The Genetic Architecture of Economic and Political Preferences

Daniel J. Benjamin[1]*, David Cesarini[2], Matthijs J.H.M. van der Loos[3], Christopher T. Dawes[4], Philipp D. Koellinger[3], Patrik K.E. Magnusson[5], Christopher F. Chabris[6], Dalton Conley[7], David Laibson[8], Magnus Johannesson[9], and Peter M. Visscher[10]

[1]*Department of Economics, Cornell University, 480 Uris Hall, Ithaca, NY 14853, USA.*

[2]*Center for Experimental Social Science, Department of Economics, New York University, 19 W. 4th Street, New York, NY 10012, USA.*

[3] *Department of Applied Economics, Erasmus Universiteit Rotterdam, Post Box 1738, 3000 DR Rotterdam, Netherlands.*

[4] *Department of Politics, New York University, 19 W. 4th Street, New York, NY 10012, USA.*

[5] *Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Box 281, SE-171 77 Stockholm, Sweden.*

[6] *Department of Psychology, Union College, 807 Union Street, Schenectady, NY 12308, USA.*

[7] *Department of Sociology, New York University, Department of Sociology, New York University, 295 Lafayette Street, 4th floor, New York, NY 10012, USA.*

[8] *Department of Economics, Harvard University, Littauer Center, 1805 Cambridge Street, Cambridge, MA 02138, USA.*

[9] *Department of Economics, Stockholm School of Economics, Box 6501, SE-113 83 Stockholm, Sweden.*

[10] *University of Queensland Diamantina Institute, Level 4, R Wing, Princess Alexandra Hospital, Ipswich Road, Woolloongabba 4102, Australia.*

Classification: Social Sciences; Economic Sciences; Social Sciences; Political Sciences; Biological Sciences; Genetics.

Manuscript information: 12 text pages, 1 Figure, 2 Tables.

*To whom correspondence should be addressed: phone: +1 607 255 2355. E-mail: daniel.benjamin@gmail.com.

Abstract: Preferences are fundamental constructs in all models of economic and political behavior and important precursors to many lifetime outcomes. Twin studies suggest that individual-level variation in preferences partly due to genetic factors, but twin-based heritability estimates remain controversial. Here, with a new sample of comprehensively-genotyped subjects with data on political and economic preferences, as well as income and educational attainment, we use genome-wide data to estimate the proportion of variation in these traits explained by common SNPs. The overall pattern of results is consistent with findings for other complex traits: (1) the estimated fraction of phenotypic variation that could ultimately be explained by dense SNP arrays is between one-quarter and one-half of the narrow heritability as estimated using twin and family studies; and (2) GWAS and prediction analyses reveal that many common SNPs with large explanatory power for these traits are unlikely to exist. These findings have implications for evaluating the extent to which the potential benefits of molecular genetic data in the social sciences will be borne out in the near future. The results are also useful for evaluating existing published associations in candidate gene studies of economic and political phenotypes. We propose some constructive responses to the inferential challenges posed by the small explanatory power of individual SNPs.

Introduction

Preferences are fundamental constructs in most theories of economic and political behavior. Economists believe that risk aversion, patience, fair-mindedness, and trust are fundamental preferences because they explain a wide range of behaviors, and political ideology plays a similar role within political science. For example, measures of risk preferences predict diverse risky behaviors, such as smoking, drinking, and holding stocks rather than bonds (1-2). Experimentally-elicited patience predicts body mass index, smoking behavior and exercise (3), as well as credit card borrowing (4). Political preferences similarly predict a wide range of political behaviors, including voting (5) and monetary campaign contributions (6), as well as campaign activities like volunteering, attending rallies, and displaying yard signs (7).

Behavior genetic studies, beginning with some classical papers on social and political attitudes (8-9), have found that some of the variation in political and economic preferences can be statistically accounted for by genetic factors (10-15). However, these conclusions continue to be contested (e.g., 16). Critics point out that the twin-based estimates of "heritability"—which compare the correlation of an outcome across monozygotic (MZ) twin pairs with that correlation across dizygotic (DZ) twin pairs—rely on strong assumptions which render it difficult to draw any definite conclusions. For example, one criticism sometimes leveled against twin studies is that the similarity of MZ twins may be inflated due to failure of the equal-environment assumption (17), a bias which could cause heritability estimates to be positive even if the true value were zero.

Other researchers have embraced the twin-based heritability estimates and argued that the next step is to use molecular data to identify the specific genetic pathways that influence behavior. If behavioral phenotypes are heritable, then associated genetic markers exist in principle and might be identified in practice. If specific genetic markers can be identified that are associated with a preference, then it might be possible to predict a given individual's preference without access to any phenotypic data (a feat that cannot be accomplished just using heritability estimates). Identifying such predictive markers may shed light on the biological pathways underlying preferences and ultimately help us better understand how genes affect outcomes (18). If a set of genetic markers is sufficiently predictive, then these markers could be used in social science research as covariates, as instrumental variables* (21-

22), or, under certain conditions, as factors for identifying at-risk populations who might benefit from additional resources.

Even if social-science traits are heritable, the extent to which the potential of molecular genetic data will be fulfilled for a given trait hinges crucially on its "molecular genetic architecture," i.e., the joint distribution of effect sizes and allele frequencies of the causal genetic markers (23-25). The molecular genetic architecture is the result of evolutionary forces, including mutation, drift, and selection. The architecture determines the difficulty with which the genetic variants associated with a trait can be identified and what sample sizes will be required. It also determines the out-of-sample aggregate predictability that can be derived from a set of SNPs considered jointly.

In this paper, we use a new sample of comprehensively-genotyped subjects from the Swedish Twin Registry who were recently administered, as part of a survey called 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, and foreign policy.

We also study educational attainment and income because much is known about their heritability not only from twin studies (26), but unlike the preference measures, also from behavior-genetic estimates that use other pedigree relationships (27-28). Educational attainment and income are 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, but our main focus is on using the whole-genome data, first, to provide new evidence regarding heritability as estimated directly from the genetic data, and, second, to learn about the genetic architecture of these traits.

First, we employ a recently-developed method (29-30) that uses the whole-genome data to estimate a lower bound of the narrow heritability of these traits. The technique—which we will call Genomic-Relatedness-Matrix Restricted Maximum Likelihood (GREML)—has been applied to height (30), intelligence (31), personality traits (32), and several common diseases (33), but never before to economic and political phenotypes.

Since these lower-bound estimates of narrow heritability do not rest on the same assumptions used in twin studies, they provide an additional source of evidence regarding heritability. The method is instead based on the assumption 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 single nucleotide polymorphism (SNP) data, unlike in behavior-genetic studies, where *expected* relatedness (known from the family pedigree) is used. Some of the concern about behavior-genetic studies is that expected relatedness could be correlated with environmental factors that are not endogenous to genotype (as defined by Jencks (18)). Since 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 estimate of heritability can be obtained by examining how the correlation in phenotype between pairs of individuals relates to the realized genetic distance between those individuals. This would be an unbiased estimator of narrow heritability if genetic distance were calculated using all the genetic variants that are causal for the phenotype. In practice, since 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 heritability (30). †

Second, we use the whole-genome data to 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 with 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 (36), height (37) and intelligence (31), but none of these methods have been applied to economic or political preferences.

Results

We began by computing the sibling correlations for all eleven variables and the re-test reliabilities for the nine preference measures. ‡ Table 1 reports the results. In total, the sample of SALTY respondents is 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. We estimate re-test reliabilities for the preference measures using data from 491 respondents who answered the survey twice. The sibling correlations for the SALTY questions on patience (38), risk aversion (25), and political preferences (39) have previously been analyzed and are reproduced here to facilitate comparison with the remaining results. The income and educational attainment variables we used have also been previously studied in partially overlapping samples (27-28). We report the correlations in educational attainment and the natural logarithm of income averaged over 1985 and 1990. The implied heritabilities of the economic preferences are typically in the vicinity of 30% and the estimates for political preferences are typically around 40%. The final column of the table shows the estimated test-retest reliability of each of the preferences phenotypes. These reliabilities are estimated from a subset of respondents who answered the survey twice.

We estimated, for each trait, the proportion of phenotypic variation accounted for by all SNPs, following the method of (30). These lower-bound heritability estimates for the nine traits are reported in Table 2. For economic preferences, only one of the four variables, trust, is significant, with the point estimate suggesting that the common SNPs explain over twenty percent of phenotypic variation (p = 0.047). The remaining effects are lower, in one case zero, and not statistically distinguishable from zero. For political preferences, three out of the five derived attitudinal dimensions are statistically significant estimates, though one of the three is only significant at the 10% threshold. These estimates are 0.203 (p = 0.079) for immigration/crime, 0.344 (p = 0.012) for economic policy, and 0.353 (p = 0.001) for foreign policy attitudes. The estimates are noisier (and the p-values tend to be higher) for phenotypes with lower re-test reliabilities. Keeping in mind that these are noisy and are lower bounds, the estimates taken as a whole 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 the genotyped individuals in the sample, not just the survey respondents. For this phenotype, the larger sample decreases the standard error of the estimates substantially, while the point estimate of 0.191 (p = 0.001) is only somewhat higher than the average point estimate for the other phenotypes. These analyses are all based on mixed-sex samples, controlling for sex, age, and the first ten principal components of the genotypic data. For our last variable, log income, we restrict the sample to males only. § Our point estimate for log income in males is 0.061 (p = 0.313). In the Supplementary Information, we report additional analyses of log income and find that when we pool men and women, the point estimate is 0.084 (p = 0.085).

We also conduct the analysis separately by chromosome, as in (30) and (31). Between unrelated individuals, realized relatedness is random and independent across chromosomes, and the expected relatedness measured from any chromosome is zero. If, rather than being concentrated in a particular location, the genetic variation that predicts a trait were uniformly distributed across the genome, then greater realized relatedness from any given chromosome will predict greater phenotypic similarly, and this association will be stronger from longer chromosomes. The bottom row of Table 2 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 8 out of 11 phenotypes, significantly so in 3 cases. Analogous positive correlations have been reported for height (30) and cognitive ability (31) and have been interpreted as evidence that the trait is highly polygenic with causal variants distributed across the genome.

Next, we examine whether we can identify individual SNPs that predict economic and political preferences. For none of the eleven 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 eleven traits. The standard diagnostic for population stratification (i.e., ethnic confounding) in GWAS is inflated test statistics in the QQ-plot (e.g., 42); there is no evidence of inflated test statistics across the traits, with estimated lambdas in the range 0.982 (income) to 1.017 (educational attainment). While this suggests that our controls for population structure worked well, it is somewhat surprising that there is no tendency for the lambdas to be larger than 1, given that some inflation is expected

under a polygenic model even without any stratification (43). In the Supplementary Information, we provide details on the full set of SNPs with *p*-values below 10^{-5} for the nine preference measures, but we are skeptical that any of these associations will be replicable, given the absence of genome-wide significant hits for any of the eleven phenotypes.

Finally, we examine the aggregate, out-of-sample predictive power of the SNPs. Following (36), 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 based on a pruned set of 107,360 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 their observed phenotype. We do not find any significant out-of-sample predictability for any of the traits, and for most phenotypes, the explanatory power of the predictor is well below $R^2 = 0.1\%$. These results are reported in the Supplementary Information.

Discussion

The data reported here reveal a number of descriptive facts about the heritability and genetic architecture of political and economic preferences. First, we report 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 MZ and DZ correlations. We obtain heritability estimates that are consistent with typical estimates previously reported for both political attitudes (10, 12) and economic preferences (11, 28, 38, 14), as well as educational attainment (44). Our estimates for income are actually a little higher than what has previously been reported in Swedish data (27). Overall, these results are consistent with the hypothesis that there exists a moderate correlation between genotype and the eleven phenotypes. None of these 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, as evidenced by their lower test-retest reliabilities.

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.

When we estimate the cumulative effect of genotyped SNPs using the method of (30), 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 non-zero and, for the better measured variables, statistically distinguishable from zero. Previous papers on height (30), intelligence (31), personality traits (32), and several common diseases (33) have found that the SNP-based heritability estimates are between one-quarter and one-half the size of the twin-study estimates. One interpretation of the gap is that genotyped SNPs tag less than half the genetic variation in those traits. The gap may also reflect an upward bias in twin-based estimates of narrow heritability estimates due to environmental confounding or non-additive variation, both of which will cause an upward bias in the estimated additive genetic proportion of variance. Consistent with the interpretation that some of the gap is due to bias, behavior-genetic heritability estimates for income and education based on non-twin siblings, for example adoptees and full siblings, are somewhat lower than those based on twins (27-28).

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. Since the economic and political preference measures have twin-based heritabilities around 0.30 (22-25, 11) and 0.40 (see Table 1), 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 standard error of 0.15 (as suggested by Table 2), then our power to statistically reject the null hypothesis of zero heritability in a one-sided test at the five percent 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 standard error of 0.15 (again as suggested by Table 2), 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 five percent level. In fact, we observe three significant associations at the five percent level and one more at the eight percent 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 (45), especially in the context of social science traits (24-25). Figure 1 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 which explain at least 1.25% of trait variation at a nominal significance level of 10^{-7}—yet no single SNP in our sample attains this level of significance in our sample. Moreover, 1.25% is an upper bound to the effect sizes we can rule out because: first, since 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; 46); and second, for many of the variables, the lowest observed *p*-value was considerably higher than 1.1×10^{-7}. To illustrate our statistical power another way, if across the eleven traits there are a total of ten independently-distributed SNPs each with R^2 of 0.75% or larger, our study was well-powered to detect at least one of them—and yet we found none. We conclude that is unlikely that many common polymorphisms with such effect sizes exist. 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. Of course, our evidence does not rule out the possibility that there exist rare variants with large effects on these phenotypes.

Fourth, the results from our prediction exercise show that in a sample of approximately 3,200 individuals, a standard polygenic risk score has negligible out-of-sample predictability. This 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 based on these estimates. Evidently a discovery sample of 2,900 individuals (about 90% of 3,200) is far too small to obtain even mildly useful predictive power for standard measures of economic or political preferences.

These findings fit in nicely with an emerging consensus in medical genetics, according to which common 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 be even more acute in the social sciences, since the phenotypes are usually several degrees removed from genes in the chain of causation (24-25). Our results suggest that much of the "missing heritability" (49)—the gulf between the cumulative explanatory power of common variants identified to date and the heritability as 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 (50) famously proposed three "laws of behavior genetics": first, all human behavioral traits are heritable; second, the proportion of variance attributable to family is smaller than the proportion attributable to genes; and third, a large portion of individual differences is explained by factors other than families 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 (51-52), 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 (53). There is direct evidence for such an architecture for intelligence (31, 51), personality (32), 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 other complex 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 of several hundred individuals (for reviews of work to date, see (54) and (25)). These samples are very small by the standards of medical genetics, often based on samples in the hundreds or less. Such studies are only adequately powered if the 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, suggest that adequate power actually requires a sample size that is yet another order of magnitude larger even than ours. Statistically significant associations obtained in a small sample should be approached with caution for two reasons: first, since most existing published studies are dramatically underpowered, the probability that an association study will detect a true signal is vanishingly small; hence if a significant association is observed, Bayesian calculations indicate that the posterior odds of a true association are low (52, 24, 25); and second, publication bias—the tendency for findings, as opposed to non-findings, 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 (55).

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 of precursors to important behaviors and outcomes, such as preferences. More speculatively, such insights may also help inspire the development of new theoretical constructs which are more closely aligned with the underlying biology than the existing concepts such as "risk aversion" or "patience" that we study here (56). 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 construct samples which are several orders of magnitude larger than those presently employed in this sort of research. Unfortunately, even if these pathways are eventually identified, our quantitative results suggest that many of the identified markers will only explain a tiny share of variance.

Another interesting potential contribution to economics and political science would be the use of genetic markers as instrumental variables in (non-genetic) empirical work. In order for the gene-as-instrument to be convincing, 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, making the marker invalid as an instrumental variable. As more is understood about the pathways through which candidate IV markers operate, researchers will be in a better position to assess the plausibility of the exclusion restriction on a case-by-case basis, depending on the research question.

A different potential use of molecular genetic data to social science would be as control variables for genetic heterogeneity in (non-genetic) empirical work, in order to reduce the variance of the error term and shrink the standard errors of coefficient estimates. For such an application to have any practical utility, the markers that are selected as controls need to explain a non-negligible 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, this does not appear to be feasible for economic and political traits.[11] It is likely that extremely large—perhaps impossibly large—samples will be required. If so, some of the most exciting possible uses of molecular genetic data in the social sciences lie many years in the future.

In summary, our molecular-genetic-based estimates of heritability partially corroborate the twin-based estimates and suggest that molecular genetic data *could* be predictive of preferences. Our other results, however, suggest that excitement about the utility of molecular genetic data in social science research likely needs to be tempered by an appreciation that much of that 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 sizeable predictive power from many SNPs considered jointly.

Rather than being destructive to the enterprise of incorporating genetic data into social science, an understanding of the molecular genetic architecture of economic and political preferences can help guide research in more productive directions. Indeed, there are several constructive and complementary responses one might imagine to the inferential challenges posed by the genetic architecture documented here. One is to undertake efforts to actually

obtain very large samples that contain both genetic and social science data. A second response is to carefully evaluate the psychometric properties of social science phenotypes to minimize attenuation bias due to error in measurement and thereby maximize power for any given sample size. In our view, the larger GREML estimates for the political preference measures relative to the less reliable economic measures illustrates these potential gains. A third suggestion is to focus on behavioral phenotypes that are more biologically proximate. One example that is a focus of some current work is smoking, a behavior for which large, replicated associations have been found with SNPs in the nicotine receptor gene *CHRNA3* (57). For biologically proximate phenotypes, it is more likely that some associations will have non-trivial effect sizes and clearer causal interpretations.

Materials and Methods

Since December 2010 10,946 Swedish Twins have been genotyped using the Illumina HumanOmniExpress BeadChip genotyping platform. We applied standard quality controls to the genetic data; see the Supplementary Information for details. In all our GWAS analyses, we control for the first ten principal components of the genotypic data, sex and age and adjust the standard errors for non-independence within family. We computed the GREML estimates using the publicly-available GCTA software (58). 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. 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; details are relegated to the Supplementary Information.

Footnotes

* For critical perspectives, see (19-20).

† Another approach is to estimate the genetic variance from within-family variation in genetic relatedness (see 35). The estimates derived from such an analysis are unbiased estimates of heritability rather than lower bounds because the identical-by-descent probabilities for all variants, including the rare ones not tagged by the genotyped SNPs or microsatellites, can be inferred if one has sibling data.

‡ Additional details on variable construction and materials and methods are available in the Supplementary Information.

§ Labor economists usually study earnings because it is taken to be a proxy for productivity. This assumption is most reasonable in prime aged males, whose attachment to the labor market tends to be strong and who typically work full time, making variation in hours worked less of a confound for measuring productivity. For a discussion of the difficulties with analyzing female labor supply and earnings, see (40).

¶The fact that we observe some GREML point estimates of zero is not surprising. Since the estimator is constrained to produce a non-negative estimate, the bound at zero will often be attained when the true population parameter is low and estimated imprecisely.

‖ While the survey measures we use here are common in economics (e.g., 1), it is also common in economics to measure preferences using laboratory tasks that attach financial incentives to performance (47). It is sometimes argued that these incentivized laboratory tasks produce measures of preferences that are more reliable and more correlated with real-world behaviour 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 (25, 48). Moreover, our conclusion that effects of individual SNPs on risk preferences are very small would hold even with measures of preferences were much more reliable than those we use. For example, suppose we could improve the reliability of one of the measures from 0.58 (the average reliability across our four economic preference measures) to 0.80. Then the upper bound of 1.25% that we calculate would imply an upper bound of 1.72% for this better-measured phenotype.

** While we have emphasized the possibility that the heritability of preferences is composed of many common SNPs of small effect, we also note that the common SNPs (the heritable variation measured on dense SNP arrays) do not tag all the heritable variation in the genome. If the twin-based estimates of heritability are correct, then rare, perhaps non-SNP, causal variants that are not in close linkage disequilibrium with the genotyped markers may explain some of the heritable variation. If such variants exist, they may have large effects. Nonetheless, since such variants will be rare, large samples will also be required for adequate power to detect those markers.

†† 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 if and how such information should be used. In principle, the information may be used to help people make more informed decisions. For example, if dyslexia could be predicted at a relatively early age, such information could in principle be used to help parents make better information about treatment strategies (24). Such potential benefits must of course be carefully weighed against the costs. For example, insurance markets may break down due to adverse selection (56) unless there are restrictions placed on the availability of genetic data.

Acknowledgements

The authors thank Jonathan Beauchamp, Sandy Jencks, and Sarah Medland for helpful comments and discussions and Cara Costich, Julia Goorin, and Michael Puempel for research assistance. The study was supported by the GenomeEUtwin project under the European Commission Programme OEQuality of Life and Management of the Living Resources of 5th 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 Jan Wallander and Tom Hedelius Foundation, and the NIA/NIH through grants P01AG005842-20S2 and T32-AG000186-23.

Table 1: Sibling Correlations

	Economic Outcomes		Economic Preferences				Political Preferences				
	Education	Income	Risk	Patience	Fairness	Trust	Imm./ Crime	Econ. Policy	Environ.	Femin.	Foreign Policy
ρ_{MZM}	0.70	0.47	0.41	0.05	0.32	0.37	0.56	0.42	0.26	0.48	0.48
	(.65-.75)	(.31-.61)	(.33-.49)	(-.04-.15)	(.23-.42)	(.29-.45)	(.49 -.62)	(.34-.50)	(.17-.34)	(.40-.57)	(.40-.54)
ρ_{DZM}	0.50	0.22	0.20	0.07	0.24	0.18	0.33	0.34	0.09	0.27	0.16
	(.42-.57)	(.11-.33)	(.12-.28)	(-.04-.20)	(.15-.33)	(.08-.27)	(.25-.42)	(.26-.42)	(.01-0.18)	(.19-.35)	(.07-.24)
ρ_{MZF}	0.74	0.37	0.35	0.16	0.24	0.33	0.60	0.45	0.34	0.41	0.47
	(.70-.78)	(.25-.50)	(.27-.43)	(.07-.25)	(.16-.32)	(.25-.40)	(.55-.65)	(.37-.51)	(.23-.46)	(.34-.48)	(.39-.53)
ρ_{DZF}	0.54	0.22	0.13	0.11	0.10	0.19	0.37	0.22	0.16	0.19	0.21
	(.49-.59)	(.14-.32)	(.05-.21)	(.03-.20)	(.03-.17)	(.11-.26)	(.29-.44)	(.14-.29)	(.09-.24)	(.12-.26)	(.14-.29)
N_{MZM}	561	560	443	491	450	458	448	448	448	448	448
N_{DZM}	614	612	477	655	482	486	487	487	487	487	487
N_{MZF}	544	538	594	700	652	659	630	630	630	630	630
N_{DZF}	845	815	636	1108	681	714	665	665	665	665	665
$\rho_{RE\text{-}TEST}$	-	-	0.71	0.40	0.57	0.63	0.86	0.84	0.62	0.78	0.72
			(.66-.76)	(.27-.52)	(.49-.65)	(.57-.69)	(.84-.88)	(.84-.88)	(.84-.88)	(.84-.88)	(.84-.88)
$N_{RE\text{-}TEST}$	-	-	475	483	469	482	471	471	471	471	471

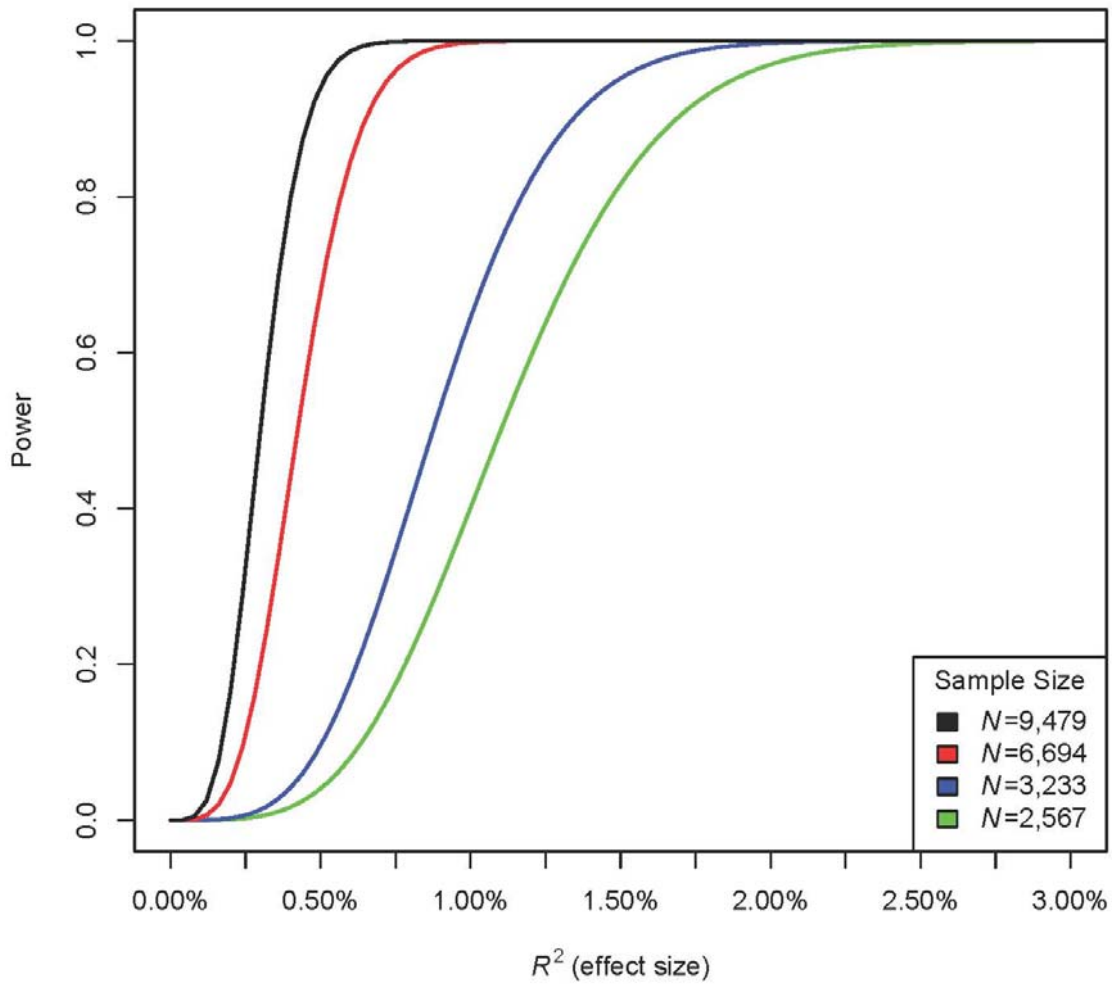
Note: This table gives the sibling and re-test Pearson correlations for the eleven phenotypes. The 95% confidence intervals given in parentheses are computed by bootstrapping with 500 draws. MZM: number of male monozygotic pairs; DZM: male dizygotic pairs; MZF: female monozygotic pairs; DZF: female dizygotic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

Table 2: GREML Analyses

	Economic Outcomes		Economic Preferences				Political Preferences				
	Education	Income	Risk	Patience	Fairness	Trust	Imm./ Crime	Econ. Policy	Environ.	Femin.	Foreign Policy
V(g)/V(P)	0.191	0.062	0.137	0.085	0.000	0.242	0.203	0.344	0.000	0.000	0.354
s.e.	0.062	0.124	0.152	0.148	0.150	0.146	0.147	0.150	0.148	0.147	0.149
p-value	0.001	0.313	0.186	0.285	0.500	0.047	0.079	0.012	0.500	0.500	0.009
N	5,682	2,866	2,327	2,399	2,376	2,410	2,368	2,368	2,368	2,368	2,368
Chrom.	0.241	0.139	0.118	-0.195	-0.111	0.460	0.118	0.496	-0.311	0.247	0.462
p-value	0.280	0.537	0.601	0.385	0.623	0.031	0.601	0.019	0.159	0.268	0.030

Note: This table reports GREML estimates for the eleven variables. 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. 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. As explained in the text, the results for income are based exclusively on men. The third row gives the *p*-value for the test of the null 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 hypothesis that, across the 22 autosomes, the correlation between chromosomal length and the proportion of variation explained by the chromosome is zero. All data are from the SALTY-Geno sample (see Supplementary Information for sample descriptions).

Figure 1: Power Analysis



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 6,694 independent observations (i.e., from unrelated individuals) and a total of 9,479 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. 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 and income, the study was well-powered to detect markers with an R^2 of 0.5%.

References

1. Barsky R-B, Juster F-T, 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: 537-579.

2. Dohmen T-F et al. (2011) Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *J Eur Econ Assoc* 9(3): 522-550.

3. Chabris C-F, Laibson D, Morris C-L, Schuldt J-P, Taubinsky D (2008) Individual Laboratory-Measured Discount Rates Predict Field Behavior. *J Risk Uncertain* 37(2-3): 237-269.

4. Meier S, Sprenger C (2010) Present-Biased preferences and credit card borrowing. *Am Econ J Appl Econ* 2(1): 193-210.

5. Jost J-T (2006) The End of the End of Ideology. *Am Psychol* 61(7): 651.

6. Morton R, Cameron C (1992) Elections and the Theory of Campaign Contributions: A Survey and Critical Analysis. *Economics and Politics* 4(1): 79-108.

7. Claassen R-L (2007) Campaign Activism and the Spatial Model: Getting Beyond Extremism to Explain Policy Motivated Participation. *Political Behavior* 29(3): 369-390.

8. Eaves, L-J, Eysenck, H-J (1974) Genetics and the development of social attitudes. *Nature* 249: 288-289.

9. Martin N-G et al. (1986) Transmission of social attitudes. *Proc Nat Acad Sci USA* 83(12): 4364-4368.

10. Alford J-R, Funk C-L, Hibbing J-R (2005) Are political orientations genetically transmitted? *Am Polit Sci Rev* 99: 153-167.

11. Cesarini D, Dawes C-T, Johannesson M, Lichtenstein P, Wallace B (2009) Experimental Game theory and Behavior Genetics. *Ann N Y Acad Sci* 1167: 66-75.

12. Hatemi P-K et al. (2007) The Genetics of Voting: an Australian Twin Study. *Behav Genet* 37(3): 435-448.

13. Hatemi P-K et al. (2011) Genome-Wide Analysis of Political Attitudes. *J Polit* 3(1):1–15.

14. Zhong S et al. (2009) The Heritability of Attitude Toward Economic Risk. *Twin Res Human Genet* 12(1) 103-107.

15. Zyphur M-J, Narayanan J, Arvey R-D, Alexander G-J (2009) The Genetics of Economic Risk Preferences. *J Behav Decis Mak* 22(4): 367-377.

16. Charney E (2008) Genes and ideologies. *Perspectives on Politics* 6(2): 299-319.

17. Stenberg A (2011) Nature or Nurture? A Note on the Misinterpreted Twin Decomposition. Swedish Institute for Social Research (SOFI) Working Paper #4/2011.

18. Jencks C (1980) Heredity, Environment, and Public Policy Reconsidered. *Am Sociol Rev* 45(5): 723-736.

19. Conley D (2009) The Promise and Challenges of Incorporating Genetic Data into Longitudinal Social Science Surveys and Research. *Biodemography Soc Biol* 55(5): 238-251.

20. Cawley J, Han E, Norton E-C (2011) The Validity of Genes Related to Neurotransmitters as Instrumental Variables. *Health Econ* 20(8): 884-888.

21. Davey Smith G, Ebrahim S (2003) 'Mendelian Randomization': Can Genetic Epidemiology Contribute to Understanding Environmental Determinants of Disease. *Int J Epidemiol* 32(1): 1–22.

22. Ding W, Lehrer S-F, Rosenquist J-N, Audrain-McGovern J (2006) The Impact of Poor Health on Education: New Evidence Using Genetic Markers. *NBER Working Paper w12304*.

23. Lander, E-S (2009) Initial impact of the sequencing of the human genome. *Nature* 470(7333), 187-197.

24. Benjamin D-J et al. (in press) The Promise and Pitfalls of Genoeconomics. *Annu Rev Econom*.

25. 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.

26. Taubman P (1976) The Determinants of Earnings: Genetics, Family, and Other Environments: A Study of White Male Twins. *Am Econ Rev* 66(5): 858-870.

27. Björklund A, Jantti M, Solon G Influences of Nature and Nurture on Earnings Variation (2005) in *Unequal Chances: Family Background and Economic Success,* eds Bowles S, Gintis H, Groves M-0 (Princeton University Press, Princeton), pp 145-164.

28. Cesarini D (2010) Family Influences on Productive Skills, Human Capital and Lifecycle Income. Mimeo. In MIT Phd Dissertation, Chapter 3.

29. Visscher P-M, Yang J, Goddard ME (2010) A Commentary on 'Common SNPs Explain a Large Proportion of the Heritability of Human Height' by Yang et al. (2010). *Twin Res Hum Genet* **13**(6): 517-524.

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

31. Davies G, Tenesa A, Payton T, Yang J, Harris S-E (2011) Genome-wide Association Studies Establish that Human Intelligence is Highly Heritable and Polygenic. *Mol Psych* doi:10.1038/mp.2011.8.

32. Vinkhuyzen A-A-E et al. (2011) Common SNPs Explain a Significant Proportion of the 'Missing Heritability' in the Personality Dimensions of Neuroticism and Extraversion. *University of Queensland Working Paper*.

33. 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.

34. Hill W-G, Weir B-S (2011) Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet Res Camb*, 93(1), 47-64.

35. Visscher et al. (2006) Assumption-Free Estimation of Heritability From Genome-Wide Identity-By-Descent Sharing Between Full Siblings. *Pubic Library of Science Genetics* 2(3):e41.

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

37. Lango Allen H et al. (2010) Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 467(7317): 832-838.

38. Cesarini D, Johannesson M, Magnusson P-K, Wallace B (2011) The Behavioral Genetics of Behavioral Anomalies. *Manage Sci* doi:10.1287/mnsc.1110.1329

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

40. Killingsworth M-R , Heckman J-J (1986) Female Labor Supply: A Survey in *Handbook of Labor Economics Vol. 1,* eds Ashenfelter O, Layard R (Amsterdam and NY: North-Holland), pp 103-204.

41. McCarthy M-I et al. (2008) Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nature Review Genetics* 9(5): 356–369.

42. Pearson T, Manolio T-A (2008) How to Interpret a Genome-wide Association Study. *JAMA* 299: 1335-1344.

43. 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.

44. Miller P, Mulvey C, Martin N (2001) Genetic and Environmental Contributions to Educational Attainment in Australia. *Econ Educ Rev* 20(3):211-224.

45. Visscher, P-M (2008) "Sizing up Human Height Variation." *Nat Genet*, **40**(5): 489–90.

46. Garner C (2007) Upward Bias in Odds Ratio Estimates from Genome-Wide Association Studies. *Genet Epidemiol* 31(4): 288–295.

47. Hertwig R, Ortmann A (2001) Experimental Practices in Economics: A Methodological Challenge for Psychologists? *Behav Brain Sci* 24(3): 383-403.

48. 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.

49. Manolio T-A et al (2009) Finding the missing heritability in complex diseases. *Nature* 461(7265): 747–753.

50. Turkheimer E (2000) Three Laws of Behavior Genetics and What They Mean. *Curr Dir Psychol Sci*. 2000 9(5): 160-164.

51. Chabris C-F., et al. (in press) Most Reported Genetic Associations with General Intelligence Are Probably False Positives. *Psych Sci*.

52. Ioannidis, J-P-A (2007) Non-Replication and Inconsistency in the Genome-Wide Association Setting. *Hum Hered*, 64(4), 203–13.

53. de Moor et al. (2010) Meta-analysis of genome-wide association studies for personality. *Mol Psych* doi:10.1038/mp.2010.128

54. Ebstein R-P, Israel S, Chew S-H, Zong S, Knafo A (2010) Genetics of Human Social Behavior. *Neuron* 65(6): 831-844.

55. Hewitt J-K (2011) Editorial Policy on Candidate Gene Association and Candidate Geneby-Environment Interaction Studies of Complex Traits. *Behav Genet* doi: 10.1007/s10519-011-9504-z.

56. Benjamin D-J et al. Genoeconomics (2007) in *Biosocial Surveys*, eds Weinstein M, Vaupel J-W, Wachter K-W (The National Academies Press, Washington), pp 304-335.

57.The Tobacco and Genetics Consortium (2010). Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 42(5): 441-449.

58. 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.