Università degli Studi Firenze
DISEI
Dipartimento di Scienze per l'Economia e l'Impresa

Working Papers - Economics

On the Evolution of Individual Preferences and Family Rules

Alessandro Cigno, Alessandro Gioffré, Annalisa Luporini

Working Paper N. 11/2019

DISEI, Università degli Studi di Firenze
Via delle Pandette 9, 50127 Firenze (Italia) www.disei.unifi.it

On the evolution of individual preferences and family rules*

Alessandro Cigno† Alessandro Gioffré‡ Annalisa Luporini§

February 26, 2019

Abstract

We study how the distribution of an inherited trait evolves through marriage if couples are formed at random. If the matching occurs across the entire population, the variance of the trait tends to diminish, and the distribution converges to a common trait. If the matching is restricted to specific subpopulations, each of these converges to a different trait. This has implications for the consequences of immigration. Using a specific model where the trait is a parameter measuring a person's taste for receiving filial attention (a good without perfect market substitutes) in old age, we also show that it may be in a couple's interest to obey a rule requiring them to give specified amounts of the good to their respective parents. The matching is random in this model because preferences are private monitoring. In the long run, if the matching extends to the entire population, either everybody obeys the rule, or nobody does. In the interim, some do, and some do not. If the matching is restricted to specific ethnic or religious groups, the population will tend to break down into a number of sharply characterized subpopulations. That may undermine social cohesion and call for policy intervention.

Keywords: Marriage, evolution, random matching, family rule, immigration

JEL classification: C78, D13, J12

1 Introduction

The way individual preferences and collective values are transmitted and evolve from one generation to the next was traditionally the concern of ethnic or religious minorities intent on preserving their identity. More recently, it has become also the concern of natives worried

*We thank participants in SIE 2018, ASSET 2018 and Nanzan University seminar for comments and suggestions. Remaining errors are the authors' responsibility.

†Department of Economics and Management, University of Florence, CESifo, IZA and CHILD. E-mail: cigno@unifi.it

‡Corresponding author. Department of Economics and Management, University of Florence, and Research Center SAFE, Goethe University Frankfurt. E-mail: alessandro.gioffre@unifi.it

§Department of Economics and Management, University of Florence, CESifo, and CHILD. E-mail: annalisa.luporini@unifi.it

about loosing their identity to immigrants. A strand of economic literature stemming from Bisin and Verdier (2001), and Tabellini (2008), assumes that optimizing parents motivated either by a paternalistic form of altruism, or by a social conscience, undertake costly actions in order to transmit their preferences or values on to their offspring. Preferences or values evolve over time as a result of social interaction among individuals who received different inputs from their respective parents. The implicit assumption underlying these models is that the parental couple think and act as if they were one person. What happens if mother and father have different preferences or values?

The issue is taken up by Cigno et al. (2017) in a more complex setting where at least some working-age individuals respond rationally not only to the economic and legal environment, but also to a family rule that is itself a collectively rational response to the environment.[1] The rule in question concerns the provision of filial attention (a good without perfect market substitutes) by working-age children to retirement-age parents. Assuming that preferences are common knowledge, the article demonstrates that a person whose preference parameters satisfy a certain condition will marry a person of the opposite sex who holds the same preferences, and that the couple thus formed will transmit their common preferences on to their children. In equilibrium, a person whose preferences are thus characterized gives attention to her or his parents while he or she is of working age, and receives attention from her or his children when he or she reaches retirement age. What happens if some dimensions of a person's preferences are not observable before marriage, and cannot thus be a criterion for a match?

In Section 2 of the present paper we look at the way the distribution of a generic trait evolves from generation to generation, first under the assumption of random matching across the entire population, and then under the assumption that matching is restricted to particular segments of it (e.g., to the original residents, or to immigrants from a particular region). In Section 3, we interpret the trait in question as a preference parameter measuring a person's

[1] The approach was originally developed by Cigno (1993, 2006) under the assumption that individuals reproduce asexually. Cigno et al. (2017) demonstrate that it works also in a world where sexually differentiated individuals marry, have children and bargain with their respective spouses over the allocation of domestic resources.

taste for filial attention, and show that this may give rise to a family rule. Assuming that this parameter is private monitoring before a marriage takes place, and that matching is consequently random (either across the entire population, or within recognizable segments of it), we also show how the distribution of this parameter and of the associated family rule evolves across generations. Our results are compared with the empirical evidence. The concluding section summarizes the results and compares them with those of others who view the same or similar issues from different standpoints.

2 Evolution of individual traits

Take a large population differentiated by sex, and by an inherited trait δ. The latter might be a genetically transmitted physical characteristic, or a cultural trait transmitted in some other way (we shall be more specific in Section 3). Assume that a person's trait is the mean of his or her parents' traits.[2] Suppose that everybody marries, and that each couple has a son and a daughter, so that the number of men and the number of women will remain the same also in subsequent generations.[3] Marriages between siblings are not allowed. Assuming, for now, that couples are drawn at random from the entire population, we show that the distribution of δ evolves from one generation to the next as a result of "mixed marriages" (marriages where the man and the woman have different δ).

In each generation, there are n men and n women, where n is a large number. In generation 0, each man (woman) is characterized by either $\delta = \delta^H$ or $\delta = \delta^L$, with $\delta^H > \delta^L$. In that generation, the number of traits is then $S(0) = 2^0 + 1 = 2$. In subsequent generations, the number of possible traits will be larger as a result of mixed marriages. In generation 1, the possible traits are δ^L, $\frac{\delta^L+\delta^H}{2}$ and δ^H. Therefore, $S(1) = 2^1 + 1 = 3$. In generation 2, the possible traits are δ^L, $\frac{3\delta^L+\delta^H}{4}$,$\frac{2\delta^L+2\delta^H}{4}$,$\frac{\delta^L+3\delta^H}{4}$and δ^H. Hence, $S(2) = 2^2 + 1 = 5$. In generation

[2]A more general assumption would be to say that a person's δ is symmetrically distributed between those of this person's parents, so that the expected value of the former is equal to the mean of the latter.

[3]In reality, the balance of the sexes is ensured by the fact that (barring selective abortion or infanticide) a child has an equal probability of being born male or female.

t, the possible traits are

$$\delta_t(j) := \frac{(2^t - j)\delta^L + j\delta^H}{2^t} = \delta^L + \frac{\delta^H - \delta^L}{2^t} j,\ j = 0, 1, \ldots, 2^t,$$

and their number is $S(t) = 2^t + 1$.

Now let n^J denote the number of persons of each sex who are characterized by the trait δ^J, $J = H, L$, in generation 0. Define $\pi_0 = (\pi_0(0), \pi_0(1)) := (1 - \pi, \pi)$ as the distribution of δ^L and δ^H, with $\pi_0(0) = \frac{n^L}{n}$, and $\pi_0(1) = \dfrac{n^H}{n}$, in generation 0. In generation t, the distribution will be

$$\pi_t = (\pi_t(0), \pi_t(1), \ldots, \pi_t(2^t)), \quad \text{with} \sum_{j=0}^{2^t} \pi_t(j) = 1 \quad \text{for all } t \geq 0.$$

Hence, the average trait of that generation will be

$$\delta_t := \sum_{j=0}^{2^t} \pi_t(j)\delta_t(j),$$

How does the distribution evolve from generation to generation? Appendix 1 demonstrates the following.

Proposition 1. *In each generation t, for n sufficiently large, the distribution of $\delta_t(j)$ converges to a binomial, with mean $(1 - \pi)\delta^L + \pi\delta^H$ and variance $\pi(1 - \pi)\frac{(\delta^H - \delta^L)^2}{2^t}$.*

Corollary 1. *As $t \to \infty$, the expected trait held by all agents is*

$$\delta^* := (1 - \pi)\delta^L + \pi\delta^H.$$

Now let $\delta^H = 1$ and $\delta^L = 0$. In the long run, everybody has the same trait, $\delta^* = \pi$, where

$$\pi = \frac{n^H}{n}.$$

How long is the long run? A sensible way to address this question is to calculate in how many generations t the standard deviation of the binomial distribution of δ will become

$\sigma \in \{0.01, 0.05\}$ for $\pi \in \{0.1, 0.5\}$. The answer is found solving the equation

$$\frac{(\delta^H - \delta^L)^2}{2^t}\pi(1-\pi) = \sigma^2 \qquad \text{for } \pi \in \{0.1, 0.2, \ldots, 0.5\}.$$

The value of t associated with each (π, σ) is shown in Table 1. Of course, the limit δ (equal to the mean of the distribution) will vary with (π, σ) too.

	$\pi = 0.1$	$\pi = 0.2$	$\pi = 0.3$	$\pi = 0.4$	$\pi = 0.5$
$\sigma = 0.01$	9.81	10.64	11.04	11.23	11.29
$\sigma = 0.05$	5.17	6.00	6.39	6.58	6.64

Table 1: Number of generations needed to reach a distribution of the population with standard deviation σ given π.

The first column of this table says that, if 10 percent of the population is initially characterized by $\delta = 1$, and the remaining 90 percent by $\delta = 0$, so that the limit value of δ is 0.1, it will take 5.17 generations for the standard deviation to become equal to 0.05, and another 4.64 generations for it to fall to 0.01. If generations overlap every 20 years, this means that it will take 130 years for approximately 68 percent of the population to have a δ comprised between 0.095 and 0.105, and more than 245 years for that same share of the population to have a δ comprised between 0.099 and 0.101 (virtually 0.1). The remaining columns show how the convergence slows down, and the limit value of δ gets closer to zero, as the initial share of individuals with $\delta = 1$ rises from one tenth to a half of the total population.

In order to better appreciate the implications of these numbers, imagine that the entire population was originally characterized by $\delta = 0$. Suppose that there is then a once-for-all influx of immigrants, equal in size to one ninth of the native population, and that all the newcomers are characterized by $\delta = 1$. After between five and ten generations, the population will be fairly homogeneous again, and its characteristics will be very similar to those of the original population. In other words, the immigrants will have been absorbed by the native population. If the number of immigrants is larger than one ninth, but no larger than one half of the native population (i.e., not so large that the immigrants outnumber the natives), it will take longer for the population to become homogeneous again, and the future inhabitants will not look much like the original ones. In other words, there will be convergence, but not absorption. Whichever is the case, however, it takes a relatively short time in evolutionary terms (between 130 and 245 years) for the descendants of the original immigrants to become virtually indistinguishable from those of the original inhabitants.

Is that what happens in reality? Galor and Özak (2016) find that a region's pre-industrial agro-climatic characteristics affect time-preference, technological adoption, education, saving, and smoking behaviour to our days. Alesina et al. (2013) find that European and US residents descending from populations who introduced the plough many thousand years ago in their countries of origin still display less equal gender norms than those descending from populations who did not implement that innovation. Hazarika et al. (2019) show that, even after ancestral plough use, the timing of the Neolithic Transition, and many other potentially confounding factors are controlled for, there are proportionately more missing women in countries where ancestral ecological endowments were poorer than in others. According to Guiso et al. (2016), Italian cities that achieved self-government in the Middle Ages display still today greater civic sense than similar cities in the same area that did not. These empirical findings do not necessarily contradict our illustrative calculations, because, in reality, immigration does not occur all at once. In our numerical examples, if a new wave of immigrants with $\delta > 1$ follows after a while the one with $\delta = 1$, the limit value of δ will obviously rise, and the convergence process consequently start all over again from a new initial distribution. For this reason, convergence may take longer than our calculations suggest. But there are also other reasons.

Diamond (1997, 2003, 2005) reports that, in a traditional society like New Guinea, if two persons meet outside their respective villages, they will start a long discussion to establish whether they have a relative or friend in common, and there is thus good reason for not trying to kill each other.[4] Bar the bloody outcome, something similar happens also in modern societies and may extend to the political sphere. Tabellini (2019) finds evidence of political opposition to immigration even when immigrants bring significant economic benefits to res-

[4]Something analogous happens in the plant kingdom. If two specimen of unrelated species are planted next to each other, each of them responds by laying down a vast amount of roots (a costly activity). By contrast, if the two are specimen of related species, they produce a smaller amount of roots and save the resources for other uses; see Callaway and Mahall (2007), and Dudley and File (2007).

idents. Add to this that, for linguistic and cultural reasons, immigrants from a part of the world find it difficult to communicate not only with the original residents, but also with immigrants from other parts of the world, and you have an explanation for the well documented phenomenon that immigrants tend to stick together and marry among themselves. It is thus not surprising that the echo of ancient events is still detectable in the actions and attitudes of today's people.

3 Family rules and individual decisions

We now examine a specific decision model where δ is a preference parameter reflecting a person's taste for filial attention.[5] In the last section, we assumed that δ is transmitted from parents to children. Is that a reasonable assumption to make in the case where δ is a preference parameter, rather than a physical characteristic? Bjorklund et al. (2006) find that parents transmit their preferences to their children even if the latter are adopted. It would thus seem that the transmission is not genetic. Cox and Stark (2005) report that preferences are transmitted by way of example. By contrast, Ottoni-Wilhelm et al. (2017) report that children can be talked into doing good deeds, but setting them a good example has limited effect. Albanese et al. (2016) find that parental influence diminishes, but does not vanish, as the children become exposed to external (school, peer group, etc.) influence. Bisin and Topa (2003) show that it is possible to discriminate empirically between the effect of the family and the effect of the outside world in preference formation. Cohen-Zada (2006) shows that some parents contrast external influence by sending their children to appropriate confessional schools. Summing up, there is evidence that preferences or values are inculcated by parents, and may be modified through socialization. As mentioned in the Introduction, Bisin and Verdier (2001), and Tabellini (2008) focus on the latter. We study the way preferences and family rules evolve through intermarriage.

We start by establishing whether the following behaviour is incentive-compatible.

Definition 1. (*Cooperative behaviour*) *The young give attention to their elderly parents.*

[5] The model is a stripped-down version of the one in Cigno et al. (2017).

This definition partitions the population into two groups: those who give attention to their parents ("cooperators"), and those who do not ("non-cooperators"). The latter may include two subgroups: the "accountable", who do not give attention to their parents when the latter were cooperators, and the "unaccountable", who do not give attention to their parents when the latter were non-cooperators. We use the term "deviator" to designate an accountable non-cooperator.

We then inquire whether the following family rule will support cooperative behavior as an equilibrium outcome.

Definition 2. (*Family rule*) *The young must provide attention to their elderly parents if the latter are not deviators.*

An implication of this rule is that a person is not obliged to give attention to a deviating parent. We will show that it is in the former's interest not to give attention if not obliged to do so, and that deviators may thus be punished by their grown-up children. If a person does not give attention to a parent who deviated from the rule, then he or she is an unaccountable non-cooperator. Conversely, if he or she does not give attention to a parent who did not deviate from the rule, then he or she is a deviator. This rule identifies two individual states, cooperator and unaccountable non-cooperator, that do not justify punishment. We will show that these states may coexist in the short run (i.e., different persons may be in different states), but only one state survives in the long run.

3.1 Basic assumptions

People live two periods. A person is young in period 1, old in period 2. The young can work and marry, the old can do neither. Let w_i and $c_{t,i}$ denote respectively the individual i's wage rate and his or her consumption of market goods (including the personal services of professional helpers) in period $t = 1, 2$.[6] Let $a^i_{k_i}$ be the amount of filial attention that i receives from $k_i = D_i, S_i$, where D_i is i's daughter and S_i is i's son, in period 2. The utility function is

$$U_i = c_{1i} + \ln c_{2i} + \max\left[0,\ \delta_i\left(\ln \beta a^i_{D_i} + \ln \beta a^i_{S_i}\right)\right], \tag{1}$$

[6] Recall that individuals can work only in period 1.

where the parameter δ_i measures i's taste for filial attention, and the constant β is a scaling factor designed to make $\ln \beta a_k^i$ positive for a_k^i sufficiently large.[7] Notice that market goods are not perfect substitutes for filial attention. Notice also that parents are not altruistic towards their children, and children are not altruistic towards their parents. Allowing for a modicum of altruism on either side would make the analysis less sharp without altering the results in any substantive way.

As in the last section, the couple formed by a particular woman f and a particular man m is a random draw either from the entire population of young men and women, or from a particular segment of it. Why random? In the present context, the answer is that, ex ante (i.e., before the couple is drawn), δ_i is private monitoring, and w_i is uncertain. Ex post, (δ_i, w_i) are common knowledge where the members of this couple are concerned, but their children's future wage rates are still uncertain. We shall assume that the individual wage rate is a random variable taking value w^H with probability ψ, and w^L with probability $1 - \psi$, where $w^H > w^L > 1$.

The (f, m) couple may either marry or split (there is no re-sampling). If they split, then $a_D^i = a_S^i = 0$. Therefore, i maximizes (1) subject to the budget constraints

$$\begin{cases} c_{1i} + s_i = w_i, \\ c_{2i} = r s_i, \end{cases}$$

where s_i is the amount of saving in period 1, and r is the interest rate. The pay-off of singlehood is then

$$R_i := \max_{s_i} \left(w_i - s_i + \ln r s_i \right) = w_i - 1 + \ln r.$$

If f and m marry, they Nash-bargain over the allocation of their time and income. Having assumed that people are not altruistic, f and m will neither give attention to their respective parents, nor get it from their children, as a present. They would do both if they and their children complied with the rule set out in Definition 2. Alternatively, they could buy it off their children. Given that the filial attention does not have a perfect market substitute,

[7] Otherwise, $\ln t_k^i$ would be negative for any t_k^i smaller than unity.

however, the children would form a cartel, and set the price so high that the entire surplus generated by the transaction would go to them.[8] Parents are thus indifferent between buying and not buying filial attention. We assume that they do not.

3.2 Bargaining in the absence of a family rule

Suppose that the (f, m) couple do not follow any family rule, and thus that $a^i_{k_i} = 0$. As there will be no re-sampling, f's (m's) best alternative to marrying m (f) is to stay single. The Nash-bargaining equilibrium then maximizes

$$N = (U_f - R_f)(U_m - R_m), \tag{2}$$

subject to f's and m's period budget constraints

$$\begin{cases} c_{1f} + s_f = w_f + T, \\ c_{2f} = rs_f, \end{cases} \qquad \begin{cases} c_{1m} + s_m + T = w_m, \\ c_{2m} = rs_m, \end{cases} \tag{3}$$

where T is defined as a transfer from m to f in period 1.

We show in Appendix 2 that the equilibrium is

$$\widehat{s}_f = \widehat{s}_m = 1, \ \widehat{T} = 0.$$

The equilibrium pay-offs are

$$\widehat{U}_i = w_i - 1 + \ln r = R_i, \ i = f, m.$$

Strictly speaking, therefore, the couple are indifferent between marrying or splitting. We

[8] Bernheim et al. (1985) argue that, as an alternative to paying cash, a parent could commit to bequeathing her entire fortune either to the child who has given her the most attention or, if that attention falls below a certain minimum, to a third party. According to this argument, the surplus would go to the parent, rather than to the children. Cigno (1991) points out, however, that the children could counter the parent's strategy by drawing-up a perfectly legal contract committing only one of them to give the parent the minimum amount of attention required to inherit the lot, and then to share the inheritance (minus a specified amount as compensation for the attention given) equally with the others. That would give the entire surplus back to the children.

assume that they marry.

3.3 Bargaining in the presence of a family rule

Alternatively, suppose that f and m comply with the rule in Definition 2. Let F_i and M_i denote, respectively, i's mother and father.[9] For the assumption that a child i's δ_i is equal to the mean of his parents,

$$\delta_i = \frac{\delta_{F_i} + \delta_{M_i}}{2}.$$

Given that the couple give attention to their respective parents, the first-period budget constraint is now

$$c_{1f} + s_f = w_f(1 - a_f^{F_f} - a_f^{M_f}) + T,$$

for f and

$$c_{1m} + s_m + T = w_m(1 - a_m^{F_m} - a_m^{M_m}), \tag{4}$$

for m. In Subsection 3.4 we will demonstrate that the amount given by $i = f, m$ to $h_i = F_i, M_i$ is

$$a_i^{h_i} = \frac{\delta^{h_i}}{w_i}. \tag{5}$$

Given that i's reservation utility is now $\widehat{U}_i = R_i$, the Nash-bargaining equilibrium then maximizes

$$N' = \left(U_f - \widehat{U}_f\right)\left(U_m - \widehat{U}_m\right), \tag{6}$$

subject to

$$\begin{cases} c_{1f} + s_f = w_f(1 - a_f^{F_f} - a_f^{M_f}) + T, \\ c_{2f} = rs_f, \end{cases} \qquad \begin{cases} c_{1m} + s_m + T = w_m(1 - a_m^{F_m} - a_m^{M_m}), \\ c_{2m} = rs_m. \end{cases} \tag{7}$$

Assuming an interior solution (or the rule would be inoperative), and recalling that k_i's

[9] Here F and M stand for female and male respectively.

wage rate is uncertain, we show in Appendix 3 that the equilibrium is

$$
\begin{aligned}
s'_f &= s'_m = 1, \\
T' &= \delta_m \left(\psi \ln \frac{\beta\delta_m}{w^H} + (1-\psi) \ln \frac{\beta\delta_m}{w^L} - 1 \right) \\
&\quad - \delta_f \left(\psi \ln \frac{\beta\delta_f}{w^H} + (1-\psi) \ln \frac{\beta\delta_f}{w^L} - 1 \right) \\
&= \delta_m \left(\ln \beta\delta_m - \bar{w} - 1 \right) - \delta_f \left(\ln \beta\delta_f - \bar{w} - 1 \right),
\end{aligned}
$$

where

$$\bar{w} := \psi \ln w^H + (1-\psi) \ln w^L.$$

In contrast with the case examined in the last subsection, where no family rule is in force, the compensatory transfer T may thus be positive, negative or zero (the spouse with the higher δ shares the benefit of obeying the rule with the spouse who has the lower δ).

$$U'_f = w_f + \delta_f \left(\ln \beta\delta_f - \bar{w} - 1 \right) + \delta_m \left(\ln \beta\delta_m - \bar{w} - 1 \right) - 1 + \ln r$$

and

$$U'_m = w_m + \delta_f \left(\ln \beta\delta_f - \bar{w} - 1 \right) + \delta_m \left(\ln \beta\delta_m - \bar{w} - 1 \right) - 1 + \ln r,$$

where U'_f and U'_m are expectations, because the children's wage rates are uncertain. Note that U'_f may differ from U'_m because w_f may differ from w_m.

A condition for it to be in the (f, m) couple's common interest to obey the rule in Definition 2 is that they would not be better-off disobeying it,

$$U'_f - \widehat{U}_f = U'_m - \widehat{U}_m = \delta_f \left(\ln \beta\delta_f - \bar{w} - 1 \right) + \delta_m \left(\ln \beta\delta_m - \bar{w} - 1 \right) \geq 0. \tag{8}$$

Since this condition may hold for some couples and not for others, it may be the case that some couples comply with the rule and some do not.[10] The latter will neither give nor receive attention. The former will do both, but the amount of attention i gives may be different from

[10] Notice that for (8) to be satisfied for some couples it must be the case that $\ln \beta\delta^H > \bar{w} + 1$. We assume that this is satisfied.

the amount i receives, because (δ_i, w_i) may differ from (δ_{h_i}, w_{h_i}) and (δ_{k_i}, w_{k_i}). Furthermore, (8) is a necessary but not a sufficient condition. For the (f, m) couple to want to obey the rule, they must expect that their children will do the same. Actually, (8) is written taking it for granted that the children will comply. Therefore, a condition analogous to (8) must be expected to hold also for each of their children,[11] for each of their children's children, and so on. In addition to (8), it must then be true that

$$\mathbb{E}_t(U'_{d_\ell} - \widehat{U}_{d_\ell} | \delta_f, \delta_m) \geq 0 \qquad \text{with } \; d_\ell \in \{\text{descendants of } (f, m)\}, \forall \ell \geq 1, \tag{9}$$

where ℓ denotes the number of generations that separate the (f, m) couple from their descendant d_ℓ. What this says is that, for each ℓ, d_ℓ complies with the rule.

We want to track how condition (9) and the distribution of δ evolve as a result of random matching. First, however, we must justify the assertion that, if a couple comply with the rule in Definition 2, the amount of attention they give to their respective parents is (5).

3.4 Optimality of the family rule

We want to show that, if the (f, m) couple comply with the rule in Definition 2, it is optimal for $i = f, m$ to give $h_i = F_i, M_i$ the amount of attention $a_i^{h_i} = \frac{\delta^{h_i}}{w_i}$, and for $k_i = D_i, S_i$ to give i the amount of attention $a_{k_i}^i = \frac{\delta^i}{w_{k_i}}$. To do so, we specialize the rule by imposing that the amount of attention due to a parent is equal to the mean of the amounts that the latter gave his or her own parents. Expressing the amount given as a function of the giver's wage rate (and writing the k index without the $i = f, m$ subscript, because f's children are also m's), we then have

$$a_k^i(w_k) = \frac{a_i^{F_i}(w_k) + a_i^{M_i}(w_k)}{2}. \tag{10}$$

[11] Given that siblings have the same preferences and that they face not only the same wage rate distribution, but also the same distribution of potential spouses's preferences and wage rates, the (f, m) couple will expect either that it is in both their children's interest to comply with the rule, or that it is in both their children's interest not to do so.

The allocation which maximizes (6) subject to (7) is (see Appendix 4)

$$\begin{aligned}
s_f &= s_m = 1, \\
T &= \frac{w_m(1 - a_m^{F_m}(w_m) + a_m^{M_m}(w_m)) - w_f(1 - a_f^{F_f}(w_f) + a_f^{M_f}(w_f))}{2} \\
&\quad + \delta_m \left[\left(\psi \ln \beta a_k^m(w^H) + (1-\psi) \ln \beta a_k^m(w^L) \right) \right] \\
&\quad - \delta_f \left[\left(\psi \ln \beta a_k^f(w^H) + (1-\psi) \ln \beta a_k^f(w^L) \right) \right].
\end{aligned}$$

After substituting for T, f's and m's equilibrium expected utilities are

$$\begin{aligned}
U_f &= U_m = U = \frac{w_m(1 - a_m^{F_m}(w_m) - a_m^{M_m}(w_m)) + w_f(1 - a_f^{F_f}(w_f) - a_f^{M_f}(w_f))}{2} \\
&\quad + \ln r - 1 + \delta_m \left[\left(\psi \ln \beta a_k^m(w^H) + (1-\psi) \ln \beta a_k^m(w^L) \right) \right] \\
&\quad + \delta_f \left[\left(\psi \ln \beta a_k^f(w^H) + (1-\psi) \ln \beta a_k^f(w^L) \right) \right].
\end{aligned}$$

Given that the rule will have been formulated before not only D_k's and S_k's, but also f's and m's, wage rates were revealed, and given that f's and m's ancestors share the same expectations regarding w_f and w_m, we find the optimal rule by maximizing the expected value of U. Using (10), this may be written as

$$\begin{aligned}
\mathbb{E}U &= \psi \frac{w^H(1 - 2a_k^m(w^H)) + w^H(1 - 2a_k^f(w^H))}{2} \\
&\quad + (1-\psi) \frac{w^L(1 - 2a_k^m(w^L)) + w^L(1 - 2a_k^f(w^L))}{2} \\
&\quad + \delta_m \left[\left(\psi \ln \beta a_k^m(w^H) + (1-\psi) \ln \beta a_k^m(w^L) \right) \right] \\
&\quad + \delta_f \left[\left(\psi \ln \beta a_k^f(w^H) + (1-\psi) \ln \beta a_k^f(w^L) \right) \right] + \ln r - 1
\end{aligned}$$

If an interior solution exists (otherwise the rule would not be obeyed), it will be (see Appendix 4)

$$a_k^i(w_k) = \frac{\delta_i}{w_k}, i = f, m, \ k = D_i, S_i.$$

Therefore, by implication,

$$a_i^{h_i} = \frac{\delta^{h_i}}{w_i}, i = f, m, h_i = F_i, M_i,$$

as assumed in the last subsection.

3.5 Evolution

We established in Section 2 that, if young men and women distinguishable by an inherited trait δ are matched at random across the entire population, the distribution of this trait will tend to converge to a single value. By contrast, if different segments of the population (the original inhabitants and their descendants, immigrants from a certain part of the world and their descendants) are initially characterized by different distributions of δ, and random matching is restricted to men and women falling in the same segment, different distributions will converge to values of δ.

In the specific model examined in the present section, δ is private monitoring before a couple is made, but it becomes common knowledge where a couple is concerned once the match is made. The couple then decide whether to marry or split. If they marry, they will obey or disobey the rule according to whether (8) and (9) are or are not satisfied. The first of these two conditions concerns the couple. The second concerns their descendants ℓ generations on. Consider, for example, the daughter's line of descendants (the same applies to the son's descendance). As ℓ goes to infinity, (9) tends to

$$\lim_{\ell \to \infty} \mathbb{E}_t(U'_{d_\ell} - \widehat{U}_{d_\ell} | \delta_f, \delta_m) = \tag{11}$$

$$\mathbb{E}_{\delta^*} \left[\delta_f \left(\ln \beta \delta_f - \bar{w} - 1 \right) | \delta_{f,1} \right] + \mathbb{E}_{\delta^*} \left[\delta_m \left(\ln \beta \delta_m - \bar{w} - 1 \right) \right] \geq 0,$$

where $\mathbb{E}_{\delta^*}$ denotes the expected value of $\ln \beta \delta_i - \bar{w} - 1$ under the assumed initial distribution of δ, and $\delta_{f,1}$ is the realization of δ_f in generation d_1. In the long run, if couples are matched at random over the entire population, everybody has the same trait δ^*, and condition (11) can then be re-written as

$$2\delta^* \left(\ln \beta\delta^* - \bar{w} - 1\right) \geq 0. \qquad (12)$$

Notice that (9) may not hold for all $t \geq 0$ and all $l \geq 0$ even if (12) is satisfied, but both these conditions will always hold if $\ln \beta\delta^L > \bar{w} + 1$. With $\ln \beta\delta^L < \bar{w} + 1$, there may exist realizations of δ_i, and (conditional and unconditional) probability distributions for which the expected value of $\ln \beta\delta_i$ is smaller than that of $(\bar{w} + 1)$ for $i = f, m$. Even if that is the case, however, (11) will converge to (12). If the latter is satisfied, there will then exist a generation $\bar{t}$ such that (9) holds for all $t \geq \bar{t}$. Summing up, it is possible that every member of each generation obeys the rule, or that nobody ever does. It is also possible that some members of a generation do not obey the rule, but some of their descendants will. In the long run, if there are no restrictions on whom one can marry, either everybody obeys the rule or nobody does. By backward induction, if nobody obeys in the long run, nobody ever does.

Now suppose that people are distinguishable by a characteristic θ. If θ denotes religion or ethnicity, it could then be the case that marrying somebody with a different θ is regarded by those who practice that religion or recognize themselves in that ethnic group as a sin or betrayal, and that it thus reduces the utility of doing so. Or it could be that a person's utility is not directly affected by the spouse's θ, but there is a different distribution of δ for each value of θ. If the probability that (8) and (9) are both satisfied increases with the spouse's θ, no young man (woman) will then accept to be paired with a young woman (man) recognizable by a θ lower than his (her) own. For either negative or positive motivations, the matching may thus be restricted to population members displaying the same θ.[12] In the long run, there would then be a different δ for each θ. In other words, the population would tend to break down in a number of sharply characterized subpopulations. This prospect may appeal to those who value identity above all else, but it could lead to sectarian conflict. A government concerned about the latter should aim at lowering barriers to cross-marriages (e.g., by investing in public education, and banning schools accessible only to those who practice a certain religion, have

[12] In the extreme case were there is only one δ for each θ, couples would then be positively assorted in δ, and the distribution of this parameter would not evolve. That is effectively the case examined in Cigno et al. (2017).

a particular ancestry, or speak a certain minority language).

4 Discussion

We have shown that, if the members of each sex are differentiated by an inherited trait, and couples are formed at random across the entire population, the variance of the trait in question will gradually diminish as a result of marriage between individuals with different traits, and that everybody will ultimately have the same trait. If people are differentiated also by another trait, and the matching is restricted to young men and women characterized by the same value of this second trait, there will be no convergence towards a common value of the first one. Different groups characterized by different values of the second trait will converge to different values of the first one. This has implications for the consequences of an inflow of immigrants who differ in some way from the original residents.

Using a specific model where the inherited trait is a parameter measuring a person's taste for receiving filial attention (a good without perfect market substitutes) in old age, and individuals are completely selfish, we have also shown that, if the parameter values characterizing a couple satisfy certain conditions, it will be in the couple's common interest to obey a rule requiring them to give specified amounts of attention to their respective parents. The amount given will depend on the receiver's taste for filial attention, and the giver's wage rate. Once again, if couples are formed at random across the entire population, the variance will gradually diminish as a result of intermarriage. In the long run, either everybody will have the same preference parameter and obey the same family rule, or nobody will. Before getting there, however, some couples may obey the rule, and others may not. It is also possible that a couple will obey the rule, but their descendants will not obey it for a few generations, and then obey it again. The reason why couples are formed at random in this model is that a person's taste for filial attention is private monitoring (and her or his wage rate is uncertain) until the couple is actually formed. If young people are differentiated not only by this unobservable trait, but also by an observable one such as ethnicity or religion, however, it is possible that the matching will be restricted to those who belong to the same ethnic group or practice the same religion. If that is the case, the population will tend to break down into a number of sharply characterized subpopulations. That may undermine social cohesion and call for policy intervention.

The present paper bears similarities to a number of other contributions to the literature that also aim to predict how preferences evolve across generations, but it differs from them in one fundamental respect. All those contributions assume that individual preferences have either a paternalistic or a pro-social component, and that the evolution of these preferences has nothing to do with sexual reproduction and intermarriage. Bisin and Verdier (2001), and Tabellini (2008), assume that parents inculcate their preferences (or "values") into their children, and that preferences evolve through socialization only. Alger and Weibull (2013) assume that preferences have an egoistic component, that by itself would lead a person to behave like "homo oeconomicus", and a "Kantian" one, that by itself would drive a person to "do the right thing" if everyone else did the same. Using the evolutionary stability notion developed in Weibull (1995), those authors show that Kantian behaviour may prevail over selfish behaviour in pairwise encounters of a non sexual nature (there is no reproduction, sexual or asexual) if the matching has a certain degree of assortativity. In our model, doing the right thing can be the equilibrium behaviour even if people are perfectly selfish, and men and women are matched at random. All these different contributions are complementary rather than antithetical. Each from its own perspective helps to explain how preferences and behaviours evolve across generations.

5 Appendix 1. Evolution

Proof of Proposition 1. In period $t = 0$, each group (male of female) is partitioned in two subgroups: $n\pi_0(0)$ individuals have the trait δ^L, while $n\pi_0(1)$ individuals have the trait δ^H. Given that siblings cannot marry each other, there are $n(n-1)$ possible couples, where $n = n^L + n^H$.

In period $t = 1$, $S(1) = 3$ traits are possible, that is, $\delta^L, \frac{\delta^L+\delta^H}{2}, \delta^H$. The probability to have a match between two L-types, which gives birth to a male and a female with the trait

δ^L, is

$$\pi_1(0) = \frac{n^L(n^L-1)}{n(n-1)} = \frac{\pi_0(0)(n\pi_0(0)-1)}{n-1} \overbrace{\approx}^{n \text{ large}} \pi_0^2(0)$$

Similarly, the probability to have a match between two H-types, which gives birth to a male and a female with the trait δ^H, is

$$\pi_1(2) = \frac{n^H(n^H-1)}{n(n-1)} = \frac{\pi_0(1)(n\pi_0(1)-1)}{n-1} \approx \pi_0^2(1)$$

Finally, the probability to have a match between a L-type and an H-type, which generates two individuals with the mixed trait $\frac{\delta^L+\delta^H}{2}$, is

$$\pi_1(1) = \frac{2n^H n^L}{n(n-1)} = \frac{2\pi_0(1)\pi_0(0)n}{n-1} \approx 2\pi_0(1)\pi_0(0).$$

At the end of period $t = 1$ there are still n males and n females (grandchildren replace grandparents), however, for each of these groups $\pi_1(0)n$ individuals will have now a trait $\delta_1(0) = \delta^L$, $\pi_1(1)n$ individuals will have a trait $\delta_1(1) = \frac{\delta^L+\delta^H}{2}$, while $\pi_1(2)n$ individuals will inherit a trait $\delta_1(2) = \delta^H$.

In period $t = 2$, $S(2) = 5$ traits are possible, that is, δ^L, $\frac{3\delta^L+\delta^H}{4}$, $\frac{2\delta^L+2\delta^H}{4}$,$\frac{\delta^L+3\delta^H}{4}$, δ^H. The probability to have a match between two L-types, which will preserve the native trait δ^L, is now

$$\pi_2(0) = \frac{\pi_1(0)n[\pi_1(0)n-1]}{n(n-1)} \approx \pi_1^2(0) = \pi_0^4(0)$$

The probability to generate a trait $\delta_2(1) = \frac{3\delta^L+\delta^H}{4}$ is the probability that a L-type meets a type with a trait $\frac{\delta^L+\delta^H}{2}$, that is

$$\pi_2(1) = \frac{2\pi_1(0)n\pi_1(1)n}{n(n-1)} \approx 2\pi_1(0)\pi_1(1) = 4\pi_0^3(0)\pi_0(1)$$

The probability to generate a trait $\delta_2(2) = \frac{2\delta^L+2\delta^H}{4}$ is the probability that two types with trait $\frac{\delta^L+\delta^H}{2}$ meet each other plus the probability that L meets H, that is

$$\pi_2(2) = \frac{\pi_1(1)n[\pi_1(1)n-1]}{n(n-1)} + \frac{2\pi_1(0)\pi_1(2)n}{n-1} \approx$$
$$\pi_1^2(1) + 2\pi_1(0)\pi_1(2) = 6\pi_0^2(0)\pi_0^2(1)$$

The probability to generate a trait $\delta_2(3) = \frac{1\delta^L+3\delta^H}{4}$ is the probability that an H-type meets a type with a trait $\frac{\delta^L+\delta^H}{2}$, that is

$$\pi_2(3) = \frac{2\pi_1(2)n\pi_1(1)n}{n(n-1)} \approx 2\pi_1(2)\pi_1(1) = 4\pi_0^3(1)\pi_0(0)$$

and finally, the probability to generate a match between two H-types, which generates again a native trait δ^H, is

$$\pi_2(4) = \frac{\pi_1(2)n[\pi_1(2)n-1]}{n(n-1)} \approx \pi_1^2(2) = \pi_0^4(1).$$

For a generic generation t, there $S(t) = 2^t + 1$ possible traits,

$$\delta_t(j) := \frac{(2^t - j)\delta^L + j\delta^H}{2^t} = \delta^L + \frac{\delta^H - \delta^L}{2^t} j, \qquad \text{with } j = 0, 1, \ldots, 2^t,$$

identified by a random variable, j, which converges, if n is large, to a binomial distribution $B(2^t, \pi)$, with mean $2^t\pi$ and variance $2^t\pi(1-\pi)$, where we have defined $\pi := \pi_0(1)$, with $\pi_0(0) = 1 - \pi$. In other words, the probability to have j in period t is $\pi_t(j)$, with

$$\pi_t(j) = \binom{2^t}{j} \pi^j (1-\pi)^{2^t - j}, \qquad j = 0, 1, \ldots, 2^t. \tag{13}$$

We now formally prove expression (13). Start by considering $t = 1$. We proceed by induction. Recall that the population is initially distributed in two groups: those with δ^L and those with δ^H, where $(\pi, 1-\pi)$ denotes the initial distribution of δ values. As already shown, in the first generation, $t = 1$, there are $S(1) = 3$ possible traits: the trait δ^L is generated with probability $(1-\pi)^2$, the trait $\frac{\delta^L+\delta^H}{2}$ with probability $2\pi(1-\pi)$, and δ^H with probability π^2. In other words, the distribution of traits in period $t = 1$ is

$$\pi_1(j) = \binom{2^1}{j} \pi^j (1-\pi)^{2^1 - j}.$$

Consider now a generic t. Assume that for the generation in t the $S(t) = 2^t + 1$ traits are distributed according to the binomial distribution (13). We want to show that for the generation in $t+1$ the $S(t+1) = 2^{t+1} + 1$ traits are also distributed according to a binomial distribution, with probabilities

$$\pi_{t+1}(j) = \binom{2^{t+1}}{j} \pi^j (1-\pi)^{2^{t+1} - j}, \qquad j = 0, 1, \ldots, 2^{t+1}.$$

By construction each trait $\delta_{t+1}(j)$ of the generation in $t+1$ is generated by mixing the traits $\delta_t(j')$ and $\delta_t(j'')$ of the previous generation in t, such that $j = j' + j''$. Therefore, the probability of generating $\delta_{t+1}(j)$ is

$$\begin{aligned}
&\sum_{\substack{j',j'' \\ j'+j''=j}} \left[\frac{\pi_t(j') n \pi_t(j'') n}{n(n-1)} \mathbf{1}_{\{j' \neq j''\}} + \frac{\pi_t(j') n [\pi_t(j'') n - 1]}{n(n-1)} \mathbf{1}_{\{j' = j''\}} \right] \\
&\overbrace{\approx}^{n \text{ large}} \sum_{\substack{j',j'' \\ j'+j''=j}} \left[\pi_t(j') \pi_t(j'') \mathbf{1}_{\{j' \neq j''\}} + \pi_t(j') \pi_t(j'') \mathbf{1}_{\{j' = j''\}} \right] \\
&= \pi^j (1-\pi)^{2^{t+1}-j} \Big[\sum_{\substack{j',j'' \\ j'+j''=j}} \binom{2^t}{j'} \binom{2^t}{j''} \mathbf{1}_{\{j' \neq j''\}} + \sum_{\substack{j',j'' \\ j'+j''=j}} \binom{2^t}{j'} \binom{2^t}{j''} \mathbf{1}_{\{j' = j''\}} \Big] \\
&= \pi^j (1-\pi)^{2^{t+1}-j} \sum_{j'=0}^{j} \binom{2^t}{j'} \binom{2^t}{j-j'} \\
&= \binom{2^{t+1}}{j} \pi^j (1-\pi)^{2^{t+1}-j}
\end{aligned}$$

where in the last line we have used the identity

$$\sum_{j'=0}^{j} \binom{2^t}{j'} \binom{2^t}{j-j'} = \binom{2^{t+1}}{j}.$$

Hence, the distribution of the random variable j that identifies each trait of the $S(t+1)$ traits of the generation in $t+1$ is also binomial, with mean $2^{t+1}\pi$ and variance $2^{t+1}\pi(1-\pi)$.

We conclude the proof by noting that each trait $\delta_t(j)$ is a linear transformation of the random variable j. Therefore, the distribution of $\delta_t(j)$ in each period t is given by a binomial distribution with mean $\delta^L(1-\pi)+\delta^H\pi$ and variance $\frac{(\delta^H-\delta^L)^2}{2^t}\pi(1-\pi)$. Clearly, as $t\to\infty$ the variance goes to zero and every individual displays the same trait $\delta^* = \frac{(\delta^H-\delta^L)^2}{2^t}\pi(1-\pi)$.

Appendix 2. Nash-bargaining without family rules

Using the FOCs for the maximization of (2),

$$\frac{\partial N}{\partial T} = (U_m - R) - (U_f - R) = 0,$$

$$\frac{\partial N}{\partial s_f} = \left(-1 + \frac{1}{s_f}\right)(U_m - R) = 0$$

$$\frac{\partial N}{\partial s_m} = \left(-1 + \frac{1}{s_m}\right)(U_f - R) = 0,$$

we find

$$\hat{s}_f = \hat{s}_m = 1 \text{ and } \hat{T} = 0.$$

Substituting $\hat{s}_f$, $\hat{s}_m$ and $\hat{T}$ into the expression for U_f or U_m gives us the equilibrium pay-offs $\widehat{U}_f = R_f$ and $\widehat{U}_m = R_m$.

Appendix 3. Nash-bargaining with family rules

Using the FOCs for the maximization of (6) subject to (7)

$$\frac{\partial N}{\partial T} = \left(U_f - \widehat{U}_f\right) - \left(U_m - \widehat{U}_m\right) = 0,$$

$$\frac{\partial N}{\partial s_f} = \left(-1 + \frac{r}{rs_f}\right)\left(U_m - \widehat{U}_m\right) = 0$$

and

$$\frac{\partial N}{\partial s_m} = \left(-1 + \frac{r}{rs_m}\right)\left(U_f - \widehat{U}_f\right) = 0,$$

we find the Nash-bargaining equilibrium for the case in which f and m obey family rules, and the solution to the Nash-maximization problem subject to these rule is interior (i.e., the amount of filial attention received by f and m is large enough to add to their utility),

$$s'_f = s'_m = 1, \; T' = \delta_m \left(\ln \beta\delta_m - \bar{w} - 1\right) - \delta_f \left(\ln \beta\delta_f - \bar{w} - 1\right).$$

T' is determined so that $\left(U'_f - \widehat{U}_f\right) = \left(U'_m - \widehat{U}_m\right)$. Substituting s'_f and T' into the expression for U'_f or U'_m, we find the equilibrium values of the f's and m's utility.

Appendix 4. Optimality of the family rule

The utilities of f and m are

$$\begin{aligned} U_f &= w_f(1 - a_f^{F_i}(w_f) + a_f^{M_i}(w_f)) - s_f + \ln r s_f + 2\delta_f \left[\left(\psi \ln \beta a_k^f(w^H) + (1-\psi)\ln \beta a_k^f(w^L)\right)\right] + T \\ U_m &= w_m(1 - a_m^{F_i}(w_m) + a_m^{M_i}(w_m)) - s_m + \ln r s_m + 2\delta_m \left[\left(\psi \ln \beta a_k^m(w^H) + (1-\psi)\ln \beta a_k^m(w^L)\right)\right] - T \end{aligned}$$

The FOC's of the maximization of (6) subject to (7) have the form as those in Appendix 3, but the solution under a generic family rule is

$$\begin{aligned} s_f &= s_m = 1, \\ T &= \frac{w_m(1 - a_m^{F_i}(w_m) + a_m^{M_i}(w_m)) - w_f(1 - a_f^{F_i}(w_f) + a_f^{M_i}(w_f))}{2} \\ &\quad + \delta_m \left[\left(\psi \ln \beta a_k^m(w^H) + (1-\psi)\ln \beta a_k^m(w^L)\right)\right] - \delta_f \left[\left(\psi \ln \beta a_k^f(w^H) + (1-\psi)\ln \beta a_k^f(w^L)\right)\right]. \end{aligned}$$

Substituting for T, s_f and s_m into (14) and (14) we obtain the equilibrium utilities

$$\begin{aligned} U_f &= U_m = \frac{w_m(1 - a_m^{F_m}(w_m) + a_m^{M_m}(w_m)) + w_f(1 - a_f^{F_f}(w_f) + a_f^{M_f}(w_f))}{2} + \ln r - 1 + \\ &\quad + \delta_m \left[\left(\psi \ln \beta a_k^m(w^H) + (1-\psi)\ln \beta a_k^m(w^L)\right)\right] + \delta_f \left[\left(\psi \ln \beta a_k^f(w^H) + (1-\psi)\ln \beta a_k^f(w^L)\right)\right]. \end{aligned}$$

The renegotiation proof family rule will maximize

$$\begin{aligned}\mathbb{E}\left(\mathbb{E}U\right) &= \psi\frac{w^H(1-2a_k^m(w^H))+w^H(1-2a_k^f(w^H))}{2} \\ &+(1-\psi)\frac{w^L(1-2a_k^m(w^L))+w^L(1-2a_k^f(w^L))}{2} \\ &+\delta_m\left[\left(\psi\ln\beta a_k^m(w^H)+(1-\psi)\ln\beta a_k^m(w^L)\right)\right] \\ &+\delta_f\left[\left(\psi\ln\beta a_k^f(w^H)+(1-\psi)\ln\beta a_k^f(w^L)\right)\right]+\ln r-1\end{aligned}$$

with respect to a_k^f and a_k^m. The FOCs for an interior solution are

$$\frac{\partial\mathbb{E}\left(\mathbb{E}U\right)}{\partial a_k^i}=-w^J+\delta_i\frac{1}{a_k^i(w^J)}=0 \qquad \text{for } J=L,H$$

which yield

$$a_k^i(w^J)=\frac{\delta}{w^J}, \qquad \text{for } J=L,H.$$

References

Albanese, G., G. De Blasio and P. Sestito (2016), My parents taught me: Evidence on the family transmission of values, *Journal of Population Economics* 29, 571-592

Alesina, A., P. Giuliano and N. Nunn. (2013), On the origins of gender roles: Women and the plough, *Quarterly Journal of Economics* 128, 469–530

Alger, I. and J. W. Weibull (2013), Homo moralis: Preference evolution under incomplete information and assortative matching, *Econometrica* 81, 2269-2302

Bernheim, B. D., A. Schleifer and L. H. Summers (1985), The strategic bequest motive, *Journal of Political Economy* 93, 1045-1076

Bisin, A. and T. Verdier (2001), The economics of cultural transmission and the dynamics of preferences, *Journal of Economic Theory* 97, 298-319

——— and G. Topa (2003), Empirical models of cultural transmission, *Journal of the European Economic Association* 1, 363-375

Bjorklund, A., M. Lindahl and E. Plug (2006), The origins of intergenerational associations: Lessons from Swedish adoption data, *Quarterly Journal of Economics* 121, 999-1028

Callaway, R. M., and B. E. Mahall (2007), Family roots, *Nature* 448, 145-147

Cigno, A. (1991), *Economics of the Family*, New York and Oxford: Oxford University Press and Clarendon Press

———, M. Komura and A. Luporini (2017), Self-enforcing family rules, marriage and the (non)neutrality of public intervention, *Journal of Population Economics* 30, 805–834

Cohen-Zada, D. (2006), Preserving religious identity through education: Economic analysis and evidence from the US, *Journal of Urban Economics* 60, 372-398

Cox, D., and O. Stark (2005), On the demand for grandchildren: Tied transfers and the demonstration effect, *Journal of Public Economics* 89, 1665–1697

Diamond, J. (1997, 2003, 2005), *Guns, Germs, and Steel: The Fates of Human Societies*, New York :W. W. Norton

Dudley, S. and A. L. File (2007), Kin recognition in an annual plant, *Biology Letters* 3, 435-438

Galor O. and Ö. Özak (2016),. The agricultural origins of time preference, *American Economic Review* 106, 3064-3103

Guiso, L., L. P. Sapienza and L. Zingales (2016), Long-term persistence, *Journal of the European Economic Association* 14, 1401-1436

Hazarika, G., C. K. Jha and S. Sarangi (2019), Ancestral ecological endowments and missing women, *Journal of Population Economics* (forth.)

Ottoni-Wilhelm, M., Y. Zhang, D. B. Estell and N. H. Perdue (2017), Raising charitable children: The effects of verbal socialization and role-modeling on children's giving, *Journal of Population Economics* 30, 189-224

Tabellini, G. (2008), The scope of cooperation: Values and incentives, *Quarterly Journal of Economics* 123, 905-950

——— (2019), Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration, Harvard Business School BGIE Unit Working Paper No. 19-005

Weibull, J. W. (1995), *Evolutionary Game Theory*, Cambridge, Massachusetts: MIT Press