I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12373

Evolution of Individual Preferences and Persistence of Family Rules

Alessandro Cigno
Alessandro Gioffré
Annalisa Luporini

MAY 2019

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12373

Evolution of Individual Preferences and Persistence of Family Rules

Alessandro Cigno
University of Florence, CESifo, IZA and CHILD

Alessandro Gioffré
University of Florence, Research Center SAFE and Goethe University Frankfurt

Annalisa Luporini
University of Florence, CESifo, and CHILD

MAY 2019

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Evolution of Individual Preferences and Persistence of Family Rules

How does the distribution of individual preferences evolve as a result of marriage between individuals with different preferences? Could a family rule be self-enforcing given individual preferences, and remain such for several generations despite preference evolution? We show that it is in a couple's common interest to obey a rule requiring them to give specified amounts of attention to their elderly parents if the couple's preferences satisfy a certain condition, and the same condition is rationally expected to hold also where their children and respective spouses are concerned. Given uncertainty about who their children will marry, a couple's expectations will reflect the probability distribution of preferences in the next generation. We show that, in any given generation, some couples may obey the rule in question and some may not. It is also possible that a couple will obey the rule, but their descendants will not for a number of generations, and then obey it again. The policy implications are briefly discussed.

JEL Classification: C78, D13, J12

Keywords: matching, evolution, family rule, care of the elderly, migration

Corresponding author:
Alessandro Cigno
Department of Economics and Management
University of Florence
Via delle Pandette 9 (D6-256)
50127 Firenze
Italy

E-mail: cigno@unifi.it

1 Introduction

The tenet underlying most of microeconomics until not very long ago was that rational individuals with given preferences and endowments optimize subject only to the law of the land. More recently, economists have started to talk of norms or rules, and to examine their implications in different contexts. An early contribution is Cigno (1993), where it is shown that individuals may be constrained by self-enforcing family rules which are themselves a collectively rational response to the economic and legal environment. This line of thought is developed in a series of papers including among others Rosati (1996), Anderberg and Balestrino (2003), and Barnett et al. (2018).[1] Caillaud and Cohen (2000) use the same approach to explain social norms. Another 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 to transmit their values on to their offspring. These values are then modified by the interaction with other individuals who received different inputs from their parents. The implicit assumption underlying all the contributions mentioned is that reproduction is asexual or, equivalently, that the parental couple think and act as if they were one person.

What happens if reproduction is the outcome of the union between two persons of different sex, and mother and father may have different preferences? Cigno et al. (2017) address the issue in the context of a model where the old derive utility from a good, filial attention, that does not have perfect market substitutes. Assuming that individual preferences are common knowledge, the article demonstrates that a young person whose preferences are compatible with the existence of a self-enforcing family rule requiring the young to give attention to their elderly parents (conditional on the latter having obeyed the same rule in their turn) will marry a young person of the opposite sex who holds the same preferences. It also demonstrates that the couple thus formed will transmit their common preferences to their children. This couple and all their descendants will give filial attention when young, and receive it when old. In the present paper, we examine the opposite case where at least one aspect of individual

[1] For empirical evidence, see Cigno et al. (2006), Galasso et al. (2009), Billari and Galasso (2014), and Klimaviciute et al. (2017).

preferences, a person's taste for filial attention, is private monitoring until a marriage takes place, and cannot thus be a criterion for marrying a person rather than another. Of course, couples may well be formed on the basis of some observable characteristic (possibly including a person's taste for other goods), but the matching will in any case be random where the taste for filial attention is *are* concerned. Using a stripped-down version of the model in Cigno et al. (2017), we demonstrate that a couple may obey the rule in question if they expect their children to do the same. The latter is uncertain because the couple do not know whether their children's marriage partners have the "right" preferences.

Assuming rational expectations, we will show that the share of the young who comply with a family rule in any generation is determined simultaneously with the next generation's preference distribution. If all persons of the same sex looked the same, and a young person could thus marry any member of the opposite sex with equal probability, everybody would eventually have the same preferences, and either everybody would then obey the same rule, or nobody would obey any. But, suppose that people are differentiated by some visible characteristic (physical appearance, language, religious practice, etc.). In the long run, if the matching is assortative in that characteristic,[2] all individuals displaying the same characteristic (but not the rest of the population) will hold the same preferences, and consequently either obey the same rule or obey none. In the short run, however, outwardly identical persons could well have different preferences. Given that the same may apply also to different generations within the same line of descent, a rule could fall in abeyance for a number of generations and then spring back to life again. Whether and how many people obey a family rule is important for policy purposes, because only such a rule will deliver filial attention to the old, and certain policies reduce the number of those who obey it.

[2] This hypothesis bears similarities to the one underlying Alger and Weibull (2013), where the matching is assumed to be assortative in a particular component of individual preferences. There, however, the purpose of the matching is not reproduction (there are no children).

2 Basic assumptions

Consider a population consisting of n men and n women, where n is a large number. Each member of this population lives two periods. A person is young in period 1, old in period 2. The young can work and marry, the old can do neither. If a young man and a young woman marry, they have a son and a daughter. Siblings are not allowed to marry. Let $c_{t,i}$ and $a^i_{k_i}$ denote, respectively, the generic individual i's consumption of market goods in period $t = 1, 2$, and the amount of attention that this person receives in period $t = 2$ from $k = D, S$, where D is i's daughter and S is i's son. The utility function is

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

where δ_i is a measure of i's taste for filial attention, and β a scaling factor designed to make $\ln \beta a^i_{k_i}$ positive for $a^i_{k_i}$ sufficiently large.[3] Notice that market goods (including the services of professional helpers) are not perfect substitutes for filial attention. Notice also that neither i nor k is altruistic. Allowing for a modicum of altruism on either side would make the analysis less sharp without altering the results in any substantive way.

Let w_i and w_k denote, respectively, i's and k's wage rates. Before i marries, δ_i and w_i are private monitoring, and w_k is uncertain. We assume that wage rates take value w^H with probability ψ,[4] and w^L with probability $(1 - \psi)$, where $w^H > w^L > 1$. When a couple marry, they observe each other's wage rate and taste for filial attention, but their prospective children's wage rates remain uncertain until the next period. Given that individual wage rates and preference parameters are private monitoring when the couples are formed, and cannot thus be a criterion for the choice of a marriage partner, we assume that a couple is a random draw from the entire population of young men and women (later in the paper we shall allow for the possibility that the matching is restricted to specified sub-populations).

Take the couple formed by a particular woman f, and a particular man m. When the couple is drawn, they may either marry or split (there is no re-sampling). If they split,

[3] Otherwise, $\ln t^i_k$ would be negative for any t^i_k smaller than unity.

[4] This probability could be contingent on educational investments carried out by either i or k_i, but allowing for that would complicate the analysis unnecessarily.

$i = f, m$ maximizes (1) with $a_D^i = a_S^i = 0$, subject to the period budget constraints

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

where s_i is the amount saved by i in period 1, and r is the interest factor. The pay-off of singlehood is then

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

If the couple marry, they Nash-bargain over the allocation of their time and earnings. Having assumed that they are not altruistic, they will neither give attention to their respective parents, nor receive it from their children as a present. They could buy it off their children. Given that filial attention does not have a perfect market substitute, however, their children would form a cartel, and set the price so high that the entire surplus generated by the transaction would go to them.[5] Parents are thus indifferent between buying and not buying filial attention. We assume that they will not. In Section 6 below we establish conditions such that a rule requiring the young to give a specified amount of attention to their elders is self-enforcing and renegotiation-proof. In the following section, however, we examine the behaviour of a couple for whom the said conditions are not satisfied.

3 Bargaining in the absence of a family rule

Suppose that the (f, m) couple are under no obligation to give attention to their elderly parents in period 1, and do not expect to receive attention from their children i period 2. As there will be no re-sampling, i's reservation utility is R_i. The Nash-bargaining equilibrium

[5] 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.

then maximizes

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

subject to $a^i_D = a^i_S = 0$, $i = f, m$, and to the 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 1 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, f and m are indifferent between marrying or splitting. We assume that they marry.

4 Family rules

According to a strand of economic literature stemming from Bisin and Verdier (2001), and Tabellini (2008), cooperative behaviour arises because well-meaning parents expend resources to instill pro-social values into their children. According to Alger and Weibull (2013), by contrast, innate individual preferences have a selfish and an altruistic component. Cooperative behaviour will tend to prevail, according to these authors, if matching is assortative in the weight of the altruistic component of individual preferences. We take a different tack. Our argument is that cooperative behaviour may arise even if individuals are entirely selfish as assumed is Section 2, simply because cooperation pays, and self-enforcing rules may thus emerge. In our specific model (but the argument applies also to other contexts), the most profitable form of cooperation is that between the young and their elderly parents, because that is the only way in which a person can secure filial attention, a good without perfect market substitutes, in old age. The rules in question here are thus family rules.

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

This definition partitions the young 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.

Could the following rule 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.*

Note that a young person is not obliged to give attention to a deviating parent. Having assumed that the young do not get direct utility from giving attention (or anything else) for free, deviating parents may thus be punished by their 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.

In the following sections, we first examine the properties of the Nash-bargaining equilibrium under the assumption that married young people give their elders certain specified amounts of attention. Then we establish conditions such that a rule requiring married young people *to* give their elders the *said* amounts of attention is self-enforcing.

5 Bargaining in the presence of a family rule

Suppose that f and m comply with the rule set out in Definition 2. Let F_i and M_i denote, respectively, i's mother and father,[6] where $i = f, m$. In general, δ_i will be a linear combination of δ_{F_i} and δ_{M_i}. To simplify, however, we assume that

$$\delta_i = \frac{\delta_{F_i} + \delta_{M_i}}{2}. \tag{4}$$

In Section 6 below we demonstrate that the amount of attention i must give $h_i = F_i, M_i$ if the latter is not a deviator is

$$a_i^{h_i} = \frac{\delta_{h_i}}{w_i}.$$

Analogously, the amount of attention that k must give i if the latter is not a deviator is

$$a_k^i = \frac{\delta_i}{w_k}. \tag{5}$$

Notice that the amount of attention i gives in period 1 may be different from the amount that he or she will receive in period 2, because δ_i may be different from δ_{h_i}, and w_i different from w_k. Notice also that, in period 1, w_k is not yet known. Therefore, a_k^i is uncertain and U_i is an expectation.

Given that the best alternative to marrying and obeying the rule is to marry and disobey it, i's reservation utility is $\widehat{U}_i$. 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

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

[6] F for mother and M for father because the former is female and the latter male.

Assuming an interior solution (or the rule would be inoperative), we show in Appendix 2 that the equilibrium is now

$$\begin{aligned} s_i' &= 1, \\ T' &= \delta_m \left(\ln \beta\delta_m - \bar{w} - 1\right) - \delta_f \left(\ln \beta\delta_f - \bar{w} - 1\right), \end{aligned} \tag{8}$$

where

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

Notice that, in contrast with the case examined in Section 3 where no family rule is in force, the compensatory transfer T may now 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). The equilibrium expected utilities are

$$U_i' = w_i + \delta_f \left(\ln \beta\delta_f - \bar{w} - 1\right) + \delta_m \left(\ln \beta\delta_m - \bar{w} - 1\right) - 1 + \ln r,\ \ i = f, m.$$

Notice also that U_f' may differ from U_m' because w_f may differ from w_m.

6 Self-enforcing rules

We must now establish conditions such that the rule laid out in Definition 2 is self-enforcing. Given (5), a necessary condition for (f, m) to want to obey the rule is that they would not be better-off disobeying it,

$$U_i' - \widehat{U}_i = \delta_f \left(\ln \beta\delta_f - \bar{w} - 1\right) + \delta_m \left(\ln \beta\delta_m - \bar{w} - 1\right) \geq 0,\ \ i = f, m. \tag{9}$$

This condition is obviously satisfied if a Nash-bargaining equilibrium exists, because the equilibrium expected utilities cannot be lower than the reservation utilities.[7] That is not enough for the rule to be self-enforcing, however, because the (f, m) couple will not obey

[7]For an equilibrium to exist and (9) to be satisfied for at least some couples, it must be the case that $\ln \beta\delta^H > \bar{w} + 1$. We assume that it is.

the rule if they do not expect their children to do the same. Since the same applies to their children, to their children's children, and so on to infinity, the (f, m) couple must then expect that a condition analogous to (9) will hold also for all their descendants and respective spouses. Again taking (5) as given, a sufficient condition is then that

$$\mathbb{E}_t(U'_{d_\ell} - \widehat{U}_{d_\ell}|\delta_f, \delta_m) \geq 0 \text{ for } t = 0, 1, 2, ..., \text{ with } d_\ell \in \{\text{descendants of } (f, m)\}, \forall \ell \geq 1, \quad (10)$$

where $t = 0$ refers to the (f, m) couple, $t = 1$ to the couples formed by their children, $t = 2$ to those formed by their children's children, and so on; ℓ denotes the number of generations separating the (f, m) couple from their d_ℓ descendant. In any given generation t, this condition may hold for a couple and not for another. Therefore, some may comply with the rule, but some may not. The latter will neither give nor receive attention.

We must now justify the assumption made in the last section, that the amount of attention due to a non-deviating parent is (5). In general, there may be more than one Nash-bargaining equilibrium, one for each specification of $a_i^{h_i}$. Which will apply? A suitable selection criterion is that $a_i^{h_i}$ must be such that the equilibrium associated with it is not Pareto-dominated by any of the alternatives.[8]

To find the undominated rule, we must first derive the solution to the maximization of (6) subject to (7) without imposing (5) . In Appendix 3 we show that this results in an equilibrium level of utility

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

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 are revealed, and given that f's and m's ancestors share the same

[8]This makes the rule renegotiation-proof, meaning that there would be no advantage for any couple in trying to change it. As pointed out in Cigno et al. (2017), a rule thus characterized is the family-level equivalent of what, in the political sphere, is called a constitution.

expectations regarding w_f and w_m, we find the undominated rule by maximizing the expected values of U_f and U_m. The latter turn out to have the same value because the probability that $w_f = w^H$ is the same as that of $w_m = w^H$(and consequently the probability that $w_f = w^L$ is the same as that of $w_m = w^L$) . The common maximand is then

$$\begin{aligned} E\left(U''\right) = \psi w^H + (1-\psi)\, w^L - \frac{\psi w^H (a_m^{F_m}(w^H) + a_m^{M_m}(w^H)) + (1-\psi)\, w^L (a_m^{F_m}(w^L) + a_m^{M_m}(w^L))}{2} \\ - \frac{\psi w^H (a_f^{F_f}(w^H) + a_f^{M_f}(w^H)) + (1-\psi)\, w^L (a_f^{F_f}(w^L) + a_f^{M_f}(w^L))}{2} \\ + \ln r - 1 + \delta_m \left[\left(\psi \ln \beta a_k^m(w^H) + (1-\psi) \ln \beta a_k^m(w^L) \right) \right] \\ + \delta_f \left[\left(\psi \ln \beta a_k^f(w^H) + (1-\psi) \ln \beta a_k^f(w^L) \right) \right],\ i = f, m. \end{aligned} \tag{11}$$

In order to simplify the problem, we specialize the rule by imposing that the amount of attention k must give i is equal to the mean of the amounts that i should have given F_i and M_i if w_i were equal to w_k. Expressing the amount given as a generic function $a_k^i(w_k)$ of the giver's wage rate, we then write

$$a_k^i(w_k) = \frac{a_i^{F_i}(w_k) + a_i^{M_i}(w_k)}{2}. \tag{12}$$

Using (12) , the maximand is now

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

If an interior solution to the maximization of $E\left(U''\right)$ with respect to a_k^f and a_k^m exists, it is (see Appendix 3)

$$a_k^i(w_k) = \frac{\delta_i}{w_k},\ i = f, m,\ k = D, S,$$

as assumed in the last section.

7 Evolution

Random matching implies that the distribution of the preference parameter δ evolves across generations. How will that affect family rules? Will the rule binding a particular couple bind also all their descendants? That is not a trivial question, because self-enforcement depends on expectations, and expectations evolve with the distribution of δ. In the present section we approach the issue under the assumption that couples are drawn at random from the entire population. In the next one we will discuss reasons why that may not be the case.

Suppose that, 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 may be larger as a result of marriages between individuals with different preferences. We have already assumed that everybody marries and that, as every married couple has a daughter and a son, there will be the same number of men and women in each generation. In generation 1, the possible values of δ are δ^L, $\frac{\delta^L+\delta^H}{2}$ and δ^H. Consequently, $S(1) = 2^1 + 1 = 3$. In generation 2, they 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 t, the possible values 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 $\delta = \delta^J$, where $J = H, L$, 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 $\pi_0(1) = \frac{n^H}{n}$, 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 value of δ characterizing that generation will be

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

How does the distribution evolve? Appendix 4 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 δ held by all agents is

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

Therefore, if we set $\delta^H = 1$ and $\delta^L = 0$, the long-run value of the preference parameter is $\delta^* = \pi = \frac{n^H}{n}$.

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 long-run value of 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.

Let us now go back to (10). As ℓ goes to infinity, (10) tends to

$$\lim_{\ell \to \infty} \mathbb{E}_t(U'_{d_\ell} - \widehat{U}_{d_\ell} | \delta_f, \delta_m) = \tag{13}$$
$$\mathbb{E}_{\delta^*}\left[\delta_f \left(\ln \beta\delta_f - \bar{w} - 1\right)\right] + \mathbb{E}_{\delta^*}\left[\delta_m \left(\ln \beta\delta_m - \bar{w} - 1\right)\right] \geq 0,$$

where $\mathbb{E}_{\delta^*}$ denotes the expected value of $\ln \beta\delta_i - \bar{w} - 1$ under the assumed initial distribution of δ. In the long run, when everybody has the same preference parameter δ^*, condition (13) can then be re-written as

$$2\delta^* \left(\ln \beta\delta^* - \bar{w} - 1\right) \geq 0. \tag{14}$$

Notice that, even if (14) is satisfied,(10) may not hold for all $t \geq 0$ and all $l \geq 0$. 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, (13) will still converge to (14). If the latter is satisfied, there will then exist a generation $\bar{t}$ such that (10) holds for all $t \geq \bar{t}$. Summing up, it is possible that every member of each generation obeys the rule, or that none 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, either everybody obeys the same rule or nobody does. By backward induction, if nobody obeys in the long run, nobody will before that. Given that only those who obey a family rule look after their elders, and having assumed that filial attention has no perfect market substitute, this implies that a society where the government taxes the young to buy professional services for the old is welfare-inferior to a society where the young provide their own services to their own parents.

8 Persistence of family rules

Now suppose that all members of a certain population are originally characterized by $\delta = 0$. Suppose that there is 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 $\delta = 1$. According to Table 1, after between five and ten generations, the population will be homogeneous again, and its common characteristics will be very similar to those that were once common to the original inhabitants. In other words, the immigrants will be absorbed by the native population. If the number of immigrants is larger than one ninth, but no larger than a 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 this case, there will be convergence, but not absorption. Whichever is the case, random matching implies that it takes a relatively short time in evolutionary terms (between 130 and 245 years) for a population to become homogeneous again. In our model, this implies that either everybody will ultimately obey a family rule, or nobody will. Is that what we observe in reality? Klimaviciute et al. (2017) report that, despite thousands of years of cross-migrations and, more recently, complete freedom of movement within the EU, working-age Greek, Italian and Spanish people spend, on average, more than 33 hours a month caring for their elderly parents, while the Danish and the Dutch spend less than 11 hours. Cigno et al. (2006) find that, after a century and a half of political union and despite intense internal migrations, the share of the Italian population whose transfer behaviour appears to be regulated by a family rule differs widely across Italian regions.[9] It would thus seem that, despite freedom of movement, matching is not completely random. But why?

Cigno et al. (2017) demonstrate that, if preferences (in particular, the taste for filial attention) were observable, it would be in the interest of a young person whose preferences are compatible with the existence of a self-enforcing family rule to seek out and marry a like-minded member of the opposite sex. As marriage partners would then be assorted according to their preferences, the latter would not evolve, and the share of the population who are

[9] More precisely, the different regional dummies have widely different effects on the probability of giving.

governed by a family rule would be constant. In the last section, we examined the opposite case where individual preferences are private monitoring before the couple are formed, and there is an equal probability of being matched with any member of the opposite sex. But suppose that the distribution of the taste-for-filial-attention parameter δ varies systematically with an observable trait θ denoting, for example, physical type, language or religious practice. If the density function of δ associated with each θ is common knowledge, and the expected value of δ is increasing in θ, a rational young person characterized by a high enough δ will then restrict his or her search to those members of the opposite sex who are characterized by a high θ. Or, if the young are too impulsive to be concerned about what will happen to them in the next period of life, it will be the old who, in their children's (but also their own) interest, try and restrict the range of persons with whom their children come into contact. Choice of school[10] and area of residence are powerful instruments for restricting that range. In the long run, there will then be a different δ for each θ, and the population will tend to break down in a number of sharply characterized subpopulations recognizable by their θ. As the (unobservable) limit value of δ varies with the (observable) value of θ, we will then find that, not only in the short but also in the long run, some θ-types will look after their elders, and some other θ-types will not. Of those who give, some will give more than others with the same wage rate. Given that the different countries forming the European Union, and the different regions forming a single country like Italy (but the same could be said of Belgium, France, Germany, Spain or the UK), are themselves different combinations of ethnolinguistic and religious groups, we have then a possible explanation for the persistent heterogeneity of care-giving patterns reported by Cigno et al. (2006), Klimaviciute et al. (2017), and many others.

9 Discussion

Assuming that a person's taste for filial attention is a linear combination of her or his parents' taste for the same, we have shown that, if the preferences and wage rates characterizing a

[10] Religious schools are an obvious example. Schools restricted to those who speak a particular minority language are another.

couple satisfy a certain condition, and the same condition is expected to hold for the couple's children and their respective spouses, it is in the couple's common interest to obey a rule requiring them to give specified amounts of attention to their respective parents conditional on the latter having obeyed the same rule in their turn. The amount to be given is increasing in the receiver's taste for filial attention, and decreasing in the giver's wage rate. Assuming that a person's taste for filial attention is private monitoring until a couple is actually formed, this taste cannot then be a criterion for forming a couple.[11] We have shown that, if a couple is a random draw from the entire population, the variance of the preference parameter in question will gradually diminish. In the long run, everybody will have the same preferences, and either everybody will obey the same family rule, or nobody will. Before getting there, however, some couples may obey the rule, and others disobey it. As this applies also to members of different generations within the same line descent, the rule may fall in abeyance for a number of generations, and then come back into force again. Alternatively, if the population consists of a number of subpopulations differentiated by a visible trait such as physical appearance, language or religious practice, and sampling is restricted to members of the same subpopulation, each subpopulation will converge to its own limit value of the unobservable preference parameter. Even in the long run, we may then find that some obey a family rule, and some obey none. Among those who obey one, the amount of filial attention given by a person may differ from that given by another, even if the two have the same wage rate, because their taste for the good in question may be different. This prediction is consistent with evidence that the average amount of filial attention given still differs widely across EU member states, and across different regions of the same member state, despite centuries of cross-migrations and decades of free movement. These issues have some policy relevance, because the market does not offer perfect substitutes for filial attention, and the old cannot buy this good from their children.[12] The government should thus refrain from forms of intervention that reduce the number of persons obeying a family rule.[13]

[11] We have assumed that the same applies to individual wage rates, but nothing of substance would change if the latter were observable ex ante, and the matching were assortative in them.

[12] Strictly speaking, as note earlier, they would have to pay such a high price, that they would be indifferent between buying and not buying.

[13] See Cigno (1993), and Cigno et al. (2017).

Our approach differs from those of others who also aim to explain how preferences, rules or values evolve across generations in that those others invariably assume that reproduction is asexual. It differs from that of Bisin and Verdier (2001), and Tabellini (2008), also in that parents do not need to inculcate what we call a rule and those authors call values into their children. If a person follows a rule, it is because it is in her or his interest to do so. Furthermore, individual preferences do not evolve because of social interaction as hypothesized by those authors, but through marriage between persons with different preferences (in reality, preferences may evolve also through social interaction, but we focus on the marriage channel). A contribution that bears some similarities to ours even though it is not concerned with reproduction is Alger and Weibull (2013). Those authors assume that preferences have a selfish component, which by itself would lead a person to behave like "homo oeconomicus", and a “Kantian” one, which 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 social or business encounters (reproduction is not on the agenda) if the matching has a certain degree of assortativity in the non-selfish component of individual preferences. In our model, by contrast, doing the right thing can be the equilibrium behaviour even if people are entirely selfish, and there is no assortativity in preferences. We regard our approach as complementary to those of others who examine the same or similar issues from different standpoints.

Appendix 1. 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 2. 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),

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

T' is determined so that $\left(U'_f - \widehat{U}_f\right) = \left(U'_m - \widehat{U}_m\right)$. Substituting s'_f, s'_m 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 3. Self-enforcement

If we do not impose (5), the utilities of f and m can be written as

$$\begin{aligned} U_f &= w_f(1-a_f^{F_i}(w_f)-a_f^{M_i}(w_f))-s_f+\ln rs_f \\ &+2\delta_f\left[\left(\psi\ln\beta a_k^f(w^H)+(1-\psi)\ln\beta a_k^f(w^L)\right)\right]+T \qquad (15)\\ U_m &= w_m(1-a_m^{F_i}(w_m)-a_m^{M_i}(w_m))-s_m+\ln rs_m \\ &+2\delta_m\left[\left(\psi\ln\beta a_k^m(w^H)+(1-\psi)\ln\beta a_k^m(w^L)\right)\right]-T \qquad (16)\end{aligned}$$

The FOC's of the maximization of (6) subject to (7) have the form as those in Appendix 2, but the solution under a generic family rule is

$$\begin{aligned} s_f'' &= s_m''=1, \\ T'' &= \frac{-w_m(a_m^{F_i}(w_m)+a_m^{M_i}(w_m))+w_f(a_f^{F_i}(w_f)+a_f^{M_i}(w_f))}{2} \\ &+\delta_m\left[\left(\psi\ln\beta a_k^m(w^H)+(1-\psi)\ln\beta a_k^m(w^L)\right)\right]-\delta_f\left[\left(\psi\ln\beta a_k^f(w^H)+(1-\psi)\ln\beta a_k^f(w^L)\right)\right].\end{aligned}$$

Substituting for T, s_f and s_m into (15) and (16) we obtain i's equilibrium utilities

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

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

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 are 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 values of U_f and U_m, which in fact turn out to be the same because the probability of $w_f=w^H$ is the same as that of $w_m = w^H$. Substituting $a_k^i(w_k)$ for $\frac{a_i^{F_i}(w_k)+a_i^{M_i}(w_k)}{2}$. and taking the expectation with respect to the wage rate of f and m, we obtain the objective function that allows to derive the renegotiation proof family rule

$$\begin{aligned} \mathbb{E}U_f'' &= \mathbb{E}U_m'' = \psi w^H \left[1 - a_k^m(w^H) - a_k^f(w^H)\right] \\ &+(1-\psi)w^L \left[1 - a_k^m(w^L)) - a_k^f(w^L)\right] \\ &+\delta_m \left[\left(\psi \ln \beta a_k^m(w^H) + (1-\psi)\ln \beta a_k^m(w^L)\right)\right] \\ &+\delta_f \left[\left(\psi \ln \beta a_k^f(w^H) + (1-\psi)\ln \beta a_k^f(w^L)\right)\right] + \ln r - 1 \end{aligned} \tag{17}$$

with respect to a_k^f and a_k^m. The FOCs for an interior solution of the maximization of (17) with respect to a_k^f and a_k^m, 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

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

10 Appendix 4. Evolution

Proof of Proposition 1. In generation $t = 0$, 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, δ^L, $\frac{\delta^L+\delta^H}{2}$, δ^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

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

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 generation $t=2$, $S(2)=5$ 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,

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

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

$$\pi_t(j) = \binom{2^t}{j} \pi^j (1-\pi)^{2^t - j}, \qquad j = 0, 1, \dots, 2^t. \tag{18}$$

We now formally prove expression (18). 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 (18). 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

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

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 $\delta^* = \frac{(\delta^H-\delta^L)^2}{2^t}\pi(1-\pi)$.

References

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

Anderberg, D. and A. Balestrino (2003), Self–enforcing intergenerational transfers and the provision of education, *Economica* 70, 55-71

Barnett, R. C, J. Bhattacharya and M. Puhakka (2018), Private versus public old-age security, *Journal of Population Economics* 31, 703-746

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

Billari, F. C. and V. Galasso (2014), Fertility Decisions and Pension Reforms. Evidence from Natural Experiments in Italy, IdEP Economic Papers 1403, USI Università della Svizzera Italiana

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

Caillaud, B. and D. Cohen (2000), Intergenerational transfers and common values in a society, *European Economic Review* 44, 1091-1103

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

——— (1993), Intergenerational transfers without altruism: Family, market and state, *European Journal of Political Economy* 9, 505-518

———, G. C. Giannelli, F. C Rosati and D. Vuri (2006), Is there such a thing as a family constitution? A test based on credit rationing, *Review of Economics of the Household* 4, 183-204

———, 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

Galasso, V., R. Gatti and P. Profeta (2009), Investing for the Old Age: Pensions, Children and Savings, *International Tax and Public Finance* 16, 538-559

Klimaviciute, J., S. Perelman, P. Pestieau and J. Schoenmaeckers (2017), Caring for dependent parents: Altruism, exchange or family norm? *Journal of Population Economics* 30, 835-873

Rosati, F. C. (1996), Social security in a non-altruistic model with uncertainty and endogenous fertility, *Journal of Public Economics* 60, 283-294

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

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