Supplement to:

The Trivers-Willard hypothesis: sex ratio or investment?

Carl Veller[1,2,4] David Haig[1] Martin A. Nowak[1,2,3,4]

Appendix S1

We show in Appendix S1 that the unique optimal sex ratio of a mother in condition c_m is male-biased (resp., female-biased) if and only if she would prefer, in terms of fitness, to have an all-male brood than an all-female brood (resp., all-female than all-male).

We first prove the reverse direction. Let the fitness sum of an all-male brood (given optimal investment), a mixed brood (ditto), and an all-female brood (ditto) be M, X, and F respectively. Suppose that the mother would prefer an all-male brood to an all-female brood: $M > F$. Denoting by p the proportion of males in the sex ratio, the problem of an optimal sex ratio is then

$$p^* := \arg\max_{p\in[0,1]} \left[p^2 M + 2p(1-p)X + (1-p)^2 F\right].$$

There are two possible cases: either the optimum is interior, i.e., $p^* \in (0,1)$, or it is boundary, i.e., $p^* = 0$ or $p^* = 1$. If it is boundary, then it must be male-biased ($p^* = 1$), since the alternative boundary solution ($p = 0$) is strictly inferior ($M > F$). A necessary condition for the optimum to be interior is $X > M$ since if this condition does not hold, the optimal sex ratio would be fully male-biased, $p^* = 1$. If interior, the first order condition

[1]Department of Organismic and Evolutionary Biology, Harvard University, Massachusetts, U.S.A.
[2]Program for Evolutionary Dynamics, Harvard University, Massachusetts, U.S.A.
[3]Department of Mathematics, Harvard University, Massachusetts, U.S.A.
[4]carlveller@fas.harvard.edu; martin_nowak@harvard.edu

is

$$\frac{\partial}{\partial p}\Big(p^2 M + 2p(1-p)X + (1-p)^2 F\Big)\Big|_{p=p^*} = 0 \qquad \Rightarrow \qquad p^* M + (1-2p^*)X - (1-p^*)F = 0$$
$$\Rightarrow \qquad p^* = \frac{X - F}{[X-M]+[X-F]} > \frac{1}{2},$$

where the inequality is a consequence of $X > M > F$. So the optimal sex ratio when $M > F$ is male-biased.

To prove the forward direction, let $p^* > 1/2$ be the unique optimal sex ratio for a mother in condition c_m. If $F > M$, then the alternative sex ratio $1-p^*$ would yield higher average brood fitness sum, since it would generate the same proportion of mixed broods as does p^*, but a higher proportion of the fitter same-sex brood (all-female). If $F = M$ and $X > M$, then the optimal sex ratio is $p^* = 1/2$, since this generates the highest proportion of mixed broods. If, instead, $F = M$ and $X \leq M$, then the optimal sex ratio is not unique. Thus, $p^* > 1/2 \Rightarrow M > F$.

By a symmetric argument, the unique optimal sex ratio is female-biased if and only if $F > M$.

Appendix S2

Here we show that the single-crossing condition, where $f_♂(c)$ starts below $f_♀(c)$ for low c, crosses $f_♀(c)$ at some unique value c^*, and lies above $f_♀(c)$ for all $c > c^*$, is in general not sufficient to guarantee that the sex ratio version of the TWH holds. Our proof is graphical, with a particular specification of $f_♂$ and $f_♀$ (Fig. A1).

Suppose, owing to convexity in the appropriate region of $f_♂$, that there is some c_m^1 such that the optimal investment decision for a mother of condition c_m^1 who has two sons is to invest asymmetrically in them, resulting in adult conditions c_1 and c_3, and thus fitness sum $f_1 + f_3$. Investing more in the son whose condition is c_3 would be unproductive, since fitness increases very little for adult conditions above this. Also for this reason, a mother who is in condition c_m^2 slightly larger than c_m^1, and who has two sons, invests such that one

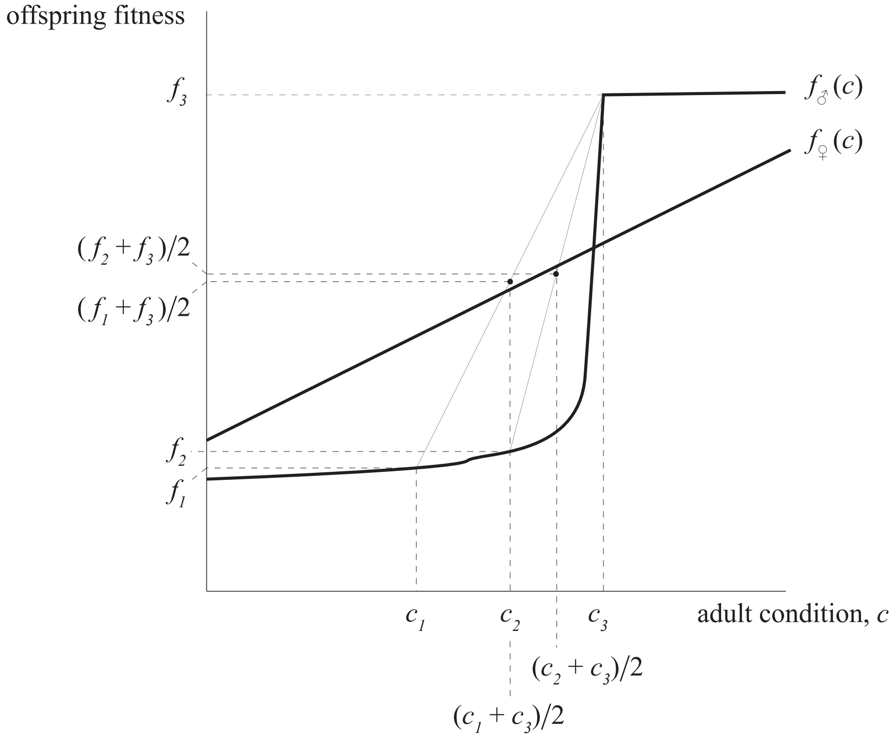

Figure A1: A graphical example of fitness functions satisfying the single-crossing conditions (SC1) and (SC2), but for which the sex ratio version of the Trivers-Willard hypothesis does not hold.

son achieves adult condition c_3, while the other receives what is left of the mother's initial investment endowment $I(c_m^2)$. With this investment, it achieves adult condition $c_2 > c_1$, and the fitness sum of the brood is $f_2 + f_3$.

Since the female fitness curve is concave in investment (being linear in adult condition, which in turn is concave in investment), the optimal investment decision for mothers in condition c_m^1 and c_m^2 who each have two female offspring is to apportion investment equitably within their respective broods. One can find functions c, consistent with the general conditions imposed upon such functions in Section 2 in the main text, such that this leaves the female offspring in conditions $(c_1 + c_3)/2$ for the mother in condition c_m^1, and $(c_2 + c_3)/2$ for the mother in condition c_m^2. In the former case, the brood's fitness sum is lower than that of a same-condition mother's male brood, since the female fitness profile passes *below* the leftmost dot marking average offspring fitness $(f_1 + f_3)/2$. But in the latter case, the brood's fitness sum is greater than that of a same-condition mother's male brood, since the female fitness profile passes *above* the rightmost dot marking average offspring fitness $(f_2 + f_3)/2$. Thus, mothers in condition c_m^1 are selected to exhibit a male-biased sex ratio, while mothers in condition $c_m^2 > c_m^1$ are selected to exhibit a female-biased sex ratio, in contravention of the sex ratio version of the TWH.

Appendix S3

Here, we provide a brief analysis of a simple two-period sequential brood model. We assume that a mother has one offspring in each of the two periods, that her condition c_m does not change across the two periods, and that, starting with investment capability $I(c_m)$, if she invests $i \in [0, I(c_m)]$ in first offspring, that she has investment capability $I(c_m) - i$ left over for the second offspring (which, when optimizing, she of course exhausts). There are two possibilities with regard to the mother's sex ratio: First, that she can alter her sex ratio across the two periods. Second, that she cannot, and must choose a sex ratio that applies to both.

Case 1: Changeable sex ratio

In the first case, where her sex ratio can change from period to period, it is never strictly preferable for a mother to have, in either period, a sex ratio that is not zero or one: instead, she deterministically sets the sequence of her offspring's sex. Here, the analysis is almost identical to the single-brood model considered in the main part of this paper. It is shown in Appendix S1 that, in the single-brood model, the optimal sex ratio is mixed (i.e., not zero or one) if, and only if, the fitness sum of a mixed brood, given optimal investment allocation, is greater than both the fitness sum of a male-only and a female-only brood (also given optimal investment allocation). Otherwise, the optimal sex ratio is biased entirely towards the sex with higher same-sex-brood fitness sum. Therefore, the links between the one-period and two-period models are the following: When the optimal sex ratio in the one-period model is completely male-biased (resp. female biased), then the optimal sequence of offspring sex in the two-period model is male-male (resp. female-female). In cases where the optimal sex ratio in the one-period model is mixed, the optimal sequence of offspring in the two-period model is male-female, or, equivalently (same fitness sum), female-male. The optimal pattern of investment is then the same as in the one-period model.

Case 2: Unchangeable sex ratio

In the second case, where the sex ratio is constrained to be fixed across the two periods, the difference between the one-period and two-period models is that, in the former, the realization of the two offspring's sex is simultaneous, and so the mother can condition her allocation of investment perfectly on it. In the latter, if her sex ratio is not zero or one, her investment decision must be made with knowledge only of the first offspring's sex.

It is clear that, for maternal conditions for which the optimal sex ratio in the one-period model is zero or one, the optimal sex ratio in the two-period model would be identical (as would the investment decisions then be).

Consider then the case where the mother's optimal sex ratio is between zero and one. First, we shall show that, for examples such as that in Section 3 of the main text (the 'classic' Trivers-Willard fitness functions), the investment version still fails.

Take the linear fitness functions from Section 3, and consider a mother with a mixed optimal sex ratio. In the one-period model, if the mother has a same-sex brood, she apportions investment equitably among the two offspring, but if she has a mixed brood, she invests more in the male offspring. In the two-period model, suppose now that the mother first has a son. Should she give it half of her investment endowment (as would be optimal if she knew that her next offspring would be a son), or should she give it the amount – more than half – that she would if she knew her next offspring would be a daughter? It seems clear that the answer will lie somewhere in between, and will depend on the mother's sex ratio.

Formally, the optimization problem facing a mother of condition c_m and with sex ratio $p \in (0,1)$ who has had a son in the first period, and must now decide how much to invest in him and how much to leave aside for her next offspring, is:

$$\begin{aligned}
i_s^* &:= \arg\max_i \; f_{♂}\big(c(c_m, i)\big) + p f_{♂}\big(c(c_m, I(c_m) - i)\big) + (1-p) f_{♀}\big(c(c_m, I(c_m) - i)\big) \\
&= \arg\max_i \; \big[\lambda c(c_m, i)\big] + p\big[\lambda c(c_m, I(c_m) - i)\big] + (1-p)\big[c(c_m, I(c_m) - i) + k\big],
\end{aligned}$$

the first-order condition for which is:

$$\begin{aligned}
& \frac{\partial}{\partial i}\Big(\big[\lambda c(c_m, i)\big] + p\big[\lambda c(c_m, I(c_m) - i)\big] + (1-p)\big[c(c_m, I(c_m) - i) + k\big]\Big)\Big|_{i=i_s^*} = 0 \\
\Rightarrow \quad & \lambda \partial_2 c(c_m, i_s^*) - p\lambda \partial_2 c(c_m, I(c_m) - i_s^*) - (1-p)\partial_2 c(c_m, I(c_m) - i_s^*) = 0 \\
\Rightarrow \quad & \lambda \partial_2 c(c_m, i_s^*) = \big[\lambda - (1-p)(\lambda - 1)\big]\partial_2 c(c_m, I(c_m) - i_s^*) \\
\Rightarrow \quad & \partial_2 c(c_m, i_s^*) < \partial_2 c(c_m, I(c_m) - i_s^*) \quad \Rightarrow \quad i_s^* > I(c_m) - i_s^*. \qquad (1)
\end{aligned}$$

Similarly, if a mother with a mixed sex ratio first has a female offspring, she will provide it with less than half of her investment endowment. So, also in the two-period model, the investment version of the TWH does not work for this set of fitness functions. Notice that the factor on the right-hand side of the third-to-last line, $\lambda - (1-p)(\lambda - 1) = 1 + p(\lambda - 1)$, is greater than its counterpart, 1, from the first-order condition of the one-period model (see Section 3 of main text), and so the optimal bias in investment is smaller here, in accordance with the heuristic argument above.

Analysis of the sex ratio version is complicated by the fact that the optimal sex ratio of a mother will depend on her optimal investment allocation across (sequential) broods, and this in turn will depend on her sex ratio. That is, the fitness sums of the various sequential broods that result from optimal investment, say MM, MF, FM, FF, are all functions of the sex ratio p. So, the problem of optimizing the sex ratio is:

$$p^* := \arg\max_p \; p^2 MM(p) + p(1-p)MF(p) + p(1-p)FM(p) + (1-p)^2 FF(p),$$

the first-order condition of which appears to be intractable (though note again that the first-order condition is relevant only for interior solutions, and so applies only to maternal conditions where, in the one-period model, the optimal sex ratio is mixed).

Nonetheless, we shall derive some preliminary results that suggest this to be the case. We focus again on the linear case from Section 3 of the main text, assume that offspring condition depends only on investment received (to simplify the notation), and specify a functional form for this relationship: $c(i) = \ln(i)$ (which satisfies the restrictions imposed in Section 2 of main text). Our strategy will be similar to that employed in Section 5 of the main text: we shall ask, starting from an even sex ratio, where selection 'points' for mothers of different conditions. In particular, we shall ask if selection points towards a male-biased sex ratio for mothers above a particular condition, and towards a female-biased sex ratio for mothers below that condition.

We noted above that the first-period investment decision of a mother depends both on the sex of the first offspring and on her sex ratio. Let $i_s(p, c_m)$ be the optimal investment in the first offspring of a mother in condition c_m with sex ratio p who has a son in the first period, and $i_d(p, c_m)$ the analogous quantity if she has a daughter instead. Then

$$
\begin{aligned}
i_s(p, c_m) \quad &= \arg\max_i \left(\lambda \ln(i) + p\left[\lambda \ln\left(I(c_m) - i\right)\right] + (1-p)\left[\ln\left(I(c_m) - i\right)\right]\right), \\
i_d(p, c_m) \quad &= \arg\max_i \left(\ln(i) + k + p\left[\lambda \ln\left(I(c_m) - i\right)\right] + (1-p)\left[\ln\left(I(c_m) - i\right)\right]\right),
\end{aligned}
$$

the first-order conditions (with respect to i) for which are

$$
\begin{aligned}
& 0 \equiv \frac{\lambda}{i_s(p, c_m)} - p\frac{\lambda}{I(c_m) - i_s(p, c_m)} - (1-p)\frac{1}{I(c_m) - i_s(p, c_m)} \\
\Rightarrow \quad & I(c_m) - i_s(p, c_m) \equiv \frac{p\lambda + 1 - p}{\lambda} i_s(p, c_m), \qquad (2) \\
& 0 \equiv \frac{1}{i_d(p, c_m)} - p\frac{\lambda}{I(c_m) - i_d(p, c_m)} - (1-p)\frac{1}{I(c_m) - i_d(p, c_m)} \\
\Rightarrow \quad & I(c_m) - i_d(p, c_m) \equiv (p\lambda + 1 - p) i_d(p, c_m). \qquad (3)
\end{aligned}
$$

Since we are considering marginal deviations from an even sex ratio, we write $p = \frac{1}{2} + \varepsilon$, and consider the case where ε is small in magnitude: $|\varepsilon| \ll 1$. The optimal investment decision functions are continuous (because the optimization problem is smooth), and so, as $\varepsilon \to 0$, $i_s(p, c_m) \to i_s(1/2, c_m)$ and $i_d(p, c_m) \to i_d(1/2, c_m)$. In this case, the expected fitness sum of a mother in condition c_m with sex ratio $p = \frac{1}{2} + \varepsilon$, with ε small, and when she is making optimal investment decisions, is approximately

$$
\begin{aligned}
F(\varepsilon, c_m) = \quad & \left(\frac{1}{2} + \varepsilon\right)^2 \Big[\lambda \ln[i_s(1/2, c_m)] + \lambda \ln[I(c_m) - i_s(1/2, c_m)]\Big] \\
+ \quad & \left(\frac{1}{2} + \varepsilon\right)\left(\frac{1}{2} - \varepsilon\right) \Big[\lambda \ln[i_s(1/2, c_m)] + \ln[I(c_m) - i_s(1/2, c_m)] + k\Big] \\
+ \quad & \left(\frac{1}{2} + \varepsilon\right)\left(\frac{1}{2} - \varepsilon\right) \Big[\ln[i_d(1/2, c_m)] + \lambda \ln[I(c_m) - i_d(1/2, c_m)]\Big] \\
+ \quad & \left(\frac{1}{2} - \varepsilon\right)^2 \Big[\ln[i_d(1/2, c_m)] + k + \ln[I(c_m) - i_d(1/2, c_m)] + k\Big].
\end{aligned}
$$

Ignoring terms of order ε^2,

$$
\begin{aligned}
F(\varepsilon, c_m) = \quad & \left(\frac{1}{4} + 2\varepsilon\right) \Big[\lambda \ln[i_s(1/2, c_m)] + \lambda \ln[I(c_m) - i_s(1/2, c_m)]\Big] \\
+ \quad & \frac{1}{4}\Big[\lambda \ln[i_s(1/2, c_m)] + \ln[I(c_m) - i_s(1/2, c_m)] + k\Big] \\
+ \quad & \frac{1}{4}\Big[\ln[i_d(1/2, c_m)] + \lambda \ln[I(c_m) - i_d(1/2, c_m)]\Big] \\
+ \quad & \left(\frac{1}{4} - 2\varepsilon\right) \Big[\ln[i_d(1/2, c_m)] + k + \ln[I(c_m) - i_d(1/2, c_m)] + k\Big],
\end{aligned}
$$

so that

$$
\begin{aligned}
\frac{\partial}{\partial \varepsilon} F(\varepsilon, c_m) = \quad & 2\Big[\lambda \ln[i_s(1/2, c_m)] + \lambda \ln[I(c_m) - i_s(1/2, c_m)]\Big] \\
- \quad & 2\Big[\ln[i_d(1/2, c_m)] + k + \ln[I(c_m) - i_d(1/2, c_m)] + k\Big]. \qquad (4)
\end{aligned}
$$

From Eqs. (2) and (3), evaluated at $p = 1/2$, this can be rewritten:

$$
\begin{aligned}
\frac{\partial}{\partial \varepsilon} F(\varepsilon, c_m) = \quad & 2\left[\lambda \ln[i_s(1/2, c_m)] + \lambda \ln\left(\frac{\lambda+1}{2\lambda} i_s(1/2, c_m)\right)\right] \\
- \quad & 2\left[\ln[i_d(1/2, c_m)] + k + \ln\left(\frac{\lambda+1}{2} i_d(1/2, c_m)\right) + k\right] \\
= \quad & 4\big(\lambda \ln[i_s(1/2, c_m)] - \ln[i_d(1/2, c_m)]\big) + A,
\end{aligned}
\tag{5}
$$

where $A = 2[(\lambda - 1)\ln(1/2) + (\lambda - 1)\ln(\lambda + 1) - \lambda \ln(\lambda)] - 4k$ is a constant.

The sex ratio version of the TWH will hold, in terms of the direction of selection from an initial even sex ratio, if there is a maternal condition c_m^* such that: (i) $\frac{\partial}{\partial \varepsilon} F(\varepsilon, c_m^*) = 0$, (ii) $\frac{\partial}{\partial \varepsilon} F(\varepsilon, c_m) > 0$ if $c_m > c_m^*$, and (iii) $\frac{\partial}{\partial \varepsilon} F(\varepsilon, c_m) < 0$ if $c_m < c_m^*$. This in turn would hold if there is a maternal condition c_m^* such that $\frac{\partial}{\partial \varepsilon} F(\varepsilon, c_m^*) = 0$, and $\frac{\partial}{\partial c_m} \frac{\partial}{\partial \varepsilon} F(\varepsilon, c_m) > 0$ for all c_m.

From Eq. (5),

$$
\frac{1}{4} \frac{\partial}{\partial c_m} \frac{\partial}{\partial \varepsilon} F(\varepsilon, c_m) = \frac{\lambda}{i_s(1/2, c_m)} - \frac{1}{i_d(1/2, c_m)},
$$

which is positive if, and only if, $\frac{i_s(1/2, c_m)}{\lambda i_d(1/2, c_m)} < 1$. From Eqs. (2) and (3), evaluated at $p = 1/2$, we have that

$$
\frac{i_s(1/2, c_m)}{\lambda i_d(1/2, c_m)} = \frac{I(c_m) - i_s(1/2, c_m)}{I(c_m) - i_d(1/2, c_m)},
$$

and from our consideration of the general investment problem above (Eq. (1)), we know that $I(c_m) - i_s(1/2, c_m) < I(c_m)/2$ while $I(c_m) - i_d(1/2, c_m) > I(c_m)/2$. Therefore, $\frac{I(c_m) - i_s(1/2, c_m)}{I(c_m) - i_d(1/2, c_m)} < 1$, whence $\frac{i_s(1/2, c_m)}{\lambda i_d(1/2, c_m)} < 1$, and so $\frac{1}{4} \frac{\partial}{\partial c_m} \frac{\partial}{\partial \varepsilon} F(\varepsilon, c_m) > 0$ for all c_m.

So, conditional on there being a c_m^* such that $4\big(\lambda \ln[i_s(1/2, c_m^*)] - \ln[i_d(1/2, c_m^*)] - k\big) + A = 0$, the sex ratio version of the TWH holds in this case.