

## **1. Actual and marginal costs and benefits**

In our model, *B* and *C* are the actual benefits and costs exchanged in an interaction. We

suppose that the dependence of *B* on *C* exhibits diminishing returns; that is, for *B* both

 positive and negative, the |*B*/*C*| ratio decreases as *C* increases. In each case the actor chooses the desired (*B*,*C*) values based on the information it has about its partner. To

check whether a current (*B*,*C*) choice for a particular actor-recipient pair is stable under

24 the action of selection, we look at the fitness of a mutant strategy playing  $(B+b, C+c)$  for

small increments *b* and *c* and these are what the inclusive-fitness increment employs. The

 ratio *b*/*c* can be approximated by the derivative d*B*/d*C* at the base point (*B*,*C*). The mutant 27 ratio  $b/c$  is the slope of the tangent line at the point and the actual ratio  $B/C$  is the slope of the secant line drawn from the origin (Fig. 1 of the main text). For any particular

29 interaction, the evolutionary equilibrium is attained at a point at which  $W_{\text{IF}} = Rb - c = 0$ , that is, where the slope  $b/c$  is equal to 1/*R*. that is, where the slope  $b/c$  is equal to  $1/R$ .

 The geometry of the curve can be used to show that such a point must be 32 convergence stable [1]. We make the argument using the upper branch of the curve  $(B >$  0). Starting at such a point, if we increase *B* and *C* (i.e. move up the curve), the *b*/*c* ratio 34 decreases (i.e. diminishing returns) so that  $b/c < 1/R$  and that implies  $Rb < c$  which, in 35 turn, implies  $W_{\text{IF}} < 0$ . Thus for a higher value of *B* and *C*, mutants with positive *b* and *c*  are less fit than residents, and selection moves *B* and *C* to lower values. A similar argument shows that a decrease in *B* and *C* favours mutants with positive *b* and *c*, moving *B* and *C* to higher values. An analogous argument holds for the lower branch  $(B < 0)$ .

## **2. Generality of Assertions**

 Our purpose here is to establish (and make precise) the three principal assertions made in 43 the paper: (1) that altruism evolves more easily among migrant than native actors; (2) that spite evolves more easily among native than migrant actors; and (3) that under some circumstances, natives will pay costs that surpass those they spitefully impose on others. We make the standard assumption on the *B*-*C* trade-off curve that the level of both altruism and spite increases with cost but exhibits diminishing returns; in terms of Fig. S1, that the top half of the curve (altruism) is increasing and concave-down and the bottom half of the curve (spite) is decreasing and concave-up. We also assume, as seems 50 reasonable, that buds are relatively small; that is, that  $q < \frac{1}{2}$ .

 For assertions (1) and (2), we compare the behaviour of a migrant and a native actor when they have the same coefficient of consanguinity (CC) with their partners. That is, we compare the behaviour of a migrant actor interacting with a signal *s* partner with 54 that of a native actor with a signal *s'* partner, where  $G_N(s') = G_M(s)$ . Let  $\rho_I(s)$  be the 55 evolutionarily stable marginal cost-benefit  $(c/b)$  ratio for a native  $(I = N)$  or migrant  $(I =$  M) in an interaction with a partner with signal *s*. These are found by inverting eqs. (7) 57 and (8) in the main text. We get:

$$
\rho_{\rm N}(s) = \frac{G_{\rm N}(s) - (1 - q)}{q} \tag{S1}
$$

$$
\rho_{\mathcal{M}}(s) = \frac{G_{\mathcal{N}}(s) - q}{1 - q} \tag{S2}
$$

60

61 Now let *τ*(*s*) be the solution *s'* to the equation  $G_N(s') = G_M(s)$ . We will show that, 62 when  $q < \frac{1}{2}$ ,

$$
\rho_{\rm N}(\tau(s)) < \rho_{\rm M}(s). \tag{S3}
$$

64

63

65 From (S1)

66

$$
q\rho_{N}(\tau(s)) = G_{N}(\tau(s)) - (1 - q)
$$
  
=  $G_{M}(s) - (1 - q)$   
=  $[(1 - q)\rho_{M}(s) + q] - (1 - q)$   
=  $(1 - q)\rho_{M}(s) - 1 + 2q$ ,

67 then

68

69  $q[\rho_N(\tau(s)) - \rho_M(s)] = (1-q)\rho_M(s) - 1 + 2q - q\rho_M(s)$ <br>70  $= (1-2q)\rho_M(s) - (1-2q)$ 70  $=(1-2q)\rho_M(s)-(1-2q)$ 

71  $=(1-2q)(\rho_M(s)-1)$ 72

73 and this is negative for  $q < \frac{1}{2}$  since  $\rho_M(s)$  is clearly less than 1 from eq. (S2). That establishes (S3). establishes (S3).

75 The  $\rho_I(s)$  are the evolutionarily stable *c/b* ratios so that their reciprocals are the *b/c*  ratios, which interpret as slopes in figure 2a of the main text. For the case of native and 77 migrant altruism  $(B > 0, C > 0$ ; upper branch of the curve in figure S1a), condition (S3) tells us that the slope of the trade-off curve at the native evolutionarily stable strategy (ESS) is greater than the slope at the migrant ESS and, since the curve is concave-down, this tells us the migrant has the higher (*B*,*C*) value and thus pays more to give a higher benefit. This supports assertion (1).



83 84

85 **Figure S1.** Three scenarios for interactions when  $K = 5$ ,  $q = 1/K$  (no budding) and  $G_N(s')$  $= G_M(s)$  along the benefit-cost trade-off curve  $C = B^2$ . Values represent evolutionarily<br>87 stable (*B*,*C*) pairs for native (blue squares) and migrant actors (red circles) for (*a*)  $G =$ stable (*B*,*C*) pairs for native (blue squares) and migrant actors (red circles) for (*a*)  $G =$ 88 0.9, (*b*)  $G = 0.5$ , and (*c*)  $G = 0.1$ . The upper branch of each curve represents altruistic 89 behaviour ( $B > 0$ ,  $C > 0$ ) and the lower branch represents spiteful behaviour ( $B < 0$ ,  $C > 0$ ) 90 0). Both types of actor can be altruistic (*a*), migrants can be altruistic whereas natives can 91 be spiteful (*b*), or both types can be spiteful (*c*).

92<br>93 A second possibility for (S3) would have migrant altruism  $(B > 0, C > 0$ ; upper 94 branch of the curve in figure S1b) and native spite  $(B < 0, C > 0)$ ; lower branch in figure 95 S1b), so that the first slope is positive and the second is negative. Observe that we cannot 96 reverse these and have migrant spite and native altruism. This gives support to both assertions (1) and (2). assertions  $(1)$  and  $(2)$ .

98 Third, for the case of native and migrant spite  $(B < 0, C > 0)$ ; lower branch of the 99 curve in figure S1c), the *c/b* ratios represented by  $\rho_I(s)$  are negative so that the native slope *b/c* in (S3) is greater than the migrant slope. For negative slopes, that means that slope  $b/c$  in (S3) is greater than the migrant slope. For negative slopes, that means that the 101 native slope is less steep than the migrant slope and hence (since the lower branch of the 102 curve is concave-up) the native ESS is the farther from the origin. Thus, the native actor 103 pays a greater cost to exact a greater harm than does the migrant, giving support to 104 assertion (2).

 Finally we look at assertion (3). We need to get hold of the set of signals *s* for which a native actor will pay a greater cost *C* than the inflicted harm –*B*. These are the 107 signals for which  $C > -B$ , which we dub 'extreme' spite, and a standard geometric argument will show that since the lower branch of the trade-off curve is concave up, it 109 must eventually cross and lie above the diagonal line  $-B = C$ . The question is whether, for large enough *s*, the native (*B*,*C*) point will actually cross that line and enter the section 111 (coloured in green in figure S2) where  $C > -B$ . Well, as *s* gets large,  $G_N(s)$  will certainly approach zero, and from eq. (S1), the slope of the curve,  $1/\rho_N(s)$  will approach  $-q/(1-q)$ . approach zero, and from eq. (S1), the slope of the curve,  $1/\rho_N(s)$  will approach  $-q/(1-q)$ . In most cases, for reasonably small *q*, we can expect this to send the point into the extreme spite zone of the curve, but examples of trade-off curves can be constructed where this will not be the case.

116





118

**Figure S2.** The case of 'extreme' spite for the benefit-cost trade-off curve  $C = B^2$ . The 120 diagonal line  $C = -B$  is drawn in black. Beyond this line, along the green portion of the 121 benefit-cost trade-off curve, actors pay costs that exceed those they impose on recipients 122  $(C > -B)$ . 123

However, for the curve  $C = B^2$  we can certainly calculate the *s* value at which this will 125 happen. For this equation:

127 
$$
\rho_{\rm N}(s) = \frac{c}{b} = \frac{\mathrm{d}C}{\mathrm{d}B} = 2B \tag{S4}
$$

126

128<br>129 and the condition  $C > -B$  becomes

130 131  $B < -1$ 

132 
$$
\rho_N(s) = 2B < -2
$$
 from (S4)

133 
$$
G_N(s) - (1-q) < -2q
$$
 from (S1)

134 
$$
P_N(s) = G_N(s) < 1-3q
$$
 from eq. (4)

135 
$$
\frac{P_N(s)}{P_M(s)} = \frac{P_N(s)}{1 - P_N(s)} < \frac{1 - 3q}{1 - (1 - 3q)} = \frac{1 - 3q}{3q}
$$

136 
$$
\frac{D_N(s)(1-q)}{D_M(s)q} < \frac{1-3q}{3q}
$$
 from eq. (6)

137 
$$
\frac{D_N(s)}{D_M(s)} < \frac{1-3q}{3(1-q)}
$$
 (S5)

 $\text{since } P_M(s) = 1 - P_N(s)$ 

 Condition (S5) gives us the critical *s*-value for what we have called extreme spite, 140 that being the score value  $s^*$  at which the ratio of the density functions  $D_1(s)$  is  $(1-3q)/(3-1)$  $3q$ ). When a native encounters a partner with a signal  $s > s^*$ , we can expect it to act with extreme spite. Since the critical ratio (1–3*q*)/(3–3*q*) decreases with *q*, the critical score *s\** increases with *q*, and thus small *q* (small relative bud size *q* or large deme size *K*) promotes extreme spite. For example, for deme size *K* = 5 and no budding (*q =* 1/5), the 145 critical ratio is  $2/12 = 1/6$  such that the signal is at least 6 times as likely to belong to a migrant than to a native. This gives a critical score of *s\** = 1.90, as illustrated in the signal distribution drawn in figure S3. This critical *s*-value is less than the migrant mean (*s* = 2); that tells us that in more than half of all native-migrant encounters, the native will pay a cost that exceeds the inflicted harm. This supports assertion (3).



 

 **Figure S3.** Example probability density distribution distributions of native (blue curve) 154 and migrant (red curve) signals. The dashed line marks  $s = 1.90$ ; more than half of all migrants bear a signal above this value. migrants bear a signal above this value. 

Figure 1 of the main text has equation  $C = B^2$  and we use this to provide a 158 numerical example. For this equation,  $dB/dC = 1/2B$  and from this we deduce that  $b/c =$ *B*/2*C* so that the secant has twice the slope of the tangent. Suppose the deme has  $K = 10$  breeders and there is no budding (*q =* 1/10) and consider a native actor that knows her partner is a migrant. According to table 1 ('marginal ESS') of the main text, she would (as a 'mutant') be willing to pay a higher cost if that could increase the harm done to the migrant by as much as one ninth of the cost increase. Thus at equilibrium, the harm done by a native actor to a known migrant should be two ninths of the cost she bears ('actual ESS'). Conversely the harm done by a migrant actor to a known native would have to be eighteen times the cost sustained by the actor. A native is thus much more likely to inflict harm on a known migrant than a migrant is to inflict harm on a known native.

 In figure S4 we use a simulation to compare the behaviour of a native and a 169 migrant actor interacting with partners with signals  $s_N$  and  $s_M$ , respectively, of<br>170 comparable consanguinity (i.e.  $G_N(s_N) = G_M(s_M)$ ). When faced with partners of comparable consanguinity (i.e.  $G_N(s_N) = G_M(s_M)$ ). When faced with partners of the same consanguinity, native actors in this simulation are less altruistic or more spiteful than

 migrant actors and, at high signal values, they evolve to pay costs that surpass those they 173 spitefully impose on their partners  $(-B < C, B < 0)$ . Conversely, when migrant actors are

174 spiteful (at small  $s_M$ ) they always pay smaller costs than those that they impose  $(-B > C$ , 175  $B < 0$ ).

## **3. Ethnicity and Hate Crime**

 'Ethnicity' can mean many things, but here we use it to mean a phenotypic marker for (broad) consanguinity. Thus, we conceive as 'ethnic' any cues to recent ancestry, language, religion, cultural heritage, or national origin [2]. We searched for literature on ethnically motivated 'hate' crimes in which both offender and victim ethnicity were reported. We found reports published by five sources from which we could glean such information: the Los Angeles County Human Relations Commission (for the years 1999 to 2012, though offender ethnicity data are available only until 2006); the Chicago Police Department (for the years 1995, 1997, 2000, 2002, 2004, 2006, 2008, and 2010); the US Federal Bureau of Investigation (for the years 1996 to 2012); the Scottish Government (for the years 2003 to 2013); and the UK Home Office, Ministry of Justice, and the Office of National Statistics (for the year 2012 only; it covers England and Wales and provides information on offender but not victim ethnicity) (3). For consistency within and across sources, we use victim and offender 'race' as our marker of ethnicity.

 For these five populations, we averaged over the available years the proportion of ethnically motivated hate crimes with ethnic minority victims and the proportion of such crimes with ethnic majority offenders. These proportions are presented in Table S1, alongside the local majority ethnic group (categorized by 'race') and the proportion of the local population represented by this group (4).

**Table S1.** Ethnic composition by population and by victim-offender status.



 Table S1 shows that members of the ethnic majority tend more often to be the offenders than the victims of ethnically motivated hate crimes whereas the opposite is true of members of ethnic minority groups. Indeed, there is a strong correlation between the proportion of ethnically motivated hate crimes committed by members of the majority 205 ethnic group and the prevalence of this group in the population  $(r = 0.91, p = .032)$ . This supports the hypotheses that (1) individuals bearing common phenotypes will tend to impose more harm on partners of rare phenotypes than the converse and (2) that this pattern is magnified by the frequency of individuals bearing the majority phenotype.

