Online Appendix

Genes, Personality, and Political Behavior: A Replication and Extension Using Danish Twins

1 Univariate ACE Model

The univariate ACE model can be represented as:

$$y_{ij} = aA_{ij} + cC_{ij} + eE_{ij} \tag{1}$$

where y is the trait of interest assumed to be expressed as deviations from zero with unit variance, A is the additive genetic component, C is the common environment component, and E is the unique environment component. All three components are assumed to be mutually independent and are each distributed as a standard normal. For a given twin pair in family j:

$$cov(y_{1j}, y_{2j}) = a^2 cov(A_{1j}, A_{2j}) + c^2 cov(C_{1j}, C_{2j}) + e^2 cov(E_{1j}, E_{2j}) \tag{2}$$

Since MZ twins share all of their genes while DZ share 50% on average, $cov(A_{1j}, A_{2j}) = 1$ for MZ twins and $cov(A_{1j}, A_{2j}) = 0.5$ for DZ twins. For both MZ and DZ twins $cov(C_{1j}, C_{2j}) = 1$ and since unique environment is by definition unique to each twin, $cov(E_{1j}, E_{2j}) = 0$. This leaves us with:.

$$cov(y_{1j}, y_{2j})_{MZ} = a^2 + c^2 \tag{3}$$

$$cov(y_{1j}, y_{2j})_{DZ} = 0.5a^2 + c^2 \tag{4}$$

Since A, C, and E are assumed to be independent, the variance for all twins is:

$$var(y_{ij}) = a^2 + c^2 + e^2 \tag{5}$$

where a^2 is the heritability, defined as the proportion of total variation accounted for by additive genetic factors, c^2 is the proportion accounted for by common environmental factors, and e^2 the proportion accounted for by unique environmental factors. The variance-covariance matrix for MZ and DZ twin pairs respectively is then:

$$\mathbf{\Omega_{MZ}} = \begin{bmatrix} a^2 + c^2 + e^2 & a^2 + c^2 \\ a^2 + c^2 & a^2 + c^2 + e^2 \end{bmatrix} \tag{6}$$

$$\mathbf{\Omega_{DZ}} = \begin{bmatrix} a^2 + c^2 + e^2 & 0.5a^2 + c^2 \\ 0.5a^2 + c^2 & a^2 + c^2 + e^2 \end{bmatrix} \tag{7}$$

2 Bivariate ACE Model

The bivariate ACE model can be represented as:

$$z_{ij} = a_{11}A^1_{ij} + c_{11}C^1_{ij} + e_{11}E^1_{ij} \tag{8}$$

$$y_{ij} = a_{21}A^1_{ij} + c_{21}C^1_{ij} + e_{21}E^1_{ij} + a_{22}A^2_{ij} + c_{22}C^2_{ij} + e_{22}E^2_{ij} \tag{9}$$

where z and y are the two traits of interest both assumed to be expressed as deviations from zero with unit variance. Following the same logic as in the univariate model, $cov(A^1_{1j}, A^1_{2j}) = cov(A^2_{1j}, A^2_{2j}) = 1$ for MZ twins and $cov(A^1_{1j}, A^1_{2j}) = cov(A^2_{1j}, A^2_{2j}) = 0.5$ for DZ twins, $cov(C^1_{1j}, C^1_{2j}) = cov(C^2_{1j}, C^2_{2j}) = 1$ and $cov(E^1_{1j}, E^1_{2j}) = cov(E^2_{1j}, E^2_{2j}) = 0$ for both MZ and DZ twins. If we stack z and y for each family such that $\mathbf{Y} = (z_{1j}, y_{1j}, z_{2j}, y_{2j})'$ the variance-covariance matrices become:

$$\mathbf{\Omega_{MZ}} = \begin{bmatrix} \mathbf{A} + \mathbf{C} + \mathbf{E} & \mathbf{A} + \mathbf{C} \\ \mathbf{A} + \mathbf{C} & \mathbf{A} + \mathbf{C} + \mathbf{E} \end{bmatrix} \tag{10}$$

$$\mathbf{\Omega_{DZ}} = \begin{bmatrix} \mathbf{A} + \mathbf{C} + \mathbf{E} & 0.5\mathbf{A} + \mathbf{C} \\ 0.5\mathbf{A} + \mathbf{C} & \mathbf{A} + \mathbf{C} + \mathbf{E} \end{bmatrix} \tag{11}$$

where:

$$\mathbf{A} = \begin{bmatrix} a^2_{11} & a_{11}a_{21} \\ a_{11}a_{21} & a^2_{21} + a^2_{22} \end{bmatrix} \tag{12}$$

with **C** and **E** analogously defined. This model, known as a Cholesky decomposition model (Martin and Eaves, 1977), assumes that the latent factors underlying z (A^1, C^1, and E^1) also influence y, however the latent factors underlying y (A^2, C^2, and E^2) do not affect z.[1]

This model assumes that the latent factors underlying the two traits are uncorrelated with one another across and within individuals. These assumptions are necessary in order for the model to be identified and leaves us with nine parameters to be estimated (a_{11},a_{21},a_{22},c_{11},c_{21},c_{22},e_{11},e_{21},e_{22}). The parameters can be estimated using the following moments from the data:

$$cov(y_{1j}, z_{2j})_{MZ} = a_{11}a_{21} + c_{11}c_{21} \tag{13}$$

$$cov(z_{1j}, z_{2j})_{MZ} = a_{11}^2 + c_{11}^2 \tag{14}$$

$$cov(y_{1j}, y_{2j})_{MZ} = a_{11}^2 + c_{11}^2 + a_{21}^2 + c_{21}^2 \tag{15}$$

$$cov(y_{1j}, z_{2j})_{DZ} = 0.5a_{11}a_{21} + c_{11}c_{21} \tag{16}$$

$$cov(z_{1j}, z_{2j})_{DZ} = 0.5a_{11}^2 + c_{11}^2 \tag{17}$$

$$cov(y_{1j}, y_{2j})_{DZ} = 0.5a_{11}^2 + c_{11}^2 + 0.5a_{21}^2 + c_{21}^2 \tag{18}$$

$$var(z_{ij}) = a_{11}^2 + c_{11}^2 + e_{11}^2 \tag{19}$$

$$var(y_{ij}) = a_{11}^2 + c_{11}^2 + e_{11}^2 + a_{22}^2 + c_{22}^2 + e_{22}^2 \tag{20}$$

$$cov(z_{ij}, y_{ij}) = a_{11}a_{21} + c_{11}c_{21} + e_{11}e_{21} \tag{21}$$

where the last three equations are the population variance and covariances. Since both variables in our bivariate models are ordinal, estimation of the thresholds and variance components is achieved following the same approach described in the univariate case.

The parameter estimates generated by this bivariate model can be used to construct quantities of interest. The genetic correlation (r_g) is obtained by dividing the genetic covari-

ance between the two traits by the square root of the product of their genetic variances:[2]

$$r_g = \frac{a_{11}a_{21}}{\sqrt{a_{11}^2(a_{21}^2 + a_{22}^2)}} \tag{22}$$

The total correlation between two traits (r), based on (21) and (22) is:[3]

$$r = r_g\sqrt{a_{11}^2(a_{21}^2 + a_{22}^2)} + r_c\sqrt{c_{11}^2(c_{21}^2 + c_{22}^2)} + r_e\sqrt{e_{11}^2(e_{21}^2 + e_{22}^2)} \tag{23}$$

where a_{11}^2 and $a_{21}^2 + a_{22}^2$ represent the heritabilities of traits z and y respectively. The product of the square root of the two univariate heritabilities and the genetic correlation is known as *bivariate heritability*. Therefore, the percentage of the total correlation accounted for by genetic factors[4] is:

$$\% \text{ Genetic} = \frac{r_g\sqrt{a_{11}^2(a_{21}^2 + a_{22}^2)}}{r} \tag{24}$$

Notice that (23) implies that if, for example, r_g has an opposite sign to that of r then % Genetic will be negative and consequently % Common Environment and % Unique Environment will add up to $> 100\%$.

3 Equal Environments Assumption

An assumption of the ACE that is the most controversial is the equal environments assumption (EEA). Identification of the univariate and bivariate ACE models requires the assumption that differences in MZ and DZ twins for a particular trait are not due to greater similarity in the exogenous environmental conditions facing MZ twins. Based on this definition, exposure to environmental conditions that are the result of genetic endowments does not violate the EEA. For example, MZ twins may be treated more similarly by others than DZ twins because they look alike. MZ twins may also select into more similar environments due to their genetic similarity.[5] However, the EEA is violated if genetic similarity affects within-pair behavior. This would be the case, for example, if due to genetic similarity an MZ twin became better at playing the piano by observing their sibling's experiences than a DZ twin would (Benjamin et al., 2012). In terms of ACE model estimates, a violation of the EEA would mean that the estimate of genetic influence was overstated and the estimate of common environmental influence was understated.

One way to theoretically avoid relying on the EEA is to analyze twins reared apart since by definition there would be no shared environmental factors. Studies of twins reared apart yield similar heritability estimates for cognitive ability and personality traits (Bouchard et al., 1990) suggesting that the EEA is valid for those traits. Scholars have attempted to test the EEA by comparing the similarity of twins as a function of perceived rather than actual zygosity. The EEA suggests that only actual zygosity should matter, which has been shown to be the case for intelligence, social attitudes, personality, and psychiatric disorders (Scarr, 1968; Scarr and Carter-Saltzman, 1979; Kendler et al., 1993; Xian et al., 2000).

Several studies published in political science journals have also asserted that heritability estimates reported for political orientations are likely inflated by, or completely the result of, EEA violations (Suhay et al., 2007; Beckwith and Morris, 2008; Charney, 2008).[6] Smith et al. (2012) attempted to test whether similar experiences and mutual influence among twin pairs inflated heritability estimates of political attitudes by analyzing whether twins

attended the same classes at school, dressed alike when growing up, shared a bedroom at home, and had the same friends growing up as well as measures of frequency of contact. However, the authors found no evidence that this was the case. Littvay (2012) also failed to find evidence of unequal MZ and DZ environments based on an biometric model that estimated both measured common environment, capturing the effect of similar experiences, as well as residual common environment. Finally, Hatemi et al. (2010), based on an extended family design, found that the influence of family environment on political attitudes among twin pairs was not significantly different from that of non-twin siblings. The authors argued that the extended family design provides a test of the EEA reasoning that if more similar treatment of MZ twins relative to DZ twins resulted in inflated heritability estimates then the same should be true for DZ twins relative to non-twin siblings.

A recently developed method of estimating heritability that avoids the EEA utilizes genotyped single nucleotide polymorphisms (SNPs) to estimate a lower-bound estimate of heritability (Yang et al., 2010, 2011; Visscher et al., 2010). Unlike twin studies, this method relies on individuals who are not in the same extended families and therefore environmental similarity is uncorrelated with genetic relatedness.[7] This technique has been utilized to investigate the genetic architecture of a wide variety of physical and psychological traits.

References

Beckwith, J. and C. A. Morris (2008). Twin studies of political behavior: Untenable assumptions? *Perspectives on Politics 6*(04), 785–791.

Benjamin, D., D. Cesarini, C. Chabris, E. Glaeser, D. Laibson, V. Guonason, T. Harris, L. Launer, S. Purcell, A. Smith, M. Johannesson, P. Magnussen, J. Beauchamp, N. Christakis, C. Atwood, B. Herbert, J. Freese, R. Hauser, T. Hauser, A. Grankvist, C. Hultman, and P. Lichtenstein (2012). The promise and pitfalls of genoeconomics. *Annual Review of Economics 4*, 627–662.

Bouchard, T., D. Lykken, M. McGue, N. Segal, and A. Tellegen (1990). Sources of human psychological differences. *Science*.

Charney, E. (2008). Genes and ideologies. *Perspectives on Politics 6*(2), 299–319.

Hatemi, P., J. Hibbing, S. Medland, M. Keller, J. Alford, K. Smith, N. Martin, and L. Eaves (2010). Not by twins alone: Using extended family design to investigate genetic influence on political beliefs. *American Journal of Political Science 54*(3), 798–814.

Kendler, K. S., M. C. Neale, R. C. Kessler, A. C. Heath, and L. J. Eaves (1993). A test of the equal-environment assumption in twin studies of psychiatric illness. *Behavior Genetics 23*(1), 21–27.

Littvay, L. (2012, Feb). Do heritability estimates of political phenotypes suffer from an equal environment assumption violation? evidence from an empirical study. *Twin Res Hum Genet 15*(1), 6–14.

Martin, N. and L. Eaves (1977). The genetic analysis of covariance structure. *Heredity 38*(1), 79–95.

Scarr, S. (1968). Environmental bias in twin studies. *Eugenics Quarterly 15*(34-40).

Scarr, S. and L. Carter-Saltzman (1979). Twin method: Defense of a critical assumption. *Behavior Genetics 9*(6), 527–542.

Smith, K., J. R. Alford, P. K. Hatemi, L. J. Eaves, C. Funk, and J. R. Hibbing (2012). Biology, ideology, and epistemology: How do we know political attitudes are inherited and why should we care? *American Journal of Political Science 56*(1), 17–33.

Suhay, E., N. Kalmoe, and C. McDermott (2007). Why twin studies are problematic for the study of political ideology: Rethinking are political orientations genetically trasmitted. Unpublished Manuscript.

Visscher, P. M., J. Yang, and M. E. Goddard (2010, December). A commentary on 'common snps explain a large proportion of the heritability for human height' by yang et al. (2010). *Twin Research and Human Genetics 13*(6), 517–524.

Xian, H., J. F. Scherrer, S. A. Eisen, W. R. True, A. C. Heath, J. Goldberg, M. J. Lyons, and M. T. Tsuang (2000, Jul). Self-reported zygosity and the equal-environments assumption for psychiatric disorders in the vietnam era twin registry. *Behav Genet 30*(4), 303–10.

Yang, J., B. Benyamin, B. McEvoy, S. Gordon, A. Henders, D. Nyholt, P. Madden, A. Heath, N. Martin, G. Montgomery, M. Goddard, and P. Visscher (2010). Common snps explain a large proportion of the heritability for human height. *Nature Genetics 42*(7), 565–569.

Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher (2011, January). Gcta: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics 88*(1), 76–82.