The genetic architecture of economic and political preferences

Daniel J. Benjamin[a,1], David Cesarini[b,c], Matthijs J. H. M. van der Loos[d], Christopher T. Dawes[e], Philipp D. Koellinger[d], Patrik K. E. Magnusson[f], Christopher F. Chabris[g], Dalton Conley[h], David Laibson[i], Magnus Johannesson[j,2], and Peter M. Visscher[k,2]

[a]Department of Economics, Cornell University, Ithaca, NY 14853; [b]Center for Experimental Social Science, Department of Economics, New York University, New York, NY 10012; [c]Research Institute of Industrial Economics, SE-102 15 Stockholm, Sweden; [d]Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, 3000 DR Rotterdam, The Netherlands; [e]Department of Politics, New York University, New York, NY 10012; [f]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE-171 77 Stockholm, Sweden; [g]Department of Psychology, Union College, Schenectady, NY 12308; [h]Department of Sociology, New York University, New York, NY 10012; [i]Department of Economics, Harvard University, Cambridge, MA 02138; [j]Department of Economics, Stockholm School of Economics, SE-113 83 Stockholm, Sweden; and [k]University of Queensland Diamantina Institute, Princess Alexandra Hospital, Woolloongabba 4102, Australia

Edited* by Eric S. Lander, The Broad Institute of MIT and Harvard, Cambridge, MA, and approved March 22, 2012 (received for review December 19, 2011)

Preferences are fundamental building blocks in all models of economic and political behavior. We study a new sample of comprehensively genotyped subjects with data on economic and political preferences and educational attainment. We use dense single nucleotide polymorphism (SNP) data to estimate the proportion of variation in these traits explained by common SNPs and to conduct genome-wide association study (GWAS) and prediction analyses. The pattern of results is consistent with findings for other complex traits. First, the estimated fraction of phenotypic variation that could, in principle, be explained by dense SNP arrays is around one-half of the narrow heritability estimated using twin and family samples. The molecular-genetic–based heritability estimates, therefore, partially corroborate evidence of significant heritability from behavior genetic studies. Second, our analyses suggest that these traits have a polygenic architecture, with the heritable variation explained by many genes with small effects. Our results suggest that most published genetic association studies with economic and political traits are dramatically underpowered, which implies a high false discovery rate. These results convey a cautionary message for whether, how, and how soon molecular genetic data can contribute to, and potentially transform, research in social science. We propose some constructive responses to the inferential challenges posed by the small explanatory power of individual SNPs.

genoeconomics | genopolitics | GCTA

There has been growing enthusiasm for the use of molecular genetic data in social science research. This enthusiasm is based on a number of potential contributions that such research could make to social science (1–3). For example, if specific genetic markers can be identified that are associated with a behavioral trait, then such predictive markers may shed light on the biological pathways underlying that trait (3, 4). If a set of genetic markers is sufficiently predictive, then these markers could be used in social science research as control variables, as instrumental variables (5, 6; for critical perspectives, see refs. 7, 8) or, under certain conditions, as factors for identifying at-risk individuals (1–3).

The extent to which this potential of molecular genetic data will be fulfilled for a given trait hinges on the trait's "molecular genetic architecture," i.e., the joint distribution of effect sizes and allele frequencies of the causal genetic variants (9). The architecture—which is the result of evolutionary forces, including mutation, drift, and selection—determines the difficulty with which the genetic variants associated with a trait can be identified and what sample sizes will be required for gene discovery. It also determines the out-of-sample aggregate predictability that can be derived from a set of genetic markers considered jointly.

Existing studies claiming to have established genetic associations with economic and political traits typically use samples of several hundred individuals, and no such study has used a sample larger than 3,000 individuals (for a recent review, see ref. 10). An implicit assumption underlying these studies is that there exist genetic variants whose effects are large enough that they can be reliably detected in samples of this size.

In this paper, we study the genetic architecture of economic and political preferences. For these traits, we ask whether the assumption of large effects of individual genetic variants is justified. We also explore the implications of the genetic architecture for efforts to realize the potential contributions of molecular genetic data in economic and political research.

We focus on preferences because they are fundamental building blocks in the models that economists and political scientists use to predict behavior. For example, measures of risk preferences predict diverse risky behaviors, such as smoking, drinking, and holding stocks rather than bonds (11, 12). Experimentally elicited patience predicts body mass index, smoking behavior, and exercise (13). Political preferences similarly predict a wide range of political behaviors, including voting (14) and monetary campaign contributions (15), as well as campaign activities like volunteering, attending rallies, and displaying yard signs (16). Behavior genetic studies, beginning with pioneering work on social and political attitudes (17, 18), have found that some of the variation in political and economic preferences can be statistically accounted for by genetic factors (19–24).

We use a new sample of comprehensively genotyped subjects from the Swedish Twin Registry. These subjects were recently administered, as part of a survey called Screening Across the Lifespan Twin survey, Younger cohort (SALTY), a rich set of questions measuring economic and political preferences. We study four fundamental economic preferences—risk aversion, patience, trust, and fair-mindedness—and five dimensions of political preferences, derived from a factor analysis of a comprehensive battery of attitudinal items. The five attitudinal dimensions are immigration/crime, economic policy, environmentalism, feminism/equality,

Author contributions: D.J.B., D. Cesarini, M.J.H.M.v.d.L., M.J., and P.M.V. designed research; D.J.B., D. Cesarini, M.J.H.M.v.d.L., C.T.D., P.D.K., P.K.E.M., C.F.C., D. Conley, D.L., and M.J. performed research; D.J.B., D. Cesarini, M.J.H.M.v.d.L., and C.T.D. analyzed data; and D.J.B., D. Cesarini, C.F.C., D.L., M.J., and P.M.V. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: daniel.benjamin@gmail.com.

[2]M.J. and P.M.V. contributed equally to this work.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1120666109/-/DCSupplemental.

and foreign policy. We also study educational attainment because, even though it is probably more distal from basic biological processes, it is available for a larger sample of genotyped individuals. For comparability with previous work and with our other estimates, we report twin-based estimates of heritability from this new sample, and we confirm moderate (30–40%) twin-based heritability estimates for these traits. However, our main focus is on using the dense single nucleotide polymorphism (SNP) data to learn about the genetic architecture of these traits.

As a first step, we use a new method (25–27) that uses the dense SNP data to estimate the proportion of variance in these traits that can be jointly explained by the genotyped SNPs. The technique—which we will call genomic-relatedness-matrix restricted maximum likelihood (GREML)—has been applied to height (25), intelligence (28, 29), personality traits (30), several common diseases (31), and schizophrenia (32), but never before to economic and political phenotypes. GREML provides a lower-bound estimate of narrow heritability that does not rest on the same set of assumptions relied on in twin studies. A key assumption behind GREML is that among individuals who are *not* in the same extended families, environmental factors are uncorrelated with differences in the degree of genetic similarity, or "relatedness." In this analysis, genetic relatedness is directly *estimated* from the SNP data, unlike in behavior genetic studies, where *expected* relatedness (inferred from the family pedigree) is used. (Another, distinct approach is to estimate the genetic variance from within-family variation in genetic relatedness; see ref. 33.) A common concern raised about behavior genetic studies of political and economic traits is that expected relatedness could be correlated with similarity in environmental factors that are not endogenous to genotype (as defined by Jencks; ref. 4). Because there is more random variation in the realized degree of genome sharing relative to the expected degree as the expected relatedness declines (34), environmental confounding is less likely to drive estimates that are based on realized relatedness among individuals whose expected relatedness is negligible.

Under the key assumption of no environmental confounding, an estimator for heritability can be obtained by examining how the correlation in phenotype between pairs of individuals relates to the realized genetic distance between those individuals. It would be an unbiased estimator of narrow heritability if genetic distance were calculated using all of the genetic variants that are causal for the phenotype. In practice, because the causal variants are not known, the SNPs typed on the genotyping chip are used to estimate genetic distance. Because these SNPs are only imperfectly correlated with the causal variants, relatedness with respect to the causal variants is measured with error. Consequently, the estimated relationship between phenotype and genetic relatedness is attenuated, and hence the estimator is a lower bound for narrow-sense heritability (25). Our GREML-based heritability estimates, although noisy, are on average about half the size of the twin-based heritability estimates. This gap may imply that genotyped SNPs tag about half of the genetic variation in these traits or that twin-based estimates of narrow heritability are biased upward (35, 36).

We next explore the molecular genetic architecture of the phenotypes. Specifically, we estimate heritability using relatedness measured separately by chromosome to test how evenly distributed the genetic effects are across the genome. We supplement these results by reporting findings from a standard genome-wide association study (GWAS) for each trait, in which individual SNPs are tested for association with the outcome of interest. Finally, we also perform a risk prediction exercise in which we randomly split the dataset into a discovery and a validation sample. We use a pruned set of SNPs from the discovery sample to build a predictor and then examine to what extent the predictor is correlated with the outcome in the validation sample. Similar approaches have been applied in the study of schizophrenia (37), height (38), and intelligence (28), but none of these methods have been applied to economic or political preferences. We find essentially no predictive power for the traits we study.

Our results paint a picture of economic and political preferences as highly polygenic traits for which individual SNPs explain only a small fraction of variance. The inferential challenges implied by this genetic architecture suggest—as we later discuss—that new approaches are needed to study genetic influences on preferences and behavior.

Results

As a preliminary, we computed the sibling correlations for all 10 variables (*SI Appendix*, Table S4). (Details on variable construction and materials and methods are available in *SI Appendix*.) The sibling correlations for the SALTY questions on patience (39), risk aversion*, and political preferences† have previously been analyzed and are reported to facilitate comparison with the GREML estimates. The implied heritabilities of the economic preferences are typically about 30% and the estimates for political preferences are typically around 40%.

We next estimated, for each trait, the proportion of phenotypic variation accounted for by measured SNPs, using the GREML estimator (25–27). These lower-bound heritability estimates for the nine traits are reported in Table 1. These analyses are all based on mixed-sex samples, controlling for sex, birth year, and the first 10 principal components of the genotypic data. (The fact that we observe some GREML point estimates of zero is not surprising. Because the estimator is constrained to produce a nonnegative estimate, the bound at zero will often be attained when the true population parameter is low and estimated imprecisely.)

For economic preferences, only one of the four variables, trust, is significant, with the point estimate suggesting that the common SNPs explain over 20% of phenotypic variation (P = 0.047). The remaining effects are lower, in one case zero, and not statistically distinguishable from zero. For political preferences, two of the five GREML estimates for the derived attitudinal dimensions are statistically significant, and one is marginally significant: 0.203 (P = 0.079) for immigration/crime, 0.344 (P = 0.012) for economic policy, and 0.354 (P = 0.009) for foreign policy. Keeping in mind that the GREML estimates are noisy and are lower bounds, taken as a whole they are consistent with low-to-moderate heritabilities for these traits. The cumulative effect of the SNPs is much more precisely estimated for educational attainment because this phenotype is available for all of the genotyped individuals in the sample. For this phenotype, our estimate is 0.158 (P = 0.004).

To help interpret our findings, we also report retest reliabilities for the preference measures using data from 491 respondents who answered the survey twice. We find that the political preferences were on average measured more reliably than the economic preferences. This finding is consistent with previously reported retest reliabilities for the risk questions* and the political preferences† on the SALTY survey. The reliabilities are given in the bottom two rows of Table 1. The GREML estimates are noisier (and the P values tend to be higher) for phenotypes with lower retest reliabilities.

We also conducted the GREML analyses separately by chromosome (as in refs. 40 and 28). Between conventionally unrelated individuals, realized relatedness is random and independent across chromosomes, and the expected relatedness measured from any chromosome is zero. If the genetic variation that predicts a trait were uniformly distributed across the genome, rather

*Beauchamp JP, Cesarini D, Johannesson M (2012) The Psychometric Properties of Measures of Economic Risk Preferences. Working Paper, Harvard University, New York University, and Stockholm School of Economics.

†Oskarsson, et al. (2010) Do Genes Mediate the Relationship between Personality and Ideology? Uppsala University Working Paper.

Table 1. GREML estimates of narrow heritability

	Economic preferences					Political preferences					
	Education	Risk	Patience	Fairness	Trust	Imm./crime	Econ. policy	Environ.	Femin./equal.	Foreign policy
V(g)/V(P)	0.158	0.137	0.085	0.000	0.242	0.203	0.344	0.000	0.000	0.354
SE	0.061	0.152	0.148	0.150	0.146	0.147	0.150	0.148	0.147	0.149
P value*	0.004	0.186	0.285	0.500	0.047	0.079	0.012	0.500	0.500	0.009
N	5,727	2,327	2,399	2,376	2,410	2,368	2,368	2,368	2,368	2,368
Chrom.	0.442	0.118	–0.195	–0.111	0.460	0.118	0.496	–0.311	0.247	0.462
P value†	0.039	0.601	0.385	0.623	0.031	0.601	0.019	0.159	0.268	0.030
ρ_{retest} ‡		0.71 (0.66–0.76)	0.40 (0.27–0.52)	0.57 (0.49–0.65)	0.63 (0.57–0.69)	0.86 (0.84–0.89)	0.85 (0.81–0.87)	0.62 (0.53–0.69)	0.78 (0.74–0.82)	0.70 (0.65–0.75)
N_{retest} ‡		475	483	469	482	471	471	471	471	471

GREML estimates for the 10 variables are reported. We estimated the matrix of genetic relatedness after omitting one twin per pair and then restricted the analyses to individuals whose relatedness did not exceed 0.025 in absolute value. Chrom. (chromosome) shows the estimated correlation between chromosomal length (measured in centimorgan) and the proportion of variation explained by relatedness estimated from that chromosome. Imm., Immigration; Econ., Economic; Environ., Environmentalism; Femin./equal., Feminism/equality.

*One-tailed likelihood ratio test of the hypothesis that the proportion of variation explained by common SNPs on the autosomes is zero.

†Test of the null hypothesis that, across the 22 autosomes, the correlation between chromosomal length and the proportion of variation explained by the chromosome is zero.

‡These rows show the estimated retest correlations (with 95% confidence intervals) and sample sizes for the retest correlations. For sample descriptions, see *SI Appendix*.

than being concentrated in a particular location, then greater realized relatedness from any given chromosome will predict greater phenotypic similarity, and this association will be stronger from longer chromosomes because longer chromosomes make up a larger fraction of the genome than shorter ones. Table 1 shows the estimated correlation between chromosomal length, measured in centimorgans, and the fraction of variance explained by the estimates of realized relatedness estimated using only data from one chromosome. The correlation is positive for 7 of 10 phenotypes, significantly so ($P < 0.05$) in four cases. Positive correlations have previously been reported for height (40), cognitive ability (28), and schizophrenia (32) and have been interpreted as evidence that the trait is highly polygenic with causal variants distributed across the genome.

Next, we tried to identify individual SNPs that predict economic and political preferences. For none of the 10 traits did we identify any SNPs that pass the conventional genome-wide significance threshold of $P < 5 \times 10^{-8}$ (41). In fact, no single SNP attains a P value lower than 10^{-7} for any of the 10 traits. The standard diagnostic for population stratification (i.e., ethnic confounding) in GWAS is inflated test statistics in the Q-Q plot (e.g., ref. 42); there is no evidence of inflated test statistics across the traits, with estimated λs (43) in the range of 0.987 (environmentalism) to 1.014 (educational attainment). While this diagnostic check suggests that our controls for population structure worked well, it is somewhat surprising that there is no systematic tendency for the λs to be larger than 1, given that some inflation is expected under a polygenic model even without any stratification (44). However, more inflation is expected in a larger sample, and λs have not been much larger in other analyses with comparable sample sizes to ours (28). In *SI Appendix*, we provide details on the full set of SNPs with P values $<10^{-5}$ for the nine preference measures, but we are skeptical that any of these associations will be replicable.

Finally, we examined the aggregate, out-of-sample predictive power of the SNPs. Following ref. 37, we first estimate the regression coefficient for each SNP in a discovery sample, composed of a randomly drawn 90% of the sample. From this set of coefficients, we form a prediction equation on the basis of a pruned set of 111,957 markers that includes only SNPs that are approximately in linkage equilibrium (to avoid double counting SNPs that are correlated with other SNPs). In a validation sample composed of the remaining 10%, we evaluate the correlation between individuals' predicted phenotype and observed phenotype. Although the correlation between the predicted and observed phenotype is positive, as expected, in 7 cases out of 10, we do not find quantitatively appreciable out-of-sample predictability for any of the traits: For most phenotypes, the explanatory power of the predictor is well below $R^2 = 0.1\%$. These results are reported in *SI Appendix*.

Discussion

The data reported here reveal a number of descriptive facts about the heritability and genetic architecture of political and economic preferences. First, we estimate sibling correlations for several traits, some of which have never before been studied in large samples, and we confirm that there is a robust separation of the monozygotic (identical twin) and dizygotic (fraternal twin) correlations. We obtain heritability estimates that are consistent with typical estimates previously reported for both political attitudes (19) and economic preferences (20, 23, 24), as well as educational attainment (45, 46). Overall, these results are consistent with the hypothesis that, for each of the 10 phenotypes, there exists a moderate correlation with genetic factors. None of the sibling correlations are adjusted for measurement error. A plausible conjecture is that the lower heritabilities of the economic preferences relative to the political preferences result from attenuation bias due to greater measurement error in economic preferences, as evidenced by the lower test–retest reliabilities of the economic preference measures.

Second, our molecular-genetic–based estimates of heritability partially corroborate the twin-based estimates and suggest that molecular genetic data *could* be predictive of preferences if the causal variants were known. When we estimate the cumulative effect of genotyped SNPs using GREML (25–27), we find that the estimated heritabilities are lower than the twin-based estimates, but the overall pattern of results suggests that point estimates are generally nonzero and, for the better measured variables, statistically distinguishable from zero. Because there is a lively debate regarding whether twin studies of political behavior have established that there is heritable variation in these traits (35), we note that our evidence for heritability is based on different assumptions than twin studies.

Previous papers on height (25), intelligence (28, 29), personality traits (30), and several diseases (31, 32) have found that the SNP-based heritability estimates are between one-quarter and one-half the size of twin-study estimates. One interpretation of the gap is that genotyped SNPs tag less than half the additive genetic variation in those traits (which would occur if causal variants are imperfectly correlated with SNPs on the SNP arrays, e.g., because their allele frequencies are lower than those of the SNPs). The gap may also reflect an upward bias in twin-based estimates of narrow heritability due to environmental confounding (35) or nonadditive variation (36).

Do economic and political preferences parallel other phenotypes in having SNP-based heritabilities that are half or less the magnitude of the twin-study estimates? If so, it would suggest that economic and political preferences have a similar genetic architecture, a similar degree of bias in twin-based estimates, or both. Because the economic and political preference measures have twin-based heritabilities around 0.30 (20) and 0.40 (19), respectively, the hypotheses of one-half magnitude would be GREML point estimates of around 0.15 and 0.20. Our evidence, considered in its entirety, is not inconsistent with these hypotheses, but the point estimates are quite noisy. An alternative approach is to examine the number of statistically significant associations. For economic preferences, if the SNP-based heritability parameter in the population is 0.15, and if sample estimates have a SE of 0.15 (as suggested by Table 1), then our power to statistically reject the null hypothesis of zero heritability in a one-sided test at the 5% level is about 26%. For political attitudes, if we assume a SNP-based heritability parameter in the population of 0.20, and we assume a SE of 0.15 (again as suggested by Table 1), then the corresponding statistical power is about 38%. If the traits are independently distributed, this calculation implies that for the nine preference variables, we should expect to observe 2.9 significant associations at the 5% level. In fact, we observe three significant associations at the 5% level and one more at the 8% level. The results, therefore, are close to what one would expect under the hypothesis that the SNP-based heritability estimates are about half the magnitude of the twin-based estimates.

Third, our analysis of individual SNPs does not reveal any associations that are significant at the conventional threshold of genome-wide significance required in genetic association studies. This is unsurprising in light of the accumulating evidence that the effects of common variants on complex outcomes are small (47), especially in the context of social science traits (1, 2). *SI Appendix*, Fig. S1 displays power calculations, given the SALTY sample size, for detecting true associations across a range of effect sizes as measured by the R^2. For the preference measures, the study was well powered to detect individual markers that explain at least 1.25% of trait variation at a nominal significance level of 10^{-7}. No single SNP in our sample attains this level of significance—the lowest P value we observe is 1.1×10^{-7}. Moreover, 1.25% is an upper bound to the effect sizes we can rule out because: first, because 1.1×10^{-7} is the smallest of many millions of P values we estimated, it almost surely capitalizes on chance to some extent and overstates the strongest genetic association in our data [the well-known "winner's curse" in statistical inference (48)]; and second, for many of the variables, the lowest observed P value was considerably higher than 10^{-7}. To illustrate our statistical power another way, if across the nine preference measures there are a total of 10 independently distributed SNPs each with R^2 of 0.75% or larger, our study had statistical power greater than 90% to detect at least one of them—and yet we found none. We conclude that it is unlikely that many common polymorphisms with such effect sizes exist. *SI Appendix*, Fig. S1 shows that the study was even better powered for educational attainment. Hence our failure to detect associations at these levels of significance indicates that true associations between common SNPs and economic and political phenotypes are likely to have very small effect sizes. [Whereas the survey measures we use here are common in economics (e.g., ref. 11), it is also common in economics to measure preferences using laboratory tasks that attach financial incentives to performance (49). As we explain in *SI Appendix*, our conclusion that the effect of individual SNPs on preferences are very small would hold even if measures of preferences were much more reliable than those we use here.] Of course, our evidence does not rule out the possibility that there exist rare variants with large effects on these phenotypes because sufficiently rare variants will have low correlations with the genotyped SNPs. Because such variants would be rare, however, a large sample would be required to detect them, as well.

Fourth, the results from our prediction exercise show that a standard polygenic risk score estimated in our sample has negligible out-of-sample predictability. This finding does not in any way contradict the results from the GREML analysis. GREML uses the measured SNPs to estimate realized relatedness between individuals, and given the large number of SNPs in a dense SNP array, realized relatedness can be estimated relatively precisely. In contrast, estimating a prediction equation that can predict well out of sample requires precise estimates of the effects of individual SNPs. In the limit of an infinite sample, it would be possible to perfectly estimate the effects of individual SNPs and thereby construct a polygenic risk score whose predictive power reaches the theoretical upper bound that is estimated by GREML. The smaller the discovery sample used to estimate the prediction equation, the noisier are the estimates of the individual SNP effects, and hence the lower will be the out-of-sample predictive power of the polygenic risk score that is constructed on the basis of these estimates. Evidently a discovery sample of 2,900 individuals (about 90% of 3,200) is far too small to obtain predictive power for standard measures of economic or political preferences.

These findings fit well with an emerging consensus in medical genetics that genetic variants that individually explain a substantial share of the variation in complex traits are unlikely to exist. If anything, the problem is likely to pose an even greater challenge in the social sciences because the phenotypes are usually several degrees removed from genes in the chain of biological causation (1–3). Our results suggest that much of the "missing heritability" (50)—the gulf between the cumulative explanatory power of common variants identified to date and the heritability estimated in behavior genetic studies—for social science traits reflects the fact that these traits have a complicated genetic architecture, with most causal variants explaining only a small fraction of the phenotypic variation. If so, then large samples will be needed to detect those variants.

Turkheimer (51) famously proposed three "laws of behavior genetics": first, all human behavioral traits are heritable; second, the proportion of variance attributable to common environment is smaller than the proportion attributable to genes; and third, a large portion of individual differences is explained by factors other than common environment and genes. We believe that there is accumulating evidence in favor of a fourth "law" regarding the molecular genetic architecture of behavioral traits: Genetic variants that are common in a population have very small individual

effects on behavioral traits. If true, this law would help explain the repeated failure to replicate initially promising candidate gene findings with large effect sizes (29, 52, 53), as well as the failure to date of genome-wide association studies to discover genetic variants associated with behavioral traits even in samples numbering tens of thousands of individuals (54). There is direct evidence for such an architecture for intelligence (28, 29), personality (30), and now economic and political preferences. Like Turkheimer's three laws, this fourth law is a summary of patterns of empirical results, not a theoretical necessity, so it could fail to hold in some specific cases, but we conjecture that it will generalize to most other behavioral phenotypes.

Our conclusions have a number of implications for research at the intersection of genetics and social science. There has recently been an explosion of reported associations in samples that are very small by the standards of medical genetics. Such studies are only adequately powered if the genetic marker's population R^2 for the trait is considerably larger than the upper bounds established by the GWAS findings reported here. Our findings, based on a sample an order of magnitude larger than most existing studies, suggest that adequate power actually requires a sample size that is yet another order of magnitude larger even than ours (1, 2, 55).

Published genetic associations with economic and political traits, even if statistically significant, should be approached with caution for two reasons. First, because most published studies are dramatically underpowered, the probability that an association study will detect a true signal is vanishingly small; hence, when a significant association is observed, Bayesian calculations indicate that the posterior odds that it is a true association are low (1, 2, 52). For example, Benjamin et al. (1) show that under reasonable assumptions about genetic effect sizes on economic traits, an observed association that is statistically significant at the 5% level in a sample of a few hundred individuals constitutes virtually no evidence in favor of a true effect—yet such reported associations are typical of most published genetic studies of social science traits. Second, publication bias—the tendency for findings, as opposed to nonfindings, to be selectively reported by researchers and selectively published by journals—are magnified in genetic association work because the typical dataset has many behavioral measures and many genetic markers (56). Testing for gene–environment interactions further compounds the problem of multiple hypothesis testing (57).

Our conclusions regarding the molecular genetic architecture of economic and political preferences also have implications for whether, how, and how soon molecular genetic information can contribute to, and potentially transform, research in social science. One possibility is that genetic associations may shed light on biological pathways through which precursors, such as preferences, lead to important behaviors and outcomes. More speculatively, such insights may also help inspire the development of new theoretical constructs that are more closely aligned with the underlying biology than the existing concepts such as "risk aversion" or "patience" that we study here (1, 3). Contributions such as these require the identification of specific genetic variants that correlate robustly with behavior. As discussed above, the results reported here suggest the need to analyze samples which are several orders of magnitude larger than those presently used in this sort of research if such markers are to be successfully identified. Unfortunately, even if these markers are eventually identified, our quantitative results suggest that many of them will only explain a tiny share of variance. Moreover, it is possible that such markers will be too far removed from the behavioral trait in the chain of causation to elucidate the biological pathway.

Another potential contribution to social science, already being actively pursued (e.g., 5, 6), is the use of genetic markers as instrumental variables in (nongenetic) empirical work. In order for the gene-as-instrument to be valid, not only must the marker be robustly associated with the "endogenous regressor," but *all* of the behaviors associated with that marker must be understood. Otherwise, if the marker has pleiotropic effects, then the exclusion restriction assumption could be violated, invalidating the instrumental variable application. As more is understood about genetic pathways, researchers will be in a better position to assess the plausibility of the exclusion restriction assumption on a case-by-case basis.

A different potential use of molecular genetic data in social science would be as control variables for genetic heterogeneity in (nongenetic) empirical work, to reduce the variance of the error term and shrink the SEs of coefficient estimates. For such an application to have practical utility, the markers that are selected as controls need to explain a nonnegligible share of the variation. Similarly, use of genetic data to target interventions requires that the aggregate predictive power of a set of genetic variants for the trait be sufficiently large. As we have shown here, given presently attainable sample sizes, the use of genetic data to predict economic and political traits does not appear to be feasible. It is likely that extremely large—perhaps impractically large—samples will be required. (Were it the case that economic behaviors, or their precursors in the form of various preferences, traits, and skills, could be predicted from molecular genetic information, it would raise a host of ethical questions about whether and how such information should be used. The potential benefits we have emphasized here must be carefully weighed against the costs, for example, discrimination based on genotype or the breakdown of insurance markets due to adverse selection.)

In summary, our molecular-genetic–based estimates of heritability partially corroborate the twin-based estimates and suggest that molecular genetic data could, in principle, be predictive of preferences. Our other results, however, suggest that excitement about the practical usefulness of molecular genetic data in social science research needs to be tempered by an appreciation that much of the heritable variation is likely explained by a large number of markers, each with a small effect in terms of variance explained. As a consequence, for economic and political preferences, much larger samples than currently used will be required to robustly identify individual SNP associations or to generate sizeable predictive power from many SNPs considered jointly.

Rather than being destructive to the enterprise of incorporating genetic data into social science, our conclusions regarding the molecular genetic architecture of economic and political preferences can help guide research in more productive directions. First, researchers could obtain very large samples that contain both genetic and social science data. Second, to minimize attenuation bias due to error in measurement and thereby maximize power for any given sample size, researchers could formulate more reliable measures of economic and political phenotypes. Third, researchers could focus on behavioral phenotypes that are more biologically proximate. One example is smoking, a behavior for which large, replicated associations have been found with SNPs in the nicotinic acetylcholine receptor gene *CHRNA3* (58). For such biologically proximate phenotypes, it is more likely that there exist genetic markers whose associations will have nontrivial effect sizes and clearer causal interpretations.

Materials and Methods

Beginning in December 2010, a total of 9,836 Swedish twins who passed initial laboratory-based quality controls were genotyped using the Illumina HumanOmniExpress BeadChip genotyping platform. We applied standard quality controls to the genetic data. In all our GWAS analyses, we controlled for sex, birth year, and the first 10 principal components of the genotypic data, and we adjusted the SEs for nonindependence within family. We computed the GREML estimates using the publicly available GCTA software (26). Before computing the matrix of genetic relatedness for the SALTY sample, we dropped one twin per pair, always the twin with a larger number of missing phenotypes. We used a relatedness threshold of 0.025. For our prediction exercise, we randomly split the sample into a 90% training sample to construct the genetic score and a 10% validation sample to examine its predictive accuracy. See *SI Appendix* for all details on the sample and methods.

ACKNOWLEDGMENTS. The authors thank the editor and three reviewers for constructive comments; Jonathan Beauchamp, Sandy Jencks, Sarah Medland, and Sven Oskarsson for helpful discussions; and Austin Bean, Cara Costich, Julia Goorin, Hazel Gunapala, Patrik Ivert, and Michael Puempel for research assistance. The study was supported by the GenomeEUtwin Project under the European Commission Programme Quality of Life and Management of the Living Resources of Fifth Framework Programme, the Swedish Research Council, the Swedish Council for Working Life and Social Research, the Swedish Foundation for Strategic Research, the Swedish Heart and Lung Foundation, the Ragnar Söderberg Foundation, the Jan Wallander and Tom Hedelius Foundation, and the National Institute on Aging/National Institutes of Health (NIH) Grants P01AG005842-20S2 and T32-AG000186-23. P.M.V. acknowledges funding from the Australian National and Medical Research Council, the Australian Research Council, and NIH Grant GM075091.

1. Benjamin DJ, et al. (2012) The promises and pitfalls of genoeconomics. *Annu Rev Econom*, in press.
2. Beauchamp JP, et al. (2011) Molecular genetics and economics. *J Econ Perspect* 25: 57–82.
3. Benjamin DJ, et al. (2007) Genoeconomics. *Biosocial Surveys*, eds Weinstein M, Vaupel JW, Wachter KW (The National Academies Press, Washington DC), pp 304–335.
4. Jencks C (1980) Heredity, environment, and public policy reconsidered. *Am Sociol Rev* 45:723–736.
5. Davey Smith G, Ebrahim S (2003) 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 32:1–22.
6. Ding W, Lehrer SF, Rosenquist JN, Audrain-McGovern J (2009) The Impact of Poor Health on Education: New Evidence Using Genetic Markers. *J Health Econ* 28:578–597.
7. Conley D (2009) The promise and challenges of incorporating genetic data into longitudinal social science surveys and research. *Biodemography Soc Biol* 55:238–251.
8. Cawley J, Han E, Norton EC (2011) The validity of genes related to neurotransmitters as instrumental variables. *Health Econ* 20:884–888.
9. Lander ES (2011) Initial impact of the sequencing of the human genome. *Nature* 470: 187–197.
10. Ebstein RP, Israel S, Chew SH, Zhong S, Knafo A (2010) Genetics of human social behavior. *Neuron* 65:831–844.
11. Barsky RB, Juster FT, Kimball MS, Shapiro MD (1997) Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *Q J Econ* 112:537–579.
12. Dohmen TF, et al. (2011) Individual risk attitudes: Measurement, determinants, and behavioral consequences. *J Eur Econ Assoc* 9:522–550.
13. Chabris CF, Laibson D, Morris CL, Schuldt JP, Taubinsky D (2008) Individual laboratory-measured discount rates predict field behavior. *J Risk Uncertain* 37:237–269.
14. Jost JT (2006) The end of the end of ideology. *Am Psychol* 61:651–670.
15. Morton R, Cameron C (1992) Elections and the theory of campaign contributions: A survey and critical analysis. *Econ Polit* 4:79–108.
16. Claassen RL (2007) Campaign activism and the spatial model: Getting beyond extremism to explain policy motivated participation. *Polit Behav* 29:369–390.
17. Eaves LJ, Eysenck HJ (1974) Genetics and the development of social attitudes. *Nature* 249:288–289.
18. Martin NG, et al. (1986) Transmission of social attitudes. *Proc Natl Acad Sci USA* 83: 4364–4368.
19. Alford JR, Funk CL, Hibbing JR (2005) Are political orientations genetically transmitted? *Am Polit Sci Rev* 99:153–167.
20. Cesarini D, Dawes CT, Johannesson M, Lichtenstein P, Wallace B (2009) Experimental game theory and behavior genetics. *Ann N Y Acad Sci* 1167:66–75.
21. Hatemi PK, et al. (2010) Not by twin alone: Using the extended family design to investigate genetic influences on political beliefs. *Am J Pol Sci* 54:798–814.
22. Hatemi PK, Dawes CT, Frost-Keller A, Settle JE, Verhulst B (2011) Integrating social science and genetics: News from the political front. *Biodemography Soc Biol* 57: 67–87.
23. Zhong S, et al. (2009) The heritability of attitude toward economic risk. *Twin Res Hum Genet* 12:103–107.
24. Cesarini D, Dawes CF, Johannesson M, Lichtenstein P, Wallace B (2009) Genetic variation in preferences for giving and risk-taking. *Q J Econ* 124:809–942.
25. Yang J, et al. (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 42:565–569.
26. Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet* 88:76–82.
27. Visscher PM, Yang J, Goddard ME (2010) A commentary on 'common SNPs explain a large proportion of the heritability for human height' by Yang et al. (2010). *Twin Res Hum Genet* 13:517–524.
28. Davies G, et al. (2011) Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol Psych* 16:996–1005.
29. Chabris CF, et al. (2012) Most reported genetic associations with general intelligence are probably false positives. *Psych Sci*, in press.
30. Vinkhuyzen AAE, et al. (2012) Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Transl Psychiatry,*10.1038/tp.2012.27.
31. Lee SH, Wray NR, Goddard ME, Visscher PM (2011) Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 88:294–305.
32. Lee SH, et al.; Schizophrenia Psychiatric Genome-Wide Association Study Consortium (PGC-SCZ); International Schizophrenia Consortium (ISC); Molecular Genetics of Schizophrenia Collaboration (MGS) (2012) Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nat Genet* 44:247–250.
33. Visscher PM, et al. (2006) Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genet* 2:e41.
34. Hill WG, Weir BS (2011) Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet Res (Camb)* 93:47–64.
35. Charney E (2008) Genes and ideologies. *Perspect Polit* 6:299–319.
36. Zuk O, Hechter E, Sunyaev SR, Lander ES (2012) The mystery of missing heritability: Genetic interactions create phantom heritability. *Proc Natl Acad Sci USA* 109: 1193–1198.
37. Purcell SM, et al.; International Schizophrenia Consortium (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460: 748–752.
38. Lango Allen H, et al. (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 467:832–838.
39. Cesarini D, Johannesson M, Magnusson PK, Wallace B (2012) The behavioral genetics of behavioral anomalies. *Manage Sci* 58:21–34.
40. Yang J, et al. (2011) Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet* 43:519–525.
41. McCarthy MI, et al. (2008) Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nat Rev Genet* 9:356–369.
42. Pearson TA, Manolio TA (2008) How to interpret a genome-wide association study. *JAMA* 299:1335–1344.
43. Devlin B, Roeder K (1999) Genomic control for association studies. *Biometrics* 55: 997–1004.
44. Yang J, et al.; GIANT Consortium (2011) Genomic inflation factors under polygenic inheritance. *Eur J Hum Genet* 19:807–812.
45. Miller P, Mulvey C, Martin N (2001) Genetic and environmental contributions to educational attainment in Australia. *Econ Educ Rev* 20:211–224.
46. Behrman JR, Taubman P (1989) Is schooling 'mostly in the genes'? Nature-nurture decomposition using data on relatives. *J Polit Econ* 97:1425–1446.
47. Visscher PM (2008) Sizing up human height variation. *Nat Genet* 40:489–490.
48. Garner C (2007) Upward bias in odds ratio estimates from genome-wide association studies. *Genet Epidemiol* 31:288–295.
49. Hertwig R, Ortmann A (2001) Experimental practices in economics: A methodological challenge for psychologists? *Behav Brain Sci* 24:383–403, discussion 403–451.
50. Manolio TA, et al. (2009) Finding the missing heritability of complex diseases. *Nature* 461:747–753.
51. Turkheimer E (2000) Three laws of behavior genetics and what they mean. *Curr Dir Psychol Sci* 9:160–164.
52. Ioannidis JPA (2007) Non-replication and inconsistency in the genome-wide association setting. *Hum Hered* 64:203–213.
53. van der Loos MJHM, et al. (2011) Candidate gene studies and the quest for the entrepreneurial gene. *Small Bus Econ* 37:269–275.
54. de Moor MH, et al. (2012) Meta-analysis of genome-wide association studies for personality. *Mol Psychiatry* 17:337–349.
55. Koellinger PD, et al. (2010) Genome-wide association studies in economics and entrepreneurship research: Promises and limitations. *Small Bus Econ* 35:1–18.
56. Hewitt JK (2011) Editorial policy on candidate gene association and candidate gene-by-environment interaction studies of complex traits. *Behav Genet* 42:1–2.
57. Duncan LE, Keller MC (2011) A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *Am J Psychiatry* 168:1041–1049.
58. Tobacco and Genetics Consortium (2010) Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 42:441–447.

Supporting Information:

"The Genetic Architecture of Economic and Political Preferences"

This document provides additional details about data and additional analyses to accompany the article "The Genetic Architecture of Economic and Political Preferences."

Contents

AI.A Extended Materials and Methods

Between December 2010 and May 2011, 9,836 Swedish Twins who passed initial, lab-based quality controls were genotyped by the SNP&SEQ Technology Platform, Uppsala, using the Illumina HumanOmniExpress BeadChip genotyping platform. A total of 79,893 SNPs were omitted because their minor allele frequency was lower than 0.01; 3,071 markers were excluded because they failed a test of Hardy-Weinberg equilibrium at $p < 10^{-7}$; and 3,922 SNPs were dropped because of a missingness greater than 3%. IMPUTE Version 2 (1) was used to impute all autosomal SNPs on HapMap, using the publicly-available phased haplotypes from HapMap2 (release 22, build 36, CEU population) as a reference panel. However, since all of our respondents were genotyped using the same genotyping platform, we drop imputed SNPs, and all of our analyses are based only on the directly genotyped markers that passed the quality controls.

As part of the quality control, one of the authors (van der Loos) conducted a careful analysis designed to detect pedigree errors. We estimated the proportion of genome-wide allele sharing among all pairs of individuals in the dataset by calculating, for a set of 547 randomly-selected markers with a minor allele frequency greater than 0.20, the mean and standard deviation of identity-by-state (IBS). For each pair of individuals in the data, we plotted the mean IBS against the standard deviation of the IBS across the markers. Visual inspection revealed a small number of DZ pairs (n = 38) with IBS equal to 2.0; we recoded these as MZ pairs. From the set of pairs coded as DZs, we also detected a small cluster of outliers (n = 5) whose mean IBS was appreciably lower than the rest of the pairs believed to be DZ twins; we randomly removed one individual per pair, treating the remaining observation as a singleton observation. Finally, we plotted the mean IBS against the standard deviation for all pairs of individuals in our data without any known relatedness. We identified one cluster of four parents and their six offspring and reclassified their relatedness accordingly. We identified all pairs of supposedly unrelated individuals whose mean IBS exceeded 1.46. We then eliminated one or several individuals from each such pair so that there were no remaining pairs without known relatedness whose mean IBS was above 1.46. This step resulted in 119 individuals being dropped. After these quality control corrections, the graph of mean IBS versus standard deviation IBS produced clusters that were consistent with the coded pedigree relationships. After quality control, the final sample is comprised of n = 9,617 individuals.

The principal components of the genotypic data were computed using EIGENSTRAT (2). We used the program *smartpca* of the EIGENSTRAT software to calculate the principal components of the genotypic data from a subsample of 6,813 unrelated individuals and to project the other individuals in the sample onto those principal components, thus obtaining the loadings of each individual on each of the top 10 principal components. We dropped individuals whose score is at least six standard deviations from the mean on one of the top ten principal components.

In all our GWAS analyses, we control for the first ten principal components of the genotypic data, sex, and birth year. The standard errors are adjusted for the non-independence of the residuals within twins using the *fastAssoc* function in the software Merlin Offline. The software assumes a behavior-genetic AE model for the distribution of the error terms within family. We did not use genomic control (3) to adjust the test statistics for inflation both because we do not find clear evidence of inflation and because any inflation we do observe could be the result of polygenic inheritance rather than population stratification (4).

For our prediction exercise, we randomly split the sample into a 90% training sample and a 10% validation sample, randomizing at the level of the family to ensure independence across

samples. We ran linkage-disequilibrium-based pruning of the validation sample using the software PLINK. Following (5), we used a linkage disequilibrium threshold of 0.2 with a 200-SNP window that slides by 25 SNP. The process started with 643,826 markers and left us with 111,957 markers which are approximately in linkage equilibrium. For each phenotype in turn, we then constructed the polygenic risk score for each individual in the validation sample. The polygenic risk score is defined as the sum of the estimated regression coefficients (from a regression of the phenotype on each SNP within the training sample) multiplied by the number of reference alleles and divided by the number of non-missing SNPs. If a genotype in the score is missing for a particular individual, then the expected value is imputed based on the sample allele frequency. Finally, within the validation sample, we regressed the phenotype on the polygenic risk score, controlling for the number of non-missing SNPs, with standard errors clustered by family.

We computed the GREML estimates using the publicly-available GCTA software (6). In all our GREML analyses, we control for the first ten principal components, sex, and year of birth. Before computing the matrix of genetic relatedness for the SALTY sample, we dropped one twin per pair, always the twin with a larger number of missing phenotypes. If the same number of observations were available for both twins, we selected one randomly. After restricting the genotypic data to the selected individuals, we performed some additional quality controls in order to follow the methodology of (7) and (8) as closely as possible: we dropped 6 SNPs because they failed a test of Hardy-Weinberg equilibrium at $p < 1\times10^{-6}$; excluded 629 SNPs with a minor allele frequency lower than 0.01; and removed 2 SNPs because of a missingness greater than 5%. 4 individuals were dropped because more than 1% of their genotypic data were missing. The pairwise genetic relationships were then estimated from 628,599 autosomal SNPs. We used the same matrix of genetic relatedness for all the SALTY phenotypes. The matrix of genetic relatedness for the larger sample with data on educational attainment was estimated in a similar way: before computing the matrix, we randomly dropped one twin per pair and performed additional quality control of the genotypic data. A total of 48 SNP were dropped because they failed a test of Hardy-Weinberg equilibrium at $p < 1\times10^{-6}$; 249 SNPs with a minor allele frequency lower than 0.01 were excluded; and 13 individuals were omitted because more than 1% of their genotypic data were missing. The matrix of genetic relatedness was then estimated from 628,922 autosomal SNPs.

AII.A. Risk Attitudes

We used three survey measures from SALTY to construct an overall index of risk attitudes. The first measure elicits the subjective level of general risk taking of the individual:

"How do you see yourself: Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Please tick a box on the scale, where the value 1 means 'unwilling to take risks' and the value 10 means 'fully prepared to take risks.'"

The second measure is the same as the first, except it is specific to financial risks:

"Are you a person who is fully prepared to take financial risks or do you try to avoid taking financial risks? Please tick a box on the scale, where the value 1 means 'unwilling to take risks' and the value 10 means 'fully prepared to take risks.'"

These two measures have been used in the German Socioeconomic Panel (9), except with a 0-10 scale rather than the 1-10 scale used in SALTY.

Our third measure of risk attitudes is adapted from the Health and Retirement Survey questions first described in (10). Subjects answer three questions about hypothetical gambles. The first question is phrased:

"Imagine the following hypothetical situation. You are the sole provider for your household, and you have the choice between two equally good jobs:

Job A will with certainty give you SEK 25,000 per month after taxes for the rest of your life.

Job B will give you a 50-50 chance of SEK 50,000 per month after taxes for the rest of your life, and a 50-50 chance of SEK 20,000 per month after taxes for the rest of your life.

Which job do you choose?"

The second and third questions are identical, except that instead of being SEK 20,000, the low-income outcome in job B is SEK 22,000 and SEK 17,000, respectively. For each subject, our third measure of risk-taking is the number of questions out of the three that the subject chose Job B (the risky job), generating a measure with a 0 (never choosing the risky job) to 3 (always choosing the risky job). (Only 2.6% of subjects behave inconsistently in the sense of choosing job B in a question but choosing job A in another question that offers more money for job B; we retain these inconsistent subjects in the sample.)

To combine our three measures of risk taking into an overall index, we calculate the z-score for each of the three measures, and the index is defined as the average z-score.

AII.B. Time Preference

To measure the rate of time preference, we used a sequence of three hypothetical choice questions from SALTY. Subjects were asked to choose between an amount of money today and a larger amount of money in one week. The first of these questions is phrased:

"Imagine that you can choose between receiving a sum of money today, or to wait and receive a larger sum in one week. Which would you choose?

SEK 5,000 today

SEK 6,000 in a week"

The second and third questions are identical, except that instead of being SEK 6,000*,* the delayed outcome is SEK 7,000 and SEK 5,500, respectively. For each subject, we calculate the number of questions out of the three that the subject chose the delayed outcome, generating a variable with a 0 (never choosing the delayed reward) to 3 (always choosing the delayed reward) scale. (Only 1.1% of subjects behave inconsistently in the sense of choosing job B in a question but choosing job A in another question that offers more money for job B; we retain these inconsistent subjects in the sample.) Our measure of patience is the z-score of this variable.

Note that the questions ask about a short time horizon (one week) to get a proxy for short-term time discounting, which appears to be distinct from long-term discounting (11) and responsible for many impatient behaviors (12). Similar questions (but with smaller real stakes) have been used by other authors (e.g., 13); see (14) for a discussion of this and other approaches to measuring time preferences.

AII.C. Trust

To measure trust, we use two questions included in SALTY. Both are taken from the National Opinion Research Center's General Social Survey (GSS), except that the SALTY questions use a 10-point scale instead of a binary scale. The first question is a classic measure of trust used widely in political science (15) and other social sciences:

"Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people? Please tick on the scale below, where the value 1 means 'need to be very careful' and the value 10 means 'most people can be trusted.'"

We also included a second question from the GSS:

"Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair? Please tick on the scale below, where the value 0 means 'would take advantage of me' and the value 10 means 'would treat me fairly.'"

For each subject, we calculate the z-score for each response to the two questions, and we use the average z-score as our index of trust. The above two trust questions have been previously used by (16); see (16) for a discussion and empirical comparison of different measures of trust.

AII.D. Fairness

To measure attitudes about fairness we used three survey questions in SALTY, which are adapted from (17) (their question 1, question 4A, and question 11B):

"A hardware store has been selling snow shovels for SEK 150. The morning after a large snowstorm, the store raises the price to SEK 200. How fair do you think that is?

Completely fair

Acceptable

Unfair

Very unfair"

"A company is making a small profit. However, due to a recession, unemployment is high, and it is easy to hire people. The company therefore decides to decrease wages and salaries by 10% for all its employees. How fair do you think that is?

Completely fair

Acceptable

Unfair

Very unfair"

"A small factory is making kitchen tables. Because of changes in the price of materials, the cost of making each table has decreased by SEK 200. But the factory does not lower its price for the tables. How fair do you think that is?

Completely fair

Acceptable

Unfair

Very unfair"

We coded the responses from 1 (completely fair) to 4 (very unfair) for each question, and we summed the answers from the three questions. We then use the z-score of this variable as our index of fairness attitudes.

AII.E. Political Attitudes

We measure political attitudes using a battery of 34 questions included in SALTY that elicit attitudes toward various policy issues. These items have also been included in other Swedish surveys, such as the Swedish Election Studies. On each question the respondents rate their attitude toward a policy on a 5-point scale from (1) “very good proposal” to (5) “very bad proposal.” The battery of questions is reproduced below.

Table S1: Political Attitudes Battery.

Below is a table with proposals some people think should be implemented in Sweden. State what you think about each of these proposals.

	Very good proposal	Fairly good proposal	Neither good nor bad proposal	Fairly bad proposal	Very bad proposal
1. Decrease the public sector	☐	☐	☐	☐	☐
2. Decrease defense expenses	☐	☐	☐	☐	☐
3. Decrease welfare benefits	☐	☐	☐	☐	☐
4. Decrease taxes	☐	☐	☐	☐	☐
5. Keep property taxes	☐	☐	☐	☐	☐
6. Sell public companies to private buyers	☐	☐	☐	☐	☐
7. Decrease income inequality in society	☐	☐	☐	☐	☐
8. Have more private companies in health care	☐	☐	☐	☐	☐
9. Decrease the influence of financial markets on politics	☐	☐	☐	☐	☐
10. Keep the “maxtaxa” in daycare	☐	☐	☐	☐	☐
11. Have more private schools	☐	☐	☐	☐	☐
12. Introduce grades earlier in school	☐	☐	☐	☐	☐
13. Increase the economic support to rural areas	☐	☐	☐	☐	☐
14. Introduce 6-hour working day for all employees	☐	☐	☐	☐	☐
15. Forbid all kinds of pornography	☐	☐	☐	☐	☐

16.	Limit the right to free abortion	☐	☐	☐	☐	☐
17.	Introduce much harder punishments for criminals	☐	☐	☐	☐	☐
18.	Strengthen animal rights	☐	☐	☐	☐	☐
19.	Sweden should in the long run carry through a nuclear phase-out	☐	☐	☐	☐	☐
20.	Stop motoring in the inner city	☐	☐	☐	☐	☐
21.	Invest more to prevent environmental damages	☐	☐	☐	☐	☐
22.	Decrease carbon dioxide emissions	☐	☐	☐	☐	☐
23.	Increase labor immigration to Sweden	☐	☐	☐	☐	☐
24.	Introduce a language test to become a Swedish citizen	☐	☐	☐	☐	☐
25.	Decrease foreign aid	☐	☐	☐	☐	☐
26.	Accept fewer refugees in Sweden	☐	☐	☐	☐	☐
27.	Increase the economic support to immigrants so that they can preserve their own culture	☐	☐	☐	☐	☐
28.	Remit debt to developing countries	☐	☐	☐	☐	☐
29.	Give private companies more freedom	☐	☐	☐	☐	☐
30.	Sweden should leave the EU	☐	☐	☐	☐	☐
31.	Sweden should introduce the EURO	☐	☐	☐	☐	☐
32.	Sweden should become members in NATO	☐	☐	☐	☐	☐
33.	Sweden should work for increased free trade all over	☐	☐	☐	☐	☐

	the world					
34.	Sweden should actively support the US war on terrorism	☐	☐	☐	☐	☐

83.4% of subjects responded to all 34 questions, and 97% responded to at least 29. We dropped the 3% of subjects who answered fewer than 29 questions, and for the subjects we retained, we replaced missing values by the question-specific sample mean. We factor-analyzed the items, retaining five factors, and performed a varimax rotation of the data (18). An inspection of the factor loadings shows that the first factor can be crudely interpreted as an "immigration and crime" factor, the second as an "economic policy factor," the third as an "environmentalism factor," the fourth as a "foreign policy factor," and the last as a "feminism and equality" factor. Although the labels of each factor do not describe all the items that load heavily on the factor, the factor structure and interpretation are similar to what has previously been reported for this scale (19-20).

Table S2: Factor Structure of Political Attitudes Battery.

	Questions With Highest Loading	*Interpretation*
Factor 1	23,24,25,26,27	Immigration and Crime
Factor 2	1,4,6,8,11	Economic Policy
Factor 3	18,19,20,21,22	Environmentalism
Factor 4	19,30,31,32,33	Foreign Policy
Factor 5	7,13,14,15,18	Feminism and Equality

Note: For each of the factors, this table show the five rotated factor loadings that are highest in absolute value.

AII.F. Educational Attainment

Our measure of educational attainment is obtained from subjects' responses on the SALT survey's questions on educational attainment. Subjects were asked "What is the highest degree that you have completed (or are in the process of completing)?" and were given twelve different choices, ranging from minimum compulsory schooling all the way up to 3 years of college or more. The responses to these questions can thus be used to define a categorical variable which takes one of 12 distinct values. To convert this variable into a continuous measure, we used a separate dataset constructed by Sandewall, Cesarini and Johannesson (21). This dataset contains both the SALT survey responses and Statistics Sweden's categorical variable measuring educational attainment according to the 1990 census. The census data are also categorical but can be used to impute years of schooling by assigning to each category the population average of the number of years of schooling, as estimated by Isacsson (22). For each of the 12 categorical responses in SALT, we thereby impute the average years of schooling of individuals in that category.

AIII.A. Sample Definitions

The first sample we use is the *SALTY sample*. SALTY is a survey administered by the Swedish Twin Registry between 2008 and 2010 to twins born between 1943 and 1958. The response rate from contacted individuals was 47.1%, leaving a sample of 11,743 subjects. Of these, 11,418 (97.2%) gave informed consent to have their answers stored and analyzed. We work with a sample comprised of 1,143 complete MZ pairs (464 of them male); 1,237 complete, same-sex DZ pairs (502 of them female); 1,114 complete, opposite-sex DZ pairs; and 4,394 singletons.

In addition, 491 respondents answered the survey twice (out of 800 contacted a second time). See (23) for an analysis of non-response both to the original survey and to the re-test survey. We refer to the *SALTY* re-test respondents as *SALTY Retest*.

The second sample is *TwinGene*, a sample of 10,946 twins born between 1926 and 1958 who agreed to provide biological specimens for genotyping. The *TwinGene* and *SALTY* samples partly overlap; 4,040 *SALTY* respondents are also in *TwinGene*. We refer to these genotyped *SALTY* twins as *SALTY-Geno.* The *SALTY* respondents who are not in *SALTY-Geno* have not been genotyped.

The economic and political preference data are available from the survey administered to the *SALTY* sample (and administered twice to the *SALTY Retest* sample). These data are *not* available for the *TwinGene* sample, except for those in *SALTY-Geno*.

Data on educational attainment is available for nearly all members of the *TwinGene* sample, as this variable is drawn from the SALT survey, which was administered by the Swedish Twin Registry beginning in 1998. Since the *TwinGene* sample is (without exception) a subset of the SALT respondents, data on educational attainment is available for almost the entire *TwinGene* sample.

AIII.B. Descriptive Statistics

Table S3: Summary Statistics

	Education	Economic Preferences				Political Preferences				
	Education	Risk	Time	Fairness	Trust	Imm./ Crime	Econ. Policy	Environ.	Femin./ Equality	Foreign Policy
SALTY	-	0.00	-0.02	-0.00	-0.00	-0.00	0.00	-0.01	-0.02	-0.01
	(-)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	(1.00)
N	-	10,788	11,187	11,045	11,237	11,000	11,000	11,000	11,000	11,000
SALTY-Geno	-	0.06	0.02	0.01	0.02	0.02	0.02	-0.04	-0.02	-0.00
	(-)	(0.98)	(0.99)	(0.98)	(0.99)	(1.00)	(0.98)	(0.99)	(0.99)	(0.97)
N	-	3,818	3,970	3,922	3,986	3,886	3,886	3,886	3,886	3,886
SALTY-Retest	-	-0.00	-0.00	0.00	0.03	-0.10	-0.03	0.06	0.11	0.01
	(-)	(0.98)	(0.94)	(0.98)	(0.94)	(0.98)	(1.00)	(0.89)	(0.97)	(0.93)
N	-	487	486	486	487	480	480	480	480	480
TwinGene	11.17	-	-	-	-	-	-	-	-	-
	(2.42)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)
N	10,877	-	-	-	-	-	-	-	-	-

Note: This table reports means, standard deviations (in parentheses), and sample sizes for each of three samples: the entire SALTY sample, the subset of the SALTY sample which has been genotyped, and the entire TwinGene sample. The TwinGene sample only party overlaps with the SALTY sample, which is why the number of genotyped subjects with information on educational attainment is larger. The economic and political variables were standardized to have mean zero and variance one using all of SALTY as a standardization sample.

Table S4: Sibling Correlations

	Education	Economic Preferences				Political Preferences				
	Education	Risk	Patience	Fairness	Trust	Imm./ Crime	Econ. Policy	Environ.	Femin./ Equality	Foreign Policy
ρ_{MZM}	0.64	0.41	0.05	0.32	0.37	0.56	0.42	0.26	0.48	0.48
	(.58-.69)	(.33-.49)	(-.04-.15)	(.23-.42)	(.29-.45)	(.49 -.62)	(.34-.50)	(.17-.34)	(.40-.57)	(.40-.54)
ρ_{DZM}	0.45	0.20	0.07	0.24	0.18	0.33	0.34	0.09	0.27	0.16
	(.39-.52)	(.12-.28)	(-.04-.20)	(.15-.33)	(.08-.27)	(.25-.42)	(.26-.42)	(.01-0.18)	(.19-.35)	(.07-.24)
ρ_{MZF}	0.65	0.35	0.16	0.24	0.33	0.60	0.45	0.34	0.41	0.47
	(.59-.69)	(.27-.43)	(.07-.25)	(.16-.32)	(.25-.40)	(.55-.65)	(.37-.51)	(.23-.46)	(.34-.48)	(.39-.53)
ρ_{DZF}	0.49	0.13	0.11	0.10	0.19	0.37	0.22	0.16	0.19	0.21
	(.44-.55)	(.05-.21)	(.03-.20)	(.03-.17)	(.11-.26)	(.29-.44)	(.14-.29)	(.09-.24)	(.12-.26)	(.14-.29)
N_{MZM}	566	443	491	450	458	448	448	448	448	448
N_{DZM}	621	477	655	482	486	487	487	487	487	487
N_{MZF}	553	594	700	652	659	630	630	630	630	630
N_{DZF}	849	636	1108	681	714	665	665	665	665	665

Note: This table gives the sibling correlations for the ten phenotypes and re-test correlations for the nine preference measures. The 95% confidence intervals given in parentheses are computed by bootstrapping with 500 draws. MZM: number of male monozygotic pairs; DZM: male dizygotic pairs; MZF: female monozygotic pairs; DZF: female dizygotic pairs. The economic and political preference data are from the SALTY sample, whereas the education data are from the TwinGene sample.

AIV.A. GREML Analyses Without Threshold for Genetic Relatedness

Table S5: GREML Analyses, No Threshold

	Education	Economic Preferences				Political Preferences				
	Education	Risk	Time	Fairness	Trust	Imm./ Crime	Econ. Policy	Environ.	Femin./ Equality	Foreign Policy
V(g)/V(P)	0.163	0.157	0.040	0.000	0.258	0.207	0.294	0.000	0.000	0.344
s.e.	0.052	0.139	0.134	0.136	0.134	0.134	0.139	0.135	0.132	0.138
p-value	<0.001	0.130	0.385	0.500	0.026	0.058	0.019	0.500	0.497	0.006
N	6,754	2,519	2,604	2,579	2,613	2,567	2,567	2,567	2,567	2,567
Chrom.	0.429	0.301	-0.170	-0.044	0.391	0.310	0.523	-0.110	0.186	0.554
p-value	0.047	0.173	0.449	0.846	0.072	0.160	0.013	0.626	0.407	0.007

Note: This table reports GREML estimates for the ten variables. We estimated the matrix of genetic relatedness after omitting one twin per pair but did not impose a relatedness threshold before estimating the model. The row Chrom. shows the estimated correlation between chromosomal length (measured in centimorgan) and the proportion of variation explained by relatedness estimated from that chromosome. The third row gives the *p*-value for the one-tailed likelihood ratio test of the hypothesis that the proportion of variation explained by common SNPs on the autosomes is zero. The sixth row gives the *p*-value for the test of the null hypothesis that, across the 22 autosomes, the correlation between chromosomal length and the proportion of variation explained by the chromosome is zero. The bottom two rows show the estimated re-test correlations.

AIV.B. Individual SNP Analyses

Table S6: Top Hits from GWA of the Preference Variables

Chr.	SNP	Effect allele	Gene	Function	Beta	SE	*p*-value	Phenotype
2	rs10204325	T			0.134	0.025	1.18×10^{-7}	Fairness
11	rs11233413	T	*RAB30*	intronic	-0.108	0.021	3.53×10^{-7}	Feminism / equality
3	rs10937540	T			-0.091	0.018	4.28×10^{-7}	Environmentalism
3	rs6775909	T			0.090	0.018	5.51×10^{-7}	Environmentalism
18	rs41418949	T	*NETO1*	intronic	-0.358	0.072	5.86×10^{-7}	Time
3	rs9821642	G			-0.092	0.019	7.13×10^{-7}	Environmentalism
2	rs436000	T			0.125	0.025	8.03×10^{-7}	Fairness
3	rs9820695	G			0.091	0.019	8.54×10^{-7}	Environmentalism
8	rs2299587	T	*PCM1*	intronic	-0.125	0.026	9.27×10^{-7}	Trust
3	rs4856162	G			0.091	0.019	1.03×10^{-6}	Environmentalism
3	rs1397924	T			-0.092	0.019	1.03×10^{-6}	Environmentalism
8	rs10112514	T	*PCM1*	intronic	-0.124	0.025	1.08×10^{-6}	Fairness
10	rs2250149	T			0.114	0.023	1.23×10^{-6}	Feminism / equality
3	rs7628767	G			0.087	0.018	1.24×10^{-6}	Environmentalism
3	rs4493441	G			0.087	0.018	1.26×10^{-6}	Environmentalism
18	rs8083633	T	*DLGAP1*	intronic	0.154	0.032	1.39×10^{-6}	Time
14	rs4902960	G	*RGS6*	intronic	0.193	0.040	1.40×10^{-6}	Environmentalism
6	rs210648	G	*DCBLD1*	intronic	-0.122	0.025	1.59×10^{-6}	Risk
10	rs2250245	G			-0.112	0.023	1.62×10^{-6}	Feminism / equality
6	rs240768	T	*ASCC3*	coding	0.267	0.056	1.77×10^{-6}	Immigration / crime
1	rs574773	T			0.197	0.041	1.78×10^{-6}	Trust
3	rs7612581	T			0.088	0.018	1.89×10^{-6}	Environmentalism
1	rs4384209	G	*PDE4B*	intronic	0.126	0.027	1.98×10^{-6}	Feminism / equality
18	rs1346987	C			0.180	0.038	1.99×10^{-6}	Immigration / crime
3	rs10511217	T			-0.092	0.019	1.99×10^{-6}	Environmentalism

3	rs12485744	T			0.091	0.019	2.14×10^{-6}	Environmentalism
17	rs2291447	T	*ACBD4*	5’ upstream	0.114	0.024	2.22×10^{-6}	Risk
4	rs12499086	G			-0.129	0.027	2.26×10^{-6}	Time
4	rs11730243	T			0.129	0.027	2.30×10^{-6}	Time
6	rs6931919	C	*ASCC3*	intronic	-0.263	0.056	2.34×10^{-6}	Immigration / crime
6	rs3213542	T	*ASCC3*	coding	-0.263	0.056	2.39×10^{-6}	Immigration / crime
11	rs4944425	C	*RAB30*	intronic	-0.099	0.021	2.59×10^{-6}	Feminism / equality
11	rs7101446	T	*SLC22A9*	intronic	-0.145	0.031	2.71×10^{-6}	Economic policy
12	rs2120771	G			-0.107	0.023	2.72×10^{-6}	Feminism / equality
22	rs4823246	G			-0.100	0.021	2.82×10^{-6}	Feminism / equality
13	rs7984869	T			0.115	0.025	2.90×10^{-6}	Fairness
14	rs12434047	G			-0.127	0.027	2.94×10^{-6}	Fairness
11	rs17504704	G	*RAB30*	intronic	-0.102	0.022	3.30×10^{-6}	Feminism / equality
3	rs13073838	T			-0.116	0.025	3.30×10^{-6}	Economic policy
18	rs8090196	T			-0.179	0.038	3.35×10^{-6}	Immigration / crime
17	rs2360111	T	*NXN*	intronic	-0.113	0.024	3.43×10^{-6}	Economic policy
4	rs1850744	T			0.305	0.066	3.56×10^{-6}	Risk
18	rs7235528	T			0.167	0.036	3.74×10^{-6}	Immigration / crime
6	rs11969893	G			0.324	0.070	4.06×10^{-6}	Immigration / crime
2	rs3789119	T	*ACOXL*	intronic	-0.136	0.029	4.17×10^{-6}	Immigration / crime
1	rs12125250	C			-0.124	0.027	4.18×10^{-6}	Foreign Policy
3	rs1682825	T			-0.118	0.026	4.28×10^{-6}	Feminism / equality
18	rs12606301	G			0.171	0.037	4.54×10^{-6}	Immigration / crime
12	rs10772939	T			-0.112	0.024	4.56×10^{-6}	Economic policy
21	rs16998084	G			-0.400	0.088	4.82×10^{-6}	Immigration / crime
10	rs4586057	T	*RPP30*	intronic	0.141	0.031	5.08×10^{-6}	Time
12	rs10748180	G	*THAP2*	3’ utr	-0.123	0.027	5.14×10^{-6}	Time
17	rs218676	G	*SLC13A5*	intronic	0.149	0.033	5.19×10^{-6}	Time
11	rs470763	G			0.113	0.025	5.24×10^{-6}	Fairness
5	rs1978633	T	*FBXL7*	intronic	-0.090	0.020	5.40×10^{-6}	Feminism / equality
18	rs7231412	T			0.170	0.037	5.67×10^{-6}	Immigration / crime

13	rs7327064	T	*SOHLH2*	intronic	0.140	0.031	5.69×10^{-6}	Foreign Policy
10	rs2904804	T	*AKR1C1*	intronic	0.117	0.026	5.69×10^{-6}	Immigration / crime
3	rs10937544	C			-0.088	0.019	5.74×10^{-6}	Environmentalism
3	rs1488193	G	*CD200R1*	intronic	-0.230	0.051	5.80×10^{-6}	Economic policy
3	rs13325751	T	*CADPS*	intronic	0.132	0.029	6.17×10^{-6}	Environmentalism
14	rs1951681	G	*AKAP6*	intronic	0.194	0.043	6.17×10^{-6}	Environmentalism
6	rs9267663	T	*EHMT2*	5’ upstream	-0.290	0.064	6.20×10^{-6}	Environmentalism
2	rs12713280	T			0.123	0.027	6.25×10^{-6}	Risk
5	rs17376026	T			-0.121	0.027	6.28×10^{-6}	Immigration / crime
9	rs2226006	T	*ASTN2*	intronic	-0.111	0.025	6.45×10^{-6}	Fairness
1	rs438895	G			0.384	0.085	6.67×10^{-6}	Education
14	rs10146615	T			-0.138	0.031	6.71×10^{-6}	Immigration / crime
3	rs16854884	C			0.103	0.023	6.74×10^{-6}	Feminism / equality
10	rs1570854	T			0.079	0.018	7.03×10^{-6}	Environmentalism
7	rs1399090	T	*PLXNA4*	intronic	-0.216	0.048	7.06×10^{-6}	Foreign Policy
7	rs2598202	T	*PLXNA4*	intronic	-0.216	0.048	7.06×10^{-6}	Foreign Policy
22	rs117294	C			0.110	0.025	7.08×10^{-6}	Fairness
22	rs138597	G	*LOC388910*	3’ downstream	0.123	0.027	7.30×10^{-6}	Time
14	rs3784178	T	*AKAP6*	intronic	-0.192	0.043	7.37×10^{-6}	Environmentalism
1	rs10754644	G			0.090	0.020	7.40×10^{-6}	Feminism / equality
3	rs13068298	T			-0.102	0.023	7.71×10^{-6}	Feminism / equality
5	rs784420	G			-0.120	0.027	7.85×10^{-6}	Trust
6	rs9364813	G			0.265	0.059	8.17×10^{-6}	Immigration / crime
3	rs10511400	T	*LRRC58*	3’ utr	0.157	0.035	8.22×10^{-6}	Feminism / equality
12	rs11178918	G			-0.184	0.041	8.32×10^{-6}	Time
2	rs12619788	G			0.112	0.025	8.33×10^{-6}	Immigration / crime
3	rs16831033	T	*LRRC58*	3’ downstream	-0.157	0.035	8.88×10^{-6}	Feminism / equality
1	rs9728717	T	*EDG7*	intronic	0.140	0.031	8.95×10^{-6}	Time
9	rs4838320	T			0.197	0.044	9.03×10^{-6}	Immigration / crime
17	rs7209847	G	*RICH2*	intronic	-0.273	0.062	9.39×10^{-6}	Risk
10	rs4282910	G	*RPP30*	5’ upstream	0.148	0.033	9.45×10^{-6}	Time

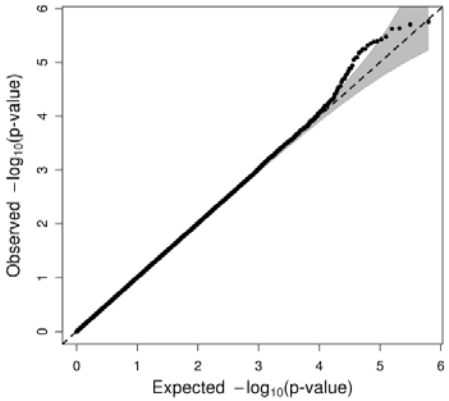
6	rs228146	C	*BMP5*	intronic	0.109	0.025	9.95×10^{-6}	Time

Note: This table reports the full set of SNPs with $p < 10^{-5}$ in the GWA analyses for political and economic preferences. Education results are not reported because these data are part of ongoing consortium meta-analyses, but none of the *p*-values are lower than 10^{-7}. For 5' upstream and 3' downstream, the SNP is located within 2 kb from the UTR start or end.

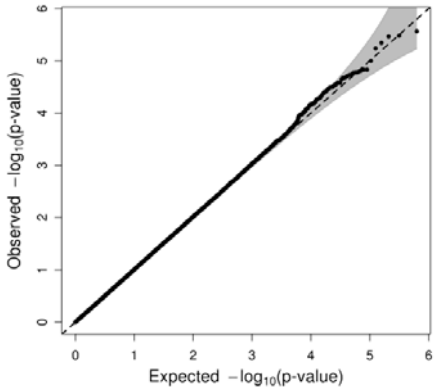
AIV.C. Q-Q Plots for the Political Preferences

Figure S1. QQ Plots for Political Preferences

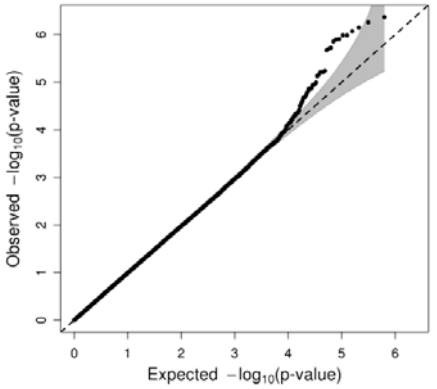
Immigration and Crime



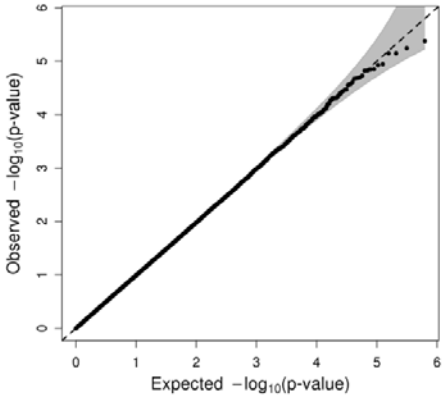
Economic policy



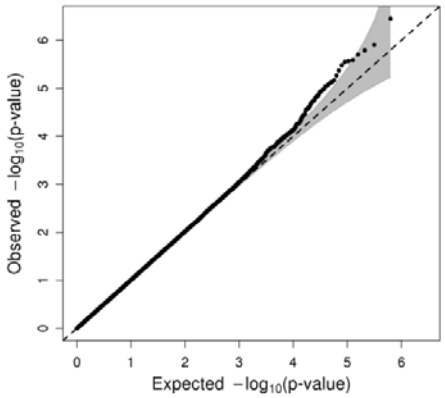
Environmentalism



Foreign policy



Feminism and Equality

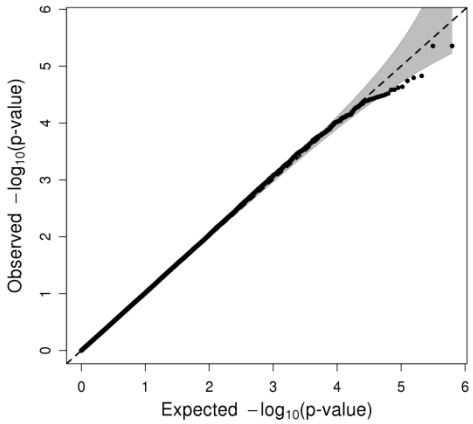


Note: These figures are quantile–quantile plots of the association analysis *p*-values from the GWA of the five political phenotypes. Genomic control was not applied.

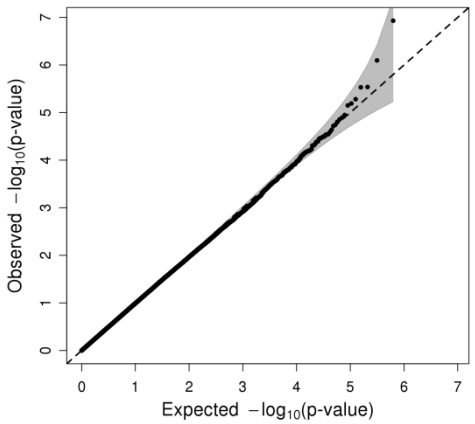
AIV.D. Q-Q Plots for the Economic Preferences and Education

Figure S2. QQ Plots for Economic Preferences and Education

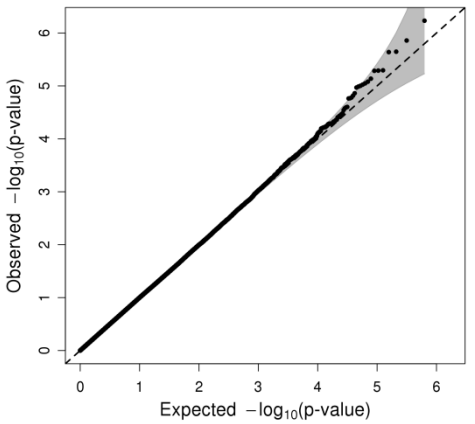
Education



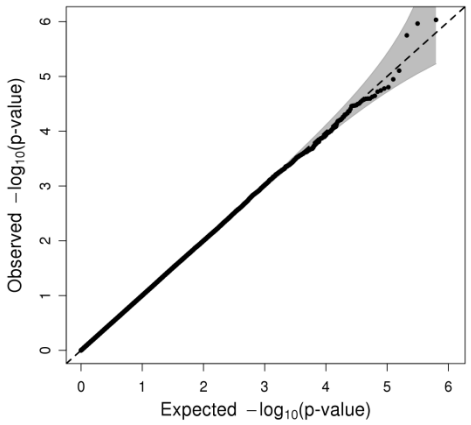
Fairness



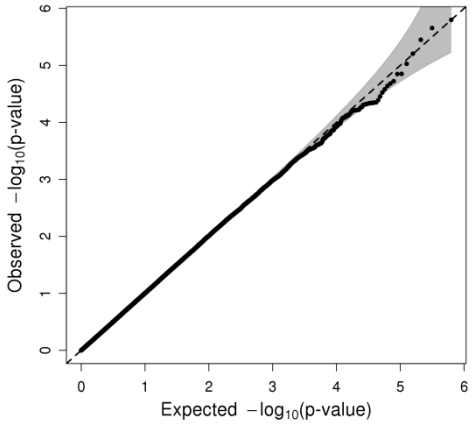
Patience



Trust



Risk

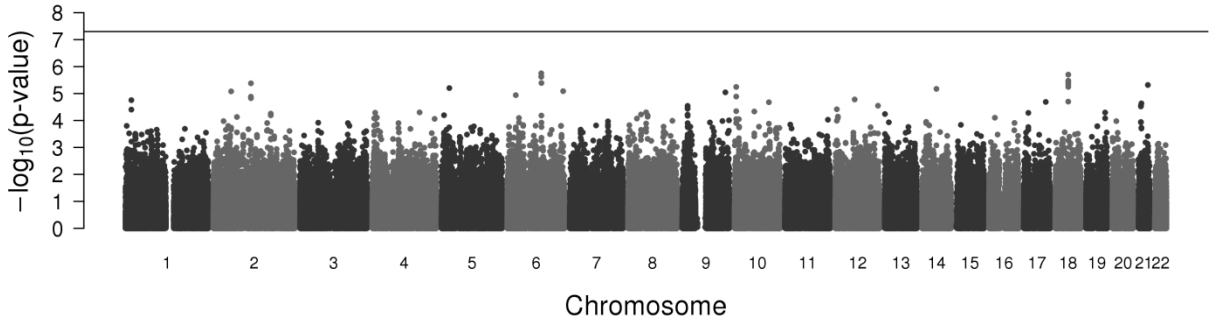


Note: These figures are quantile–quantile plots of the association analysis *p*-values from the GWA of the economic phenotypes. Genomic control was not applied.

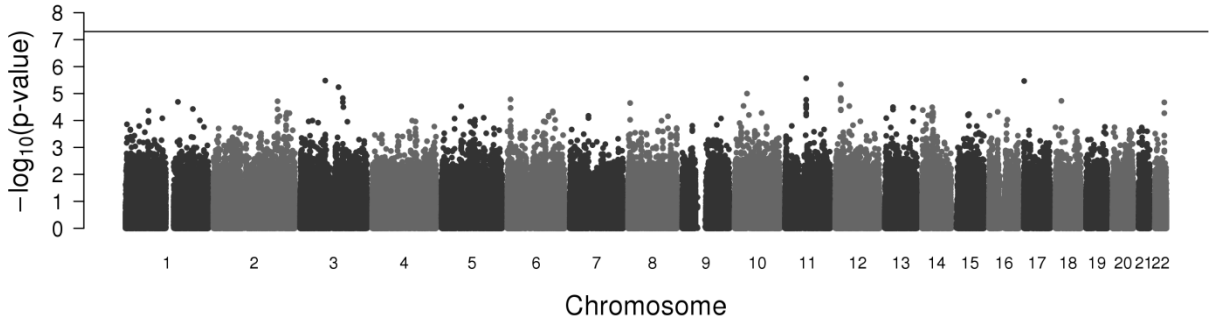
AIV.E. Manhattan Plots for the Political Preferences

Figure S3. Manhattan Plots for Political Preferences

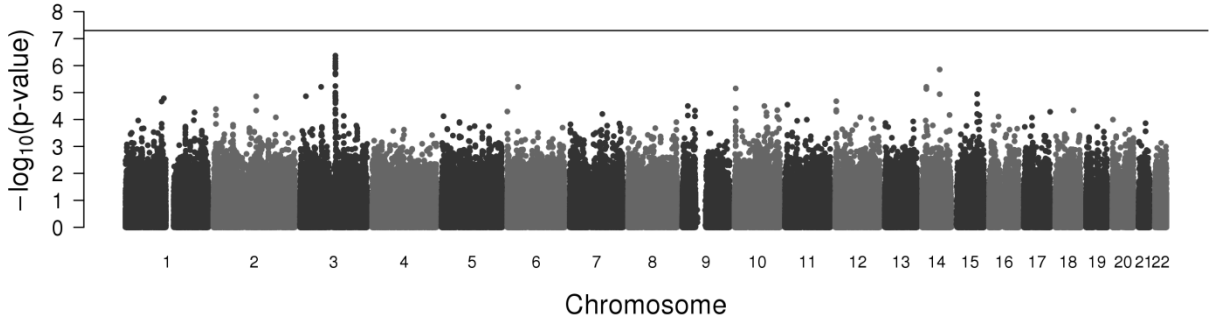
Immigration and Crime



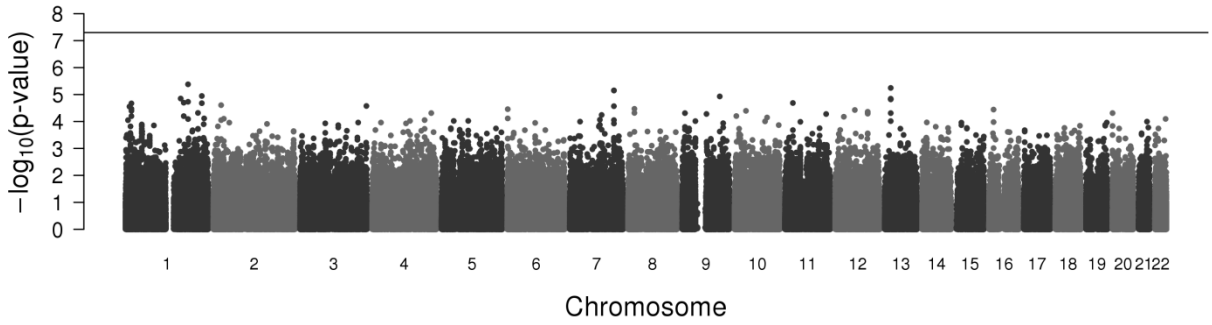
Economic policy



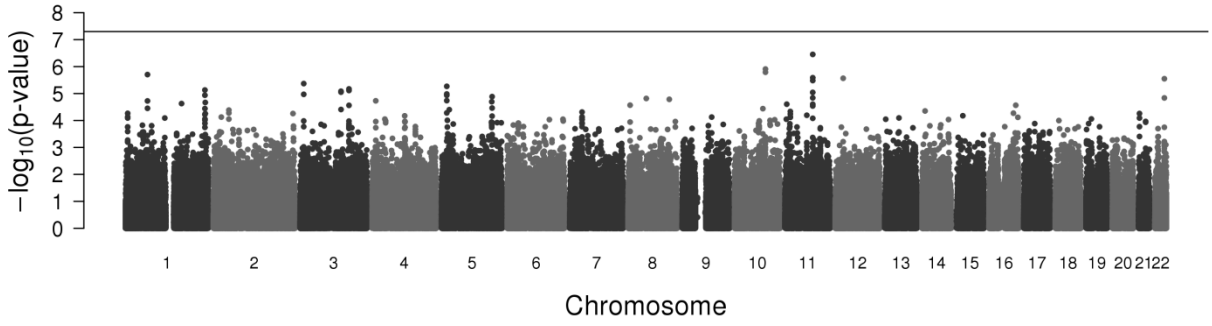
Environmentalism



Foreign policy



Feminism and Equality

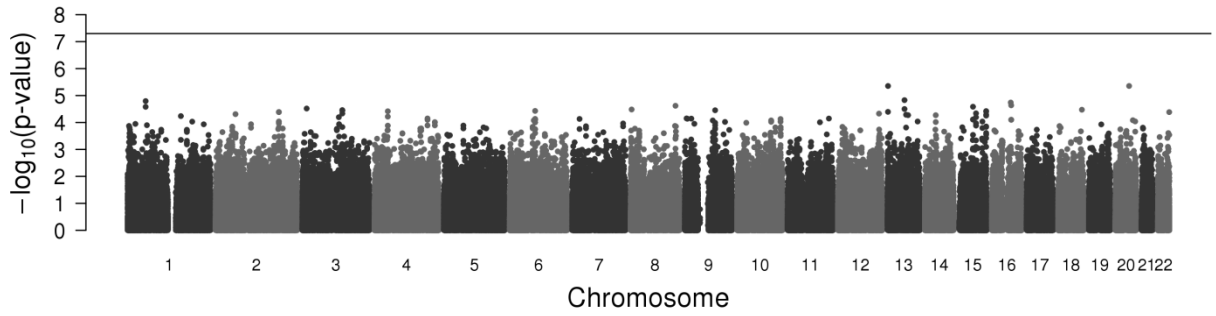


Note: Manhattan plots for the political phenotypes. The horizontal line at 7.3 corresponds to genome-wide significance ($p = 5 \times 10^{-8}$).

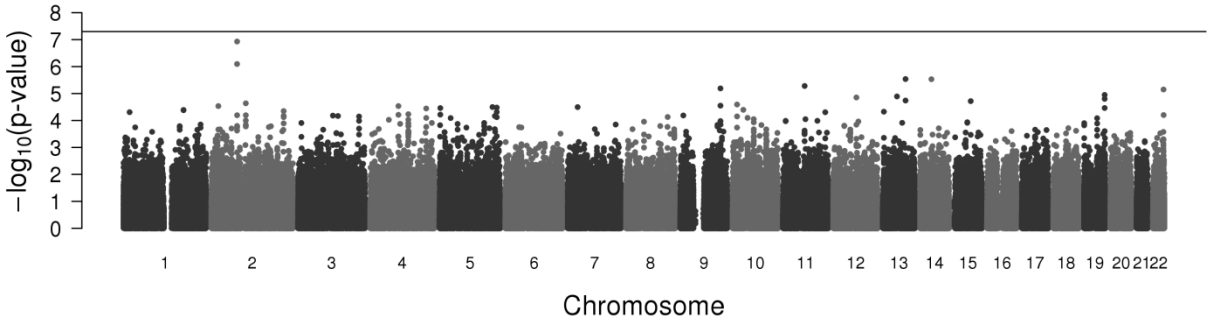
AIV.F. Manhattan Plots for the Economic Preferences and Education

Figure S4. Manhattan Plots for Economic Preferences and Education

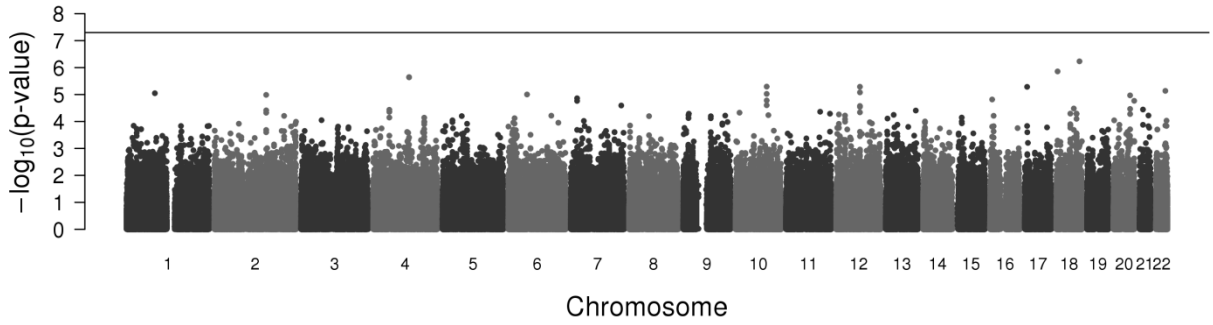
Education



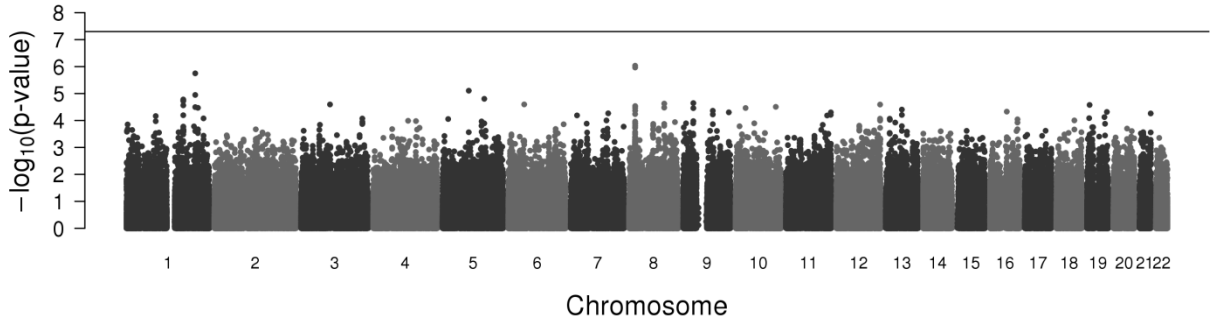
Fairness



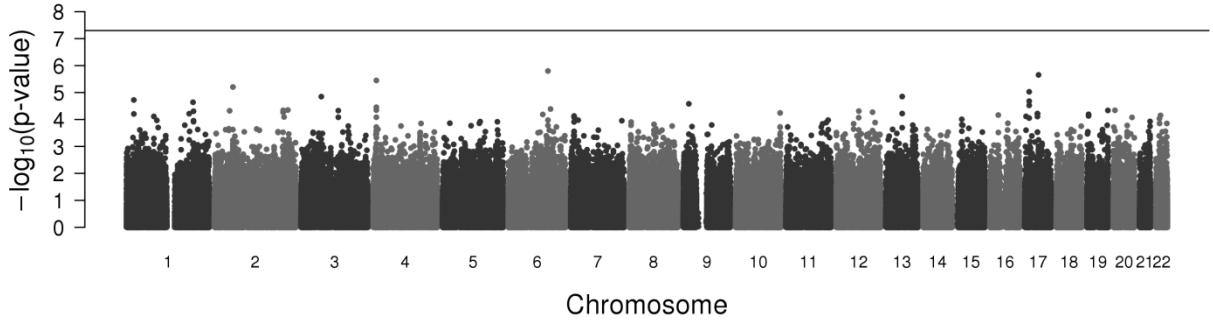
Patience



Trust



Risk



Note: Manhattan plots for the economic phenotypes. The horizontal line at 7.3 corresponds to genome-wide significance ($p = 5 \times 10^{-8}$).

AIV.G. Results of Prediction Analysis

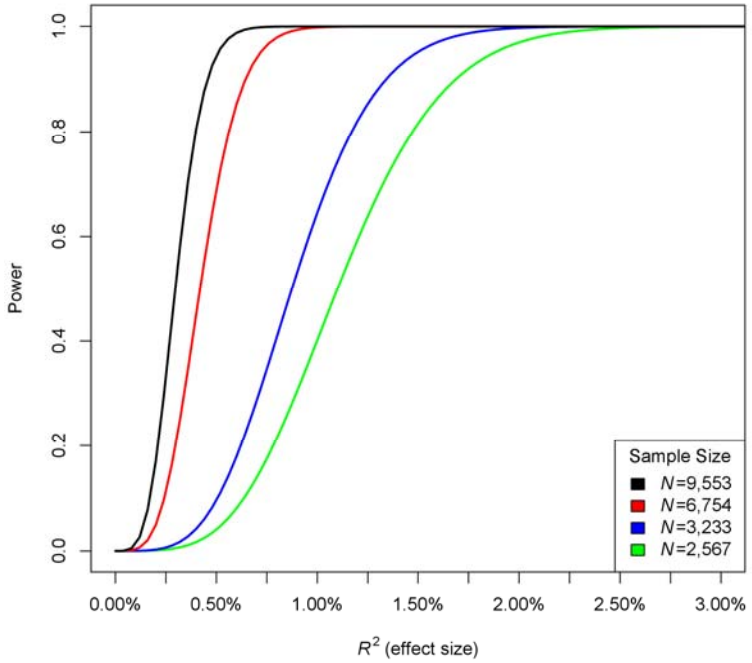
Table S7: Prediction Analysis

	Education	Economic Preferences				Political Preferences				
	Education	Risk	Patience	Fairness	Trust	Imm./ Crime	Econ Policy	Environ.	Femin./ Equality	Foreign Policy
R^2	0.0001	0.0001	0.0106	0.0002	0.0006	0.0058	0.0024	0.0014	0.0081	0.0002
N	928	312	322	321	327	321	321	321	321	321
Direction	-	-	+	+	+	+	+	-	+	+
p-value	0.723	0.922	0.098	0.816	0.674	0.183	0.374	0.567	0.125	0.794

Note: This table reports the results from the prediction analyses. Following (5), we constructed a polygenic risk score for each phenotype by splitting the sample into a 90% discovery sample and a 10% validation sample. The row *N* gives the number of individuals in the validation sample. The score is defined as the sum of the estimated regression coefficients multiplied by the number of reference alleles , divided by the number of non-missing SNPs. If a genotype in the score is missing for a particular individual, then the expected value is imputed based on the sample allele frequency. Finally, we regressed the dependent variable on the score, controlling for the number of non-missing SNPs. The R^2 is the incremental R^2 obtained when adding the score as a regressor. Direction describes whether the regression coefficient on the genetic score has the same (+) or opposite (-) sign in the validation sample as in the discovery sample. The *p*-value is obtained from the Wald test of the null hypothesis that the coefficient on the polygenic risk score is zero. All data are from the SALTY-Geno sample.

AIV.H. Power Analysis for GWAS

Figure S5: Power Analyses



This figure shows how the power to detect a marker at a nominal significance level of 10^{-7} as a function of sample size and the fraction of variance (R^2) explained by the marker. This *p*-value threshold was selected because no single SNP attained this level of nominal significance in any of the analyses. For educational attainment, there were 6,754 independent observations (i.e., from unrelated individuals) and a total of 9,553 observations – the true power therefore lies somewhere in between the two lines shown. For the political preference measures, we had 2,567 independent observations and a total of 3,233 observations (and similar sample sizes for the economic preference measures). The true power again lies somewhere between the lines shown. Even for the preference variables, where the sample size is smaller, the study was well-powered to detect a marker with an R^2 of 1.25% at a nominal significance level of 10^{-7}. The fact that we did not observe any associations at this level of significance suggests that it is unlikely that common variants with effects of that magnitude exist. For several of the traits, the lowest *p*-values observed were considerably higher than 10^{-7}. Hence, the 1.25% estimate of the upper bound is conservative. For educational attainment, the study was well-powered to detect markers with an R^2 of 0.5%.

AIV.I. Statistical Power and Reliability

It is sometimes argued that incentivized laboratory tasks, where financial incentives are attached to respondents' choices, produce measures of preferences that are more reliable and more correlated with "real-world" behavior than the survey measures. In fact, however, existing work does not support the hypothesis that such incentivized measures of risk aversion or other preferences are measured more reliably than survey-based measures (24-25). Moreover, our conclusion that effects of individual SNPs on risk preferences are very small would hold even if measures of preferences were much more reliable than those we use here.

To quantify this claim, let *T* denote the true value of the phenotype, i.e., its value if it were measured without error. Define R^2 as the squared correlation between the genetic marker and T. Rather than observing *T*, suppose that we observe *T* plus classical measurement error ε: ε is distributed normally with mean zero, is independent of *T*, and is drawn independently on each occasion when the phenotype is measured. As is standard, define the reliability, ρ, of the trait as the ratio of the variance in the true variable, $Var(T)$, to the variance in the observed variable $Var(T) + Var(\varepsilon)$. It is easy to show that the squared correlation between the genotype and the measured phenotype, $T + \varepsilon$, is given by ρR^2. In our data, we estimate ρ using the re-test observations. Across our four economic preference measures, the average reliability is 0.58. Suppose that we could improve the reliability of a measure from 0.58 to 0.80 (by shrinking the variance of ε). Then the upper bound of 1.25% ($= (0.58)R^2$) that we calculate would imply an upper bound of 1.72% ($= (0.80)R^2$) for this better-measured phenotype. Since our other phenotypes have higher re-test reliabilities than our economic measures, an analogous calculation would imply smaller upper bounds for those phenotypes.

AV. References

1. Howie B-N, Donnelly P, Marchini J (2009) A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association of Studies. *PLoS Genet* 5(6): e1000529.

2. Price A-L et al. (2006) Principal Components Analysis Corrects for Stratification in Genome-wide Association Studies. *Nat Genet* 38(8): 904–909.

3. Devlin B, Roeder K (1999) Genomic Control for Association Studies. *Biometrics* 55(4): 997–1004.

4. Yang J, Weedon M-N, Purcell S, Lettre G, Estrada K (2011a) Genomic inflation factors under polygenic inheritance. *Eur J Hum Genet* 19: 807–812.

5. Purcell S-M et al. (2009) Common Polygenic Variation Contributes to Risk of Schizophrenia and Bipolar Disorder. *Nature* 460(7256): 748–752.

6. Yang J, Lee S-H, Goddard M-E, Visscher P-M (2011b) GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 88(1): 76–82.

7. Yang J et al. (2010) Common SNPs Explain a Large Proportion of the Heritability for Human Height. *Nat Genet* 42(7): 565–569.

8. Lee S-H, Wray N-R, Goddard M-E, Visscher P-M (2011) Estimating Missing Heritability for Disease from Genome-wide Association Studies. *Am J Hum Genet* 88: 294–305.

9. Dohmen T et al. (in press) Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences. *J Eur Econ Assoc*.

10. Barsky R-B, Thomas Juster F, Kimball M-S, Shapiro M-D (1997) Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. *Q J Econ* 112(2): 537–579.

11. Laibson D-I (1997) Golden Eggs and Hyperbolic Discounting. *Q J Econ* 112(2): 443–477.

12. DellaVigna S (2009) Psychology and Economics: Evidence from the Field. *J Econ Lit* 47: 315–372.

13. Benjamin D-J, Brown S-A, Shapiro J-M (2006) Who is 'Behavioral'? Cognitive Ability and Anomalous Preferences. *Cornell University Working Paper*.

14. Frederick S, Loewenstein G, O'Donoghue T (2001) Time Discounting and Time Preference: A Critical Review. *J Econ Lit* 40(2): 351–401.

15. Nannestad P (2008) What Have We Learned About Generalized Trust, If Anything? *Annual Review of Political Science* 11: 413–436.

16. Glaeser E, Laibson D, Scheinkman J, Soutter C (2000) Measuring Trust. *Q J Econ* 115(3): 811–846.

17. Kahneman D, Knetsch J-L, Thaler R (1986) Fairness as a Constraint on Profit Seeking: Entitlements in the Market. *Am Econ Rev* 76(4): 728–741.

18. Kaiser H-F (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23(3): 187–200.

19. Oskarsson S et al. (2010) Ideology and Vote Choice. *Uppsala University Working Paper*.

20. Statistics Sweden (2008) General Elections in 2006. Part 4. Special Studies (in Swedish). Statistics Sweden: Örebro.

21. Sandewall Ö, Cesarini D, Johannesson M. (2012) Co-Twin Methodology and the Return to Schooling – Testing a Critical Assumption. *New York University Working Paper*.

22. Isacsson G (2004) Estimating the Economic Return to Educational Levels Using Data on Twins. *Journal of Applied Econometrics* 19(1): 99–119.

23. Cesarini D, Johannesson M, Magnusson P-K, Wallace B (2011) The behavioral genetics of behavioral anomalies. *Manage Sci* doi 10.1287/mnsc.1110.1329.

24. Beauchamp J-P, Cesarini D, Johannesson M (2011) The Psychometric Properties of Measures of Economic Risk Preferences. *Working Paper, Harvard University, New York University, and Stockholm School of Economics*.

25. Lönnqvist J-E, Verkasalo M, Walkowitz G, Wichhardt P-C (2011) Measuring Individual Risk Attitudes in the Lab: Task or Ask? An Empirical Comparison. *CGS Working Paper, University of Cologne*.