Showing that you care: The evolution of health altruism

Robin Hanson*

Department of Economics, George Mason University, MSN 1D3, Carow Hall, Fairfax VA 22030, United States

Received 24 August 2007; accepted 25 August 2007

Summary

Human behavior regarding medicine seems strange; assumptions and models that seem workable in other areas seem less so in medicine. Perhaps, we need to rethink the basics. Toward this end, I have collected many puzzling stylized facts about behavior regarding medicine, and have sought a small number of simple assumptions which might together account for as many puzzles as possible.

The puzzles I consider include a willingness to provide more medical than other assistance to associates, a desire to be seen as so providing, support for nation, firm, or family provided medical care, placebo benefits of medicine, a small average health value of additional medical spending relative to other health influences, more interest in public that private signals of medical quality, medical spending as an individual necessity but national luxury, a strong stress-mediated health status correlation, and support for regulating health behaviors of the low status. These phenomena seem widespread across time and cultures.

I can explain these puzzles moderately well by assuming that humans evolved deep medical habits long ago in an environment where people gained higher status by having more allies, honestly cared about those who remained allies, were unsure who would remain allies, wanted to seem reliable allies, inferred such reliability in part based on who helped who with health crises, tended to suffer more crises requiring non-health investments when having fewer allies, and invested more in cementing allies in good times in order to rely more on them in hard times.

These ancient habits would induce modern humans to treat medical care as a way to show that you care. Medical care provided by our allies would reassure us of their concern, and allies would want you and other allies to see that they had pay enough to distinguish themselves from posers who didn’t care as much as they. Private information about medical quality is mostly irrelevant to this signaling process.

If people with fewer allies are less likely to remain our allies, and if we care about them mainly assuming they remain our allies, then we want them to invest more in health than they would choose for themselves. This tempts us to regulate their health behaviors. This analysis suggests that the future will continue to see robust desires for health behavior regulation and for communal medical care and spending increases as a fraction of income, all regardless of the health effects of these choices.

© 2007 Elsevier Ltd. All rights reserved.

* Tel.: +1 703 993 2326; fax: +1 703 993 2323.
E-mail address: rhanson@gmu.edu
Introduction

Health economists’ workhorse model has long been "medical care insurance". That is, individuals can ex ante prefer insurance, to pay for expensive medical care to get them well should they get sick. And such insurance may require state intervention to mitigate market failures [5,78,80]. This standard framework has illuminated may aspects of health policy.

This framework, however, has trouble accounting for a disturbingly wide range of health policy phenomena, many of which are reviewed below. While many auxiliary assumptions have been suggested to explain such policy puzzles, dissatisfaction with these alternatives has led many health economists to conclude that an important explanation of behavior in health and health policy is "philanthropic externalities" [77], i.e., the fact that "individuals derive utility from knowing that other (sick) individuals are receiving medical care" [36].

The idea that people care about the outcomes of others is widely considered plausible, and has inspired researchers to look at both how such altruism might have evolved [87,31,10] and how it might in general lead to counter-intuitive outcomes [11,62]. Researchers have also considered the implications of altruism for many aspects of family behavior, such as bequests and fertility.

The health policy implications of altruism have, however, not yet been explored in much detail. That is, there are many possible "altruists," depending on which people and outcomes the altruist cares about, and researchers have yet to look in much detail at which kinds of altruists are theoretically and empirically plausible. That is, which types of altruists can both account well for observed behavior in health and health policy, and fit well with what we know about the behavior and environment of our hunter-gatherer and primate ancestors, where such altruism presumably evolved?

This paper begins to explore one possible set of answers to this question. While only some of these answers seem original, they together seem to offer a simple and unified synthesis of diverse phenomena.

In particular, we explore the evolutionarily-plausible assumptions that our ancestors

1. cared more about their social allies, especially those with more and better other allies,
2. suffered more crises when they had few allies (i.e., were of low status), crises being events where the appropriate response diverts energies from investing in health, and
3. were unsure about who would remain a long-time ally, with some often knowing things others did not about the chances that associates would remain allies.

These assumptions have many implications. For example, a person B considering how much to invest in health would weigh both the chance that he would end up with many allies (and become high status), and the chance he would end up with few allies (and become low status). By assumption two, the better he thought his chance of ending with many allies, the more sense it would make to invest in health. He might invest via self-care, reduced risk-behaviors, or a reduced stress response.

An associate A of B, however, would place less weight on what happens when B ends up with few allies. After all, in this case, A also probably not be B’s ally, and by assumption one A would then care less about B. Thus A would prefer that B invest more in health, compared to what B would choose for himself. This divergence in perspectives would be especially strong when B had an especially high chance of ending up with few allies. Our assumptions therefore predict paternalistic altruistic preferences about health, with paternalism especially strong toward the low status.

If A is considering how much to care for an injured or sick B, she will consider the chance p that they will remain allies. Since the value to A of a healed B increases with chance p, A will naturally offer more care when this chance is higher. By assumption three, however, B and other observers can then use A’s level of care as a signal of what A knows about the chance p of remaining allies. For example, more care will persuade B that he is more likely to remain an ally of A, and hence is more likely to be of high status. This can convince B to invest more in health.

Person A might know things about either A or B’s loyalty or desirability as an ally. Since A would typically like others to believe in a high chance p of remaining allies, A will over-care in order to credibly signal p. Thus our assumptions predict excessive health care due to efforts to signal social solidarity, and they predict a comforting placebo effect from the appearance of care.

The health-care behavior of humans today may still reflect a genetic inheritance of tendencies toward once-adaptive behaviors, even if humans today are not aware of the origins or ancient function of their current behaviors.

If so, the assumptions above may explain the following modern behavior:
Paternalistic health-favoring regulation of behaviors, especially toward the low status.

Support for national, not international, health insurance, independent of market failures.

A strong influence of social status on health, mediated by care, behavior, and stress.

Genuine concern mixed with self-serving efforts to be seen as helping.

A near-zero marginal health-value of medical care, and a placebo benefit of apparent care.

If we further assume that for our ancestors, desirability or loyalty as social allies increased with age, we can also explain an especially low marginal health-value of medicine for older people. Finally, if we assume that the value of allies, relative to other resources, increased with increasing material wealth, we might also explain the apparent "luxury" nature of both medical care and leisure. We might thus account for the increasing fraction of our resources devoted to health care.

After a more detailed examination of these health policy puzzles, we will discuss how our assumptions fit with what we know about the behavior of our ancestors, present some simple formal models, and finally review how our assumptions may explain these policy puzzles.

Health policy puzzles

Health altruism and paternalism

Several health policy puzzles surround the ways in which health behavior and care seems to be treated differently from other consequences and industries.

National health insurance (NHI) was begun in Germany in the late 1800s, and similarly in Japan in 1911, apparently to gain allegiance from workers unhappy with industrialization [36]. Since then something like NHI has long attracted widespread political support. This support is especially striking when compared to the relatively weak support for international health insurance, for nationalization of most other industries, or for redistribution of food, housing, clothing, etc. Health care seems different somehow.

Contagion externalities were once a favorite justification for NHI, but appeals to contagion have faded over the generations as contagion appears to have become a minor health concern. A common recent justification for NHI is adverse selection, i.e., that people who know they are low risk under-insure to persuade insurance companies to offer them low rates. Even if health insurance markets suffers from a serious adverse selection problem, however, this would only seem to justify a requirement that everyone purchase a minimum amount of long-term catastrophic health insurance. Furthermore, the empirical evidence seems contrary to the adverse selection in insurance hypothesis. Simple adverse selection predicts that those with a higher risk of illness will more fully insure. When insurance companies are free to price based on what they know about customers, however, the correlation between insurance level and risk (both real and perceived) goes the other way; risker people buy less, not more, insurance [50,14,18,19].

Further evidence against a simple market failure explanation of NHI support is found in the fact that positive opinions about the nature of the health care market do not seem to predict normative positions on NHI, not among physicians, economic theorists, or health economists [35]. Support for NHI insurance appears to instead be a matter of values.

A related phenomena is the widespread opinion that the rich should not get more medical care than the rest of us, i.e., that "income should not determine access to life itself" [36]. Interestingly, people given fruit to divide up divide among themselves divide it more equally when told that the fruit is a health aid, instead of something that tastes good [103].

Another difference between health and other areas is an apparently high level of paternalism, i.e., an unwillingness to defer to individual judgments regarding tradeoffs between health and other considerations. Examples include professional licensing of physicians, regulations of foods, drug, and medical devices, and safety rules in transportation, consumer appliances, and the workplace. These limits on health choices contrast with the wide freedom most of us enjoy regarding most other types of personal consequences.

Consumer ignorance is often suggested as an explanation for such paternalism, but in theory instead of banning products a trusted regulator need only tell consumers what they know, such as via a "would have banned" label. Consumer ignorance, by itself, is thus not a sufficient explanation, though information asymmetries can magnify other reasons for paternalism [47].

Health paternalism seems particularly strong toward low status individuals. For example, great concern is expressed about the hard-to-clearly-

---

1 Similar independence has been found between positive opinions of labor and public economists and their related policy positions [37].
document risk to babies from teen pregnancy [64], while little concern is expressed about the clearly-documented and substantial risks to babies from pregnancies of women over the age of 40. Great concern is expressed about liquor stores in poor neighborhoods, but not about the even larger liquor sales in rich neighborhoods.

As another example, blacks are 13% of US monthly drug users, about the same as their population fraction, but get 74% of drug-crime prison sentences [69]. Similarly, in Massachusetts those in the poorest zip codes are between 2.6 and 16.5 times more likely to end up in treatment for drug abuse than those in the richest zip codes, and yet are 54 times more likely to end up in prison for drug crimes [15].

A related health policy phenomena is strong focus of public health researchers on health outcomes, to the exclusion of other outcomes which people trade off against health, such as cost, fun, appearance, etc. For the most part, only health consequences are examined. Public health also seems to pay disproportionate attention to health of the low status. Note also that while one often hears messages encouraging people to eat right, exercise, sleep enough, etc., one rarely hears messages encouraging people to live a little and take more risks.

Similarly, it is notable that while there are many charities devoted to helping with health crises, few charities are devoted to helping with other sorts of crises with similar magnitude utility hits, such as divorce, falling out of love, unemployment, failed careers, breakup of friendships, etc. A further complication comes from the observation that while some charity behavior is outcome-oriented, much other charity behavior seems oriented more to creating the appearance of charity efforts [49].

Finally, it seems to me that politicians and others considered for positions of influence in health policy are frequently selected in part for how much they care about health. In contrast, it does not seem to matter much whether people who regulate electric utilities, for example, care much about electricity.

A straightforward, if apparently ad hoc, explanation for most of the above phenomena is that we care about others within our nation, that we tend to care about their health more than their happiness, and that this tendency is especially strong for low status people.

Some researchers have suggested that we explain some of these phenomena using simple altruism without paternalism. For example, simple altruism can lead to under-investment by recipients if donors cannot commit [16], or to under-insurance by recipients if collective action among donors is only possible before risks are realized [20]. There seems, however, to be little reason to think of NHI as an investment, or to assume post-realization collective action on health is substantially harder than early collective action.

**Status and health**

Another striking puzzle is that high status people tend to be much healthier than others. While health influences status to some degree, most of the influence seems to go from status to health (though there are doubters [90]). Furthermore, while there are declining health returns to status, the health-status relation continues to be strong all the way up the status ladder, even after one controls for lower status people’s weak tendency to get less medical care, and stronger tendency to engage in more health-risking behaviors [1,32]. For example, a recent study of 3600 US adults over eight years found mortality rates varying by a factor of 2.8 with income, even after controlling for age, sex, exercise, crowding, smoking, alcohol, weight, and education [59]. (These other controls varied mortality rates by respective factors of 40, 2.9, 2.4, 1.5, and 1.3, with the rest being insignificant.)

Identifying the causal paths relating status and health has proved difficult, however. For a while it seemed that social support, i.e., friendly contacts and relationships, were a key element in the causal chain, especially for men [52]. An influential study found, however, that living in a poverty area increased mortality rates by a factor of 1.5, even controlling for social support, income, education, access to medical care, and unhealthy behaviors [44].

Several studies have suggested that a reduced sense of control is central, finding social support to be irrelevant after controlling for factors like authority and skill discretion at work [67,68,66]. Other studies have, however, found a sense of control to be irrelevant [51]. For example, among 1800 US bus drivers, job control was irrelevant after controlling for age, sex, income, education, marriage, weight, family history, fitness, alcohol, and caffeine [2]. Also it seems that status is men is more related to work while status for women is more related to relations at home [82].

A further puzzle is the apparently very low impact of information on health-risking behaviors. For example, 13,000 middle-age men at high risk for heart attack were randomly assigned usual care or special counseling about hypertension, smoking and diet. No significant mortality benefit was seen after 7 years [41] and after 16 years there was only...
a marginally significant (6% level) benefit \[42\]. A similar lack of effect was found in counseling for low weight babies \[70\] and smoking \[6\].

A perhaps related puzzle is the placebo effect, whereby health improves from physically inactive treatments. For example, in double-blind clinical trials the placebo effect seems to be 75% of the effect of common anti-depressive medications, and much of the remaining 25% may be due to patient ability to discern “real” drugs from placebos via their larger side-effects \[58\]. The relation between health and status has remained strong for centuries across diverse societies, even as causes of death and illness have varied radically. The causal pathways thus seem to be many and varied, resisting simple descriptions of a canonical causal path. One of the few general explanations that has been offered is that those who discount the future more are less likely to invest in either health or career advancement, and so are more likely to be both sick and poor \[35\].

**Medicine and health**

Publication selection bias makes it hard to be sure, but the vast medical literature on randomized clinical trials certainly suggests that medical care has health benefits, at least when best practice is applied to patients deemed most likely to benefit. This leaves open, however, the question of the average benefit of typical practice on typical patients, especially since the vast majority of medical treatments have yet to be carefully studied with clinical trials.

Perhaps the most striking puzzle in health policy is the apparent lack of an aggregate empirical relation between medical care and health. Observed variations in medical care typically have an insignificant effect on average population health, even when looking at large data sets, sets larger than those which convinced most researchers of the reality of many other influences on health.

One of the first studies on the aggregate health effects of medicine found mortality variations across the 50 US states were unrelated to health care spending, given various controls \[7\]. A recent comparison of 21 developed countries also found national life expectancy did not vary significantly with medical care spending, after controlling for income, education, unemployment, animal fat intake, smoking, and consumption of pharmaceuticals\footnote{Pharmaceutical consumption was surprisingly effective, however, at an estimated $20,000 cost per life-year gained.} \[54\].

The most definitive data on this topic comes from the RAND Health Insurance Experiment, which for three to five years in the mid 1970s randomly assigned two thousand non-elderly US families to either free health care or to plans with a substantial copayment. Those with free care consumed on average about 25–30% more health care, as measured by spending. They went to the doctor and hospital more often, and as a result suffered one more restricted activity day per year, when they could not do their normal activities. The extra hospital visits were rated by physician reviewers to be just as medically appropriate, and to treat just as severe a stage of disease, as the other hospital visits.

Those with free care obtained more eyeglasses, and had more teeth filled. Beyond this, however, there was no significant difference in a general health index, which was the designed outcome measure. There was also no significant difference in physical functioning, physiologic measures, health practices, satisfaction, or the appropriateness of therapy. Blood pressure may have been reduced, but the point estimate was that this produced a 1% reduction in average future mortality rates, which translates to roughly seven weeks of life \[13,65,76\]. And this estimate was not significantly different from no effect.

Having failed to find an aggregate benefit of medical care, many have sought to find benefits for identifiable subpopulations. The international comparison cited above, for example, found that lagged medical care did seem to improve infant morality \[54\]. And while the RAND experiment described above found no mortality benefit to children, it did suggest lower blood pressure in those especially at risk, though the statistical tests reported did not correct for having search effects \[13,65,76\].

One recent study of US Medicaid clients found a significant effect on infant mortality, as low as a $2 million cost per life saved for certain targeted populations (and much more for others) \[23\]. Recent studies of Medicaid children conflict; one finds a significant $2 million cost per life saved \[22\], while another finds “at best weak support” \[56\].

A recent study of geographic variations in Canadian medical spending reported substantial health effects \[21\]. And an older study estimated large benefits for the elderly, with a 10% increase in US Medicare spending reducing elderly mortality by 3–4% \[46,45\]. Usually, however, the $25,000 spent on care in last year of life \[38\] (27% of US Medicare spending \[63\]) is suspected of being the least effective on mortality. For example, large medically-unexplained variations in Medicare spending across
hospital regions [100,101] were recently used to estimate the regional mortality benefit from spending $1000 more in the last six months of life. The benefit is bounded (at the 95% c.l.) to be less than 0.1% in general, and less than 1% for a subpopulation with certain specific conditions (such as heart attack) [89].

New heart attack treatments are among the most celebrated of recent medical innovations, and both medical spending and mortality improvements have increased more than average among the heart attack population. Assuming all this added mortality reduction is due to the added spending implies a low cost of about $10,000 per life-year [26,25], which compares favorably to typical value estimates of $75,000–175,000 per life-year [94]. Similar large benefits come if we assume all improvements in post attack mortality are due to medicine [24]. Also, assuming medical care is the cause of all heart attack mortality reduction not attributable to changes in identifiable risk factors such as blood pressure and smoking implies a large medical health benefit [53]. There is, however, no obvious reason to make these generous assumptions [81].

The most sophisticated statistical analysis to date, of 800,000 Medicare patients, estimated that adding a heart attack catheterization capability to a hospital costs $70,000 per life year. This estimate, however, was only marginally significant (7% level) [71].

The above studies are mostly about mortality. What about quality of life? Unfortunately, most of the value in estimated quality-adjusted life years (QALY) is in raw lifespan, and most changes in QALY over time have also been due to lifespan changes [28,27]. It is thus hard to see how there could be substantial QALY improvements without there being lifespan improvements.

The above studies are mostly about the marginal value of the last one third of spending. What about the average value of the first two thirds of spending? Both life expectancy and medical spending have increased in the last four decades, and a recent analysis [28] calculates that this extra spending was worth it if at least 30% of the increase in US lifespan was due to the increase in US spending (assuming a $100,000 life-year value). It is far from clear, however, that medicine can claim this much credit. An optimistic accounting of the benefits of specific treatments attributes only five years of the 40 or more years of added lifespan over the last two centuries to medicine [17].

An average value effect was sought in the study mentioned above of US Medicare spending in the last six months of life. The coefficient of a squared term in spending was small and insignificant, however, even though spending in the sample varied by a factor of two [88,89]. Also, the RAND experiment described above found that, compared to the other care, the extra care which insurance induces was just as medically appropriate, had just as severe diagnoses, and was just as often in the hospital [93,79].

Shamans and doctors have long been in demand, even though the common wisdom among medical historians today is that such doctors did very little useful on average until this century [36]. The studies above suggest that much the same story may still apply to doctors today, at least regarding the medical care that some people now get and others do not.

One common explanation for the low marginal value of health care is health insurance. Health insurance is endogenous, however, and there are several ways to insure against unusual large events while retaining incentives to attend to the costs of frequent small events. Catastrophic insurance can be combined with medical savings accounts, for example, or one can subscribe to an HMO which refuses to cover care of questionable value.

After a brief recent period of HMO cost reduction, however, consumers do not appear much interested in further HMO care cuts. HMO market shares have stagnated compared to more generous plans, and most political discussions are now about increasing, not decreasing, coverage. Similarly, hospice and advance directives seem to save at most 10–17% of expenses in the last six months of life [30]; even when we give up on preventing death we spend almost as much to comfort the dying. This tendency is sometimes attributed to psychosis, such as an inability to “let go”.

The lack of interest in reducing apparently-useless health care spending seems all the more puzzling in light of the 14% of US GDP now spent on medical care [55]. This high spending level is projected to increase even further, as health spending appears to be a luxury good at the national level [57], even though it appears to be a necessity at the individual level [40]. For example, a recent study of OECD nations from 1974 to 1987 found that health spending rose with GDP per capita to the power of 1.27, a relation that accounts for about two thirds of increased spending, the rest of which is attributed to a time trend [39]. (Contrary results, however, have also been found [12].)

The small effect of medicine on health seems related to several more puzzles. One is that consumers seem very unresponsive to information given to them privately about hospital quality. For example, only 8% of 784 patients about to undergo car-
were rarer, and labor was more physical. They were adapted to an environment where such foods were too much salt or fat in terms of preferences which profitably understand current tendencies to eat. And we can spend more time raising the young [3]. And we can mate females live longer than males when females women live longer than men because in general primate females live longer than males when females spend more time raising the young [3]. And we can profitably understand current tendencies to eat too much salt or fat in terms of preferences which were adapted to an environment where such foods were rarer, and labor was more physical.

The small health effects of medicine also raises the question of why exactly lifespans have increased so dramatically. Over the last century, age-specific mortality rates have fallen at a steady exponential rate across developed countries, without noticeable changes due to major medical and public health innovations [98,61]. Improvements in sanitation are often given great credit, but no effect on mortality has been found among individual variations in water source and sanitation, even among high mortality populations [60].

The evolution of health altruism

Perhaps our distant ancestors can tell us something interesting about modern health behaviors. Many aspects of human behavior are surely local cognitive adaptations to local modern circumstances, while many other aspects are culturally evolved and inherited adaptations to the modern world. However, much of our strategies of adaption, and the menu of behaviors these adaptations choose from, surely reflect our genetic inheritance. About half of the variation in personalities, for example, can be attributed to genetic variation.

By most accounts, the last ten thousand years of human agriculture have been too short for genetic selection to have had much effect. To understand our genetic behavioral inheritance, we must therefore look to the behavior and environments of our hunter-gatherer, primate, and even more distant mammal ancestors. Fortunately, over the last few decades we have learned a great deal about these ancestors from the lives, behaviors, and environments of the our living primate relatives, and the few remaining isolated human hunter-gathers.

We have, for example, learned a great deal about our preferences and strategies of mate choice [33,9], and about how time preferences vary with age and gender [48]. In health, we have learned that women live longer than men because in general primate females live longer than males when females spend more time raising the young [3]. And we can profitably understand current tendencies to eat too much salt or fat in terms of preferences which were adapted to an environment where such foods were rarer, and labor was more physical.

Evolved health care

We have also learned some things about health care in mammals, primates and hunter-gatherers. Whales and dolphins, for example are reluctant to abandon disoriented associates, and as a result sometimes beach themselves as a group. Elephants try to hold up dying associates, and mourn and cover them when they die. Macaque primates born without hands or feet have survived to raise their own healthy children, due in part to extra food and protection from associates. Neanderthals with dwarfism and paralysis of the limbs also survived into adulthood. Chimps are less aggressive toward injured associates, and have even been seen faking injuries to avoid fights [99].

Sick chimps who pose a threat of contagion, however, are not treated so kindly. During an epidemic of poliomyelitis, for example, partially paralyzed chimps were treated with fear and hostility and attacked as if they were no longer in the group [99].

Human hunter-gatherers care for sick or injured associates. Among the Ache of Paraguay, for example, people get sick or injured with a median period of 30 days, and remain sick for a median of 3.5 days, though occasionally people are sick for months or years. 26% of such events were injuries, and 13% were_birthings. Sick people are given an average of 3.3 food items from an average of 2.4 family members and associates, and those who give a larger fraction of their production to sick associates will get more food from others when they are sick [43].

Aid for injury is treated similarly to other forms of aid, such as sharing meat from an hunt, participating in a work party to build a hut, or joining in an attempt to avenge the killing of an associate. Failing to help is interpreted as being less loyal to a group, coalition, or partner. The net effect of all this mutual aid is that household consumption can be remarkably well insured against shocks to individuals, though those low in status (e.g., the landless in village India) seem less well insured [96].

Tit-for-tat reciprocity seems at risk of being less effective in ensuring cooperation for severe sicknesses or injuries [95]. In response, hunter-gatherers try to acquire a reputation for generosity or unique abilities which others will miss if they are gone. That is, they try to induce in others a true concern for their welfare. Those with unique abilities have more prestige and those with more prestige receive more care from associates. And to ensure that there are enough group members to shoulder the burden of aiding an injured member,
groups seem to be larger than they would need to be to insure against other risks [92,91].

The main known exceptions to hunter-gatherers caring for associates occur in situations of extreme deprivation. Starving hunter-gatherers have been seen delighting in the suffering of associates, with young people laughing while stealing food from the mouths of their elders [99].

**Contingent altruism**

Health behaviors are often described in terms of simple altruism. But does that make evolutionary sense? While it is hard to see how a simple “promiscuous” altruism could be selected for, it is easier to see how we could have evolved preferences favoring good outcomes for those who share our genes [10]. It also seems that we have evolved some forms of “reciprocal” altruism [97], favoring good outcomes for those who are likely to favor us in a similar way.

Theoretically, we can understand how selection could favor an altruism which is contingent on some characteristic or behavior which indicates reciprocity in equilibrium. For example, altruism toward physical neighbors can be favored when neighbors are naturally likely to be related [31]. Similarly, selection can favor altruists over selfish agents when the altruists vary between acting nicely or spitefully depending the fraction of altruists among those with which they interact [87].

Reciprocal or contingent altruism seems to have been amply verified in numerous empirical animal studies. For example, competition within groups of chimps is largely a matter of two allies ganging up on a third. Studies have shown clear correlations; if A helps B in a fight against X, then B is likely to help A against Y. Retaliation is also observed; if A helps X against B, B is likely to help Y against A [99].

Humans even seem to have special cognitive modules for detecting “cheaters” [9]. We seem have evolved to pay close attention to identifying those who will or won’t help us, and to treat our allies better than others.

**Groups as correlated allies**

Altruism is often described in terms of groups, instead of in terms of allies. Human allegiances and “morals” are clearly oriented to groups; it has typically been acceptable to kill and enslave “them” in ways that it almost never is for “us”. All human societies also seem to have a sense of belonging to and a need for acceptance by a group [99]. Human groups also have complex nested and overlapping structures. Those in the same nation tend to be allies against those in other nations, and similar trends hold for companies, schools, towns, clubs, gangs, and families. The concept of a group also seems highly salient to primates and many other animals, who also can have complex nestings of groups and coalitions.

Instead of thinking of groups as a new concept distinct from allies, it seems sufficient for the purposes of this paper to think of groups as correlated social allies. That is, all else equal, two members of the same group are more likely to treat each other as allies than members of different groups. If someone is an ally of one group member, they are more likely to be an ally of other group members. Conversely, if rejected by one member, they are more likely to be rejected by others. These correlations can be induced by alliances; it is awkward to have a friend of a friend be an enemy. Correlations can also be induced by shared preferences, and by shared information about the desirability or loyalty of individuals as allies.

This concept of groups fits comfortably with group nesting. If two members of the same group are in different internal coalitions, they are likely weaker allies than two members in the same coalition.

**Status as more better allies**

Social status is another important concept for understanding humans and other animal behavior [34]. Higher status animals reproduce more, and most animals clearly strive for higher status. But what exactly is status?

For the purposes of this paper, it seems sufficient to think of status in higher primates, such as chimps and humans, as having more and better social allies. Higher status animals may just be animals who are considered more desirable as allies. Such animals naturally have more and stronger alliances, and are in more and larger coalitions. This view helps us integrate the views that the purpose of alliances is to increase status [102], and that status is a measure of one’s value to allies [91]. It also helps us understand why low status group members are defended less often against outsiders such as predators. This view also makes sense of status in higher primates status being more about social skills and coalition building than physical ability [3]. This in turn helps explain why older primates tend to have higher status, even as their physical ability wanes. With time, senior males develop secure alliances, an “old boys network”, that keeps the strong but less organized young at bay [99].
Higher status animals seem to be more valued by their associates. For example, people are more generous toward high status and discriminate against low status people [8]. And high status animals are often valued as allies because they tend to be the most generous in sharing food with others [99]. Some birds even fight over the right to give away food, or to help their group by watching for predators. This has been interpreted as a way to signal physical ability to potential mates; those who can get enough to eat while doing these things are likely to be stronger [104].

Primate group leaders are especially valued by group members because one of their main roles is to keep the local peace by serving as an arbitrator and police. Great suffering can result from leaders who fail in this role, and such leaders don’t usually stay leaders for long [33].

Biological bases for some of these status behavior correlations have been found. For example, higher status primates and humans tend to have more serotonin in their blood, and serotonin seems to relax people, making them more gregarious and socially assertive [102].

Uncooperative strategies among the low status

Low status primates, who have few and poorer allies, might naturally choose less ally-based strategies for getting what they want. If such strategies make them more risky as allies, this could reduce their value as allies even more.

For example, a new rival for leadership or a new immigrant into a group can disrupt a primate group for months. This is good for rivals but bad for other group members. Thus the leadership status quo in a group tends to benefit the young and weak, making the leader a better ally than the rival, and immigrants worse allies than other group members [99].

The vast majority of human violent crimes are done by young adult males, both in the modern world and in hunter-gatherer societies [99]. Wars today tend to be started in places where there are many young (age 15–29) men relative to older men in the population, and in societies with more polygyny. Periods of Portuguese global expansion also correlated with periods of more young landless nobility [75,74].

We expect status to be more important for men than women, since male reproductive success varies by larger factors [102]. And in the last 160 years, aggressors prevailed in most wars. This all suggests that wars tend to be started by low status young men seeking mates or resources to attract mates. The impact of such behavior is large. Among isolated hunter gatherers today, such as the Yanomamo Indians, 30% of adult male deaths are from violence [99]. And since typical hunter-gather should have been less isolated, and thus more war-like, ancestral death rates were likely even higher.

It also seems plausible that low status females, who would typically be paired with low status males in human-like pair bonding, would have more of a reason to “cheat” on their mates, in order to get better quality genes. In one study of chimps, half of the children were fathered by males outside the local group [3]. Such cuckoldry may well have been a serious concern.

Inducing low self-esteem in humans seems to make them more likely to cheat at cards [4]. All these observations suggest that our low status ancestors may in general have tended to engage more in non-cooperative strategies, such as theft, rape, and cuckoldry. If so, that would give our ancestors all the more reason to value high status over low status allies.

Status and the relative value of health

For most primates, low status is less healthy. For example, in most primates one sex migrates into a new group as it comes of age. At first a newcomer baboon has very low status. As a result it loses fights, is pushed around, has its food stolen, is infested with parasites because no one will groom it, doesn’t know the group’s signals to warn of predators, and isn’t helped by others against predators. As time goes on, however, a newcomer may gain allies, first for grooming, then for other activities [84].

To deal with being attacked by a predator, most mammals invoke the same generic “stress response” system. This system rapidly mobilizes energy by temporarily halting or curtailing systems for energy storage, body growth, digestion, reproduction, and immunity. It also sharpens most senses while dulling the sense of pain. And it can induce defecation to reduce excess body mass [85].

Invoking the stress response tends to help in dealing with a short term crisis, but it hurts long term health. If frequent crises lead to frequent invocations of the stress response, a body can wear down, making it harder to deal with new crises.

Glucocorticoids are steroid hormones released by the stress response, and low ranking baboons, as well as those with few friends, have higher levels of glucocorticoids. This is not just because low ranking baboons suffer more crisis events. Primates seem to induce the stress response just by thinking about or anticipating something stressful. (Consider changes in your body when you watch a scary
movie, or imagine giving a speech.) Presumably, invoking a stress response in anticipation of an event like a predator’s attack aids primates in watching for and avoiding such events.

Glucocorticoid levels have been observed to drop quickly in baboons when an unanticipated event raises their social status, such as when the accidental death of a leader allows a rival to replace him. This suggests that primates base their stress levels in part on their social status level, since status predicts the chances of meeting a stressful event. High glucocorticoid levels have also been observed in newly dominant baboons who have good reason to fear their dominance will not last. This suggests that stress levels are based more fundamentally on anticipation of future status levels [83,85].

Lower stress does not correlate with higher status in all primates. The correlation can go the other way when low status primates are infrequently harassed by high status primates, or if punishments for “cheating” are mild. Humans, however, both modern and hunter-gatherer, seem to follow the usual status-stress correlation. In humans, stress seems to be reduced by social supports and a belief in control over one’s life. Modern humans of lower status not only suffer more undesirable crisis events, but they are also more strongly affected by them emotionally [72].

Formal models

The above qualitative descriptions of our ancestors can be embodied in formal models, models which should allow us to more easily see the implications of our assumptions.

Social status and health

Let us first consider paternalism in health altruism. Assume each person has health $h$ and miscellaneous remaining resources $r$. Also assume the personal payoff of a combination $h, r$ depends on whether the person will become or remain in, or be left out, of a certain group or coalition that this person is now associated with. In vs. out is intended as a simplified description of high vs. low status, since we are treating high status as essentially being in more coalitions and alliances. The payoff to being in is $u(h, r)$, while the payoff to being out is $\bar{u}(h, r)$.

Consider someone who is uncertain about being in or out, and who makes choices between health and other resources in anticipation of both possibilities. If she assigns a probability $p$ to being in, her expected payoff is

$$ Eu(h, r, p) = pu(h, r) + (1 - p)\bar{u}(h, r). $$

If the functions $u, \bar{u}$ are increasing and strictly convex, then so is $Eu$ (in $h, r$ for $p \in [0, 1]$). In this case, for any bounded convex set $S$ of possible $(h, r)$, there is a unique best choice $(\hat{h}(p), \hat{r}(p)) = \arg\max_{(h, r) \in S} Eu(h, r, p)$.

Let us assume $u_1/u_2 > \bar{u}_1/\bar{u}_2$, so that health is more valuable, relative to other resources, for those who are in. This implies $h'_p > 0$, i.e., those who are more likely to be in choose more health, at the expense of other resources.

Paternalistic altruism

An altruist is a person $i$ whose utility $U^i$ depends not only on her personal payoff $u^i$, but also on the payoff $u^j$ of others. But which others? In principle, $i$ might care about each of her allies in proportion to the strength of their alliance and the value of each person as an ally.

For simplicity, however, we will here assume consider a single group, and have $i$ care about $j$ only if they are both in this group. This models a polar form of correlation among alliances, where losing one person as an ally means you lose all allies. We expect similar, if muted, results from weaker ally correlations.

Let us specifically consider a group-contingent altruist $i$, for which each state $x$ has state-dependent utility

$$ U^i_x = u^i_x + \sum_{j \neq i} u^j_x \times \begin{cases} W & \text{if } i, j \text{ both in the group at } x \\ w & \text{otherwise} \end{cases}, $$

where $W > w$.

Now what are the preferences of a group-contingent altruist $A$ regarding the tradeoff between health and other resources $h, r$ of a person $B$? Since $A$ puts more relative weight than $B$ on states where $B$ is in, and since health is more valuable in such states, $A$ thus puts more relative weight on health and lower status predicts the chances of meeting a crisis. The correlation can go the other way when low status primates are infrequently harassed by high status primates, or if punishments for “cheating” are mild. Humans, however, both modern and hunter-gatherer, seem to follow the usual status-stress correlation. In humans, stress seems to be reduced by social supports and a belief in control over one’s life. Modern humans of lower status not only suffer more undesirable crisis events, but they are also more strongly affected by them emotionally [72].

Formal models

The above qualitative descriptions of our ancestors can be embodied in formal models, models which should allow us to more easily see the implications of our assumptions.

Social status and health

Let us first consider paternalism in health altruism. Assume each person has health $h$ and miscellaneous remaining resources $r$. Also assume the personal payoff of a combination $h, r$ depends on whether the person will become or remain in, or be left out, of a certain group or coalition that this person is now associated with. In vs. out is intended as a simplified description of high vs. low status, since we are treating high status as essentially being in more coalitions and alliances. The payoff to being in is $u(h, r)$, while the payoff to being out is $\bar{u}(h, r)$.

Consider someone who is uncertain about being in or out, and who makes choices between health and other resources in anticipation of both possibilities. If she assigns a probability $p$ to being in, her expected payoff is

$$ Eu(h, r, p) = pu(h, r) + (1 - p)\bar{u}(h, r). $$

If the functions $u, \bar{u}$ are increasing and strictly convex, then so is $Eu$ (in $h, r$ for $p \in [0, 1]$). In this case, for any bounded convex set $S$ of possible $(h, r)$, there is a unique best choice $(\hat{h}(p), \hat{r}(p)) = \arg\max_{(h, r) \in S} Eu(h, r, p)$.

Let us assume $u_1/u_2 > \bar{u}_1/\bar{u}_2$, so that health is more valuable, relative to other resources, for those who are in. This implies $h'_p > 0$, i.e., those who are more likely to be in choose more health, at the expense of other resources.

Paternalistic altruism

An altruist is a person $i$ whose utility $U^i$ depends not only on her personal payoff $u^i$, but also on the payoff $u^j$ of others. But which others? In principle, $i$ might care about each of her allies in proportion to the strength of their alliance and the value of each person as an ally.

For simplicity, however, we will here assume consider a single group, and have $i$ care about $j$ only if they are both in this group. This models a polar form of correlation among alliances, where losing one person as an ally means you lose all allies. We expect similar, if muted, results from weaker ally correlations.

Let us specifically consider a group-contingent altruist $i$, for which each state $x$ has state-dependent utility

$$ U^i_x = u^i_x + \sum_{j \neq i} u^j_x \times \begin{cases} W & \text{if } i, j \text{ both in the group at } x \\ w & \text{otherwise} \end{cases}, $$

where $W > w$.

Now what are the preferences of a group-contingent altruist $A$ regarding the tradeoff between health and other resources $h, r$ of a person $B$? Since $A$ puts more relative weight than $B$ on states where $B$ is in, and since health is more valuable in such states, $A$ thus puts more relative weight on health and lower status predicts the chances of meeting a crisis. The correlation can go the other way when low status primates are infrequently harassed by high status primates, or if punishments for “cheating” are mild. Humans, however, both modern and hunter-gatherer, seem to follow the usual status-stress correlation. In humans, stress seems to be reduced by social supports and a belief in control over one’s life. Modern humans of lower status not only suffer more undesirable crisis events, but they are also more strongly affected by them emotionally [72].

Formal models

The above qualitative descriptions of our ancestors can be embodied in formal models, models which should allow us to more easily see the implications of our assumptions.

Social status and health

Let us first consider paternalism in health altruism. Assume each person has health $h$ and miscellaneous remaining resources $r$. Also assume the personal payoff of a combination $h, r$ depends on whether the person will become or remain in, or be left out, of a certain group or coalition that this person is now associated with. In vs. out is intended as a simplified description of high vs. low status, since we are treating high status as essentially being in more coalitions and alliances. The payoff to being in is $u(h, r)$, while the payoff to being out is $\bar{u}(h, r)$.

Consider someone who is uncertain about being in or out, and who makes choices between health and other resources in anticipation of both possibilities. If she assigns a probability $p$ to being in, her expected payoff is

$$ Eu(h, r, p) = pu(h, r) + (1 - p)\bar{u}(h, r). $$

If the functions $u, \bar{u}$ are increasing and strictly convex, then so is $Eu$ (in $h, r$ for $p \in [0, 1]$). In this case, for any bounded convex set $S$ of possible $(h, r)$, there is a unique best choice $(\hat{h}(p), \hat{r}(p)) = \arg\max_{(h, r) \in S} Eu(h, r, p)$.

Let us assume $u_1/u_2 > \bar{u}_1/\bar{u}_2$, so that health is more valuable, relative to other resources, for those who are in. This implies $h'_p > 0$, i.e., those who are more likely to be in choose more health, at the expense of other resources.
Relative to a possibility set $S$, we thus have $h'(p) > h'(q)$. meaning that altruist $A$ prefers $B$ to have more health than $B$ would choose for herself. If $A$ is spiteful enough, $A$ can even prefer $B$ to have more health than the maximum $B$ would ever choose, when she felt sure to be in. Thus by combining group-contingent altruism with health being more valuable for those who are in, we get paternalistic preferences regarding the health of others, a paternalism which is stronger regarding the low in status.

**Stress**

An important example of a choice between health and other resources is stress, or more precisely the invocation of the biological stress response.

Let each person privately chooses a stress level $s$, which influences personal payoffs via

$$u(h, r, s) = v(H(h, s), R(r, s)),$$

where $v_H, v_R, H, R, R > 0$ and $H < 0$. Here added stress hurts health but aids in obtaining other resources. Person $B$ will choose stress $s(h, r, p)$ to satisfy

$$0 = EU_s(h, r, p) = pu_s(h, r, s) + (1 - p)u_s(h, r, s),$$

(assuming $EU_s < 0$). Group-contingent altruist $A$, however, would prefer $B$ to chose $s(h, r, p) < s(h, r, p)$, where $p > p$. Since $s_p < 0$, $A$ prefers $B$ to be less stressed than $B$ would choose for herself.

Similar results should hold for other health-risking choices which trade health for other benefits, if health is more relative valuable for those of high status.

**Signaling models**

Let us now consider the possibility that some people may know things that others do not about the chance that a person will remain in the group. In such a situation, the amount of care given to an injured or sick person may be interpreted as a signal of this hidden information.

Consider a group-contingent altruist $A$ who has an opportunity to care for an injured or sick person $B$, i.e., a person with low $h_B$. For simplicity let us assume $w = w$ and $w = 0$. Let $q$ be the probability that $A$ will be in, and let $p$ be the probability that $B$ will be in, given that $A$ will be in. If $c$ is the amount of $A$’s remaining resources that she now devotes to improving $B$’s health, we can (ignoring stress) write $A$’s expected utility as

$$EU^A = Eu^A(h, r_A - c, q) + wp u^B(h_B + c, r_B).$$

When information is symmetric, making signaling impossible, $A$ will choose care $c$ to satisfy a first-order condition (FOC)

$$Eu^A = wp u^B.$$  

That is, $A$ weighs $A$’s resource cost of care (on the left) against the direct health benefit of care to $B$ (on the right), a benefit discounted by $A$’s level of altruism and the probability both $A$ and $B$ will remain in the same group.

When information is asymmetric, so that $A$ knows things that others do not, $A$’s level of care $c$ can be interpreted as a signal of that hidden information. There are many possible things that $A$ could know, each of which could contribute a signaling effect to the choice of care $c$. For simplicity, we will now consider a number of such effects in isolation. That is, for each thing $A$ might know, we will assume that $A$ has one-dimensional private information $\theta$ about only that thing. We will also assume that in equilibrium care $c$ fully reveals this private information to observers.

We will thus consider a series of standard one-dimensional signaling models. (General results for one-dimensional signaling models are reviewed in the Appendix). In each case care $c$ will be taken as the signaling action, and the hidden type $\theta$ will be varied to explore various possible sources of signaling effects.

**Showing that I am able**

One simple thing care $c$ might signal is her ability to care $r$. While there are many reasons why $A$ might want to show that she is highly able, we will here focus on wanting $A$’s group to keep her in it. That is, let us assume that at some point after care $c$ is given (and observed), $A$ will be retained or excluded from the group depending on how the group perceives $A$’s remaining resources $r$. In particular, assume that a reservation ability $r_0$ will be drawn from a c.d.f. distribution $F(r)$, and that $A$ will only be retained if perceived ability $\bar{r}$ exceeds this cutoff, i.e., if $\bar{r} > r_0$. We thus have $q = F(\bar{r})$.

A separating equilibrium where care $c$ signals privately-known ability $r$ satisfies a FOC,

$$Eu^A = wp u^B + [Eu^A + wp u^B]F'(r)/c'(r),$$

where $c(r)$ is the equilibrium care given $r$, and is determined as the solution to this differential equation. Here $A$’s cost of care (left term) is weighed not only against $B$’s health improvement (middle term), but also against the value to $A$ of an improved chance of being in the group (right term), here induced by a higher perceived ability.
Showing that I am staying

Ability to care can be considered one aspect of desirability, which can be signaled more directly. For example, assume that there will be a c.d.f. distribution \( F(d) \) over the group’s reservation desirability \( d_0 \). If so, \( A \)’s private information on desirability \( d \) is equivalent to \( A \) having private information on \( q = F(d) \). To make a separating equilibrium reasonable, assume also that there is only a probability \( b < 1 \) that the group will be able to and need to take care \( c \) as a signal of \( A \)’s desirability \( d \). That is, with probability \( 1 - b \), the group will either not observe care \( c \), or will reliably observe \( d \) in some other way.

Private information on \( A \)’s chance \( q \) of staying in can come not only from knowing \( A \)’s desirability, but also from knowing \( A \)’s loyalty. For example, we could assume that at some point after care \( c \) is given (and observed), \( A \) will choose to stay in or leave the group, depending on whether the value \( e \) of an outside option is better than \( A \)’s reservation value \( e_0 \) of staying. If the outside option value is drawn from a c.d.f. distribution \( F(e) \), then private information on reservation value \( e_0 \) is equivalent to private information on probability \( q = 1 - F(e_0) \). Alternatively, knowledge about \( A \)’s chance \( q \) of staying in could come not only from knowledge about \( A \)’s desirability and outside option reservation value, but also from knowing about the group’s reservation desirability and the actual outside options value.

Regardless of where private information on chance \( q \) comes from, a separating equilibrium where care \( c \) signals privately-known \( q \) satisfies a FOC,

\[
E_{u_c} = wq u_h^* + b(E_{u_c} + wp u_h^*)/c'(q),
\]

where \( c(q) \) is the equilibrium care given \( q \). Here \( A \)’s cost of care is weighed against \( B \)’s health improvement and the value to \( A \) of an improved chance of being in the group.

Showing that you’re staying

Care \( c \) can also signal things \( A \) knows that \( B \) does not about the probability \( p \) that \( B \) will remain in the group. As with knowledge of \( A \)’s chance \( q \), knowledge of \( B \)’s chance \( p \) could come from knowing about \( B \)’s desirability, about the group’s reservation desirability \( d_0 \), about the value to \( B \) of an outside option, or about \( B \)’s reservation value for outside options.

If \( B \) can make choices which trade health against other resources, such as choosing a stress level \( s \), altruist \( A \) can want to choose care \( c \) in order to persuade \( B \) that \( p \) is high. This would persuade \( B \) to invest more in health, for example by choosing a lower stress level.

A separating equilibrium where care \( c \) signals a privately-known chance \( p \) that \( B \) will be in, in order to induce a lower stress \( s \), satisfies a FOC

\[
E_{u_c} = wq u_h^* + wq u_h^* s_p/c'(p),
\]

where \( c(p) \) is the equilibrium care given \( p \). Here \( A \)’s cost of care is weighed against both \( B \)’s direct health improvement, and health improvements in \( B \) due to persuading \( B \) to lower his stress level.

As \( p \to 1 \), we have \( u_h^* \to 0 \), and so the signaling value of care disappears. That is, paternalism disappears when \( B \) is almost sure to be in; \( A \) then has few complaints about \( B \)’s choice of stress \( s \), and so has little reason to persuade \( B \) to reduce her stress. Thus there should be a wide plateau of high \( p \) types who get similar levels of care, reminiscent of the common feeling that “the rich should not get more health care than the rest of us”. And “social solidarity” signaling should be targeted more at those of middle and low status.

Finally, imagine that \( A \) might only give \( B \) “placebo” help. That is, with probability \( g \) care \( c \) would be given, and with probability \( 1 - g \) no care would be given. If \( B \) did not know which possibility applied, the FOC would become

\[
E_{u_c} = wq u_h^* + (wp/g) u_h^* s_p/c'(p).
\]

This form displays a “placebo effect”. Even when no care is actually given, the appearance of care induces \( B \) to increase his health by reducing his stress.

Signaling modifiers

We can identify several simple parameters that would strengthen or weaken the incentive to signal via care \( c \). For example, the FOC for the case where \( A \) has private information on \( q \) can be elaborated to

\[
E_{u_c} = wq u_h^* + (b/fg) c'(E_{u_c} + wp u_h^*)/c'(q).
\]

Here, we consider incentives to offer to provide care \( c \), where \( g \) is the probability than an offer to help is taken up. There is a chance \( b \) that an offer to help is observed by an audience who will not otherwise find out the information being signaled. And \( f \) is the fraction of time before the next signaling opportunity that \( B \) will need help. That is, \( f \) is how often \( h_B \) will be low enough for \( c > 0 \) to make sense. The expectation \( \delta' \) on the right is \( f \) times payoffs when \( h_B \) is low and \( c > 0 \), and \( 1 - f \) times payoffs in states where \( h_B \) is high and \( c = 0 \). (We ignore time discounting.)
Thus, we can see that signaling incentives to offer to help, relative to direct incentives to help, are stronger when health crises are rarer, when more other people are also likely to offer to help, when offers are more likely to be observed, and when observers are more likely to be ignorant. Without a formal analysis, we can also surmise that incentives to signal via care are stronger when there are fewer similarly effective other opportunities to signal loyalty or desirability as an ally.

Similar factors are relevant when A is choosing a political ideal point c to use when voting on how much health care to give N members of some group, such as a company, church, or nation. If there are M similar other donors in the group who would similarly pay for this care, the FOC is

$$(1 + wqM)\text{Eu}_A^c = wqp(1 + M) \text{u}_R^h + (\beta/\lambda)\text{Eu}_R^d + wp(M \text{u}_M + N \text{u}_N)/c(q).$$

Here $\beta$ is the chance that associates who don’t already know q will observe A’s ideal point, and $\lambda$ is the chance that A is pivotal in the election. Signaling incentives for choosing c (right hand term) are thus more likely to dominate instrumental incentives (middle term) when the chance of being pivotal in an election is smaller than the chance that ignorant observers will learn about one’s political ideal point. This is a plausible description of modern elections.

Finally, note that signaling incentives increase as the utility difference between being in or out increases relative to other margins. That is, the signaling incentive is mostly due to $\text{Eu}_p = u(h,r) - u(h,r)$, and so if this utility difference increases relative to the other margins $u_h, u_r, u, u_r$, the signaling incentive for care should become larger compared to the direct instrumental incentive for care.

**Signaling limits**

Over-care is a feature common to all of the above models. This is because everyone wants to be thought of as someone who will remain in an alliance. The worst possible type has some low chance of long-standing alliance, and chooses some low level of care, while all other types over-care to distinguish themselves from this worst type.³

There may well be social benefits from helping people to learn about how their associates rate in terms of their loyalty and desirability as allies. If, however, such benefits are less than the social losses from signaling via over-care, one might consider discouraging such signals. If enforceable, this might be done via a maximum allowed care c. Alternatively, one might prefer a tax $\tau$ on care, so that expected utility becomes

$$\text{Eu}^h = \text{Eu}^d(h_A, r_A - (1 + \tau)c, q) + wqp \text{u}_A^d(h_B + c, r_B).$$

Note that the above discussion of voting over care suggests there may be difficulties with using democracy to imposed such limits or taxes; voters may actually choose subsidies over taxes.

Note also that such limits or taxes on health care are the opposite of those suggested by the standard model of adverse selection in health insurance. In that model, the signaling action is the fraction of medical expenses insured, and the type signaled is the risk level for events which will trigger expensive care. Those with lower risk signal that fact by agreeing to accept less than full insurance. In this standard model, it is insurance subsidies or minimum insurance levels that might mitigate signaling losses.

**Explaining health policy puzzles**

Even if our qualitative descriptions of ancestral incentives are reasonable, and even if our formal models capture their relevant essence, there remains the issue of how well they can really explain various puzzling features of modern health care behavior. Which modern behaviors would result from given inherited behavioral strategies can depend on subtle detail about how such behavior was encoded and cued.

For example, if our evolved taste for fat depended in a detailed enough way on relevant contextual features, like the amount of fat available and how physical our labor is, there would be no obvious reason to expect us to now have an excessive taste for fat. We only expect a mismatch if the contextual features that evolution found to be sufficient to distinguish between ancestral environments are in fact insufficient to distinguish modern environments from similar ancestral ones. For example, the physicality of labor might not have varied enough then to make it worth having tastes for fat depend on that.

Given this ambiguity, the following discussion of how the models above may explain current health puzzles must be consider somewhat speculative.

³ One can imagine models in which people attempt to signal how unlikely they are to ally with someone. Soldiers at war, for example, might signal to comrades how little empathy they had for enemy civilians by treating them especially roughly. Such models, however, seem an unlikely basis for most health care.
We seem to have inherited a disposition to classify associates by their strength of alliance with us, and to be genuinely concerned about the outcomes of strong allies, conditional on their remaining our allies. We expect "groups", i.e., correlations among who is allied with whom, and we expect many things to change with one's "status", i.e., with one's number and quality of allies. In particular, we expect more crisis events if we end up as low status, and when we expect more crises, we should set our stress levels higher, and invest less in health. We also prefer an associate to act as if she were confident of becoming high status, because she is less likely to remain an ally we care about if she ends up as low status. We thus prefer our associates to pick lower stress levels and to invest more health than they would choose for themselves.

In addition to genuine concern for allies, we also want to create the impression that our alliances are strong and lasting. We want everyone to think that we and our allies will end up as high status. Since those who were sure that their alliances would soon end would have much less reason to care for sick allies, we distinguish ourselves from those types by giving sick allies lots of care. We give enough care that those who are less optimistic about alliance outcomes do not find it worth their while to mimic us, even if they could thereby fool many observers. Though we may not be conscious of our motivations, this additional care is done in large part for appearance sake; we may be more concerned with been seen as putting in effort than with whether effort actually helps.

In principle there are many channels to signal allegiance. Primates and hunter-gathers would groom each other, share food, provide lodging for travelers, help in work parties such as hut building, host leisure parties, and help to avenge the killing of an ally. In many of these areas, however, self-help and impersonal markets have displaced ancient gift-exchanges; mirrors allow self-grooming and you can buy food at a market. Also, aid given in frequent small amounts could mainly only signal short-term allegiance; an ally who intended to betray you would likely keep grooming you until the last day. Thus, large infrequently-needed aid, such as for severe illnesses or revenge killings, should have had a unique ability to signal long-term allegiance. And since the modern legal systems have limited our ability to signal allegiance via revenge, health care seems one of the few remaining of our ancient ways to signal long-term allegiance.

We may thus purchase lots of health insurance for our family, and push for lots of care for our dying parents, in order to show how much we care about our family. That is, we can't stand to be thought of as the sort of uncaring heel who wouldn't try everything possible. But we are not very concerned about private signals about the quality of medical care, since our unconscious goal is mainly the appearance of effort. We do respond much more, however, to publicly visible quality signals that our intended audience would likely see as well. The marginal health-value of medicine may therefore be low, both because we spend more than is useful and because we have little incentive to privately monitor quality.

Our perception of "tribe" is plastic in many ways. Feelings of "us" vs. "them" are triggered by families ties, and probably also by distinguishing ethnicity, race, speech, and dress. Rulers have for millennia attempted to induce citizen loyalty by having citizens think of the nation as a tribe. National wars have likely entrenched this association, since among our ancestors who you went to war along side was likely a very strong signal of who was in your tribe. National health care was initially intended to trigger ancient dispositions to gratitude, and thereby induce citizen loyalty. Once nations became thought of as tribes, citizens and politicians supported national health insurance in order to show that they care about sick citizens of their nation, and to show other citizens that the nation is loyal to them. Thus the primary function of national health insurance may be to show social solidarity, rather than to respond to any failure in the market for health care. And since the world is not (yet) thought of as a tribe, there is little support for international health insurance.

Larger more ethnically diverse nations, such as the United States, might have a weaker sense of nation as group, and hence have less support for national health insurance. The main exception might be for high status people most seen as needing care, i.e., the elderly and Medicare. In the absence of national health insurance, corporations may like to offer health insurance to induce loyalty of employees. If wars cement the notion of nation as tribe, national health insurance may be most likely to arise just after a severe war such as world war two, in nations most severely effected by that war.

We may have always cared more about the health of associates than they themselves care, relative to their other resources, but perhaps hunter-gather societies offered few opportunities to regulate the health behaviors of associates. Today, however, we have stronger governments which are able to ban many health-risking foods, drugs, devices, and activities. And so we do ban. More
bans are applied to low status people; we are all disposed to trust them less to act as if they were confident they will become the high status allies we want them to assume they will be. Note that low status people today need not be worthy of this distrust; we could be disposed to distrust them regardless of how they act.

The relation between health and social status may be more direct and causal than between health and related factors like income, social supports, or a sense of control. It may be that we evolved to choose our stress level directly from an estimate of our social status, an estimate which combines cues such as our strength, material wealth, number and quality of associates, and the types of relationships we have with each associate (e.g., the level of control in that relationship). A self-estimation of social status may also be a primary input into choices of other health investments or health-risking behaviors.

To some degree we may provide health care in order to induce a "placebo effect" of less stress in those cared for. Those who are assured by our efforts that they will remain allies may unconsciously choose to invest more in health, such as by choosing lower stress levels. And since we prefer low status people to invest more in health, we have an incentive to give these assurances. Since we approve of the health choices of high status people, however, we have little marginal incentive to assure them of their status. Thus we feel that the rich, i.e. high status folks, do not need much more care than the "rest of us", i.e., middle status folks.

Our incentive to induce this placebo effect is stronger if the status-stress relation is frozen at high levels. That is, if our unconscious subsystems for choosing stress levels based on status-estimates are relatively hard-coded and impervious to conscious modification, we may be invoking much higher levels of stress response than are appropriate for a world with as few physical predators and other dangers as ours. The near-zero marginal health value of medicine also suggests that if there is a substantial positive placebo effect of such care, the other marginal health effects of medicine must be negative.

We can also perhaps understand why we keep spending more and more on health care. The primary ancestral function of leisure seems to have been social bonding; "partying" cemented social ties. Since leisure seems to be a luxury in the modern world, receiving a larger fraction of resources as people get richer, we can guess that for our ancestors investments in social status were relatively more important as wealth increases. If the primary function of health care spending is also to cement social relations, we can understand why health care spending also seems to be a luxury at the national level. We might also make sense of the demographic transition, reductions in surviving children per parent with increasing wealth, if the social status of one's children also becomes relatively more important with increasing wealth.

The basic idea here seems to be that allies were our ancestors' primary long term capital good, beyond health and children. In good times they invested in collecting and cementing allies, and in bad times they drew on those allies to help them survive. If poverty makes investing in social status and alliances less important, then we might expect a breakdown of status systems and altruism toward marginal allies in situations of extreme deprivation. And in fact, starvation can induce people to laugh at the suffering of associates, and can lead the young to steal food from the mouths of elders who would in some other situation have treated as high status [99]. In some ways morals may be luxuries which the very poor can not afford.

If the allocative benefits of learning who is really likely to be of high status outweigh the signaling losses from excessive care, we might want to encourage health care spending. If jockeying for status is mostly a zero-sum game, however, we might want to discourage such signaling by taxing or limiting health care spending. This is the opposite of what is suggested by standard models of adverse selection in health insurance. If people attempt to signal allegiance via their votes and policies, however, subsidizes and minimum spending levels may be more likely.

Acknowledgements

Appendix. Review of one-dimensional signaling

Most formal models of continuous signaling behavior are models of separating equilibria where a one dimensional action signals a one-dimensional type. The models in this paper are no exception. We now review general results for such models.

Assume agents vary according to their type \( \theta \), which is distributed on \([\underline{\theta}, \overline{\theta}]\) according to a c.d.f. \( F(\theta) \). Assume each agent chooses an action \( a \) from the real line, and this results in a utility \( U(a, \theta, \theta') \) when her type is \( \theta \), and she is perceived by observers to be of type \( \theta' \).

If observers have full information about \( \theta \), agent utility is always \( U(a, \theta, \theta) \), and so the agent’s equilibrium action choice \( a(\theta) \) maximizes this expression. If observers do not directly observe \( \theta \), however, they will interpret the action \( a \) as a signal of type \( \theta \). And if the equilibrium \( a(\theta) \) is monotonic, \( a \) will completely reveal \( \theta \), resulting in a separating equilibrium. In this case, the agent in essence chooses perceived type \( \theta \) to maximize \( U(a(\theta), \theta, \theta) \). In equilibrium, we must have \( \theta = \theta' \), resulting in utility \( U(a(\theta), \theta, \theta) \).

Table 1 compares the full information case to the case of signaling via a separating equilibrium. The first order conditions (FOC) show that in the signaling case, the agent not only considers the direct costs and benefits of an action, but also the degree to which the action will influence perceptions of the agent’s type. The second-order condition (SOC) is also changed.

Assuming \( \mathcal{U}_3 > 0 \), the integrals equations shown for action and utility are of the form \( g(\theta') = \int_{\underline{\theta}}^{\theta} G(\theta') \, d\theta \) for \( \theta' \in [\underline{\theta}, \overline{\theta}] \). In both cases the boundary condition is that \( a(\theta) \) satisfies the FOC for the full information case. That is, the worst possible type, who will not escape being seen as such, ignores the value of signaling when choosing her action.

The action integral for the signaling case bears no obvious relation to the action integral in the full information case, indicating that actions which serve as signals need not be at all close to the actions which would be taken for direct benefits. The utility integral shows that signaling is expensive; the cost of signaling reduces equilibrium utility by exactly eliminating the local benefit of being perceived to be a better type.

Signaling can serve valuable sorting functions, such as assigning better skilled people to more important jobs. In the absence of such sorting functions, however, the utility losses from signaling can in principle be mitigated by imposing limits or taxes on signals. Utility can be improved, for example, by imposing a maximum action \( \tilde{a} \), requiring \( a < \tilde{a} < \bar{a} \). In this case types \( \theta \) will not fully separate, as all types \( \hat{\theta} \) in some \([\theta_1, \bar{\theta}]\) will choose the same action \( \hat{a} \), where \( \theta_1 \) satisfies something like \( U(a(\theta_1), \theta_1, \theta_1) = U(\hat{a}, \theta_1, [\theta_1, \bar{\theta}]) \).

References

[52] Hunink Maria, Goldman Lee, Tosteson Anna, Mittleman Murray, Goldman Paula, Williams Lawrence, et al. The


