Published in final edited form as:

Curr Biol. 2020 October 05; 30(19): 3856–3861.e1. doi:10.1016/j.cub.2020.07.040.

The coupon collection behavior in human reproduction

Erping Long^{1,2}, Jianzhi Zhang^{1,3,*}

¹Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109, USA

²State Key Laboratory of Ophthalmology, Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou 510060, China

³Lead Contact

SUMMARY

There is evidence in humans for genetic influences on the probability (P_{boy}) that a birth yields a boy instead of a girl [1–6], suggesting a potential variation of P_{boy} among families. To quantify this variation, we analyze the survey data from over 300,000 UK Biobank participants primarily born between 1940 and 1970 [7]. Surprisingly, the proportion of male children in a family, or sex ratio (SR), has a significantly smaller among-family variation than expected under a uniform P_{boy} . We propose that this phenomenon results from reproductive behaviors reflecting a preference for having children of both sexes, much like the coupon collector's problem in probability theory where collecting a complete set of distinct coupons is considered a win. We find that the observed deficit in SR variation is explainable by 3.3% of “coupon-collecting” families. Consistently, significantly more families than expected have all children of the same sex except for the child born last. This trend is more pronounced in the late than the early half of the families in the data, suggesting an increasing popularity of this behavior. Analysis of a Dutch genealogical dataset spanning the past four centuries reveals higher-than-expected SR variations over much of the history; only after 1940 did the SR variation drop below the expectation. We conclude that a significant fraction of couples now exhibit the coupon collection behavior in reproduction such that SR is more homogeneous among families than expected by chance.

RESULTS

It is not uncommon to hear that all four or five children of a family are of the same sex, posing the question of whether these families have a particularly low or high P_{boy} relative to 0.5. In fact, as early as 1889, Geissler reported a significant excess of families with children

*Correspondence: jianzhi@umich.edu.

AUTHOR CONTRIBUTIONS

EL and JZ conceived the project, designed the study, and wrote the paper. EL performed the study and analyzed the data.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

DECLARATION OF INTERESTS

The authors declare no competing interests.

of only or predominantly one sex [8], although both confirmatory and contradictory results have subsequently been published [9–12]. More recently, increasing evidence suggests not only a variable P_{boy} among families but also a genetic component in the variance of P_{boy} [3–6] (but see [13, 14]). Furthermore, mutations affecting P_{boy} are known. For example, female carriers of *BRCA1* and *BRCA2* mutations tend to have more female than male children [1, 2]. Given these results, we attempt to assess the variation of P_{boy} among a large number of families.

Sex ratio is more homogeneous among families than the null expectation

The SR of the children in a family, referred to as the SR of the family, is an unbiased estimate of the family's P_{boy} . However, due to the generally small number of children per family, SR is an imprecise estimate of P_{boy} . Nevertheless, the standard derivation (SD) of SR among families is expected to be higher when P_{boy} varies among families than when it is constant, so the SD can be used to assess the among-family variation in P_{boy} .

The UK Biobank comprises genotype and phenotype data of about 0.5 million British participants [7]. Specifically, the data include the sex of each participant as well as the respective numbers of the participant's male and female siblings, permitting computing SR for the set of the participant and his/her siblings (referred to as a family). We excluded participants with invalid or unreliable sibling records, or genetic kinship to other participants in the Biobank (see STAR Methods), to reach the study set of 332,788 unrelated participants with sibling records (including those with no sibling). The P_{boy} from the 332,788 families combined is 0.5032, and the SD of SR among families is 0.329. To obtain the expected SD under the null hypothesis of a uniform P_{boy} across families, we randomly shuffled individuals among families without altering the number of children in each family and then computed the null SD . This process was repeated 1000 times. The distribution of the 1000 null SD values has a mean of 0.332 and a standard deviation of 0.00024 (Figure 1A). The SD observed from the Biobank is 12.5 standard deviations away from its expectation and is smaller than all 1000 null SD values (Figure 1A). Furthermore, the fraction of girl-only families in the Biobank is significantly lower than expected from the above shuffled families (Figure 1B). The same is true for boy-only families (Figure 1C). Thus, our analysis found no evidence for SR overdispersion (i.e., P_{boy} variation) among families. Rather, it revealed even more homogenous SR values across families than the random expectation under a uniform P_{boy} .

The higher-than-expected homogeneity in SR is explainable by the coupon collection behavior in reproduction

What could have caused the unexpectedly high homogeneity of SR among families? We hypothesize that parents prefer having children of both sexes such that they continue to reproduce until they have children of both sexes. This is analogous to the coupon collector's problem in probability theory [15], where one continues to buy a particular brand of cereal, each package of which contains one randomly placed coupon, until he or she collects all different types of coupons in the set. The hypothesized reproductive behavior is expected to render the SR of a family less extreme, which should reduce the among-family variation in SR . Our hypothesis is consistent with reports that Europeans express in interviews their

desires to have children of both sexes [16, 17] and the observation in Denmark, Sweden, and Finland that the probability for a mother to have a third child is higher when her first two children are of the same sex than when they are of different sexes [18–20]. However, it is still unknown to what extent the hypothesized coupon collection behavior influences the SD of SR and whether this behavior is changing over time. Note that coupon collection is different from the male-preferring stopping (MPS) [21]. In MPS, a family will not stop reproducing until a predetermined number of boys are born. MPS cannot explain the simultaneous deficiencies of girl-only families (Figure 1B) and boy-only families (Figure 1C), so cannot explain our observations.

To assess the impact of the hypothesized coupon collection reproductive behavior on the SD of SR , we simulated child births in each of the 332,788 Biobank families by strictly following the coupon collection behavior and using $P_{\text{boy}} = 0.5032$. That is, reproduction stops in a family when and only when the family has had at least one boy and at least one girl. We found that the simulated families have 3 children on average and have a SD of SR equal to 0.168. For comparison, we randomly shuffled the children among these simulated families without altering the number of children in each family and repeated this process 1000 times. The distribution of the 1000 SD values from the shuffled families has a mean of 0.311 and a standard deviation of 0.00029. Hence, the SD of SR of the families simulated under the coupon collection behavior is 493 standard deviations smaller than the expectation without this behavior (Figure 2A). In other words, the coupon collection behavior can in principle explain the observed deficit in the SD of SR among Biobank families.

In the above simulation, we assumed that all families strictly follow the coupon collection reproductive behavior and observed a substantial deficit in the SD of SR (Figure 2A). In reality, the difference between the observed SD and expected SD is much smaller (Figure 1A), suggesting that only a fraction of families exhibit this behavior. To estimate this fraction, we first randomly selected one shuffled dataset in Fig. 1A with the SD of SR equal to 0.332, the mean value from the 1000 sets of shuffled families. We then randomly chose a certain fraction (f) of families, identified among them those that did not conform to the coupon collection behavior, and adjusted the children in each of them. Specifically, for families that did not stop reproducing after having had children of both sexes, we removed the additional children; and for families with children of only one sex, we added children one at a time with $P_{\text{boy}} = 0.5032$ until the family has children of both sexes. We then recomputed the SD of SR among the 332,778 families. Clearly, the SD of SR gradually declined as f rose, and the SD reduced to the Biobank observed value of 0.329 when f reached 3.3% (i.e., 11,000 families) (Figure 2B). Because the above adjustment altered the number of children per family and potentially the null SD , we randomly shuffled the children among these 11,000 families and found that the SD among all 332,778 families returned to 0.3318, which is close to the original level (0.332). This result confirms that enforcing the coupon collection behavior in 3.3% of families can explain the observed deficit in SD of SR . Repeating the entire estimation 100 times, we found that the mean f equals 3.29%, with the 95% confidence interval from 3.28% to 3.31%.

Evidence for the coupon collection behavior in reproduction

While the above section demonstrated that the coupon collection reproductive behavior can explain the observed deficit in the SD of SR among Biobank families, it does not provide direct evidence for this behavior. Our hypothesis of the coupon collection behavior predicts that, compared with the random expectation represented by the preceding shuffled families, there is an excess of families with at least one boy but only one girl and similarly an excess of families with at least one girl but only one boy. Our observations from the Biobank support these predictions (Figure 3A, B).

A more definitive test of the coupon collection hypothesis is to examine whether a couple are more likely to cease reproduction after they have had children of both sexes than when they have had the same number of children but all of the same sex. The Biobank sibling record does not provide birth order data except for the information of whether the participant is the youngest child in a family. We thus focused on the subset of families where the participant is the youngest in a family and all of his or her siblings are of the same sex. We first stratified these families by the number (n) of children per family and then divided them into two groups for each $n \geq 2$. In each family belonging to group I, all children are of the same sex except for the child born last, who has the opposite sex. In each family belonging to group II, all children are of the same sex. For example, when $n = 3$, every group I family has an ordered sex configuration of MMF (first child is male, second child is male, and third child is female) or FFM, whereas each group II family has an ordered sex configuration of MMM or FFF. Under the null hypothesis that the reproductive behavior is independent of the sexes of the existing children, the expected fraction of families belonging to group I is a function of n and P_{boy} that can be computed from a binomial distribution (see STAR Methods). For example, the expected fraction equals 0.5 regardless of n when $P_{\text{boy}} = 0.5$ and is very close to 0.5 for any realistic n when $P_{\text{boy}} = 0.5032$ (Figure 3C). The coupon collection reproductive behavior should increase the number of families belonging to group I relative to that belonging to group II, so should increase the fraction of group I families. Indeed, this fraction is significantly greater than the null expectation when $n = 2, 3$, or 4 (Figure 3C). When $n = 5$ or 6, the fraction is still greater than the null expectation, although the difference is no longer significant because of the much reduced sample sizes as n increases (Figure 3C). To examine whether the above pattern has changed over time, we divided all of the above families into two birth cohorts by the birth year of the child born last in the family (i.e., the Biobank participant). Because the Biobank participants were primarily born between 1940 and 1970, the older cohort consists of families in which the child born last in each family was born before 1955 while the younger cohort consists of families in which the child born last in each family was born on or after 1955. Compared with the older cohort, the younger cohort has a significantly higher fraction of group I families (Figure 3D), suggesting a rising popularity of the coupon collection reproductive behavior in the three decades considered.

The coupon collection reproductive behavior in the Netherlands

To confirm the coupon collection reproductive behavior and its rising popularity in other populations, we examined a family tree-based genealogical database that contains 241,000 Dutch families over four centuries, with known number and genders of children of each

family [22] (see STAR Methods). We first focused on a subset of 12,547 families in which the youngest child in each family was born between 1940 and 1970, matching the birth year range in the UK Biobank data. The P_{boy} from these families combined is 0.522, and the SD of SR among families is 0.336. Using the same approach employed in the analysis of the UK Biobank data, we found this observed SD to be significantly smaller than that expected under equal P_{boy} among families ($P = 0.011$).

To investigate potential changes of SD over time in the Netherlands, we first divided the 241,000 families into 30-year birth cohorts according to the birth year of the youngest child in each family (1600–1629, 1630–1659, etc.). We excluded families in which the youngest child was born after 1995, because these families may not have completed reproduction by the end of data collection in 2017. We also combined the birth cohorts with fewer than 10,000 families. In the end, we analyzed six birth cohorts (1600–1819, 1820–1849, 1850–1879, 1880–1909, 1910–1939, and 1940–1995). We found that the observed SD s of SR are significantly greater than the null expectation in the first four cohorts that spanned over 300 years, fluctuating in a narrow range between 0.001 and 0.002 above the null expectation (Figure 4). In the fifth cohort, however, the observed SD departed from this range and was no longer significantly above the null expectation, suggesting a downward move of SD from its historic average (Figure 4). In the sixth cohort, the observed SD became significantly below the null expectation. In fact, the absolute value of the SD 's deviation from its null expectation is greater in the sixth cohort than in all previous cohorts, signaling a profound difference in the among-family variation of SR after 1940 when compared with the preceding 3.4 centuries (Figure 4). There were no corresponding changes in the population level sex ratio over the same time (Figure S1), suggesting that our observation is unrelated to potential population level sex ratio changes. Thus, the analysis of the Dutch data confirms the finding from the UK Biobank and suggests that the coupon collection reproductive behavior is a relatively recent phenomenon.

DISCUSSION

In this work, we attempted to assess the among-family variation in P_{boy} using the large survey data of the UK Biobank. Contrary to our expectation, SR is significantly less variable among families than expected under a uniform P_{boy} . However, this finding does not reject the hypothesis that P_{boy} varies among families, because it is possible that forces increasing the SR variance (e.g., P_{boy} variation) are overwhelmed by forces reducing the SR variance (e.g., the coupon collection behavior). Indeed, SR was significantly greater than the null expectation for the first three quarters of the last four centuries documented by a Dutch genealogical dataset. Regardless, the observation of a significantly lower SR variance than the expectation under a uniform P_{boy} in modern populations substantially complicates the estimation of the present among-family variation in P_{boy} and the potential genetic influence.

We proposed and provided evidence that the deficit in the SR variation among families in the UK Biobank is caused by the coupon collection reproductive behavior and estimated that 3.3% of families exhibit this behavior. Note that this estimate was obtained under the null model of a uniform P_{boy} among families. If P_{boy} actually varies among families, which is supported by some studies [1–6] and our analysis of the Dutch data (Figure 4), the true

fraction of families following the coupon collection behavior should exceed 3.3%. Furthermore, reproductive decisions are influenced by many factors such that, even when a family prefers to have children of both sexes, it may not be able to demonstrate the coupon collection behavior, for example, because of financial, childcare, or physiological constraints. Hence, the fraction of families that prefer to have children of both sexes is likely substantially greater than 3.3%.

The reason behind the coupon collection behavior in reproduction and the generality of this finding across different societies are unclear. It is notable that our finding is mainly based on the British people represented by the UK Biobank, where the vast majority of participants were born between 1940 and 1970. Furthermore, the coupon collection behavior is more prevalent in the later cohort than the earlier cohort in the Biobank. Analysis of a smaller dataset from the Netherlands that spans four centuries shows that the *SD* of *SR* was steadily and significantly above the null expectation from the beginning of the 17th century to the early 20th century, at which time it started to decline. Only after 1940 did the *SD* become significantly below its null expectation. We thus speculate that the coupon collection reproductive behavior originated from society-wide improvement in gender equality and appreciation of gender diversity [23]. Furthermore, the development of modern contraception methods in the 20th century may have helped to realize the coupon collection behavior—stopping reproduction after having had children of both sexes [24]. In the future, it would be interesting to expand our analysis to more countries and different times that represent various degrees of gender equality and/or levels of appreciation of gender diversity. For example, it is well known that, in some Asian countries, especially in the countryside, boys are preferred over girls [25] and antenatal murders of girls occur [26]. The coupon collection behavior is likely rare in such societies.

It is worth noting that the mean number of children born to a woman in UK has now declined to around 1.7 [27]. Because the expected number of children per family is 3 under the coupon collection behavior, it would be interesting to study whether this decline in family size signals other interests triumphing the desire of having children of both sexes. Large surveys of families younger than those in the UK Biobank will likely inform whether the prevalence of the coupon collection reproductive behavior is still increasing in the UK and other developed nations.

STAR METHODS

RESOURCE AVAILABILITY

Lead Contact—Further information and requests should be directed to and will be fulfilled by the Lead Contact, Jianzhi Zhang (jianzhi@umich.edu).

Materials Availability—There is no material associated with this study.

Data and Code Availability—The UK Biobank data used are available at <http://www.ukbiobank.ac.uk/about-biobankuk/> and the Dutch data are available at <https://hdl.handle.net/10411/WJN9GR>. Custom code used in the analysis is available upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

UK Biobank participants—The UK Biobank data comprise about 0.5 million participants aged from 40 to 70 years, recruited between 2006 and 2010 in 22 assessment centers throughout the UK, and followed up for a variety of health conditions from their recruitment date till February 17, 2016 or their date of death [7]. Informed consent was obtained from all participants [7]. The current study was approved by UK Biobank (reference no. 48678), and the analyses presented were based on data from 488,377 individuals accessed through the UK Biobank on April 23, 2019.

Dutch family records from a genealogical database—The Dutch records are based on marriages conducted between 1600 and 1999 in the Netherlands and include information on the birth years and sexes of all children of each couple [22]. This dataset is derived from a genealogical database constructed from family tree files contributed by users of www.genealogieonline.nl for scientific research. The contributed records were error-checked before being included in the final database. Duplicate families were identified by the year of marriage and the first seven letters of the couple's surname; only one randomly picked family was retained among duplicate families. Each child had a unique identifier to avoid duplication.

METHOD DETAILS

We first excluded individuals with invalid or unreliable sibling records, or genetic kinship to others in the UK Biobank, to reach the study set of 332,788 unrelated individuals with known sibling records. For each of these participants, his/her sex, the number of full brothers, the number of full sisters, and the number of elder siblings were recorded in the Biobank. We used the above information to compute the *SR* of each Biobank family including the participant and his/her siblings. In addition, when the number of elder siblings equals the total number of full brothers and sisters, the participant is the youngest child in the family, so we know the sex of the youngest child. For other families, we do not generally know the sex of the youngest child.

Let us consider the subset of families where the participant is the youngest in a family and all of his or her siblings are of the same sex. We first stratify these families by the number (n) of children per family and then divide them into two groups for each $n \geq 2$. In group I, each family has all children of the same sex except for the child born last, who has the opposite sex. In group II, each family has all children of the same sex. The expected number of families belonging to group I, relative to the total number of families of the two groups is $\left[P_{\text{boy}}^{n-1}(1 - P_{\text{boy}}) + (1 - P_{\text{boy}})^{n-1}P_{\text{boy}} \right] / \left[P_{\text{boy}}^{n-1} + (1 - P_{\text{boy}})^{n-1} \right]$, which equals 0.5 when $P_{\text{boy}} = 0.5$. Influences of preferences regarding the sex of children on family size and *SR* were previously studied theoretically [28].

QUANTIFICATION AND STATISTICAL ANALYSIS

Because the entire study is based on quantification and statistical analysis, they have been described in the main text and Method Details.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Corry Gellatly for assistance in accessing and interpreting the Dutch genealogical database. Members of the Zhang lab and three anonymous reviewers provided valuable comments. This work was supported in part by U.S. National Institutes of Health research grant GM103232 to J.Z.

REFERENCES

1. de la Hoya M, Fernandez JM, Tosar A, Godino J, Sanchez de Abajo A, Vidart JA, Perez-Segura P, Diaz-Rubio E, and Caldes T (2003). Association between BRCA1 mutations and ratio of female to male births in offspring of families with breast cancer, ovarian cancer, or both. *JAMA* 290, 929–931. [PubMed: 12928470]
2. Moslehi R, Singh R, Lessner L, and Friedman JM (2010). Impact of BRCA mutations on female fertility and offspring sex ratio. *Am. J. Hum. Biol* 22, 201–205. [PubMed: 19642207]
3. Gellatly C (2009). Trends in population sex ratios may be explained by changes in the frequencies of polymorphic alleles of a sex ratio gene. *Evol. Biol* 36, 190–200.
4. Trichopoulos D (1967). Evidence of genetic variation in the human sex ratio. *Hum. Biol* 39, 170–175. [PubMed: 6056273]
5. Curtsinger JW, Ito R, and Hiraizumi Y (1983). A two-generation study of human sex-ratio variation. *Am. J. Hum. Genet* 35, 951–961. [PubMed: 6614009]
6. Ryu D, Ryu J, and Lee C (2016). Genome-wide association study reveals sex-specific selection signals against autosomal nucleotide variants. *J. Hum. Genet* 61, 423–426. [PubMed: 26763874]
7. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature* 562, 203–209. [PubMed: 30305743]
8. Geissler A (1889). Beiträge zur Frage des Geschlechterverhältnisses der Geborenen. *Ztschr. d. K. Sächs. Bureaus* 35, 1–24.
9. Edwards AW (1961). A factorial analysis of sex ratio data. *Ann. Hum. Genet* 25, 117–121. [PubMed: 13889267]
10. Gini C (1908). Il sesso dal punto di vista statistico: le leggi della produzione dei sessi, (Milan: Sandron).
11. Rife DC, and Snyder LH (1937). The distribution of sex ratios within families in an Ohio city. *Human Biol.* 9, 99–103.
12. Myers RJ (1949). Same sexed families. *J. Hered* 40, 268–270. [PubMed: 15393920]
13. Catalano R, Casey JA, and Bruckner T (2020). A test of oscillation in the human secondary sex ratio. *Evol Med Public Health* 2020, eoaa012.
14. Zietsch BP, Walum H, Lichtenstein P, Verweij KJH, and Kuja-Halkola R (2020). No genetic contribution to variation in human offspring sex ratio: a total population study of 4.7 million births. *Proc. Biol. Sci* 287, 20192849. [PubMed: 32070249]
15. Isaac R (1995). *The Pleasures of Probability*, (New York: Springer-Verlag).
16. Mills M, and Begall K (2010). Preferences for the sex-composition of children in Europe: A multilevel examination of its effect on progression to a third child. *Pop. Stud* 64, 77–95.
17. Hank K, and Kohler H-P (2000). Gender preferences for children in Europe: Empirical results from 17 FFS countries. *Demogr. Res* 2, 133–144.
18. Hank K (2007). Parental gender preferences and reproductive behaviour: a review of the recent literature. *J. Biosoc. Sci* 39, 759–767. [PubMed: 17156588]
19. Jacobsen R, Moller H, and Engholm G (1999). Fertility rates in Denmark in relation to the sexes of preceding children in the family. *Hum. Reprod* 14, 1127–1130. [PubMed: 10221253]

20. Andersson G, Hank K, and Vikat A (2007). Understanding parental gender preferences in advanced societies: Lessons from Sweden and Finland. *Demogr. Res* 17, 135–156.
21. Yamaguchi K (1989). A formal theory for male-preferring stopping rules of childbearing: sex differences in birth order and in the number of siblings. *Demography* 26, 451–465. [PubMed: 2792480]
22. Gellatly C, Störmer C, Boele A, and De Moor T (2017). Life histories of persons marrying, between 1600 and 1999, and dying, in the Netherlands [GO924], 10.34894/MILU3S, DataverseNL, V1.
23. Anderson BS, and Zinsser JO (1988). *A History of Their Own: Women in Europe from Prehistory to the Present*, (New York: Harper and Row).
24. Jütte R (2008). *Contraception: A History*, (Malden, MA: Polity Press).
25. Rajan S, Nanda P, Calhoun LM, and Speizer IS (2018). Sex composition and its impact on future childbearing: a longitudinal study from urban Uttar Pradesh. *Reprod Health* 15, 35. [PubMed: 29486802]
26. Atif K, Ullah MZ, Afsheen A, Naqvi SA, Raja ZA, and Niazi SA (2016). Son Preference in Pakistan; A Myth or Reality. *Pak J Med Sci* 32, 994–998. [PubMed: 27648055]
27. Office for National Statistics (2018). Births in England and Wales: 2018, <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthsummarytablesenglandandwales/2018>
28. Sheps MC (1963). Effects on family size and sex ratio of preferences regarding the sex of children. *Popul. Stud* 17, 66–72.

Figure 1. Sex ratio (SR) is more homogeneous among families than expected under a uniform P_{boy} .

(A) Frequency distribution of the standard deviation (SD) of SR from 1000 sets of randomly shuffled UK Biobank families (blue bars). In each shuffled set, all individuals are shuffled among families while all family sizes are unaltered. The red arrow indicates the observed SD from the actual data.

(B) Frequency distribution of the number of families with $SR = 0$ from the above 1000 sets of randomly shuffled families (blue bars). The red arrow indicates the observed number, which is smaller than that in 97.3% of sets of shuffled families.

(C) Frequency distribution of the number of families with $SR = 1$ from the above 1000 sets of randomly shuffled families (blue bars). The red arrow indicates the observed number, which is smaller than that in all sets of shuffled families.

Figure 2. The higher-than-expected among-family homogeneity in SR is explainable by the coupon collection reproductive behavior.

(A) The coupon collection behavior reduces the SD of SR among families. The red arrow indicates the observed SD among 332,788 families simulated by strictly following the coupon collection reproductive behavior. Blue bars show the frequency distribution of the SD of SR in 1000 sets of families randomly shuffled from the above simulated families.

(B) Relationship between the fraction (f) of families following the coupon collection behavior in reproduction and the SD of SR among families. The red horizontal line shows the observed SD of SR in the Biobank, while the red arrow indicates the corresponding f .

Figure 3. Evidence for the coupon collection behavior in reproduction.

(A) Frequency distribution (blue bars) of the number of families with at least one boy but only one girl in 1000 sets of randomly shuffled Biobank families. The red arrow indicates the observed number in the actual data.

(B) Frequency distribution (blue bars) of the number of families with at least one girl but only one boy in 1000 sets of randomly shuffled Biobank families. The red arrow indicates the observed number in the actual data.

(C) The relative fraction of group I families (blue area) exceeds the expectation under no coupon collection reproductive behavior, indicated by the red dotted line, whereas the relative fraction of group II families (grey area) is smaller than the expectation. For each bar, numbers of families considered are indicated inside the bar, the ordered sex configurations for the grey and blue areas are shown above and below the bar, respectively, and the family size (n) considered is indicated below the bar. P -values are from binomial tests.

(D) Numbers of group I and group II families in the early and late halves of the data. The ratio of the number of group I families to that of group II families (the third row) is significantly different between the two halves of the data ($\chi^2 = 7.75$, $P = 0.005$).

Figure 4. Changes in the *SD* of *SR* among Dutch families over the last 400 years.

Shown is the difference between the observed *SD* and the expected *SD* determined from shuffled families. The error bar depicts the standard deviation of this difference based on 1000 sets of shuffled families. *, $P < 0.05$; **, $P < 0.001$. P -value is determined by the proportion of 1000 sets of shuffled families with larger (the first five bins) or smaller (the last bin) *SD* than observed. The red line indicates zero difference. See also Figure S1.

KEY RESOURCES TABLE

RESOURCE	SOURCE	IDENTIFIER
Data		
Family records in UK Biobank	[7]	http://www.ukbiobank.ac.uk/about-biobank-uk/
Dutch genealogical database	[22]	https://hdl.handle.net/10411/WJN9GR