SOME USES OF MODELS OF QUANTITATIVE GENETIC SELECTION IN SOCIAL SCIENCE

MICHAEL D. WEIGHT[1] AND HENRY HARPENDING

Department of Anthropology, University of Utah, Salt Lake City, UT, USA

Summary. The theory of selection of quantitative traits is widely used in evolutionary biology, agriculture and other related fields. The fundamental model known as the breeder's equation is simple, robust over short time scales, and it is often possible to estimate plausible parameters. In this paper it is suggested that the results of this model provide useful yardsticks for the description of social traits and the evaluation of transmission models. The differences on a standard personality test between samples of Old Order Amish and Indiana rural young men from the same county and the decline of homicide in Medieval Europe are used as illustrative examples of the overall approach. It is shown that the decline of homicide is unremarkable under a threshold model while the differences between rural Amish and non-Amish young men are too large to be a plausible outcome of simple genetic selection in which assortative mating by affiliation is equivalent to truncation selection.

Introduction

In genetics quantitative traits like stature, body weight, milk yield in cattle, yield per acre of a field crop or the outcomes of pencil and paper tests for cognitive ability or personality, are traits that can be measured on a continuous scale. The theory of the evolution of quantitative traits under selection is well established in plant and animal breeding and in evolutionary biology (Lande, 1976; Falconer, 1981; Walsh & Lynch, 2015). The origin of quantitative genetic theory is R. A. Fisher (1919), who modelled a quantitative trait as the additive effects of a large number of genetic loci, allowing for dominance within each locus, and random environmental effects. Fisher's paper ended the controversy between Mendelians and Biometricians, the two competing approaches to understanding the transmission of heritable traits in the early twentieth century. Many regard this paper, after Darwin's *Origin of Species*, as the beginning of modern evolutionary biology (Provine, 2001).

In practice Fisher's basic model is simple and robust despite potential complications and elaborations. Sufficient conditions for the key results are the stochastic

---

[1] Corresponding author. Email: michael.weight@utah.edu

independence of genotypes and environments, the absence of dominance and epistasis (non-additive interactions between genes at different genomic locations) and a linear regression of genotypic values on phenotypes (Nagylaki, 1992). Under these assumptions the breeder's equation relates the change in the average value of some quantitative trait in response to selection as:

$$r = h^2 s.$$

In this equation r, the response to selection, is the difference between the average phenotype of parents and the average phenotype of offspring after a single generation of selection. The term h^2 is the narrow-sense heritability, or the fraction of phenotypic variance in the population attributable to the additive effects of genes. Finally, the selection differential, s, denotes the difference between the average phenotype in the population and the average phenotype of parents of the next generation.

Given two terms of the breeder's equation, one can estimate the third. Agriculture favours estimating heritability through selection experiments. Breeders or experimenters fix the selection differential by choosing parents from the population under study. The response r is obtained by measuring the offspring after selection. The heritability of the trait h^2 immediately follows from the breeder's equation. Such experiments cannot be done in the social sciences. Instead, heritability is estimated from similarities among relatives.

Complications can occur in the simple theory. For example, epistatic interactions between loci violate the simple set of assumptions. Remarkably, these tend to cancel each other over time. Crow (2010) suggested that natural selection generally favours changes that lead to the simple additive model given by the breeder's equation.

Complexities have cast doubt over human heritability estimates. For example, a standard source of data has been pairs of identical and fraternal twins. Heritability estimates then assumed that identical twins separated at birth are genetically identical and raised in random environments. Challenges to this assumption arise from many sources: shared uterine environment of identical twins before birth, non-random placement of adopted twins and others (Kamin & Goldberger, 2002).

Data from the Human Genome Project render many criticisms of twin studies moot. Identical twins were of interest because they shared more or less the same genome, but suspicions of shared environment persisted. Today the state of technology is to type approximately a million single nuclear polymorphisms (SNPs) in a large sample of subjects and measure the relationship between individuals by the number of SNPs they share. Realized kinship to the order of second cousins and lower is assessed for tens of thousands of typed subjects. This technology renders studies of real pedigree relationship nearly obsolete while simultaneously validating the conclusions about human heritability estimates that had been so widely denounced in the twentieth century (see Visscher *et al.*, 2008; Yang *et al.*, 2010; Davies *et al.*, 2011; and Lee & Chow, 2014, for further discussion).

Thought experiment

Assortative mating can be equivalent to directional selection. Assortment for a trait generates either distinct groups or a continuum: those with more of the trait at one

extreme and those with less of the trait at the other, within a population. To illustrate this consider a simple thought experiment about assortative mating for stature in a population. Assign average stature of 175.26 cm for males and 162.56 cm for females, each with a standard deviation of 7.62 cm. These figures are close to observed values in North America. Suddenly assortative mating by height occurs such that those taller than the population mean mate only with others also above the mean while those shorter than the mean mate only with others shorter than the mean.

First, it is assumed that the differences between male and female stature reflect fixed differences in the development of the two sexes. The data can then be normalized for each sex and stature in standard deviation units (SDs) measured. The pooled population has a mean of zero and a standard deviation of 1.0. The new units, called Z-scores, are obtained by subtracting the population mean from everyone's stature (175.26 cm from each male, 162.56 cm from each female) and dividing by the standard deviation of 7.62 cm. After imposition of the mating rule the population separates into two populations: talls and shorts. With this normal distribution assumption the overall normal distribution has been split into two distributions. The mean of a left half-normal distribution is −0.8 SDs and that of the right half + 0.8 SDs. The difference in population means is therefore 1.6 SDs, or 12.19 cm in the original units.

The new subpopulations mate at random within themselves. The two sets of parents differ by 1.6 SDs. After reproduction, assuming an additive heritability of 0.80 for stature, the group difference will be, following the breeder's equation, $0.8 \times 1.6 = 1.28$ SDs, corresponding to $1.28 \times 7.62 = 9.75$ cm. After the first generation of reproduction the genotypic group means differ by 9.75 cm without further movement between groups.

Before the new mating system the overall variance of stature was 1.0. After a single generation of the new mating system the between-subpopulation variance is 0.64^2 or 0.41 on the standard deviation scale. The fraction of stature variance that is between groups is now 41%. The variance of stature within each of the two subpopulations is the complement or 0.59. Back on the scale of centimetres, the standard deviation within populations is the square root of 0.59 (0.773), corresponding to 5.9 cm.

Richard Lewontin (1972) famously estimated the fraction of genic variance among continental human groups to be about 15% for a collection of presumably neutral genetic markers. That estimate of the ratio of between-group to total variance has remained constant as the number of loci available to estimate it grew from dozens to hundreds of thousands. After a single generation of this thought experiment the stature variance between groups at genetic loci influencing stature is between two and three times as great as the overall neutral genic variance among major human populations. Yet there is absolutely no selection in this thought experiment. Everything is selectively neutral. On the other hand, from the perspective from within subpopulations, there has been strong selection imposed by the mating rule allocating persons to groups according to stature. The mechanism of selecting according to a threshold is called truncation selection. It is, for example, the most common selection rule in animal and plant breeding. Here, the threshold was at the mean height of the overall population.

It is worth pointing out that after assortative mating the genetic variance within each group is somewhat smaller than what it otherwise could be as a result of negative linkage disequilibrium between causal sites (Bulmer, 1980). This may seem somewhat paradoxical because of the overall positive linkage disequilibrium between causal sites

in the total population (including both talls and shorts) induced by the assortative mating itself (Fisher, 1919; Crow & Felsenstein, 1968; Nagylaki, 1982). The reader may consult Lee and Chow (2013) for an intuitive explanation of the so-called Bulmer effect of negative disequilibrium. If random mating prevails within each group after assortment, then the Bulmer effect decays; the genetic variance within each group will thus increase. The most important feature of the genetics to keep in mind is that the classical increase in the overall genetic variance due to assortative mating, first described by Fisher himself, corresponds to the divergence in means between the two groups formed by the assortment process.

The mating rule in this thought experiment about the effects of truncation selection acted on a biometric trait. Suppose, instead, individuals in the population were similarly partitioned according to a threshold along a behavioural trait. Bailey (1998) pointed out that behavioural traits like religiosity, extraversion, neuroticism and even divorce are all moderately heritable. Indeed, Eric Turkheimer in an essay titled 'Three laws of behaviour genetics and what they mean' (Turkheimer, 2000) gives 'all human behavioural traits are heritable' as his first law of behaviour genetics. Rejecting the null hypothesis that behaviours are not heritable is therefore no longer of much interest (Bailey, 1998; Turkheimer, 2011). The genetic effects of the mating rule in the thought experiment on behavioural traits therefore mirror those illustrated using height. Despite this, the social sciences have generally neglected quantitative genetic models even though many other quantitative traits like IQ (Herrnstein & Murray, 1994; Jensen, 1998), the propensity to develop mental illnesses like bipolar disorder and schizophrenia (Owen & O'Donovan, 2003; Barnett & Smoller, 2009; Wray & Visscher, 2010), aggression and antisociality (Barker *et al.*, 2009) and other behavioural traits have significant social and economic consequences.

If the rule had its origin as a purely culturally transmitted preference then the effect on the genome would be exactly the same as those described in the thought experiment. Purely learned preferences or socially imposed assortments would 'pick up' any relevant genes, giving an identical genetic effect. Pervasive social transmission of these preferences would therefore probably amplify the heritability of behavioural traits. Consider new popular social traits that rapidly spread like a new religious movement, tofu consumption or crack cocaine use. Initially there may be no genetic basis for the cultural success of these traits, but suppose the trait fostered reproductive success. Standing genetic variation favouring the trait is then subject to selection. Consequently, the traits would become more heritable as initially rare genes present within the population became more common.

An older example of this effect is due to Waddington (1956). He found that flies treated with ether were prone to develop a second thorax. Without the ether treatment the trait, bithorax, was absent. However, bithorax began to appear even in the absence of ether after 20 generations of selection for flies vulnerable to the ether treatment. Alleles that slightly increased the probability of bithorax, but not enough to bring out this phenotype without the large effect of the ether treatment, were apparently positively selected during this experiment. Selection then raised their frequencies high enough to trigger bithorax in the absence of ether. The observed heritability of bithorax increased through a process that Waddington called *genetic assimilation*. Though superficially like Lamarckian inheritance of acquired characteristics, genetic assimilation has a well-studied quantitative genetic basis.

Genetic accommodation (West-Eberhard, 2003; Braendle & Flatt, 2006) is a generalization of genetic assimilation. Even if the hypothetical examples of tofu or crack fostered reproductive success through purely social mechanisms, individuals whose biochemistry was congenial to the new trait would participate more easily and their genotypes would increase in frequency. The preference would become heritable even though it was not initially. The idea that cultural adaptation renders genetic adaptation irrelevant is widespread in anthropology. Instead, genetic accommodation suggests that the pervasiveness of cultural transmission in our species has increased overall rates of genetic adaptation (see e.g. Hawks *et al.*, 2007; Cochran & Harpending, 2009; Laland *et al.*, 2010; Richerson *et al.*, 2010, for further discussion of cultural pressures increasing the rate of human genetic adaptation).

The Old-Order Amish

Assortative mating

Are there concrete ethnographic examples of assortative mating about a trait analogous to height with the above kind of strong effect? There are several candidates. Cochran *et al.* (2006) proposed that a similar mechanism operated in the evolution of European Ashkenazi Jews during the Middle Ages: the Ashkenazim were a population closed to immigration, but with the possibility of emigration, that experienced strong selection for economic and managerial success. The Parsi of India, a closed class of managers, scholars, musicians and other economically and culturally successful individuals (Nelson, 2012), are another distinct population in which assortment may have been important in their evolution. Recently Charles Murray in *Coming Apart* (2013) proposed that assortative mating by education, socioeconomic status and IQ is splitting America into distinct and different classes – the 'coming apart' of the title.

The Old-Order Amish, an Anabaptist sect in North America, also follow the pattern of negligible inward, but significant outward, gene flow, making them a potential candidate for selection by assortment. A model is described in which there is an underlying trait that we call AQ (Amish Quotient) by analogy with IQ derived from cognitive tests. It is postulated in this model that AQ predicts the probability of an individual remaining in the community when adolescents decide to remain and be baptized or to decline membership and emigrate. Specifically the measure of AQ is a personality measure associated with the plain and simple lifestyle of the Amish, postulated to indicate an individual's underlying affinity for membership in the community. Under the proposed mechanism of selection the AQ within the Amish population would increase each generation as those with lower AQs are truncated each generation through 'boiling-off' or defection. It is therefore predicted that declining boiling-off rates over time (see below) are a consequence of this selection mechanism increasing overall population AQ each generation within the Amish population.

The model performs well relative to alternative models of Amish defection and retention (see Ericksen *et al.*, 1980; Meyers, 1991, 1994; and Greksa, 2002, for review). While Old-Order endogamy is well described and understood, the trade-offs and constraints of defection are unclear. Reductions in defection rates over the last century (Meyers, 1991, 1994; Hostetler, 1993; Greksa, 2002) need explanation. Previous models of Amish defection emphasize rapid population growth in regions with finite resources,

particularly expensive farm land, and the resulting trade-off between culturally sanctioned high fertility and the ability of families to support their own farms (Markle & Pasco, 1977; Ericksen *et al.,* 1979, 1980; Wasao & Donnermeyer, 1996). Other models focus on the effects of occupational shifts consequent to economic constraints on farming. These models predict that increasing exposure to non-Amish lifestyles may tempt youth into defection or interfere with parents' ability to transmit Amish identity (Meyers, 1991; Greksa, 2002). Unfortunately predictions of these more familiar approaches to Amish defection are difficult to quantify and corroborate.

Endogamy

As the most traditional and 'plain' Anabaptist group, the Old-Order Amish are pacifists who reserve baptism for believing adults and have strict shunning practices of deviant church members (Hostetler, 1977, 1993; Roth, 2002). Such non-hostile ethnocentrism and social solidarity introduce a near absolute boundary that culturally uncouples the Amish from the 'English', their word for non-Amish (Huntington, 1957; Savells, 1988). Familiar Amish norms like carrying out intragroup discussions in Pennsylvania Dutch, favouring horse-and-buggy over automobiles, traditional 'humble' clothing, absolute pacifism and forsaking appliances like telephones, televisions, refrigerators and indoor toilets, are means to signal this uncoupling from their broader society (Fuchs *et al.*, 1990; Kraybill, 2001). Genealogical and genetic analysis reveals the Amish have been significantly isolated since Old-Order founding populations immigrated in the 18th and 19th centuries (Hostetler, 1993; Kraybill, 2001; Pollin *et al.*, 2007; Hurst & McConnell, 2010; Lee *et al.*, 2010). Arguably the Old-Order Amish are the most culturally and reproductively isolated Anabaptist group in the United States (Nolt, 1992; Hostetler, 1993; Kraybill, 2001).

Following this isolation, the Old-Order substitute communal aid for secular public support. Initiatives like welfare, Social Security, life insurance, barn insurance (Hurd, personal communication), government agricultural subsidies or public schools (if Amish schools are available) are therefore prohibited or strongly discouraged (Hostetler, 1969, 1993; Fuchs *et al.*, 1990; Hurst & McConnell, 2010; Kraybill, 2011) in the interests of communal self-reliance. Amish substitutes for institutional support are primarily mediated through kin networks so extended families bear the emotional and financial burden of the welfare of their members (Hurd, 1983; Hostetler, 1993; Hurst & McConnell, 2010). Children, especially males, rely on kin support to start businesses and farms when they begin their own families (Ericksen *et al.*, 1979, 1980; Hurd, 1981; Wasao & Donnermeyer, 1996). Highly fertile Amish mothers draw on relatives to help with their large families (Hewner, 1998; Kraybill, 2001). Costs of disabled, ill, and senior members are mitigated by family and community support (Hewner, 1998). The Amish are a healthy population with low incidence of infectious disease, low infant and adult mortality rates, high fertility and high standards of living (McKusick *et al.*, 1964a, b; Cross, 1969; Cross & McKusick, 1970; Fuchs *et al.*, 1990; Hewner, 1998; Greksa, 2002).

The 16PF

Wittmer (1970) reported the results of a personality questionnaire administered to 25 Amish and 25 non-Amish, 18- to 20-year-old men from Daviess County, Indiana.

The questionnaire was a widely used personality assessment called 16PF (for sixteen personality factors) developed in the 1940s by Raymond Cattell and others at the University of Illinois (Cattell & Eber, 1972). It went through several revisions over the years. Each form has 180 questions that are scored to yield sixteen numbers, or factors, supposed to represent the subjects' sixteen dimensions of personality. These are listed in Table 1 along with the adjectives in common use that supposedly describe each trait.

No use is made of the details of the traits in this paper: the sixteen numbers for each person in the study are simply regarded as self-reports of interesting characteristics. The test and the scoring system are designed to maximize the independence of the factors so the scores are not redundant indicators of the same trait. In some of the literature factor B is not reported nor used since it assesses intelligence rather than a personality trait. Here B is included since these data are treated in the same way as genetic markers like SNPs would be treated, i.e. the authors are agnostic about the interpretation and meaning of the factors.

To make sense out of the array of scores for the Amish and non-Amish sample a conventional method for data reduction is used called principal components analysis (PCA). This provides a least-squares best two-dimensional picture of differences among subjects and differences (i.e. correlations) among variables.

The process used is analogous to that of estimating IQ. An important finding from cognitive testing is that cognitive ability is essentially one-dimensional. For example, people with high mathematical ability also have large vocabularies. One empirical result of this is that almost anything constructed to test ability functions as a suitable proxy for a proper IQ test. SATs, GREs, MCATs and AFQTs are, for example, nearly interchangeable with IQ tests (Frey & Detterman, 2004; Beaujean *et al.*, 2006). In addition, the United States General Social Survey (GSS) surveys a large sample of Americans every few years. The results are conveniently available on the web (General

Table 1. Factors assessed in the 16F questionnaire

Trait label	High	Low
A	Outgoing, Warmhearted	Reserved, Detached
B	Abstract thinking, Fast learner	Concrete thinking, Slow learner
C	Unemotional, Calm	Emotional, Changeable
E	Assertive, Dominant	Humble, Co-operative
F	Cheerful, Lively	Sober, Taciturn
G	Conscientious, Persistent	Expedient, Undisciplined
H	Venturesome, Socially bold	Shy, Retiring
I	Tough-minded, Self-reliant	Tender-minded, Sensitive
L	Suspicious, Sceptical	Trusting, Accepting
M	Imaginative, Bohemian	Practical, Conventional
N	Shrewd, Discreet	Forthright, Straightforward
O	Guilt-prone, Worrying	Resilient, Self-assured
Q1	Radical, Experimental	Conservative, Traditional
Q2	Self-sufficient, Resourceful	Group-dependent, Affiliative
Q3	Controlled, Compulsive	Undisciplined, Lax
Q4	Tense, Driven	Relaxed, Tranquil

Social Survey, 2012). Each round of the survey presents each respondent with ten words, asks for their definitions, and then tabulates the variable WORDSUM that varies from 0 to 10 for that subject. WORDSUM is a useful proxy for IQ in the sense that differences among groups along the variable replicate reported differences among the same groups (Jensen, 1998) from more elaborate tests. The PCA of the results from Wittmer's study may reveal an analogous one-dimensional personality trait corresponding to the AQ.

Scattergrams along principal components (groups) and principal co-ordinates (traits) from the 16PF results are shown in Figs 1 and 2. Figure 1 shows patterns among group means, the Amish and non-Amish from Wittmer's study, along with a collection of other populations culled from the literature. Details are given in the caption of Fig. 1. The scatter of traits in the left panel is not very informative. There is no clear pattern, presumably reflecting the design goal of independence among the sixteen factors. The scatter among groups, in the right panel, is more interesting. There are three groups that are on the edge of the scatter: Amish, Chinese nurses and candidates for high-level management positions in the UK. The outstanding (AN) and ordinary (EN) Chinese nurses (see Zhang *et al.*, 2013, for data) are nearly identical to each other on these first two dimensions. Generic English (UK), according to data taken from Bartram (1992), are more similar to Indiana rural young men than they are to high-level managers in the UK. Figure 1 shows a broad perspective on group differences, especially the difference between the British norms and the non-Amish Indiana rural males.
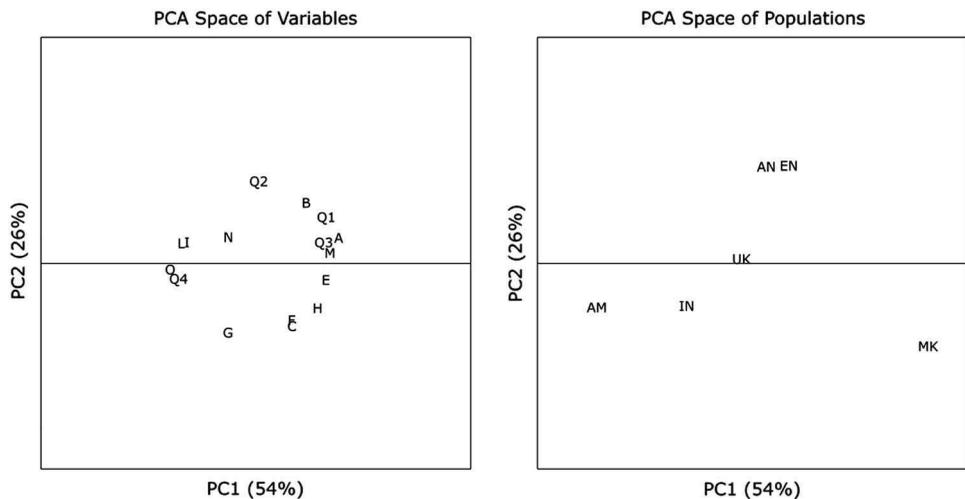

Fig. 1. Principal components analysis (PCA) of 16PF score differences among groups of 18- to 20-year-old Indiana men. The left panel portrays similarities (i.e. correlations) among the sixteen traits listed in Table 1. The right panel portrays differences among means of several large samples. AM are young Amish males from Indiana; IN are non-Amish males from the same Indiana county; AN and EN are outstanding and ordinary Chinese nurses, respectively; MK are candidates for high-level manager jobs in the UK; and UK are British norms of the 16PF test. Notice the lack of pattern in the left panel, indicating the relative independence of the sixteen traits. Several of the labels have been nudged on the figure to avoid overlaps.
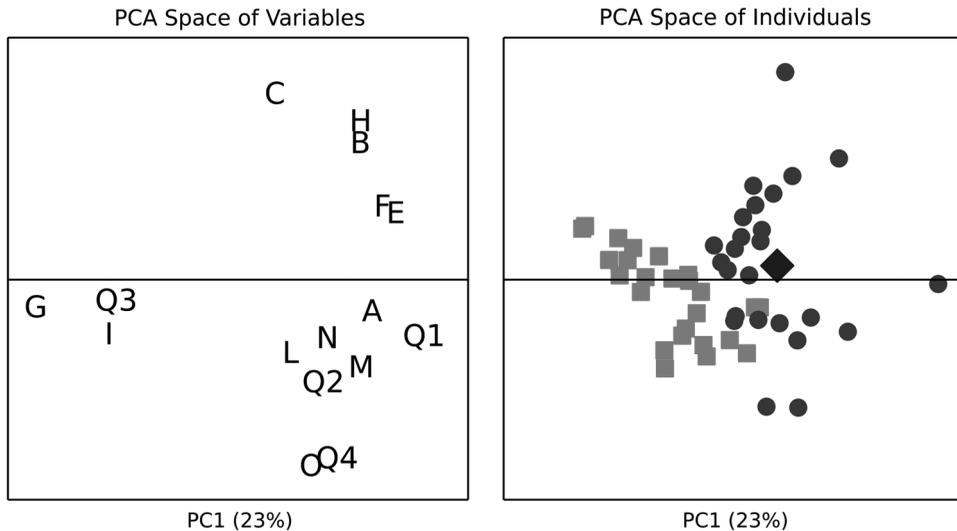
Fig. 2. Principal components analysis (PCA) of individual 16PF scores of 18- to 20-year-old Indiana men. The left panel portrays similarities among traits, and the right panel among individuals along with the overall UK population. Trait labels are given in Table 1. Amish young men are shown as circles, non-Amish as squares and the large diamond shows the centroid of the UK standard for comparison. The two panels are dual, so proximity between a trait in the left panel and a subject in the right indicates a high value of that trait in that subject. For example the Amish subjects have high values of 16PF trait G, which is called 'conscientiousness' and 'persistence' in the literature. Non-Amish young Indiana males have relatively higher scores on trait Q1, called 'radical' and 'experimental'.

The 16PF test itself is proprietary and widely used in personnel selection. Comparable data in the literature are scarce, accounting for this strange choice of Chinese nurses and UK managers. They were simply chosen by convenience to convey a sense of the magnitude of group differences. Even more dismaying is that a standard way to present 16PF results is to normalize everyone in a sample to a fixed mean. Consequently much of the published 16PF data are not comparable between groups.

The interest in Fig. 2 is to show individual differences among young rural Indiana men. The scatter shows a nearly complete separation of the Amish and non-Amish along the horizontal axis (the first principal component). The UK mean is shown for comparison in the right panel as a diamond. Remarkably, the overall UK mean is near the centre of the non-Amish scatter, implying that they are not very different. The Amish are strongly different. In a genetic model the differences are partially the accumulated effects of Amish voluntary endogamy along with boiling-off of the young Amish in each generation who decide to leave the Amish community.

The scatter of traits in the left panel of Fig. 2 helps identify those traits distinguishing the Amish and non-Amish along the horizontal axis. Traits G, Q3 and I on the extreme left of the horizontal axis are in the space of those with the highest AQ. The 16PF scores show that Amish describe themselves, albeit indirectly, as conscientious and persistent (G),

tough-minded and self-reliant (I), and controlled (Q3). Correspondingly they describe themselves as low in Q1 and E, scoring respectively as conservative/traditional and humble/co-operative. These imputed self-characterizations agree remarkably well with stereotypes of the Amish that they hold themselves according to ethnographies. The authors are agnostic about the meaning and uses of the 16PF and other similar tests except as indirect self-descriptions by subjects.

A model of (self-)selection

Under this model the horizontal axis in Fig. 2 is a numeric scale of a trait AQ identifying where an individual is in a suite of traits differentiating young Amish men from their non-Amish neighbours. Movement out of the Amish community is open to anyone, especially around adolescence. What would be the consequences for the Amish community if individuals on the low end of AQ were especially prone to leaving the population?

It is informative to bring a quantitative genetics model to the phenomenon as a well understood baseline for evaluating competing models. In this spirit it is *assumed* that AQ is a partially heritable quantitative trait. There are enough data on behavioural or psychological trait heritabilities to sensibly estimate the additive heritability of AQ, written h^2, at 0.5. It can be seen below that the conclusions are unaltered if the heritability is set to either a plausible limit of 0.25 or 0.75. Boiling-off rates vary among communities but 0.10 (i.e. 10% of youth choose to leave per generation) is a reasonable contemporary value.

Following standard practice in quantitative genetics it is now assumed that AQ has an underlying normal or Gaussian distribution in the Amish population. The 10% boiling-off rate is equivalent to losing the bottom 10% of that distribution each generation. It is also assumed that the boiling-off mechanism is equivalent to truncation selection, but other selection schemes will have some equivalent alternative truncation point.

With 10% with the lowest AQ leaving each generation, the average AQ of those who remain is 0.20 standard deviations if the population mean before emigration was 0. In other words this amount of self-selective emigration leaves behind a population whose average AQ is 0.2 standard deviations greater than it was before the emigration. Assume that mating occurs exclusively within the Amish. Fixing the heritability of AQ at 0.5 means that the offspring will regress halfway back to their parental mean.

The next generation of children will therefore have an average AQ of 0.10 standard deviations greater than their parents did before emigration. The process of selective emigration repeats each generation, increasing the mean Amish AQ by one-tenth of a standard deviation per generation. With 25 years per generation, 'Amishness' will increase by a full standard deviation in ten generations, or 250 years, unless changes in gene frequency lead to changes in heritability. This is substantial social evolution on a time scale of a few centuries.

The modelled AQ of an individual is simply the value of that individual along the *x*-axis of the right panel of Fig. 2. The computed standard deviations of AQ for the Amish and non-Amish subjects are 1.33 and 1.31 respectively, giving the within-group standard deviation to be 1.32 units. Interestingly the variances within groups, i.e. the squared within-group standard deviations, differ very little. It is not an unusual finding

in experiments that selection has only small effects on the variance of a selected population. Falconer and Mackay (1995) discussed the outcome of experiments at length in their chapter 12.

The means of the two groups are -1.4 (Amish) and 1.4 (non-Amish) units. Their difference, corresponding to the distance between the two group means, is 2.8 units. Our model implies a difference of this magnitude requires 28 generations, or 700 years, which is substantially longer than the Amish have been in the United States. This may suggest that the heritability of AQ is higher than the assumed 0.5, that the boiling-off rate was higher in the past (for which there is support), or that the process has been going on longer than ten generations. This last possibility is unlikely. Currently no Old-Order exist in Europe. Persecution and political sanctions in the 17th and 18th centuries had spread European Amish families too far apart to form cohesive communities, forcing early European Amish to eventually assimilate with non-Amish (Gascho, 1937; Hostetler, 1977; Crowley, 1978; Nolt, 1992; Roth, 2002). Thus the magnitude of Amish endogamy was probably negligible in Europe compared with Amish endogamous mating in the United States. If correct, this suggests the Old-Order tenure in Europe did not contribute to the differences reported here.

Taking the results shown in Fig. 2 at face value implies the between-group difference (2.8 standard deviations) is too large to be attributable to just simple genetic change with a predicted difference of 1.0 standard deviations, or 1.5 standard deviations if heritability is assumed to be as high as 0.75. The best conclusion from comparing the data with the simple genetic model is that other social forces, e.g. family transmission, are mostly responsible for the difference between the Amish and non-Amish young men.

Homicide decline in European history

Much of quantitative genetics is about some measurable metric character in a population. A convenient and useful way to study many discrete characters is to treat them as threshold characters that only manifest when an underlying hidden quantitative trait exceeds some threshold. Cleft palate, for example, is a failure of the palate to close at the midline during intrauterine development. There is an underlying metric trait in the model that is the strength of the mechanism leading to midline fusion. Fusion fails for trait values below a threshold strength of this mechanism. The model further specifies what the incidence should be in relatives of varying degree. It predicts cleft palate risk in relatives of those affected quite closely (Fraser, 1970).

Cloninger *et al.* (1975b) proposed a threshold model for the prevalence of sociopathy and hysteria in humans. In their model both disorders reflect the same underlying heritable trait but the sociopathy liability threshold in females is higher than the male threshold. They consider hysteria to be a less deviant form of sociopathy in women whose liability is above the male, but below the female, threshold. Sociopathy prevalence in the general population, the higher proportion of sociopathy among relatives of female sociopaths, and the clustering of both male and female sociopaths within families closely fit the predictions of the model (Cloninger *et al.*, 1975a, b).

Eisner (2001) tabulated the historical decline in homicide rates, excluding war deaths, in several European countries from the fourteenth to the twentieth century. A typical

finding is that homicide rates fell from 40 per 100,000 people in 1300 to 1 per 100,000 in 1900, a forty-fold decline. This striking social change seems 'large'. A threshold model is considered to sharpen this intuition.

This simple model posits that the propensity to murder is a normally distributed trait in the population. Individuals with high propensity will murder while those with low propensity will not. Following Eisner the threshold for homicide is placed at 40/100,000. Tables of the standard normal distribution show that this threshold corresponds to the right tail of the propensity distribution beyond 3.35 standard deviations. After 600 years the rate declined to 1 in 100,000, corresponding to the tail beyond 4.26 standard deviations. If the whole distribution moved with no change in the variance, as expected with the simple Fisher model, the shift was just $4.26 - 3.35 \approx 0.9$ standard deviation. On a stature scale in which the standard deviation is 7.26 cm this corresponds to a change in mean stature of 6.53 cm in 600 years. This amount of change in 600/25 or 24 generations amounts to just over 1/4 cm per generation. This is in the order of known changes in stature in European history. From this perspective the decline in homicide is not especially remarkable at all and plausibly reflects genetic change. This is not to claim that genetic change accounts for most of the decline in homicide. However, it does suggest that this model is an informative yardstick for measuring how substantially homicide rates declined.

Modifying the assumptions of this plausible quantitative genetics model yields some insight into its robustness. Assume, for example, that murderers have a variable number of victims and that the average number of victims is two. Since the average number of homicides per murderer is two, the thresholds are estimated by a decline from 20 to 0.5, rather than 40 to 1, per 100,000 people. These numbers correspond to 3.54 and 4.42 standard deviations on the right of normal curves. The change is therefore about 0.9 standard deviation – the same result found in the simple model. The qualitative conclusion of unremarkable change in 600 years remains the same.

A perhaps more realistic elaboration of the model is that murderers murder other murderers. The homicide rate under these conditions reflects not the density, but the square of the density, of murderers assuming a mass action rate of random encounters between murderers. In this case the change in the square root of the homicide rates is looked at so that the tails of this distribution in 1300 and 2000 are respectively at 2.73 and 2.05 standard deviations to the right of the mean. This elaborated model predicts a shift in the distribution of just 0.7 standard deviation, even less than in the simple model. It is concluded that the simple quantitative genetic model is fairly robust to perturbations. Frost and Harpending (2014) examined English homicide and execution rates, in addition to death rates of murderers awaiting trial, during this period in more detail. They also concluded that genetic transmission could account for much of the social change.

Conclusions

In this paper an index was constructed from responses to a personality questionnaire that distinguishes Old-Order Amish from their non-Amish neighbours. The associated genetic model posited that the differences between the two populations are a

consequence of selective 'boiling-off' of Amish individuals over time. This process is equivalent to truncation selection from the viewpoint of the Amish population since those who found the community least congenial left. This example illustrates that groups in particular demographic or occupational niches with reproductive assortment can undergo genetic change that significantly differentiates them from other groups.

The data suggest that the simple model incompletely accounts for the observed group differences. Even if the heritability of AQ is high, differences accumulate under genetic transmission at about 0.1 standard deviations per generation. The difference between the group means is 2.8 standard deviations. If this difference were a consequence of purely genetic change such a difference would require 28 generations, or 56 generations with a more reasonable heritability estimate of 0.25. Historical evidence suggests that the pattern is no older than ten generations. The implication is that the cultural difference between the Amish and their neighbours reflects mostly strong cultural transmission. Unfortunately, previous cultural transmission models of Amish defection proposed by Markle and Pasco (1977), Ericksen *et al.* (1979, 1980), Meyers (1991, 1994), Wasao and Donnermeyer (1996) and Greksa (2002) give only plausibility arguments that are difficult to quantify and falsify. On the other hand, quantitative genetic theory is well established in agriculture and other areas of biology like biometrics and conservation biology. The present study illustrates the utility of quantitative genetic models as a useful yardstick for describing social traits and evaluating transmission models, suggesting that quantitative genetics deserves more attention and use in human evolutionary biology.

The historical decline in European homicide rates since the fourteenth century was also considered. Gregory Clark (2007) described this decline as part of a larger heritable change in Europe leading up to the Industrial Revolution. Several variants of a simple model of selection against homicide indicate such genetic change is entirely plausible.

Acknowledgments

References

Bailey, J. M. (1998) Can behaviour genetics contribute to evolutionary behavioural science? In Crawford, C. & Krebs, D. (eds) *Handbook of Evolutionary Psychology: Ideas, Issues, and Applications*. Psychology Press, pp. 211–234.

Barker, E. D., Larsson, H., Viding, E., Maughan, B., Rijsdijk, F., Fontaine, N. & Plomin, R. (2009) Common genetic but specific environmental influences for aggressive and deceitful behaviours in preadolescent males. *Journal of Psychopathology and Behavioural Assessment* **31**, 299–308.

Barnett, J. H. & Smoller, J. W. (2009) The genetics of bipolar disorder. *Neuroscience* **164**, 331–343.

Bartram, D. (1992) The personality of UK managers: 16PF norms for short-listed applicants. *Journal of Occupational and Organizational Psychology* **65**, 159–172.

Beaujean, A. A., Firmin, M. W., Knoop, A. J., Michonski, J. D., Berry, T. P. & Lowrie, R. E. (2006) Validation of the Frey and Detterman (2004) IQ prediction equations using the Reynolds Intellectual Assessment Scales. *Personality and Individual Differences* **41**, 353–357.

Braendle, C. & Flatt, T. (2006) A role for genetic accommodation in evolution? *Bioessays* **28**, 868–873.

Bulmer, M. G. (1980) *The Mathematical Theory of Quantitative Genetics*. Oxford University Press, New York.

Cattell, R. B. & Eber, H. W. (1972) *The Sixteen Personality Factor Questionnaire (16PF)*. Institute for Personality and Ability Testing, Champagne.

Clark, G. (2007) *Genetically Capitalist? The Malthusian Era and the Formation of Modern Preferences*. URL: http://faculty.econ.ucdavis.edu/faculty/gclark/papers/Capitalism%20Genes.pdf.

Cloninger, C. R., Reich, T. & Guze, S. B. (1975a) The multifactorial model of disease transmission: II. Sex differences in the familial transmission of sociopathy (antisocial personality). *British Journal of Psychiatry* **127**, 11–22.

Cloninger, C. R., Reich, T. & Guze, S. B. (1975b) The multifactorial model of disease transmission: III. Familial relationship between sociopathy and hysteria (Briquet's syndrome). *British Journal of Psychiatry* **127**, 23–32.

Cochran, G., Hardy, J. & Harpending, H. (2006) Natural history of Ashkenazi intelligence. *Journal of Biosocial Science* **38**, 659–693.

Cochran, G. & Harpending, H. (2009) *The 10,000 Year Explosion: How Civilization Accelerated Human Evolution*. Basic Books, New York.

Cross, H. (1969) Some demographic characteristics of Amish society. In Hostetler, J. (ed.) *Conference on Child Socialization*, pp. 1–26.

Cross, H. E. & McKusick, V. A. (1970) Amish demography. *Biodemography and Social Biology* **17**, 83–101.

Crow, J. F. (2010) On epistasis: why it is unimportant in polygenic directional selection. *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 1241–1244.

Crow, J. F. & Felsenstein, J. (1968) The effect of assortative mating on the genetic composition of the population. *Eugenics Quarterly* **15**, 85–97.

Crowley, W. K. (1978) Old Order Amish settlement: diffusion and growth. *Annals of the Association of American Geographers* **68**, 249–264.

Davies, G., Tenesa, A., Payton, A., Yang, J., Harris, S. E., Liewald, D. *et al.* (2011) Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Molecular Psychiatry* **16**, 996–1005.

Eisner, M. (2001) Modernization, self-control and lethal violence. The long-term dynamics of European homicide rates in theoretical perspective. *British Journal of Criminology* **41**, 618–638.

Ericksen, E. P, Ericksen, J. A. & Hostetler, J. A. (1980) The cultivation of the soil as a moral directive: population growth, family ties, and the maintenance of community among the Old Order Amish. *Rural Sociology* **45**, 49–68.

Ericksen, J. A., Ericksen, E. P., Hostetler, J. & Huntington, G. (1979) Fertility patterns and trends among the Old Order Amish. *Population Studies* **33**, 255–276.

Falconer, D. S (1981) *Introduction to Quantitative Genetics*, 2nd edition. Longman, London.

Falconer, D. S. & Mackay, T. F. C. (1995) *Introduction to Quantitative Genetics*. Longman, London.

Fisher, R. A. (1919) On the correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* **52**, 399–433.

Fraser, F. C. (1970) The genetics of cleft lip and cleft palate. *American Journal of Human Genetics* **22**, 336–352.

Frey, M. C. & Detterman, D. K. (2004) Scholastic assessment or *g*? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science* **15**, 373–378.

Frost, P. & Harpending, H. C. (2014) Western Europe, state formation, and genetic pacification. *Evolutionary Psychology: An International Journal of Evolutionary Approaches to Psychology and Behaviour* **13**, 230–243.

Fuchs, J. A., Levinson, R. M., Stoddard, R. R., Mullet, M. E. & Jones, D. H. (1990) Health risk factors among the Amish: results of a survey. *Health Education & Behaviour* **17**, 197–211.

Gascho, M. (1937) *The Amish Division of 1693–1697 in Switzerland and Alsace*. Mennonite Historical Society, Lancaster.

General Social Survey (2012) *General Social Survey*. National Opinion Research Center, University of Chicago. URL: http://www3.norc.org/GSS+Website/.

Greksa, L. P. (2002) Population growth and fertility patterns in an Old Order Amish settlement. *Annals of Human Biology* **29**, 192–201.

Hawks, J., Wang, E. T., Cochran, G., Harpending, H. C. & Moyzis, R. K. (2007) Recent acceleration of human adaptive evolution. *Proceedings of the National Academy of Sciences of the USA* **104**, 20753–20758.

Herrnstein, R. J. & Murray, C. (1994) *The Bell Curve: The Reshaping of American Life by Differences in Intelligence*. Free Press, New York.

Hewner, S. J. (1998) Fertility, migration, and mortality in an Old Order Amish community. *American Journal of Human Biology* **10**, 619–628.

Hostetler, J. A. (1969) *Educational Achievement and Life Styles in Traditional Society: The Old Order Amish*. US Department of Health, Education, and Welfare, Office of Education, Bureau of Research, Washington, DC.

Hostetler, J. A. (1977) Old Order Amish survival. *Mennonite Quarterly Review* **55**, 352–361.

Hostetler, J. A. (1993) *Amish Society*. Johns Hopkins University Press, Baltimore.

Huntington, G. E. (1957) Dove at the window: a study of an Old Order Amish community in Ohio. Dissertation, Yale University, New Haven.

Hurd, J. P. (1981) Mate choice among the Nebraska Amish of Central Pennsylvania. Dissertation, Pennsylvania State University, State College Pennsylvania.

Hurd, J. P. (1983) Kin relatedness and church fissioning among the 'Nebraska' Amish of Pennsylvania. *Biodemography and Social Biology* **30**, 59–66.

Hurst, C. E. & McConnell, D. L. (2010) *An Amish Paradox: Diversity and Change in the World's Largest Amish Community*. Johns Hopkins University Press, Baltimore.

Jensen, A. R. (1998) *The g Factor: The Science of Mental Ability*. Praeger, Westport.

Kamin, L. J. & Goldberger, A. S. (2002) Twin studies in behavioural research: a skeptical view. *Theoretical Population Biology* **61**, 83–95.

Kraybill, D. B. (2001) *The Riddle of Amish Culture*. Johns Hopkins University Press, Baltimore.

Laland, K. N., Odling-Smee, J. & Myles, S. (2010) How culture shaped the human genome: bringing genetics and the human sciences together. *Nature Reviews Genetics* **11**, 137–148.

Lande, R. (1976) Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314–334.

Lee, J. J. & Chow, C. C. (2013) The causal meaning of Fisher's average effect. *Genetics Research* **95**, 89–109.

Lee, J. J. & Chow, C. C. (2014) Conditions for the validity of SNP-based heritability estimation. *Human Genetics* **133**, 1011–1022.

Lee, W. J., Pollin, T. I., O'Connell, J. R., Agarwala, R. & Schäffer, A. A. (2010) PedHunter 2.0 and its usage to characterize the founder structure of the Old Order Amish of Lancaster County. *BMC Medical Genetics* **11**, 68–81.

Lewontin, R. C. (1972) The apportionment of human diversity. *Evolutionary Biology* **6**, 381–398.

McKusick, V. A., Hostetler, J. A. & Egeland, J. A. (1964a) Genetic studies of the Amish. *Bulletin of the Johns Hopkins Hospital* **115**, 203–222.

McKusick, V. A., Hostetler, J. A., Egeland, J. A. & Eldridge, R. (1964b) The distribution of certain genes in the Old Order Amish. *Cold Spring Harbor Symposia on Quantitative Biology* **29**, 99–114.

Markle, G. E. & Pasco, S. (1977) Family limitation among the Old Order Amish. *Population Studies* **31**, 267–280.

Meyers, T. J. (1991) Population growth and its consequences in the Elkhart-LaGrange Old Order Amish settlement. *Mennonite Quarterly Review* **65**, 308–321.

Meyers, T. J. (1994) The Old Order Amish: to remain in the faith or to leave. *Mennonite Quarterly Review* **68**, 378–395.

Murray, C. (2013) *Coming Apart: The State of White America, 1960–2010*. Random House, New York.

Nagylaki, T. (1982) Assortative mating for a quantitative character. *Journal of Mathematical Biology* **16**, 57–74.

Nagylaki, T. (1992) *Introduction to Theoretical Population Genetics*. Springer, Berlin.

Nelson, D. (2012) India's dwindling Parsi population to be boosted with fertility clinics. *The Telegraph*. URL: www.telegraph.co.uk/news/worldnews/asia/india/ 9612009/Indias-dwindling-Parsi-population-to-be-boosted-with-fertility- clinics.html.

Nolt, S. M. (1992) *A History of the Amish*. Good Books, Intercourse.

Owen, M. J. & O'Donovan, M. C. (2003) Schizophrenia and genetics. In Plomin, R., Defries, J. C., Craig, I. W. & McGuffin, P. (eds) *Behavioural Genetics in the Postgenomic Era*. American Psychological Association, Washington, DC, pp. 463–480.

Pollin, T. I., McBride, D. J., Agarwala, R., Schäffer, A. A., Shuldiner, A. R., Mitchell, B. D. & O'Connell, J. R. (2007) Investigations of the Y chromosome, male founder structure and YSTR mutation rates in the Old Order Amish. *Human Heredity* **65**, 91–104.

Provine, W. (2001) *The Origins of Theoretical Population Genetics*. University of Chicago Press, Chicago.

Richerson, P. J, Boyd, R. & Henrich, J. (2010) Gene–culture coevolution in the age of genomics. *Proceedings of the National Academy of Sciences of the USA* **107** (Supplement 2), 8985–8992.

Roth, J. D. (2002) *Letters of the Amish Division: A Sourcebook*. Mennonite Historical Society, Goshen.

Savells, J. (1988) Economic and social acculturation among the Old Order Amish in select communities: surviving in a high-tech society. *Journal of Comparative Family Studies* **19**, 123–135.

Turkheimer, E. (2000) Three laws of behaviour genetics and what they mean. *Current Directions in Psychological Science* **9**, 160–164.

Turkheimer, E. (2011) Still missing. *Research in Human Development* **8**, 227–241.

Visscher, P. M, Hill, W. G. & Wray, N. R. (2008) Heritability in the genomics era–concepts and misconceptions. *Nature Reviews Genetics* **9**, 255–266.

Waddington, C. H. (1956) Genetic assimilation of the bithorax phenotype. *Evolution* **10**, 1–13.

Walsh, B. & Lynch, M. (2015) *Evolution and Selection of Quantitative Traits*. University of Arizona. URL: http://nitro.biosci.arizona.edu/zbook/NewVolume/_2/newvol2.html.

Wasao, S. W. & Donnermeyer, J. F. (1996) An analysis of factors related to parity among the Amish in northeast Ohio. *Population Studies* **50**, 235–246.

West-Eberhard, M. J. (2003) *Developmental Plasticity and Evolution*. Oxford University Press, New York.

Wittmer, J. (1970) Homogeneity of personality characteristics: a comparison between Old Order Amish and Non-Amish. *American Anthropologist* **72**, 1063–1068.

Wray, N. R. & Visscher, P. M. (2010) Narrowing the boundaries of the genetic architecture of schizophrenia. *Schizophrenia Bulletin* **36**, 14–23.

Yang, J., Benyamin, B, McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R. *et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* **42**, 565–569.

Zhang, L., Liu, B., Ren, H., Liu, Y. F. & Zhang, Y. (2013) The personality profile of excellent nurses in China: the 16PF. *Contemporary Nurse: A Journal for the Australian Nursing Profession* **43**, 219–224.