001), which is implemented by the fact that $w_s^m < w_s^f$ and $w_y^m < w_y^f$. In the case of w_a such an interpretation is not possible, given that the equations for men differ in this aspect. We based the exact magnitude of these parameters on preliminary model explorations in which we systematically varied all parameters orthogonally to each other.

–Table 3 here–

To assess how well the benchmark model approximates observed patterns of assortative mating in terms of education, age, and earnings prospects, we relied on empirical data provided by the European Social Survey (ESS)⁹ and the EU-SILC.

The ESS provides information about respondents' current marital status, their education (measured in ISCED categories), the education of their partner, their age, the age of their partner, and whether they were ever divorced. Based on this, we were able to inspect patterns of assortative mating in terms of education and age in couples in which at least one partner (the respondent) was married for the first time at the time of the survey. That is, we focused on respondents who were never divorced before and classified their educational attainment similarly to the classification use for initializing agent cohorts. Based on this, we categorized couples into three types (homogamic, female hypergamic, female hypergamic) and compared the shares of these couples in a given cohort to their shares in the corresponding cohorts in our simulation outcomes. To align our simulation model as closely as possible to the time frame that the ESS covers, we simulated the period 1940 and 2010 (the time the last wave considered here was collected). Next to this, we also used the ESS to assess the average age difference in within couples (his age minus her age).

⁹ Waves 1 to 5, collected between 2002 and 2010.

As discussed earlier, the EU-SILC provides detailed information about the income of respondents and their partners. This enabled us to assess assortative mating in terms of earnings prospects within married couples (operationalized as discussed above). Yet, when interpreting this result, we need to keep in mind that the EU-SILC does not provide information about the order of the marriage at the time of the survey (i.e. first marriage vs. higher order), which makes it impossible to only focus on respondents who were married for the first time at the time of the survey. In both datasets, we focused on respondents who were born in the cohorts (1940-1950], (1950-1960], and (1960-1970].

Our model implements general tendencies in human mate search and we treated each country as an independent marriage market. Across countries, we should thus expect a good fit between simulated and observed outcomes. However, within countries historical idiosyncrasies might have affected observed mating patterns, next to these general tendencies. To avoid over-fitting of our model based on such idiosyncrasies, we focused on averaged outcomes across countries, rather than on outcomes country by country. Per country, we conducted one simulation run.

Results

Figures 2 to 7 show the outcomes that our model generated, in direct comparison to the patterns observed in the ESS. Figures 2, 3, and 4 show the average shares of female hypergamic, homogamic, and female hypogamic couples. The observed data show that there was a trend of decreasing hypergamy to the benefit of homogamy and female hypogamy. In general, the shares of heterogamic couples are much lower, and the shares of homogamic couples are much higher, than our model based on random matching (*R-model*) would predict. The version of our model that contained preferences for education, age, and earnings potentials (*SYA-model*) captures these trends well with only small deviations from the observed data. Also the model in which only preferences for education matter comes close to

the observed patterns, but tends to overestimate the share of homogamic couples and tends to underestimate the share of female hypogamic couples. By contrast, both the model that only includes age preferences and the model that only includes preferences for earnings potentials provide predictions that are close that that of the *R-model*. Taken together, the results suggest that according to our model, preferences for age and earnings potential might have contributed to the patterns of educational assortative mating that we can observe.

–Figures 2, 3, and 4 here–

Figure 5 shows the results regarding average age differences within couples. In general, the observed age difference is much lower than what the *R-model* would suggest (about 6 to 8 years lower). The reason for this high difference in the *R-model* is the fact that agents who lose their partner due to death at a late age have little difficulty in finding a partner who is willing to marry them. Given that younger agents are relatively less likely to currently be in a marriage than older agents (simply because they had not had enough time yet to find a partner), younger agents are relatively more available for older agents for remarriage, which leads to a large age difference (remember that we focus on first marriages among women; the age difference estimated in the *R-model* is almost the same from the perspective of men who marry for the first time). The figure also suggests that again the *SYA-model* is able to emulate observed patterns of age homogamy, with an average overestimation of about .6 years. Similar to the case of educations, the model that includes preferences on all dimensions provides a better fit than the model only includes preferences for age. This suggests that preferences for education and earnings potential might have contributed to observed patterns of age homogamy.

–Figures 5 here–

Figure 6 suggests a more complex pattern when it comes to mating patterns in terms of earnings potentials. The figure shows that in the EU-SILC data, over time homogamy on this

dimension has decreased to the disadvantage of women. The simulation model that includes preferences on all three dimensions (*SYA-model*) captures this trend well for the cohorts of women born between 1950 and 1970. However, for the cohort born between 1940 and 1950, our model suggests a trend opposite to the observed trend (i.e. increasing homogamy). Except for the model that only includes preferences for age (*A-model*), also all reduced models fail to capture the observed trend.

Taken together, our simulation model is well able to reproduce observed patterns of assortative mating when it comes to education and age. However, when it comes to earnings potentials, improvements seem possible.

CONCLUSION

In this paper, we have developed an agent-based computational model that enables us to study the dynamics that generate observed patterns in assortative mating. The results of our computational experiments with this model suggest that patterns of assortative mating in terms of education and age might be partially the combined product of preferences education, age, and earnings prospects of prospective partners among individuals.

REFERENCES

- Björklund, A. (1993). A Comparison Between Actual Distributions of Annual and Lifetime Income: Sweden 1951-89. *Review of Income and Wealth*, 39(4), 377–386.
- Blomquist, N. S. (1981). A Comparison of Distributions of Annual and Lifetime Income: Sweden around 1970. *The Review Of Income and Wealth*, 27(3), 243–264.
- Blossfeld, H.-P. (2009). Educational Assortative Marriage in Comparative Perspective. *Annual Review of Sociology*, 35(1), 513–530. doi:10.1146/annurev-soc-070308-115913
- Bosworth, B., Burtless, G., & Steuerle, E. (2000). Lifetime Earnings patterns, the Distribution of Future Social Security Benefits, and the Impact of Pension Reform. *Social Security Bulletin*, 63, 74–98.
- Boxer, C. F., Noonan, M. C., & Whelan, C. B. (2013). Measuring Mate Preferences: A Replication and Extension. *Journal of Family Issues*, (May).
doi:10.1177/0192513X13490404
- Buss, D. M., Abbott, M., Angleitner, A., Asherian, A., Biaggio, A., Blanco-Villasenor, A., ... Yang, K.-S. (1990). International Preferences in Selecting Mates: A Study of 37 Cultures. *Journal of Cross-Cultural Psychology*, 21(1), 5–47.
doi:10.1177/0022022190211001
- Buss, D. M., & Barnes, M. (1986). Preferences in Human Mate Selection. *Journal of Personality and Social Psychology*, 50(3), 559–570. doi:10.1037//0022-3514.50.3.559
- Buss, D. M., & Schmitt, D. P. (1993). Sexual Strategies Theory: An Evolutionary Perspective on Human Mating. *Psychological Review*, 100(2), 204–232.
- Buss, D. M., Shackelford, T. K., Kirkpatrick, L. A., & Larsen, R. J. (2001). A Half Century of Mate Preferences: The Cultural Evolution of Values. *Journal of Marriage and Family*, (May), 491–503.

- England, P., & McClintock, E. A. (2009). The Gendered Double Standard of Aging in US Marriage Markets. *Population and Development Review*, 35(4), 797–816.
- Esteve, A., García-Román, J., & Permanyer, I. (2012). The Gender-Gap Reversal in Education and its Effect on Union Formation: The End of Hypergamy? *Population and Development Review*, 38(3), 535–546.
- French, R. M., & Kus, E. T. (2008). KAMA: A Temperature-Driven Model of Mate Choice Using Dynamic Partner Representations. *Adaptive Behavior*, 16(1), 71–95.
doi:10.1177/1059712307087598
- Haider, S., & Solon, G. (2006). Life-Cycle Variation in the Association between Current and Lifetime Earnings. *The American Economic Review*, 96(4), 1308–1320.
- Kalmijn, M. (1998). Intermarriage and Homogamy: Causes, Patterns, Trends. *Annual Review of Sociology*, 24(24), 395–421. doi:10.1146/annurev.soc.24.1.395
- Kalmijn, M., & Flap, H. (2001). Assortative Meeting and Mating: Unintended Consequences of Organized Settings of Partner Choice. *Social Forces*, 79(4), 1289–1312.
- KC, S., Barakat, B., Goujon, A., Skirbekk, V., Sanderson, W. C., & Lutz, W. (2010). Projection of Populations by Level of Educational Attainment, Age, and Sex for 120 Countries for 2005-2050. *Demographic Research*, 22, 383–472.
doi:10.4054/DemRes.2010.22.15
- Klevmarcken, N. A. (1982). On the Stability of Age-Earnings Profiles. *Scandinavian Journal of Economics*, 84(4), 531–554.
- Li, N. P. (2007). Intelligent Priorities: Adaptive Long- and Short-Term Mate Preferences. In G. Geher & G. Miller (Eds.), *Mating Intelligence - Sex, Relationships, and the Mind's Reproductive System* (pp. 105–116). New York: Psychology Press.

- Li, N. P., Bailey, J. M., Kenrick, D. T., & Linsenmeier, J. a. W. (2002). The Necessities and Luxuries of Mate Preferences: Testing the Tradeoffs. *Journal of Personality and Social Psychology*, 82(6), 947–955. doi:10.1037//0022-3514.82.6.947
- Li, N. P., & Kenrick, D. T. (2006). Sex Similarities and Differences in Preferences for Short-Term Mates: What, Whether, and Why. *Journal of Personality and Social Psychology*, 90(3), 468–89. doi:10.1037/0022-3514.90.3.468
- Li, N. P., Yong, J. C., Tov, W., Sng, O., Fletcher, G. J. O., Valentine, K. a, ... Balliet, D. (2013). Mate Preferences Do Predict Attraction and Choices in the Early Stages of Mate Selection. *Journal of Personality and Social Psychology*. doi:10.1037/a0033777
- Lutz, W., Cuaresma, J. C., & Sanderson, W. (2008). The Demography of Educational Attainment and Economic Growth. *Science (New York, N.Y.)*, 319(5866), 1047–8. doi:10.1126/science.1151753
- Lutz, W., Goujon, A., KC, S., & Sanderson, W. C. (2007). Reconstruction of Populations By Age, Sex and Level of Educational Attainment for 120 Countries for 1970-2000. *Vienna Yearbook of Population Research, 2007*, 193–235. doi:10.1553/populationyearbook2007s193
- Mare, R. D. (1991). Five Decades of Educational Assortative Mating. *American Sociological Review*, 56, 15–32. doi:10.2307/2095670
- Rosen, S., & Taubman, P. (1982). Changes in Life-Cycle Earnings : What Do Social Security Data Show? *THE Journal of Human Resources*, 17(3), 321–338.
- Schofer, E., & Meyer, J. W. (2005). The Worldwide Expansion of Higher Education in the Twentieth Century. *American Sociological Review*, 70(6), 898–920. doi:10.1177/000312240507000602

- Schwartz, C. R. (2010). Pathways to Educational Homogamy in Marital and Cohabiting Unions. *Demography*, 47, 735–753. Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3000057/>
- Schwartz, C. R. (2013). Trends and Variation in Assortative Mating: Causes and Consequences. *Annual Review of Sociology*, 39, 451–470.
- Simão, J., & Todd, P. M. (2002). Modeling Mate Choice in Monogamous Mating Systems with Courtship. *Adaptive Behavior*, 10(2), 113–136.
doi:10.1177/1059712302010002003
- Simão, J., & Todd, P. M. (2003). Emergent Patterns of Mate Choice in Human Populations. *Artificial Life*, 9, 403–417.
- Skopek, J., Schulz, F., & Blossfeld, H.-P. (2010). Who Contacts Whom? Educational Homophily in Online Mate Selection. *European Sociological Review*, 27(2), 180–195.
doi:10.1093/esr/jcp068
- Todd, P. M., & Billari, F. C. (2003). Population-Wide Marriage Patterns Produced by Individual Mate-Search Heuristics. In F. C. Billari & A. Prskawetz (Eds.), *Agent-Based Computational Demography - Using Simulation to Improve our Understanding of Demographic Behaviour* (pp. 117–138). Physica Verlag.
- Todd, P. M., Billari, F. C., & Simão, J. (2005). Aggregate Age-at-Marriage Patterns from Individual Mate-Search Heuristics. *Demography*, 42(3), 559–574.
- Todd, P. M., & Miller, G. F. (1999). From Pride and Prejudice to Persuasion: Satisficing in Mate Search. In G. Gigerenzer, P. M. Todd, & The ABC Research Group (Eds.), *Simple Heuristics That Make Us Smart* (pp. 287–308). New York: Oxford University Press.
- Van Bavel, J. (2012). The Reversal of Gender Inequality in Education, Union Formation and Fertility in Europe. *Vienna Yearbook of Population Research*, 10, 127–154.
doi:10.1553/populationyearbook2012s127

TABLES AND FIGURES

Table 1 Distribution of educational degrees for Belgium based on the IIASA/VID

reconstruction/predictions

Belgium		Men				Women			
Year	Birth cohort	1 NoEdu	2 ISCED 0 and 1	3 ISCED 2 to 4	4 ISCED 5 and 6	1 NoEdu	2 ISCED 0 and 1	3 ISCED 2 to 4	4 ISCED 5 and 6
1970	(1935-1940]	0	0.29	0.50	0.22	0	0.33	0.55	0.12
1975	(1940-1945]	0	0.24	0.53	0.23	0	0.28	0.59	0.14
1980	(1945-1950]	0	0.19	0.56	0.25	0	0.21	0.62	0.17
1985	(1950-1955]	0	0.14	0.58	0.28	0	0.15	0.65	0.20
1990	(1955-1960]	0	0.10	0.60	0.30	0	0.11	0.66	0.23
1995	(1960-1965]	0	0.09	0.57	0.34	0	0.08	0.65	0.28
2000	(1965-1970]	0	0.06	0.59	0.35	0	0.06	0.53	0.41
2005	(1970-1975]	0	0.06	0.57	0.36	0	0.07	0.50	0.43
2010	(1975-1980]	0	0.06	0.57	0.38	0	0.06	0.47	0.46
2015	(1980-1985]	0	0.05	0.56	0.39	0	0.06	0.45	0.50
2020	(1985-1990]	0	0.04	0.54	0.41	0	0.05	0.43	0.52
2025	(1990-1995]	0	0.04	0.53	0.43	0	0.04	0.41	0.55
2030	(1995-2000]	0	0.04	0.52	0.45	0	0.04	0.39	0.58
2035	(2000-2005]	0	0.03	0.50	0.46	0	0.03	0.37	0.60
2040	(2005-2010]	0	0.03	0.49	0.48	0	0.03	0.35	0.62
2045	(2010-2015]	0	0.02	0.47	0.50	0	0.02	0.34	0.64
2050	(2015-2020]	0	0.02	0.46	0.52	0	0.02	0.32	0.66

Table 2 Number of observations in the EU-SILC data that satisfy selection criteria

country	gender	education	(1950,1955]	(1955,1960]	(1960,1965]	(1965,1970]	(1970,1975]	country	gender	education	(1950,1955]	(1955,1960]	(1960,1965]	(1965,1970]	(1970,1975]
AT	1	1	139	1538	1888	1764	362	IT	1	1	776	5560	6934	6698	1441
		2	<u>49</u>	379	472	492	<u>97</u>			2	<u>99</u>	722	728	848	244
	2	1	218	1760	2267	2045	499		2	1	998	6896	8678	8011	1750
		2	<u>36</u>	369	399	451	140			2	117	763	1133	1242	375
BE	1	1	154	1373	1915	1831	698	LT	1	1	<u>65</u>	1155	1339	1033	247
		2	<u>60</u>	674	1051	1206	592			2	<u>13</u>	284	296	182	<u>88</u>
	2	1	180	1519	2140	1839	776		2	1	<u>57</u>	1147	1312	1011	210
		2	<u>65</u>	818	1334	1395	684			2	<u>27</u>	578	593	413	161
BG	1	1	<u>0</u>	440	554	554	242	LU	1	1	149	860	1056	1283	301
		2	<u>0</u>	<u>55</u>	104	<u>96</u>	<u>35</u>			2	<u>55</u>	284	480	570	211
	2	1	<u>0</u>	490	565	589	234		2	1	162	1120	1273	1266	299
		2	<u>0</u>	116	154	177	<u>85</u>			2	<u>59</u>	276	411	496	204
CZ	1	1	<u>51</u>	1135	1548	1581	696	LV	1	1	<u>35</u>	885	1054	989	315
		2	<u>7</u>	278	320	328	109			2	<u>8</u>	140	167	166	<u>55</u>
	2	1	<u>59</u>	1368	1800	2005	794		2	1	<u>52</u>	1002	1171	991	302
		2	<u>4</u>	220	339	298	131			2	<u>20</u>	365	452	437	120
DE	1	1	101	597	583	465	<u>0</u>	PL	1	1	191	3474	3502	3249	861
		2	100	434	461	369	<u>0</u>			2	<u>23</u>	383	431	483	206
	2	1	167	864	1036	825	<u>0</u>		2	1	290	4140	3755	3320	912
		2	<u>93</u>	528	536	429	<u>0</u>			2	<u>41</u>	773	922	903	330
EE	1	1	151	1614	2042	1599	498	PT	1	1	171	1255	1464	1242	263
		2	<u>37</u>	451	452	328	116			2	<u>29</u>	150	160	135	<u>35</u>
	2	1	169	1586	1934	1750	485		2	1	202	1513	1647	1507	254
		2	<u>86</u>	848	1226	858	270			2	<u>32</u>	210	297	233	<u>70</u>
ES	1	1	415	3697	5359	4901	2036	RO	1	1	<u>0</u>	705	1020	1427	683
		2	125	1388	2192	2142	1087			2	<u>0</u>	123	146	237	127
	2	1	566	4813	6632	5682	2126		2	1	<u>0</u>	844	1219	1639	798
		2	134	1424	2654	3034	1467			2	<u>0</u>	105	179	260	156
FR	1	1	350	2558	2987	2555	600	SI	1	1	226	2883	3194	2626	703
		2	107	642	955	943	372			2	<u>25</u>	540	477	465	192
	2	1	431	2980	3389	2834	606		2	1	223	3347	3989	2728	657
		2	107	850	1148	1320	442			2	<u>27</u>	647	843	733	309
IE	1	1	<u>0</u>	430	771	699	400	UK	1	1	<u>54</u>	819	1109	1066	312
		2	<u>0</u>	191	390	457	369			2	<u>35</u>	453	604	617	238
	2	1	<u>0</u>	629	1070	911	477		2	1	<u>90</u>	1037	1437	1339	406
		2	<u>0</u>	252	508	598	399			2	<u>58</u>	558	858	882	269

Note: The values in the column 'gender' refer to male (1) and female (2) individuals. The values in the column 'education' refer to individuals with less than tertiary education (1) and tertiary education (2). Underlined values indicate cells that were excluded from the analysis.

Table 3 Parameters per model version

	<i>R</i>	<i>S</i>	<i>Y</i>	<i>T</i>	<i>SYT</i>
w_s^m	0	0.75	0	0	0.75
w_s^f	0	1.5	0	0	1.5
w_y^m	0	0	1	0	0.75
w_y^f	0	0	1.5	0	1.5
w_a^m	0	0	0	0.5	1.5
w_a^f	0	0	0	4.5	4.5

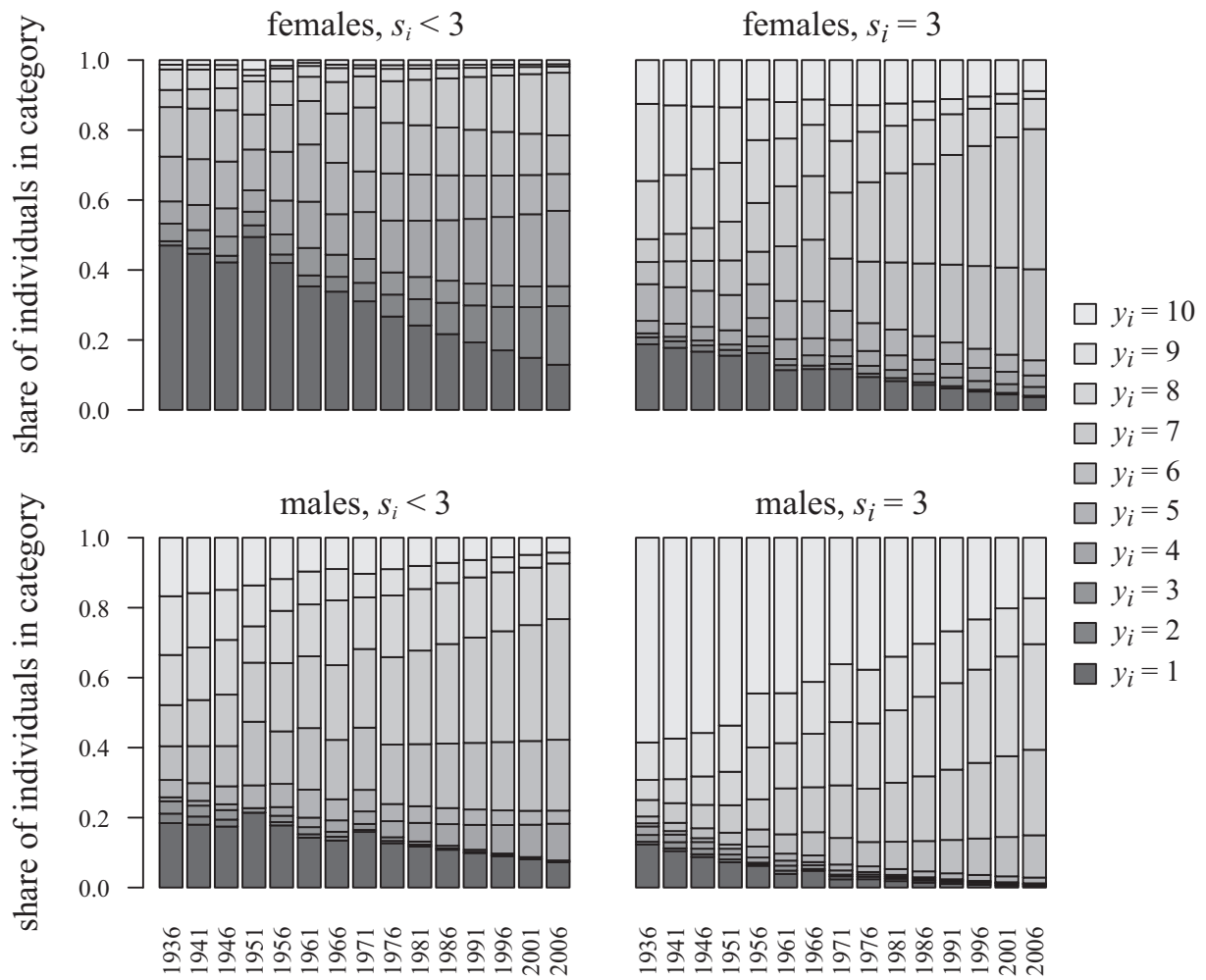


Figure 1 Distribution of women and men with different educational degrees across earnings prospect categories (y_i) in Belgium; observed and projected.

Note: Years indicate the lower boundary of the five-year birth cohort that is included

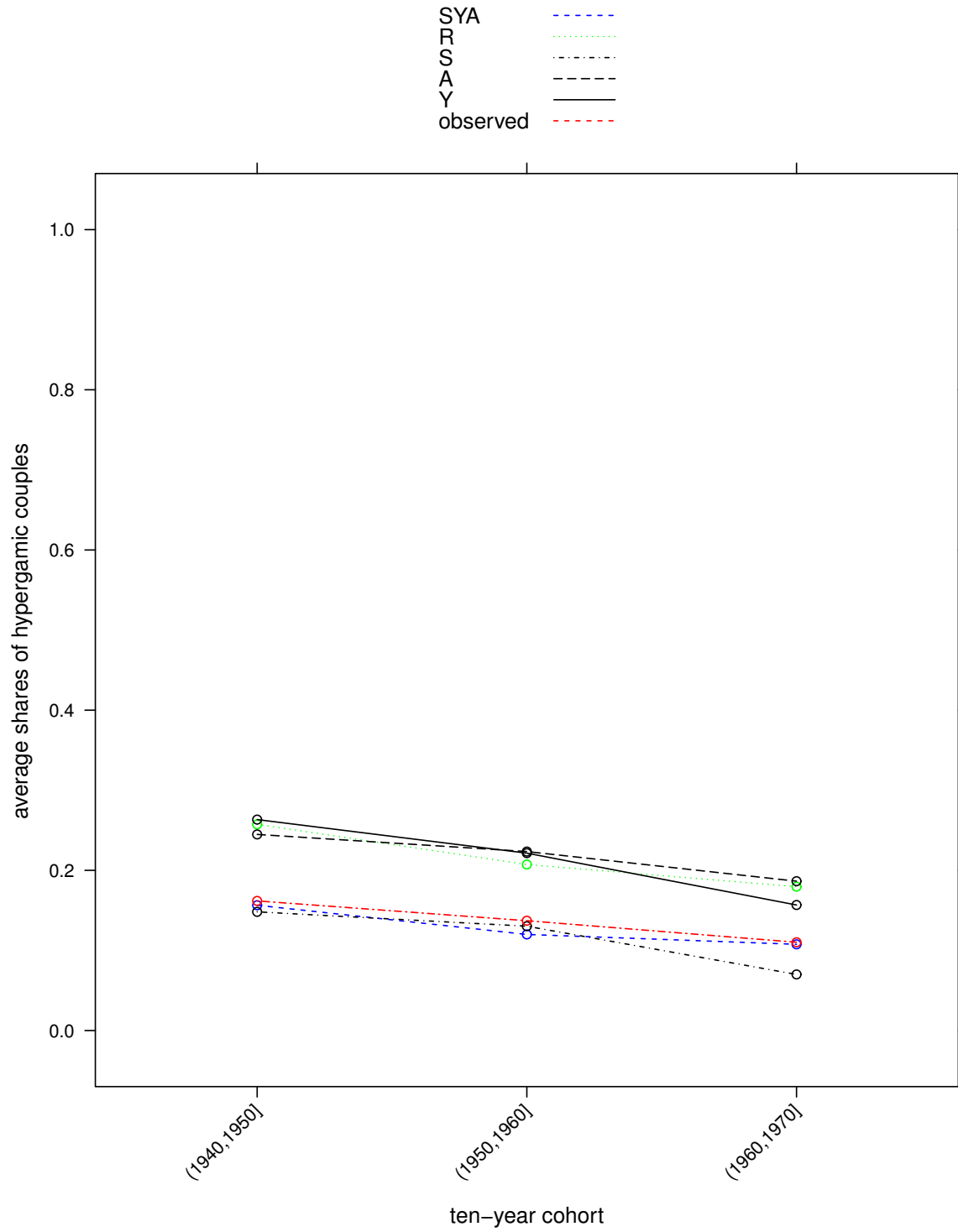


Figure 2 Comparison of shares of female hypergamic couples as observed in the ESS and as obtained in different versions of the computational model.

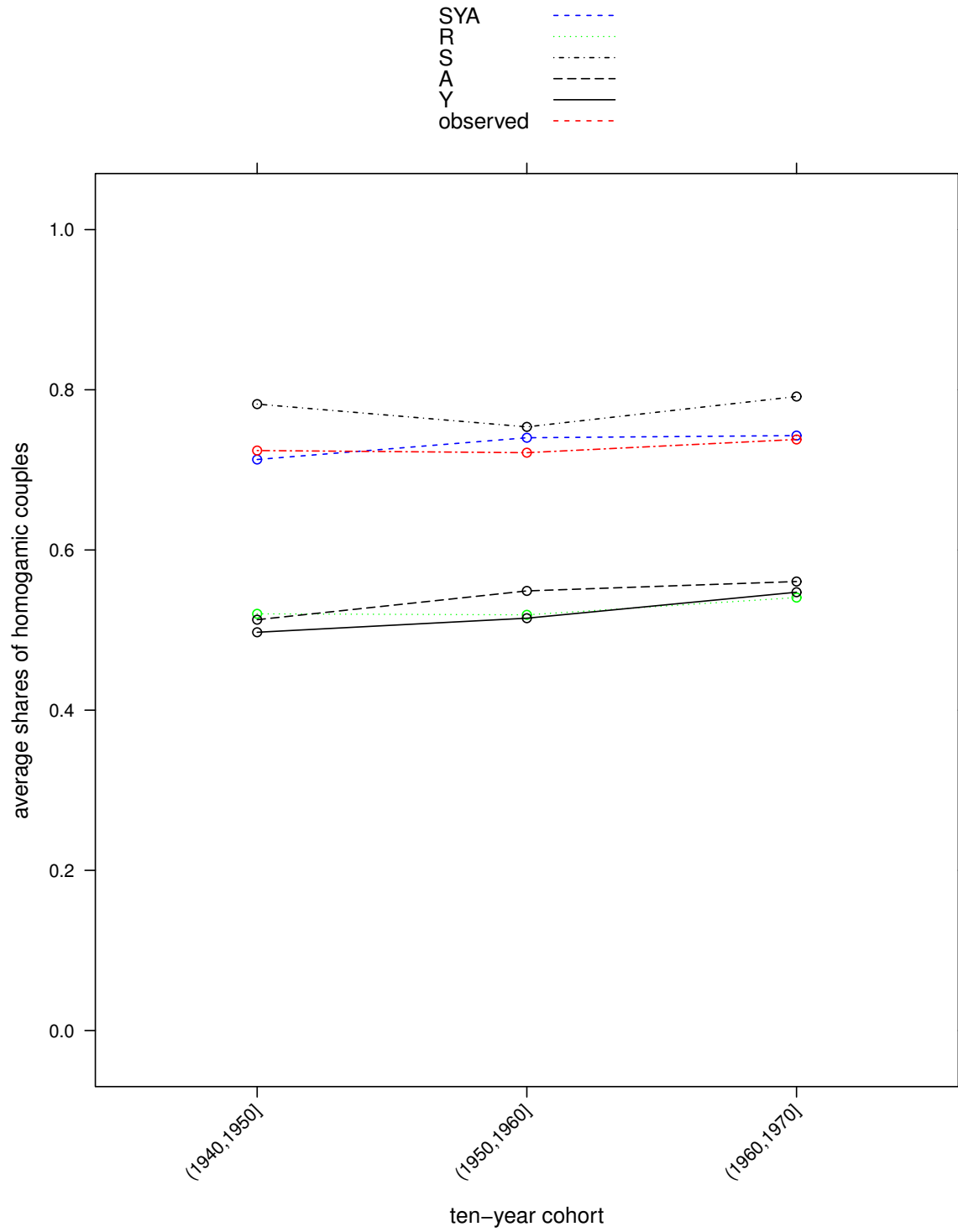


Figure 3 Comparison of shares of homogamic couples as observed in the ESS and as obtained in different versions of the computational model.

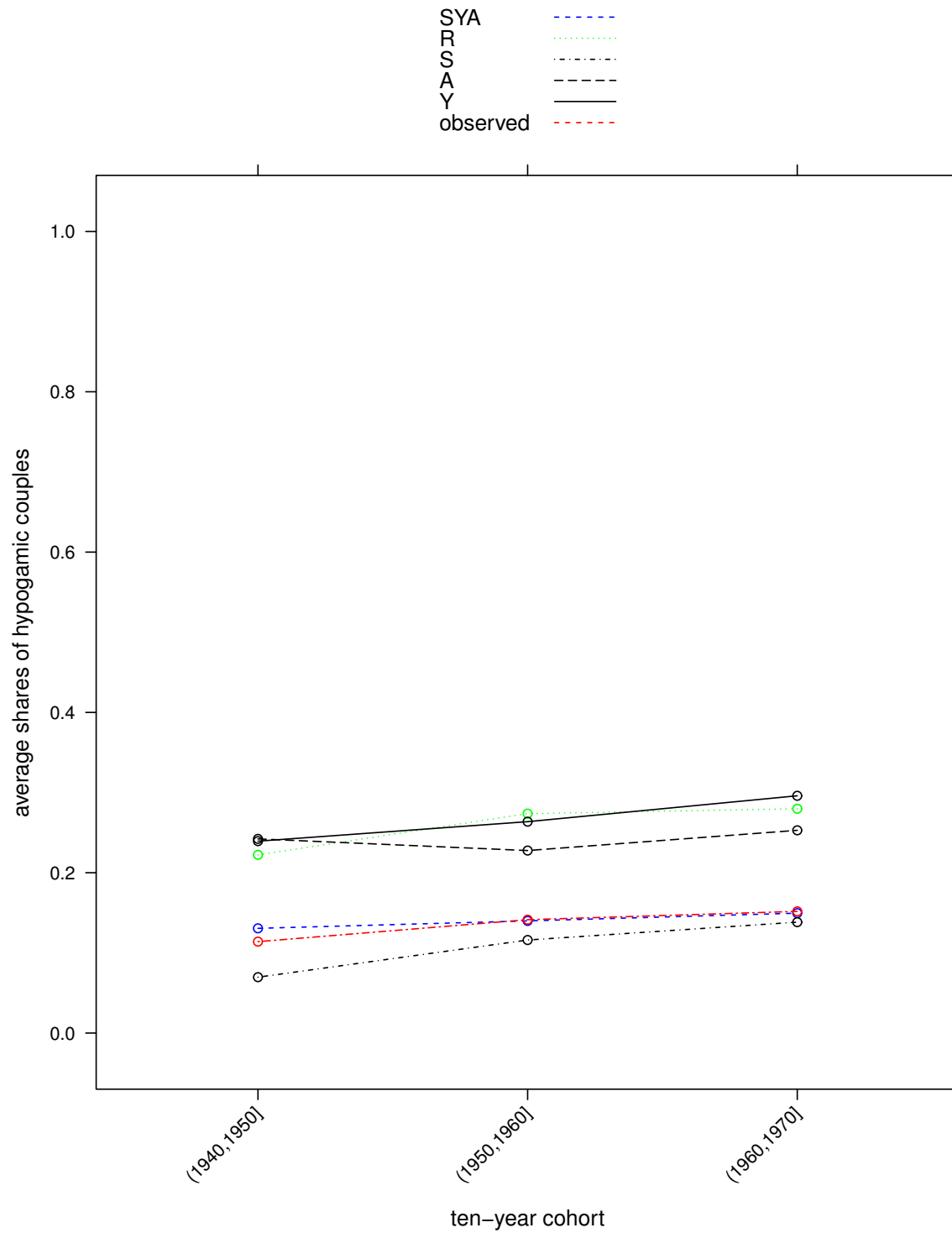


Figure 4 Comparison of shares of hypogamic couples as observed in the ESS and as obtained in different versions of the computational model.

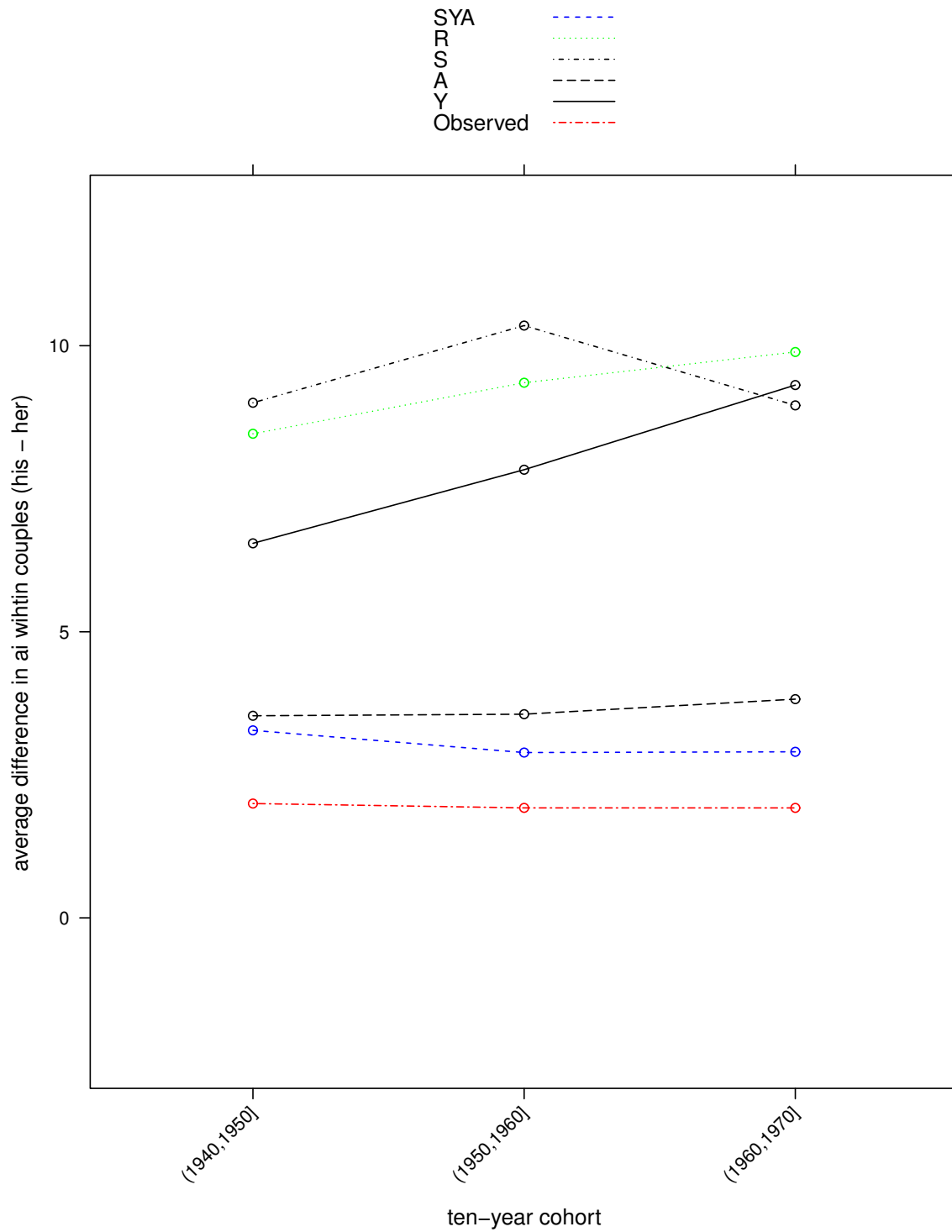


Figure 5 Comparison of average age differences within couples as observed in the ESS and as obtained in different version of the computational model.

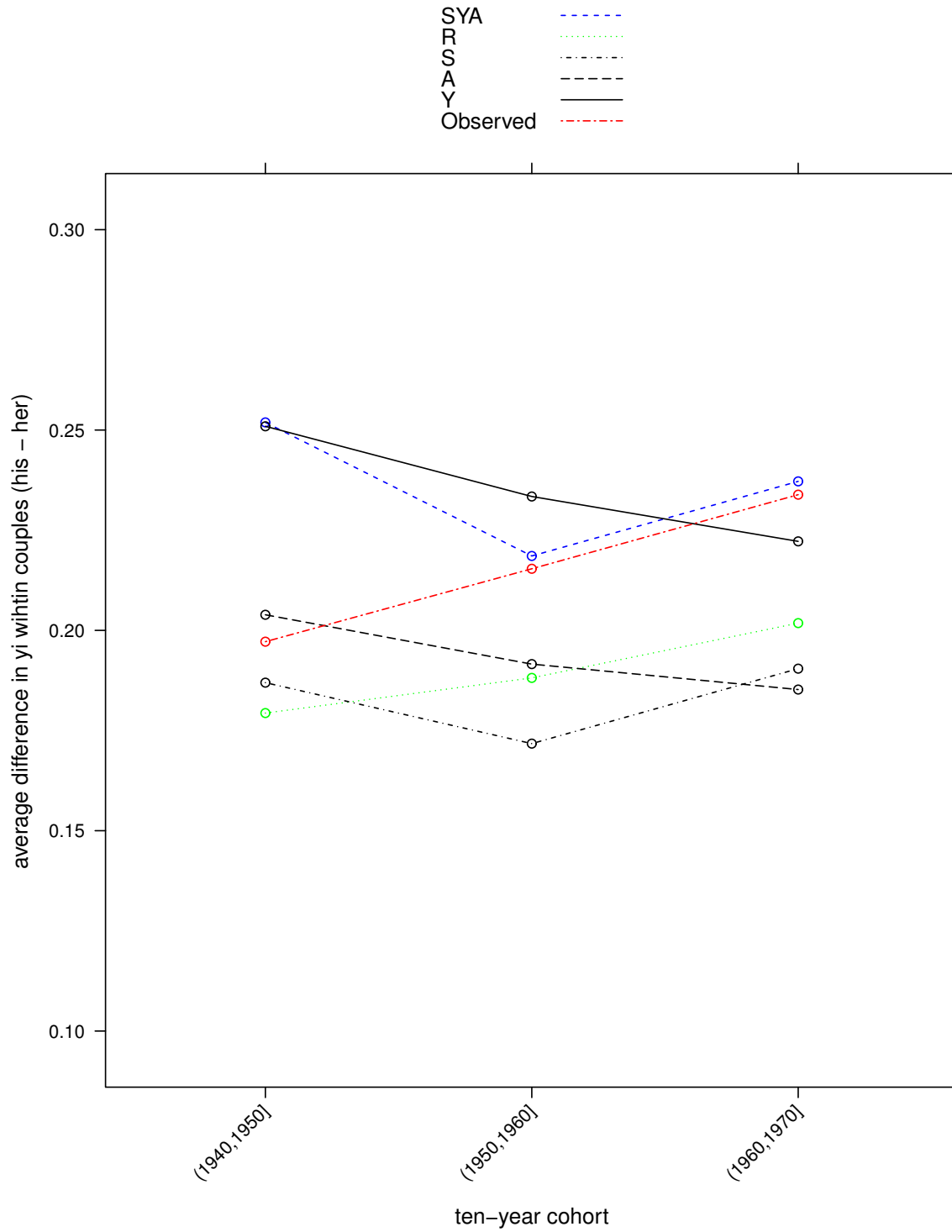


Figure 6 Comparison of average differences in earnings difference within couples as observed in the ESS and as obtained in different version of the computational model.