Preferences and skills: four studies into unobserved human nature and its implications

Nils Örjan Sandewall

A thesis submitted to the Department of Economics of the London School of Economics and Political Science for the degree of Doctor of Philosophy

February 2009

UMI Number: U615948

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

Dissertation Publishing

UMI U615948
Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

THESES
F
9080

Libra
British Library
and Econom

1203000

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without the prior written consent of the author.

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

Abstract

This thesis contains four studies of generally unobservable skills and preferences of relevance to economic behaviour.

The first chapter examines the validity of the assumption of equal latent ability among monozygotic twins. An influential literature has employed the schooling decisions of twins to estimate returns to schooling. Using a unique dataset including IQ test scores, income and two measures of schooling for 1780 monozygotic twins, within-pair differences in measured IQ are found to be a significant predictor of both income and schooling differences.

In the second chapter, the effect of general cognitive ability on proposer and responder behaviour in the ultimatum game is examined, using a large and representative sample of 895 individuals. No effects are found on proposer behaviour, and only small, but statistically significant, effects are found on responder behaviour.

In the third chapter, a sample of almost 30,000 mono- and dizygotic twins is used to study the heritability of financial risk-taking. Investment decisions made by virtually all Swedish adults regarding mandatory pension savings are taken as a field experiment to infer risk preferences. Standard techniques from

behaviour genetics are used to partition variation in risk-taking into environmental and genetic components. The results suggest that genetic variation is an important source of individual heterogeneity in financial risk-taking and that the frequently reported parent-child associations in attitudes toward risk are, at least in part, genetically mediated.

In the fourth and final chapter, the robustness of recent results indicating a large role for genes in determining variation in the propensity for self-employment is examined empirically using a novel dataset, and a large but imprecisely estimated difference between women and men is found.

Contents

Abstract **3**

Acknowledgements **12**

Introduction **14**

1 The co-twin methodology and returns to schooling - an empirical critique **18**

1.1 Introduction 18

1.2 Previous co-twin studies of the returns to schooling 22

1.3 Empirical framework 24

1.3.1 An augmented co-twin model 24

1.3.2 Two simple tests of the basic twin assumption 26

1.4 Data 27

1.4.1 IQ data 29

1.4.2 Representativeness 30

1.4.3 Comparability with previous studies 32

1.5 Results 34

1.5.1 Two simple tests 34
1.6 Robustness checks 35
1.6.1 Misclassification 36
1.6.2 IQ construction 36
1.6.3 Choice of instrument 37
1.6.4 Full-time threshold 37
1.6.5 Results 38
1.7 Evaluating some additional model restrictions 38
1.7.1 Is IQ a (nearly) perfect measure of ability? 39
1.7.2 Is schooling as familial as ability? 40
1.8 Discussion 41
1.9 Conclusion 43
1.10 Tables 44

2 Smart or selfish? - the role of cognitive behaviour in the ultimatum bargaining game 50

2.1 Introduction 50
2.2 Background 52
2.2.1 The ultimatum bargaining game 52
2.2.2 IQ 56
2.3 Data 58
2.3.1 Participants 58
2.3.2 Ultimatum game test procedure 58
2.3.3 Cognitive ability 59

2.3.4 Representativeness 60

2.4 Results 61

2.4.1 Observed behaviour 61

2.4.2 Regression analysis of the role of IQ 62

2.4.3 Group comparisons 65

2.5 Discussion 67

2.6 Conclusion 70

2.7 Tables 71

2.8 Figures 81

3 Is financial risk-taking behaviour genetically transmitted? 83

3.1 Introduction 83

3.2 Data 87

3.2.1 The Swedish Pension Reform 87

3.2.2 Portfolio Risk Data 88

3.2.3 The Swedish Twin Registry 89

3.3 Method 90

3.4 Results 94

3.4.1 Relationship to Previous Findings 96

3.5 Robustness and Generalizeability 97

3.5.1 Representativeness and Generalizeability 97

3.5.2 Equal Environment Assumption 99

3.5.3 Reciprocal Influences 100

3.5.4 Misclassification and Measurement Error 102

3.6 Heritability and the Biological Basis of Risk-Taking 103

3.7 Conclusion 106

3.8 Tables 108

4 The heritability of entrepreneurship - men are not like women 114

4.1 Introduction 114

4.2 Theoretical framework 117

4.3 Data 120

4.4 Results 123

4.5 Assortative mating 124

4.6 Discussion 125

4.7 Conclusion 130

4.8 Tables 132

Conclusion 141

Bibliography 143

List of Tables

1.1 Sample representativeness with respect to income, education, and IQ 44

1.2 Probit of being above threshold on IQ, age, education (extended main sample) 45

1.3 Comparability with previous literature 46

1.4 Two simple tests of within-pair ability differences 48

1.5 Two simple tests of within-pair ability differences, robustness checks 49

2.1 Comparison of demographic characteristics 71

2.2 Summary statistics of behaviour 72

2.3 Expected payoff from minimum acceptable offer (MAO) 73

2.4 Expected payoff from offer 74

2.5 Responder behaviour - results from regressing MAO and expected payoff from MAO on IQ and covariates 75

2.6 Proposer behaviour - results from regressing offer and expected payoff from offer on IQ and covariates 76

2.7 Results from regressing behaviour on IQ and age - without possible outliers w r t behaviour variables 77

2.8 High-IQ behaviour 78

2.9 Narrowly "selfish rational" behaviour 79

2.10 Probit analysis of "narrowly rational" (SPNE) behaviour 80

3.1 Within-pair correlations 108

3.2 Summary statistics for risk measures 109

3.3 Results of the ACE model 110

3.4 Background variables 111

3.5 Robustness checks of the ACE model 112

4.1 Self-employment in sample and in population 132

4.2 Comparison of demographic characteristics 133

4.3 Within-pair polychoric correlations 134

4.4 Parameter estimates from ACE model - Women 135

4.5 Parameter estimates from ACE model - Men 136

4.6 Confidence intervals of difference in heritability between men and women 137

4.7 Parameter estimates from ACE model - Women. Alternative assumptions on assortative mating 138

4.8 Parameter estimates from ACE model - Men. Alternative assumptions on assortative mating 139

4.9 Confidence intervals of difference in heritability between men and women. Alternative assumptions on assortative mating 140

List of Figures

2.1 Distribution of offers 81

2.2 Distribution of minimum acceptable offers 82

Acknowledgements

To my supervisor, Professor Steve Pischke, for great support in writing this thesis - I owe him a lot for his accurate comments and endless patience with the frequent stream of revised chapters.

To David Cesarini who inspired me to return to academia and without whom this project could never have been carried out.

To peers, and in particular - Erik Lindqvist for his sanity and crisp research mentality, Björn Wallace for his devotion to the world of research, and Tino Sanandaji for his enthusiasm and focus on the facts.

To faculty - Professors Johannesson, Van Reenen, and Henrekson, all of whom have contributed in vital ways.

To seminar participants at the London School of Economics, the Research Institute for Industrial Economics, and Stockholm University; and to Jonas Vlachos for encouraging comments way back in 2002, when they were much needed.

To NERA and my colleagues there, for support and a positive attitude.

To Tekn Dr Markus Wallenbergs Stiftelse, Hedelius-Wallander Stiftelse, Lars-Åke Nilssons Stipendiefond, NERA, and ESRC for financial support.

Most of all, to my family - To Davinia, my wonderful wife who has supported me along the way and who has endured tribulations too many to mention over the last five years; to Erik, Margareta, and Anton, my parents and brother, for believing in me; and to my daughter Emilia, whose smiling face has been a constant light in this tunnel.

Introduction

Skills and preferences

This thesis examines certain generally unobservable skills and preferences: how they interrelate with commonly observed economic outcomes and with each other, and from where they might derive.

In each of the first two chapters, I take IQ as a measure of some underlying property of the individual, although the interpretation of IQ differs substantially across these two chapters. In chapter one, the property of interest is latent wage-earning ability and specifically the component of it which carries any correlation with schooling. In other words, ability as in "the ability bias of returns to schooling" is under examination. IQ is taken as one possibly relevant measure of this underlying skill. However, IQ may also reflect unrelated differences in preferences for pursuing higher education, or reflect differences in wage-earning capacity which is unrelated to schooling. Two different sets of assumptions regarding the skills and preferences composing the IQ measure are proposed, and the parameter of interest is tested under each set.

In the second chapter, the main property under analysis is rationality, or the skill of making decisions which maximise the likelihood of attaining one's objectives, and IQ is taken as a plausible, although imperfect, measurement of this. Rationality is proposed as an alternative to social preferences in explaining commonly observed behavioural "anomalies" in a standard experiment, the ultimatum bargaining game, and this proposition is then evaluated empirically.

In the last two chapters, the properties under examination are risk preferences and propensity for entrepreneurial activity. In these cases, the distinction between skills and preferences is not a major turning point of the analysis. Instead, the analysis focuses on the origin of these properties, and specifically on the genetic basis of observed variation in the sample.

Chapters

The first chapter is a methodological contribution to the literature on estimating the returns to schooling. An influential literature has employed the schooling decisions of monozygotic twins to mitigate problems of unobserved ability. If ability is equal within twin pairs, then the within-pair estimator identifies returns to schooling in spite of heterogeneous ability in the population. In the first chapter, an empirical assessment of the equal abilities assumption is provided using a unique dataset including IQ test scores, income, and two measures of schooling for 1780 monozygotic twins. Within-pair differences in measured IQ are found to be a significant predictor of income and schooling differences, and inclusion of IQ in the wage equation is shown to considerably reduce the estimated within-pair effect of schooling on income. These findings cast doubts on the validity of using the co-twin methodology when estimating returns to schooling and challenge the usefulness of the empirical results derived from this literature.

In the second chapter, I examine the effect of IQ on proposer and responder behaviour in the ultimatum game, using a large and representative sample of 895 individuals. I find no effects on proposer behaviour and statistically significant, but small, effects on responder behaviour. These findings remain when controlling for a range of demographic variables, and when juxtaposing individuals from the top 5% of the population IQ distribution, and the rest of the sample. Furthermore, the 2% of subjects who played the subgame perfect Nash equilibrium of the standard ultimatum game model did not differ significantly on neither IQ nor any of the demographic variables examined.

The third chapter reports results from a study of the heritability of financial risk-taking. This chapter was coauthored with David Cesarini (MIT), Magnus Johannesson (Stockholm School of Economics), and Björn Wallace (Stockholm School of Economics), my share being approximately 35%, apportioned evenly across all parts of the work. This study is based on a sample of almost 30,000 mono- and dizygotic twins. Following a major pension reform in the fall of the year 2000, virtually all Swedish adults had to simultaneously make a financial decision affecting post-retirement wealth. We take this event, known as the Big Bang of the Swedish financial sector, as a field experiment to infer risk preferences. We then apply standard techniques from behaviour genetics to partition variation in risk-taking into environmental and genetic components. Our findings suggest that genetic variation is an important source of individual heterogeneity in financial risk-taking and that the frequently reported parent-child associations in attitudes toward risk are, at least in part, genetically mediated.

Finally, in the fourth chapter, the heritability of entrepreneurship is examined. Recent results from a sample of female twins indicate a large role for genes in determining variation in the propensity for self-employment. Using a twice as large Swedish sample of twins, almost 4000 pairs in total, of which more than one third are men, I estimate heritability in self-employment for men and women separately. I am able to replicate previous results for women, but for men I find low or no heritability. Instead, my estimates suggest that common environment is a much more important source of variation in entrepreneurship for men.

Chapter 1

The co-twin methodology and returns to schooling - an empirical critique

1.1 Introduction

There is a widely shared view that estimates of the marginal returns to schooling will be biased unless proper account is taken of heterogeneities in latent ability. If the propensity to invest in further years of education is also directly related in a positive way to the ability to earn wages, then this will cause an upward bias in estimates of the effect of an additional year of schooling on wages (see *e.g.* Card (1999)). A number of approaches to removing this endogeneity have been proposed. One influential strand of the literature has exploited within-family variation in general, and variation within monozygotic (MZ) twin pairs in particular, to control for unobserved ability. Under the identifying assumption that ability is common among siblings of the type at hand, this allows for consistent estimates, as long as problems of measurement errors in the schooling variable can be dealt with adequately. Especially with regards to MZ twins, the attraction of the equal ability assumption is easy to understand. MZ twins are the result of a fertilized egg splitting in two shortly after conception, resulting (after about 38 weeks) in two genetically identical individuals. Furthermore, MZ twins (or "identical" twins, as they are often referred to) are typically raised by the same parents, go to the same school, and are influenced by the same peer groups when growing up.

The idea that the latent wage earning ability of two individuals in such a pair would be virtually identical is not hard to accept, a priori. However, identical ability begs the question of what causes observed within-pair differences in schooling, as standard optimising models predict that two identically able individuals would choose the same level of schooling (as in *e.g.* Ashenfelter and Rouse (1998)). Any observed variation in schooling must then be explained by "optimising errors", or differences in preferences for schooling which do not affect wage earning ability. Hence, it is assumed that differences in schooling across the population are caused by ability differences, but that this is not true within twin pairs.

A further potential problem with the co-twin methodology was demonstrated by Griliches (1979); Although twins may have very similar levels of ability, the observed similarities in years of schooling and income are also large. Therefore, even though within-pair differences are purged from most of the heterogeneities in ability, they also lack most of the useful variation in schooling and income.

Griliches (1979) noted that when the degree of twin similarity is the same for ability and for schooling, first-differencing contributes nothing in terms of removing ability bias. This critique has been further developed by Bound and Solon (1999), who also point out that *a priori* the relationship between the degrees of similarity in ability and schooling, respectively, is not clear.

Finally, whereas the two above lines of critiques question the benefit from using twins data, the potentially most serious concern raised relates to measurement error in schooling and whether it can properly be instrumented for. As was acknowledged by one of the first authors to apply this methodology (Taubman (1976)), differencing within pairs reduces the amount of true variation without reducing measurement error to the same extent, and hence serves to amplify the effects of mismeasurements in schooling. Furthermore, even with valid instruments for number of years spent in an educational facility, this quantity may not perfectly reflect true education, a distinction pointed out at least as early as in Griliches (1977). Studies of such heterogeneities in the production function for human capital abound, see *e.g.* Sacerdote (2001), on peer group effects and Rivkin *et al.* (2005) on teacher quality. As the data of this study present limited opportunity to examine the issue of mismeasured schooling, the twin methodology will be given the benefit of the doubt; The assumption of perfectly instrumented schooling will be maintained, and focus directed towards the source of the alleged benefits from using twins data - the equal or virtually equal ability within twin pairs.

To address the question of ability differences, official registry data on income and schooling is combined with self-reported data on schooling from two surveys aimed at all Swedish twins born 1950-1975. Uniquely, this dataset is then linked to IQ test scores from compulsory military conscription, performed before or around the end of secondary school. Using this dataset, the usefulness of the assumption of equal ability within pairs is investigated. It is found that within-pair differences in IQ are significantly associated with income even when accounting for differences in schooling, that within-pair differences in IQ have a statistically and economically significant effect on within-pair differences in schooling, and that inclusion of IQ reduces within-pair estimates of returns to schooling by about 15% across various specifications and variable definitions. These results cast doubts on the co-twin approach to estimating the returns to schooling.

The structure of this chapters is as follows: In the next section, empirical findings from the co-twin literature on estimating the returns to schooling are reviewed briefly. After this, in Section 1.3, an empirical framework for examining the equal abilities assumption is outlined. Section 1.4 contains a presentation of the data, followed by results from the main analysis in Section 1.5, and robustness checks in Section 1.6. The consistency of the data with two additional restrictions is considered and rejected in Section 1.7. A discussion of the main findings is provided in Section 1.8, after which Section 1.9 concludes.

1.2 Previous co-twin studies of the returns to schooling

Jere Behrman and Paul Taubman (Behrman and Taubman (1976), Taubman (1976)) pioneered the use of data on twins for studying the returns to schooling. Examining within-pair differences in annual earnings and schooling among male twin veterans in the NAS-NRC dataset, Taubman (1976) found evidence of substantial upward ability bias in traditional cross-sectional estimates of the returns to schooling. Taubman's (1976) estimates decreased from 8.8% to 4.8% when moving from regression on the cross-section to within-pair estimation, despite correcting for an assumed 10% measurement error in the schooling data. The results in Behrman and Taubman (1976) imply similarly that standard OLS estimates are considerably upward biased.

The co-twin approach experienced a revival in the 1990's, following the innovation by Ashenfelter and Krueger (1994) to collect data on both own schooling and co-twin's schooling from each individual in the sample. Having two measures of schooling, they then use the first-difference of schooling reported by one member of a pair as an instrument for the first-difference reported by the other member. If measurement errors are uncorrelated, this allows for a correction of the problem of measurement error in the schooling variable. Under the crucial assumption of equal abilities within pairs, their approach thus provides a consistent estimate of the returns from schooling.

Ashenfelter and Krueger's (1994) within-pair IV estimates were, surprisingly enough, considerably higher than standard least squares estimates on the cross-section. However, later studies strongly suggest that these initial results were due to an anomalous sample, as analyses of extensions of this sample produced within-pair IV estimates that were not higher than conventional cross-sectional estimates (Ashenfelter and Rouse (1998), Rouse (1999)). These later findings are consistent with most other co-twin studies (Miller *et al.* (1995), Isacsson (1999), Behrman and Rosenzweig (1999), Bonjour *et al.* (2003)), who likewise find only a small upwards ability bias.[1]

Two recent additions to the co-twin literature are Isacsson (2004) and Zhang *et al.* (2007). Isacsson (2004) has the benefit of working with a representative dataset comprising education and income data for a very large number of Swedish monozygotic twins born 1926-1958, 2609 pairs in total, and is therefore able to provide precise estimates of non-linearities in returns to schooling, and to allow for non-classical errors in the measurements of schooling. Zhang *et al.* (2007) analyse a dataset of 914 pairs of Chinese monozygotic twins and find that the returns to schooling during the Cultural Revolution (defined as 1966-1976 in their study) was roughly the same as that of later cohorts. In both these studies, the implied ability bias in cross-sectional estimates is positive.

[1]For two older summaries of the literature on returns to schooling using twins data, see Card (1999) and Behrman and Rosenzweig (1999).

1.3 Empirical framework

1.3.1 An augmented co-twin model

Consider the following simple model of wage determination, drawing on *e.g.* Card (1999):

$$y_{ij} = \alpha_y + \beta S_{ij} + \gamma A_{ij} + u_{ij} \tag{1.1}$$

Where y_{ij}, S_{ij} and A_{ij} are income in natural logarithms, years of schooling, and ability, respectively, for individual i of twin pair j, and where the ordering of both individuals and pairs is random. Returns to schooling, β, and partial returns to ability, γ, are assumed to be equal across individuals. Let latent ability, A, be defined widely enough to allow S and u to be independent, and be measured in standard deviations about the population mean. Finally, α_y varies with a quadratic in the age of the individual, to capture experience and cohort-specific effects. Furthermore, assume the following causal model of schooling:

$$S_{ij} = \alpha_S + \delta A_{ij} + \epsilon_{ij} \tag{1.2}$$

Where ϵ is a summary measure of all determinants of schooling which are exogenous to the unobservables of the wage equation. Extend this exogeneity to apply across twins within a pair, so that $Corr(A_{ij}, \epsilon_{kj}) = 0$ and $Corr(u_{ij}, \epsilon_{kj}) = 0$, $\forall\, i, j$. Specify the sign of ability such that $\delta > 0$. To capture cohort-specific effects, the intercept again varies with a quadratic function of age.

Let the ability of a twin be statistically related to the ability of his co-twin in the following manner:

$$A_{1j} = \phi A_{2j} + \alpha_{1j} \tag{1.3}$$

Here, ϕ is the correlation between the abilities each twin and his co-twin, and α_{1j} is uncorrelated with A_{2j} by construction. Equivalently, ϕ is the share of variance in ability explained by a variance factor common to both twins. Furthermore, assume that differences in ability within pairs are independent of all other errors (u, ϵ, and τ (below))

The main identifying assumption of the literature on estimating the returns to schooling using variation within twin pairs, is that twins have identical latent abilities such that $A_{1j} = A_{2j}$. In the above framework, this translates to assuming $\phi = 1$, which in turn implies $Var(\alpha) = 0$ due to the random ordering of twins. Under $\phi = 1$, consistent estimates of β can be obtained by estimating the model in first-differences:

$$\Delta y_j = \beta_{FD}\Delta S_j + \gamma_{FD}\Delta A_j + \Delta u_j \tag{1.4}$$

Where $\Delta y_j \equiv y_{1j} - y_{2j}$ and similarly for the explanatory variables. Since ΔA_j is a zero vector under the standard twin assumption, the within-pair difference in income can simply be regressed on the within-pair difference in schooling;

$$\Delta y_j = \beta_{FD}^{-}\Delta S_j + \Delta u_j^{-} \tag{1.5}$$

This is the basic idea behind all within-pair estimators in the literature. The aim of this study is to determine whether $\phi = 1$. For this purpose, consider IQ measured at around the age of 18, and specify its relationship with ability as follows:

$$T_{ij} = \pi A_{ij} + \tau_{ij} \tag{1.6}$$

Where τ_{ij} is independent of A_{ij}. Let T_{ij} be measured in standard deviations about the population mean, and assume $\pi > 0$.

Finally, let y_1 refer to own income, as opposed to y_2 for co-twin's income, and similarly for S, A, T, u, ϵ, and τ, so that $(y_1)_{ij} = (y_2)_{kj}; \forall i \neq k$. When not specified, as above, y refers to own income, y_1.

1.3.2 Two simple tests of the basic twin assumption

Auxiliary assumptions A

Assume $\sigma_{u_1\tau_1} = \sigma_{u_1\tau_2} = 0$

Estimate:

$$\Delta y_j = \beta \Delta S_j + \lambda_1 \Delta T_j + \Delta u_j^* \tag{1.7}$$

Where the error term is:

$$\Delta u_j^* = -\lambda_1(\Delta T_j - \Delta A_j) + \Delta u_j \tag{1.8}$$

If $\phi = 1$, then $\Delta A_j = 0$ and $\Delta T_j = \Delta \tau_j$, and consequently $\lambda_1 = 0$. Furthermore, β and λ_1 are consistently estimated since $\lambda_1(\Delta T_j - \Delta A_j) = 0$, and hence independent of ΔS_j and of ΔT_j. The distribution of $\hat{\lambda}_1$ is different under the null and the alternative hypothesis. It follows that $\hat{\lambda}_1$ is a valid test statistic for the null hypothesis that $\phi = 1$. Measurement error in schooling can be dealt with using an alternative measure of schooling, the approach championed in this literature by Ashenfelter and Krueger (1994), assuming the necessary correlation and exogeneity conditions hold.

Auxiliary assumptions B

Alternatively, assume $\sigma_{\epsilon_1\tau_1} = \sigma_{\epsilon_1\tau_2} = 0$, and relax the above assumptions on $\sigma_{u_1\tau_1}$ and $\sigma_{u_1\tau_2}$. Let λ_2 be defined by the following estimating equation:

$$\Delta S_j = \lambda_2 \Delta T_j + \Delta\epsilon_j^* \tag{1.9}$$

Where, analogously, the error term is:

$$\Delta\epsilon_j^* = -\lambda_2(\Delta T_j - \Delta A_j) + \Delta\epsilon_j \tag{1.10}$$

If $\phi = 1$, then $\Delta A_j = 0$ and $\Delta T_j = \Delta\tau_j$, and consequently $\lambda_2 = 0$. Furthermore, λ_2 is consistently estimated since $\lambda_2(\Delta T_j - \Delta A_j) = 0$, and hence independent of ΔT_j. The distribution of $\hat{\lambda}_2$ is different under the null and the alternative hypothesis. It follows that $\hat{\lambda}_2$ is a valid test statistic for the null hypothesis that $\phi = 1$, under this alternative restriction on the error terms.

1.4 Data

The dataset links information from the Swedish Twin Registry with data from Statistics Sweden and Swedish enlistment records. The Swedish twin registry contains virtually all twins born in Sweden from 1926 onwards, and is kept mainly for the purpose of performing epidemiological studies (see Lichtenstein (2006) for a description of the Swedish twin registry). The survey data used in this chapter was collected in 1998-2002 (the SALT survey) from twins born 1950-1958 (the first cohort), and in 2005-2006 (the STAGE survey) from twins born in 1959-1975 (the second cohort). Response rates were 74% and 60%, respectively.

Only data on monozygotic twins (about $1/4^{th}$ of the sample) is used, where zygosity has been determined by the Swedish Twin Registry using a battery of questions relating to physical similarity. The validity of this method of determining zygosity has been repeatedly estimated to be 95-98% (Lichtenstein *et al.* (2002)).

The data contains two measures of educational achievement. One is a self-reported measure from the survey data collected by the Swedish Twin Registry. The other is based on administrative data from 2005. The self-reported data consists, for the early cohort, of an indicator of highest attained qualification, and for the late cohort, of total years of schooling at the different levels of the education system. For the early cohort, years of schooling are assigned based on the standard years of schooling associated with the degree in question. The administrative data contains highest degree attained. Years of schooling based on the survey data are used as the explanatory education variable, with degree dummies based on administrative sources as instruments.

Data on income consists of yearly taxable earnings in 2005 as reported by employers to the tax authorities. In the main specification, only pairs where both twins in a pair had earnings exceeding 70,000 Swedish krona (approximately

£5000) are included, in an attempt to capture only individuals working full-time so that income more or less corresponds to hourly earnings. The practice of either excluding data not corresponding to full-time work or using information on hourly wages is followed by practically all previous studies of the returns to schooling using twins back to at least Ashenfelter and Krueger (1994)(Ashenfelter and Krueger (1994), Ashenfelter and Rouse (1998), Behrman and Rosenzweig (1999), Bonjour *et al.* (2003), Isacsson(1999), Isacsson(2004), Miller *et al.* (1995), Rouse (1999), Zhang *et al.* (2007)).[2] As the data includes IQ measures for individuals born starting in 1950, income is measured at the ages 30 through 55, depending on birth year.

1.4.1 IQ data

Data on IQ is based on enlistment records. Virtually all Swedish men in the age group underwent aptitude testing for compulsory military service at around age 18, and the dataset of this chapter contains results on the four component tests of the intelligence test used by the Enlistment Agency. As the normal school starting age in Sweden is seven, the average individual in the main sample would have taken the test at least two years prior to finishing education. The older part of the sample was tested using Enlistment battery 67 (EB67) which came into use in 1967, and the younger part received a similar test, Enlistment battery 80 (EB80). Both tests were designed to measure a general ability factor based on spatial intelligence, verbal ability, logical reasoning, and technical comprehension. Test scores from

[2] It can also be noted that due to the logarithmic transformation, some outliers in the full dataset are more than 10 standard deviations lower than the average.

earlier versions of the Enlistment battery test (such as EB44 and EB49) had been explicitly interpreted as IQ scores, and although this interpretation was not formally carried over to the EB67 and EB80, their component tests were nevertheless very close to those of a standard IQ test and provide a good measure of general intelligence (Carlstedt (2000)). Test scores are normalised by year using all observations in the dataset for which there are test scores, and the sum across test scores is then used as the raw IQ measure[3]. This raw measure is then normalised against all observations in the dataset, to allow for an approximation of population standard deviations to be used as the metric for IQ.

Using IQ test scores which were gathered not in a school environment, but under the considerably different conditions of military conscription, reduces the risk that the test scores pick up factors related to, *e.g.*, a general affinity with school-like tests that yet do not translate into wage earning capacity. Using the terminology of the empirical framework outlined above, the risk that $\sigma_{\epsilon_1 \tau_1} \neq 0$ is reduced.

1.4.2 Representativeness

In Table 1.1, the main sample is compared to the national average with regards to income, education, and marital status. Income is slightly higher, although by less than one tenth, and education is higher by 0.8 years[4]. IQ test results are slightly higher than the average for the norm group, *i.e.* all 12366 twins in the

[3] Assigning equal weights to each sub-test is in accordance with the standard practice of the Swedish Armed Forces.

[4] This figure differs from Table 1.1 due to rounding.

relevant cohorts who answered the SALT or STAGE survey and for whom there is IQ data, by about $1/8^{th}$ of a standard deviation.

The total sample size was determined as follows: Out of the 31824 respondents to the STAGE and SALT surveys in our cohorts, 3522 were male monozygotic twins of which 2753 had data on education from both administrative and survey data. Of these, 2353 had non-missing income, and 2288 had an income above 70000 SEK, the cut-off used to eliminate observations whose income unambiguously did not derive from full-time employment. Among these, 2129 individuals had valid IQ test scores from enlistment data[5]. Finally, 1780 of these observations were from complete pairs of twins, *i.e.* where the co-twin was also in the sample.

To evaluate whether using an income threshold to exclude observations clearly not in full-time employment results in a non-representative sample, an extended sample was created using the same procedure as for the main sample, but without any income threshold. Probit regressions were then run with the full-time proxy as the dependent variable, and various combinations of IQ, education and age as independent variables. The results are presented in Table 1.2. Only age has a statistically significant (positive) effect, with an average marginal effect of 0.001, implying that an age increase of ten years corresponds to an increase in the probability of being above the threshold of 1% point. The coefficients on IQ and education are small and statistically insignificant, regardless of whether included together or separately.

[5]For individuals born in two out of the 26 years of birth in the sample, 1950 and 1960, the IQ data contains only about 1/10th of the expected number of observations.

1.4.3 Comparability with previous studies

The main purpose of this study is to generalise not from a sample of twins to a population of non-twins, but rather from one sample of twins to other samples of twins. Therefore, it is important to know how representative my dataset is of the datasets of twins used hitherto. Table 1.3 compares parameters from my dataset to parameters reported previously in the literature.

The first two parameters concern similarity between twins. In my data, measured years of schooling correlate 0.73 between a twin and his co-twin, a figure in line with what has been reported in the literature. Furthermore, results on test scores correlate 0.82, which again is a standard degree of similarity.

The following two parameters concern the structure of the measurement errors in reported years of schooling. In my sample, the reliability ratio[6] is 0.88, which is very similar to those reported in previous twin studies. The reliability ratio of the within-pair differences is 0.65, which is closer to the lower than to the higher estimates reported in Ashenfelter and Krueger (1994) and Ashenfelter and Rouse (1998). The observed within-pair reliability ratio in my data is also close to that expected based on the cross-sectional reliability ratio and the twin correlation in schooling, as reported above. If all measurement errors are classical, the thus imputed within-pair reliability ratio would be 0.58.[7]. Note also that the cross-sectional reliability ratio of 0.88 implies, under classical errors, a within-pair

[6]With classical measurement errors, the reliability ratio is the square root of the R^2 from a regression of measured years of schooling on its instruments. If there is only one instrument, this is equivalent to the correlation, as stated in the table.

[7]The imputed within-pair reliability ratio is $(r - Corr(S_1, S_2))/(1 - Corr(S_1, S_2))$, where r is the reliability ratio based on cross-sectional measures.

correlation in schooling of 0.82 (0.73/0.88) when correcting for measurement errors. Recall that as shown by Griliches (1979), co-twin estimators are less biased than cross-sectional estimators if and only if ϕ is greater than the similarity in schooling, *i.e.* in this dataset 0.82.

The final four parameters concern impacts on wages (in logarithms), and as such we would expect them to vary depending on institutional factors in the countries where they are measured. The first is a simple cross-sectional estimate of the returns to schooling, when correcting for measurement error in years to schooling. My estimates are slightly lower to those found in studies from US and UK, but slightly higher than those of Isacsson (1999) using Swedish twins. However, Isacsson's (1999) sample includes both men and women, whereas my estimates are for men only. As is clear from the following parameter, my data yields larger differences between within-pair estimates and cross-sectional estimates than what is commonly found in twin studies. Notice however that the result from Isacsson (1999) was constructed using an imputed within-pair measurement error, and as such is not strictly comparable to the other figures which apply instrumental variables techniques to correct for measurement error.

The final two parameters concern the relationship of IQ with labour market outcomes, and are not specifically based on data on twins. The total predictive effect of IQ on income (in logarithms) in my data is 0.16, *i.e.* an increase in IQ of one standard deviation corresponds to an increase in income by about 16%. The rough estimate reported by Bowles and Gintis (2002) based on a meta-study of 24 studies on US data is 0.266. This discrepancy corresponds reasonably to differences in income dispersion between US and Sweden, as reported in Gottschalk and Smeeding (1997). Finally, measured schooling correlates about 0.51 with measured IQ, a figure roughly in line with the average of 0.55 reported by Neisser *et al.* (1996) in an authoritative report on the state of intelligence research. It should be noted that the latter figure is based on IQ test scores from early years, mainly primary school. The fact that the correlation with schooling is lower in my data suggests that simultaneity in test scores is not a major concern.

1.5 Results

1.5.1 Two simple tests

The partial effect of IQ in the within-pair wage equation (Auxiliary assumptions A)

Columns 1-2 of Table 1.4presents results from within-pair regressions of income on schooling. The coefficient on IQ in the second column demonstrates that within-pair differences in IQ have a direct effect on income differences, and that this effect is statistically significant and large. The magnitude of the coefficient implies that a twin with an IQ one population standard deviation higher than his co-twin, has an income which is on average 7.4% higher than his co-twin, when controlling for schooling. The coefficient on schooling drops from 3.4% to 2.9%, or about $1/7^{th}$.[8] If we believe the restrictions underpinning this test ($\sigma_{u_1\tau_1} = \sigma_{u_1\tau_2} = 0$), we must

[8]It should be noted that since T_{ij} is an imperfect measure of ability, the estimated returns to schooling are biased and inconsistent when $\phi \neq 1$, i.e. when the equal ability assumption is violated.

reject $\phi = 1$.

The impact of IQ in the within-pair schooling equation (Auxiliary assumptions B)

Column 3 of Table 1.4 contains the results from the second simple test of $\phi = 1$, based instead on $\sigma_{\epsilon_1 \tau_1} = \sigma_{\epsilon_1 \tau_2} = 0$. The estimated within-pair effect of IQ is statistically significant and large, with a difference of one population standard deviation corresponding to a difference of 0.52 years of schooling[9]. Under these alternative test assumptions as well, $\phi = 1$ is rejected.

Conclusion

If $\phi = 1$, the data is only consistent with the model if differences within twin pairs in the errors in test scores, $\Delta\tau$, are related both to $\Delta\epsilon$ and to Δu. In other words, under the standard twin assumption of equal ability, differences in test scores, ΔT, must be correlated both with $\Delta\epsilon$ (unobservable preferences for schooling which are not directly related to wage earning capacity), and with Δu (unobservable capacity to earn wages which is not related to schooling), yet be uncorrelated with ΔA (actual unobservable ability).

1.6 Robustness checks

There are a number of legitimate concerns which may be raised with regards to the above findings. In this section, four such issues are presented briefly in turn,

[9]This can be compared with the sample standard deviation of schooling of 2.6 years.

along with some efforts taken to examine the sensitivity of the above findings with regards to these issues. This is followed by a summary of the results of these robustness checks.

1.6.1 Misclassification

Some of the twins in the sample may have been misclassified as monozygotic twins despite being in fact dizygotic twins. If ability differences are for some reason relatively less familial (*i.e.*, compared to the family share of variance of the exogenous determinant of schooling) in dizygotic twins, this will cause the above findings to be overstated. To examine this issue, the 5% of pairs which were the most dissimilar with respect to IQ were dropped and the main equations were re-estimated. This is a conservative test in that no more than 2-5% of monozygotic twins are normally misclassified as dizygotic using the type of classification algorithm employed by the Swedish Twin Registry (Lichtenstein *et al.* (2002)). It should also be noted that all twin studies referred to in the literature review above have employed similar classification algorithms as does the Swedish Twin Registry.

1.6.2 IQ construction

To examine the sensitivity of my findings to variations in the construction of the aggregate test score, a so called factor "g", *i.e.* the first principal component, was calculated from the four subtests of the military IQ test. This measure was standardised by year against all twins for whom there was data on IQ, and used as an alternative measure of IQ.

1.6.3 Choice of instrument

As a further robustness check, the roles of instrument and regressor were reversed for the two sources of schooling data. As the administrative data, which were used as instruments in the main analysis, consist of dummy variables for highest degree attained, they were converted into years of schooling using population averages estimated by Isacsson (2004)[10].

1.6.4 Full-time threshold

Finally, the sensitivity of the main results to variations in the threshold on yearly earnings was examined, by applying alternative thresholds of 40,000 and 180,000 Swedish krona (about £3500 and £15000, respectively). Regarding the lower threshold, it should be noted that it corresponds to a full-time hourly wage of about £1.50, *i.e.* impossibly low in the context of the industrialised world. Furthermore, because of the logarithmic conversion of wages, the 18 observations below the lower threshold are between 4 and 10 standard deviations away from the mean (in a sample of around 2000). The lower threshold is indeed very low for the purposes of approximating a full-time proxy.

[10]Isacsson (2004) examined a representative sample with high quality data on years of schooling and regressed this on the same type of administrative data that are used in this paper.

1.6.5 Results

Table 1.5 presents results from the two simple tests of the equal ability assumption, under the above five alternative samples and variable specifications. In all cases, the coefficient on IQ, T_{ij}, is statistically and economically significant in both the wage equation (Column 2 of estimated coefficients) and the schooling equation (Column 3 of estimated coefficients). Exclusion of the 5% of pairs where twins are the most dissimilar with respect to IQ is the modification which affects the results the most, yielding the highest estimated effect of IQ on income (0.100), and the lowest estimated effect of IQ on schooling (0.322). However, neither of these results is far from the other four robustness outcomes or the main results. Excluding the 5% most dissimilar pairs also has as its effect that the coefficient on schooling in the wage equation, when including IQ, becomes borderline statistically insignificant although its magnitude remains in line with the other cases.

1.7 Evaluating some additional model restrictions

This section expands on the previous analysis by considering whether there are additional restrictions on the variance-covariance matrix of the errors which are consistent with the data but which have not been imposed thus far.

1.7.1 Is IQ a (nearly) perfect measure of ability?

The precision with which the test score, T, proxies for ability, A, can be crudely evaluated by considering whether the variance of τ is zero:

$$\sigma_\tau^2 = 0 \tag{1.11}$$

A simple way of testing this is to consider the best linear within-pair predictor of schooling using test scores, as presented in Table 1.4[11]:

$$\Delta S_j = \lambda_3 \Delta T_j + \Delta\epsilon_j^* \tag{1.12}$$

Where, by construction, $E(\Delta T_j, \Delta\epsilon_j^*) = 0$.

If $\sigma_\tau^2 = 0$, then $\lambda_3 = \delta$, and an efficient estimator of λ_3 is given by the best cross-sectional predictor of schooling:

$$S_{ij} = \alpha_S + \lambda_4 T_{ij} + \epsilon_j^* \tag{1.13}$$

Hence, if $\sigma_\tau^2 = 0$, then $\lambda_3 = \lambda_4$, where λ_4 is efficient. Under the alternative hypothesis that $\sigma_\tau^2 \neq 0$, only λ_3 is consistent. Applying random effects GLS, $\hat{\lambda}_4 = 1.24$ (0.06) (standard error within parenthesis) to be compared to $\hat{\lambda}_3$ which is 0.52 (0.11)[12].The null hypothesis of $\sigma_\tau^2 = 0$ can be rejected at the 0.1% level

[11]Note that although their estimating equations are identical, λ_3 is not in general equal to λ_2, since the maintained assumptions on the errors when defining λ_2 are more strict. λ_3, being only a linear predictor, is unbiased and consistent regardless of the assumptions on the error terms.

[12]This is numerically identical to $\hat{\lambda}_2$ as reported in Table 1.3 although the interpretations are different.

using a standard Hausman (1978) test.

1.7.2 Is schooling as familial as ability?

As pointed out earlier, Griliches (1979) established that if the proportion of the variance in ability which is explained by a common family component is identical to the degree of within-pair similarity in the non-ability determinants of schooling (ϵ), then the within-pair estimator of returns to schooling is equally biased as the cross-sectional estimator. In the above presented framework, this condition amounts to the restriction that:

$$\phi = Corr(\epsilon_1, \epsilon_2) \tag{1.14}$$

In analogy with the previous section, consider the best linear within-pair predictor of income using schooling:

$$\Delta y_j = \lambda_5 \Delta S_j + \Delta u_j^* \tag{1.15}$$

Where, by construction, $E(\Delta S_j, \Delta u_j^*)$.

If $\phi = \sigma_{\epsilon_1 \epsilon_2}$, then $\lambda_5 = (\beta + \gamma r)$, where r is the coefficient on schooling when regressing ability on schooling and a quadratic in age. In other words, as an indicator of β, λ_5 is equally biased as the standard cross-sectional OLS estimator of income on schooling in the cross-section. Consequently, an efficient estimator of λ_5 is given by the best cross-sectional predictor of income:

$$y_{ij} = \alpha_y + \lambda_6 S_{ij} + u_{ij}^* \tag{1.16}$$

Hence, if $\phi = \sigma_{\epsilon_1\epsilon_2}$, then $\lambda_5 = \lambda_6$, where λ_6 is efficient. Under the alternative hypothesis that $\phi \neq \sigma_{\epsilon_1\epsilon_2}$, only λ_5 is consistent.

Using a random effects IV estimator to account for measurement error in schooling yields $\hat{\lambda}_6 = 0.068$ (0.005)[13], to be compared with the fixed effects estimator $\hat{\lambda}_5$, which as reported in Table 1.4 is 0.034 (0.012). The null hypothesis of $\phi = \sigma_{\epsilon_1\epsilon_2}$ can be rejected at the 0.2% level using a standard Hausman (1978) test.

1.8 Discussion

The main result of the previous sections is that the assumption of equal ability within pairs of monozygotic twins is violated in my sample. There are indications to this effect even in the previous co-twin literature on returns to schooling. For example, Ashenfelter and Rouse (1998) report that for pairs where two twins had obtained different levels of schooling, 11% stated a reason which could be directly interpreted as indicating an ability difference (such as "one twin was better at books"). Similarly, Bonjour *et al.* (2003) provide evidence that ability differences do matter, as out of the 38% of twins who went to different classes, half indicated ability differences as the reason. In both these studies however, these findings are interpreted as providing support for the idea that ability differences are relatively

[13] Note that the 7.2% reported in 1.3 is not based on a random effects estimator.

unimportant. The only directly comparable finding that I am aware of is in Griliches (1979), who reports a regression coefficient of just 0.13 for the within-pair effect of one standard deviation in IQ on years of schooling, based on a small sample of just 76 pairs of male monozygotic twins. In my much larger and more representative sample, this figure is 0.52.

It has been conceded that although we do not know whether ability is more familial than is schooling, within-pair estimates can still be used as an upper bound on the returns to schooling, under the assumption that ability bias is positive as is commonly thought (Bound and Solon (1999))[14]. Given that within-pair IV estimates are generally lower than the cross-sectional OLS estimates, co-twin estimates then contain information allowing us to tighten the bounds on the possible values of the returns to schooling. The central premises of this type of bounds argument, that ability bias can be taken to be positive *a priori* and that the suitability of an identification method therefore can be determined on the basis of the results it provides - if lower than OLS, then accept as an improvement - is dubious from a methodological perspective. Furthermore, such reasoning naturally rests on the assumption that the instruments for schooling are valid. As noted in the introduction, this is far from innocuous, and the potential reduction in bias must be weighed carefully against the plausibility of this assumption.

A crucial question is to what extent these results generalise to other countries. Returns to schooling vary across countries (Harmon *et al.* (2003)), and in principle so could the importance of ability bias. However, the argument in

[14]Note, however, that Bound and Solon (1999) do not strictly advocate the co-twin approach, but merely point out this logical implication.

this chapter rests not mainly on the magnitudes of the estimates from the wage equation, but on (i) the direct within-pair association of IQ with wages holding schooling constant, and (ii) on the significant within-pair associations between IQ and schooling. As long as within-pair dynamics are not substantially different in Sweden compared to other countries, these empirical findings constitute a general methodological argument against using within-twin differences for the estimation of returns to schooling, and in favour of placing greater weight on estimates based on alternative identification approaches.

1.9 Conclusion

Using a unique dataset of 890 pairs of male monozygotic twins' schooling, income and IQ, strong evidence is found against the usefulness of the assumption of equal ability within twin pairs. It is found that within twin pairs, differences in IQ make a significant contribution to differences in earnings as well as education, and that inclusion of IQ in the wage equation causes within-pair point estimates for the returns to schooling to decline considerably.

These findings cast doubt on the usefulness of the estimates derived from the co-twin literature, and provide an argument in favour of placing relatively greater credence to alternative approaches to identifying the returns to schooling.

1.10 Tables

Table 1.1: Sample representativeness with respect to income, education, and IQ

	Main sample	Population
Income	360	338
	228	-
Schooling	12.9	12.0
	2.6	-
IQ	0.12	0.0
	0.92	1.0
Age	42.9	
	7.6	
Obs	1780	-

Notes: Income is in 1000 SEK (6 SEK appr 1 USD)

IQ is measured in standard deviations, standardised against the group of all twins who answered the STAGE or SALT survey, and for whom there is conscription IQ data (12366 observations in total).

Population data is from Statistics Sweden

Population data is for only men, as in main sample

Schooling data for population is for cohorts born 1952-1972

Income data for population is for cohorts born 1950-1970

Schooling data for sample is based on self-reported education

Table 1.2: Probit of being above threshold on IQ, age, education (extended main sample)

Dep. variable: full-time proxy			
IQ	0.062	-0.000	
dF/dx	0.004	-0.000	
	(0.82)	(-0.10)	
Edu_{SURVEY}	-0.04		-0.03
dF/dx	-0.003		-0.002
	(-1.57)		(-1.34)
Age	0.01	0.02*	0.01
...dF/dx	0.001	0.001*	0.001
	(1.32)	(1.97)	(1.51)
Observations	1862	1862	1862
Pseudo-R2	0.013	0.008	0.012

Notes: Administrative measure of schooling used as instrument for self-reported years of schooling

Based on sample of individuals in pairs where both twins have data on schooling, IQ, and income

t-statistics within parentheses

Table 1.3: Comparability with previous literature

	Sample	Literature	Country/ies	Sources
Co-twin similarity				
$Corr(S_i^* S_k^*)$	0.73	0.66	US	AK1994
		0.75	US	AR1998
		0.70	Australia	M1995
$Corr(T_i T_k)$	0.82	0.86		BM1981
Instruments				
$Corr(S_X^* S_{Inst}^*)$	0.88	~0.90	US	AK1994
		0.92	US	AR1998
		0.88	Australia	M1995
		0.88	Sweden	I1999
$Corr(\Delta S_X^* \Delta S_{Inst}^*)$	0.65	0.57-0.83	US	AK1994
		0.62-0.76	US	R1999
Labour market				
β_{IV} (twins)	7.2%	~8%	US, UK.	C1999, B2003
		6.4%	Australia	M1995
		5.2%	Sweden	I1999
$\beta_{FE,IV}$	3.4%	~7%	US, UK	C1999, B2003
		4.5%	Australia	M1995
		(4.2%)	Sweden	I1999
$\delta \ln y_i / \delta T_i$	0.16	0.27	US	BG2002
$Corr(S_i^*, T_i)$	0.51	0.55	US	N1996

Notes: Abbreviations of sources: AK1994 - Ashenfelter and Kruger (1994); AR1998 - Ashenfelter and Rouse (1998); M1995 - Miller et al. (1995); BM1981 - Bouchard and McGue (1981); I1999 - Isacsson (1999); R1999 - Rouse (1999); C1999 - Card (1999); B2003 - Bonjour et al. (2003); BG2002 - Bowles and Gintis (2002); N1996 - Neisser et al. (1996).

Card (1999), Bowles and Gintis (2002) and Neisser et al. (1996) are all surveys or meta-analyses.

The correlation between noisy schooling and IQ from the literature is based primarily on childhood IQ scores.

Isacsson (1999) analyses men and women jointly, using income from 1987-1993

Within-pair estimates for Isacsson (1999) are based on imputed within-pair reliability ratios to correct for measurement error, and as such are not strictly comparable to the other within-pair figures in the table.

The sample "correlation" of schooling and instrument for schooling was derived as the square root of R^2 when regressing self-reported years of schooling on the set of administrative schooling dummies used as instruments.

Table 1.4: Two simple tests of within-pair ability differences

	FE, IV	FE, IV	FE
	y_{ij}	y_{ij}	S_{ij}
S_{ij}	0.034**	0.029*	-
	(0.012)	(0.012)	
T_{ij}		0.074**	0.517**
		(0.028)	(0.115)
Observations	1780	1780	1780
Groups	890	890	890

Notes: Standard error within parentheses

* 5%; ** 1%

y_{ij} is the natural logarithm of total income from employment in 2005

S_{ij} is self-reported years of schooling

Adminstrative dummies for highest degree attained used as instruments for years of schooling

T_{ij} is results from IQ test at around age 18, in standard deviations

Table 1.5: Two simple tests of within-pair ability differences, robustness checks

		FE,IV $\ln(y_{ij})$	*FE,IV* $\ln(y_{ij})$	*FE* S_{ij}	*Pairs*	*# Obs*
Excl. 5%	S_{ij}	0.030*	0.025	-	846	1692
		(0.013)	(0.013)			
	T_{ij}		0.100**	0.322*		
			(0.034)	(0.137)		
Alt. IQ	S_{ij}	0.034**	0.028*	-	890	1780
		(0.012)	(0.012)			
	T_{ij}		0.076**	0.522**		
			(0.028)	(0.115)		
Regr $\leftrightarrows$ *Instr*	S_{ij}	0.035**	0.031**	-	890	1780
		(0.011)	(0.012)			
	T_{ij}		0.071*	0.600**		
			(0.028)	(0.124)		
Low threshold	S_{ij}	0.031*	0.026*	-	900	1800
		(0.013)	(0.013)			
	T_{ij}		0.068*	0.521**		
			(0.029)	(0.114)		
High threshold	S_{ij}	0.036**	0.023*	-	791	1582
		(0.010)	(0.011)			
	T_{ij}		0.059*	0.430**		
			(0.024)	(0.122)		

Notes: Standard error within parentheses

* 5%; ** 1%

y_{ij} is the natural logarithm of total income from employment in 2005

S_{ij} is self-reported years of schooling

Adminstrative dummies for highest degree attained used as instruments for years of schooling

T_{ij} is results from IQ test at around age 18, in standard deviations

Chapter 2

Smart or selfish? - the role of cognitive behaviour in the ultimatum bargaining game

2.1 Introduction

In the ultimatum game, one agent (the "proposer") gets to propose how to allocate a sum of money between himself and another agent (the "responder"). The responder then accepts or rejects this offer. If he accepts, the proposer and the responder are paid in accordance with the proposed split. If the responder rejects, both players earn nothing. With self-interested income-maximising players, the game has a unique subgame perfect equilibrium, where the proposer offers zero or the smallest possible positive amount (in a discrete choice space) and the responder accepts this offer. Early experimental work documented that proposers typically offer a significant share of the pie, the modal offer is a fifty-fifty split, and responders routinely reject offers perceived as unfair (Guth *et al.* (1982)). These findings have subsequently been corroborated in different subject pools, with different stakes and in more elaborate settings (See Camerer (2003) for an extensive survey).

The observed outcome of the ultimatum bargaining game - that substantial shares are offered by the proposer, and that low offers generally are rejected - stands in stark contrast to the sub-game perfect equilibrium predicted if agents are assumed to be materially self-interested and rational. In this chapter I examine the extent to which departures from the rationality assumption can explain the typically observed behaviour in ultimatum bargaining games. Examining data from a sample of 895 individuals, aged 23-48 and drawn from a population register of twins, I estimate the effect of IQ on behaviour in the ultimatum game and explore the role of background variables. The explanatory value of IQ is marginal for all four aspects of behaviour examined, and only statistically significant[1] in two out of four cases.

Notably, deviations from the subgame perfect equilibrium both in terms of proposer behaviour and responder behaviour remain qualitatively unchanged even among individuals with IQ more than two standard deviations larger than the population average. This indicates that failure of the assumption of rationality is unlikely to be the driver behind these results, and constitutes evidence in favour of preference-based explanations as examined in the literature.

[1] Unless otherwise stated, statistical significance will be evaluated at the 5% level throughout this paper.

In the next section, I provide a brief background on the ultimatum game, and on the literature relating IQ to various behavioural departures from standard self-interested rationality. This is followed by a description of the data in Section 2.3. Section 2.4 contains results from my empirical analysis followed, in Section 2.5, by a discussion of my findings. Section 2.6 concludes.

2.2 Background

2.2.1 The ultimatum bargaining game

As noted above, the generally observed outcome of the ultimatum bargaining game - that substantial shares are offered by the proposer, and that low offers normally are rejected - constitutes a marked deviation from the sub-game perfect equilibrium predicted if agents are assumed to be motivated exclusively by material self-interest. In this context, it should be pointed out that the sub-game perfect equilibrium is not the obvious prediction, and that allocating large shares to the responder is consistent with a set of standard Nash equilibria. Nevertheless, the commonly observed rejections of offers in the second stage are not consistent with any equilibrium where such agents' preferences are bounded by the standard assumption of non-satiation.

One common interpretation of the observed behaviour is that people are concerned with the fairness of an outcome, and are willing to accept a loss for themselves in order to punish behaviour which is perceived as unacceptable. The term "strong reciprocity" has been proposed to describe the tendency of an agent to reward good behaviour and punish bad behaviour, the latter even at a net expected cost to the agent himself (see Fehr *et al.* (2002)).

This interpretation has inspired more elaborate games, with results generally conforming to the patterns predicted by the idea of "strong reciprocity". For instance, Fehr *et al.* (1997) design an experimental labour market where workers can choose to adhere to an effort level stipulated in a simple contract offered by the employer, or to provide more or less effort. Shirking leads randomly to a monetary punishment with a certain probability. In the final stage, employers then choose whether to punish or reward the worker by multiplying the worker's payoff by a factor between 0 and 2. The further away from one is this factor, the more costly is the action to the employer. The experimental evidence shows that workers reward generous contracts (*i.e.*, those providing a higher job rent) by shirking less often, and that employers are willing to take a cost to themselves to reward (punish) high (low) effort. Fehr *et al.* (1997) also find that workers anticipate the reciprocity of the employers, by providing lower effort when the final punishment/reward stage is removed.

The results from this extended literature seem to indicate that people, even when interacting with completely anonymous counterparts where reputational concerns should be non-existent, nevertheless exhibit strongly reciprocal social preferences. A number of authors (Bowles and Gintis (2002), Fehr and Fischbacher (2003), Fehr and Gächter (2002), Gintis (2000)) have even speculated that this demonstrates that humans possess innately other-regarding preferences which developed under conditions of group selection, although the empirical and logical basis of such claims have been criticised fiercely (Burnham and Johnson (2005)

A general alternative to the theory of "strong reciprocity" is that people are suspicious about the experimental situation and care about reputation. (see Levitt and List (2007) for a discussion of the roles of reputation and anonymity in experiments). This reputation-management may be conscious or sub-conscious, and may involve acquired behaviour or tendencies hard-wired into our neural circuitry (see, *e.g.*, Burnham and Johnson (2005))

From the perspective of the researcher, who knows that there are no reputational effects arising from the experimental situation, we might conceive of such tendencies as a form of failure of the rationality assumption; According to these theories, individuals are not acting in their objectively best interests, although they would if they just knew how to. As a corollary, individuals should not act as if motivated by social preferences when placed in an analogously one-shot situation, but with which they are familiar.

Some support for this idea comes from List (2006), who evaluates the gift-exchange hypothesis using a field experiment on traders of sports cards in a market, and compares it to the results from a similar laboratory experiment. In the field experiment, experimenters make either a high or a low offer to actual sports card traders in an actual market. The experimenters do not reveal their status, so traders are not aware of these offers being part of an experiment. When traders are local, they tend to reward high offers with a higher quality card, but this effect disappears when traders are not from the same region ("nonlocals"), *i.e.* when the probability of repeated interactions is low. When using standard laboratory conditions, List (2006) nevertheless finds the standard outcome of positive reciprocity in the form of gift-exchange. These results suggest that reputational effects, and not social preferences, are the cause of the observed gift-exchange in that particular market.

Interpreting rejections in the ultimatum game in terms of a conscious or subconscious trade-off based on the cost of violating social norms, provides one possible explanation for why IQ would matter in the experimental situation. If IQ measures a general problem-solving ability, it is reasonable to think that individuals with high IQ would also be better at correctly interpreting the cues of the experimental situation and at understanding that the probability and expected effect of their behaviour becoming known is negligible.

Another, simpler, way of conceiving of how a failure of the rationality assumption to hold might provoke "anomalous" behaviour, of the kind generally observed in the ultimatum game, is if higher cognitive ability facilitates effortless understanding of the payoff structure of the game as explained to the subject. Whereas, with enough deliberation, most people can plausibly be expected to understand what the best strategy is for the responder in the ultimatum game, it is not obvious that test subjects always invest the cognitive effort necessary.

Oechssler *et al.* (2008) examine the impact of temporary emotional states on responder behaviour in the ultimatum bargaining game, and how the differential effect relates to scores on a "cognitive reflection test", CRT (Frederick (2005)). This three-item test measures the extent to which an individual tends to think twice when faced with a task of a cognitive nature. Respondents are asked questions which have one intuitive, but wrong, answer, but where the correct answer is relatively straight-forward as long as the individual overcomes the first instinctive answer. For example, if a ball and a bat cost \$110 together, and a bat costs \$100 more than a ball, how much does a ball cost? The quick intuitive answer is \$10, whereas the correct answer is \$5. In the standard treatment, Oechssler *et al.* (2008) report no appreciable difference between the rejection rates of reflective decision-makers (2 or 3 correct answers on the CRT) and impulsive decision-makers (0 or 1 correct answer on the CRT). Furthermore, Oechssler *et al.* (2008) find that in the standard implementation where subjects are paid in money, rejection rates do not significantly decrease even when responders are given, 24 hours later, a chance to revise their action.

This distinction between decision-making under short and long time-frames has also been made by Gneezy and List (2006), although with the opposite result. Gneezy and List (2006) use two field experiments to test the gift-exchange hypothesis, a form of positive reciprocity, in the field. They find that although workers are willing in the short-term to reward higher remuneration with higher effort, this effect subsided after a few hours. The authors speculate that this may be due to the difference between "hot" decision-making, as in experiments, and "cold" decision-making, as in tasks in the field that have a longer duration.

2.2.2 IQ

There is a small but growing literature in economics attempting to view preferences and behaviour through the prism of cognitive ability. Dohmen *et al.* (2007)

and Benjamin *et al.* (2006) study the correlation between measures of cognitive ability and various measures of risk preferences, and find that risk aversion tends to associated with lower scores on an IQ test. Burnham *et al.* (2008) examine behaviour in the p-beauty game, and find that subjects with higher scores on a standard IQ test make substantially lower guesses, and are also closer to the revenue-maximizing guess. (In the case of the p-beauty game, this is simply p times the average guess. See Camerer (2003) for an explanation of the p-beauty game and the dominating theories concerning behaviour in this game). In the simultaneous prisoner's dilemma game, Kanazawa and Fontaine (2007) find that defectors have on average about one standard deviation higher IQ than cooperators. In a sequential prisoner's dilemma game, Burks *et al.* (2008) find that in the second stage, subjects with higher IQ tend to be more likely to defect when the opponent has defected in the first stage, and to be more likely to cooperate when the opponent has cooperated.

As of yet, there is however no large study examining the impact of measured IQ on behaviour in the ultimatum bargaining game. Brandstatter and Guth (2002) examine correlations between self-assessed cognitive ability and ultimatum game behaviour in a sample of just 26 subjects and find no association. Self-assessed measures, however, contain considerable noise and the use of such measures, rather than conventional test results, give rise to severe interpretational issues due to the potential interrelationship between ultimatum game behaviour and biases in self-perception.

2.3 Data

2.3.1 Participants

The experimental data was gathered by a team of researchers[2] in two rounds of experimental sessions undertaken in 2006 and 2008, respectively, in 15 Swedish cities and towns (Stockholm, Gothenburg, Malmö, Borlänge, Helsingborg, Jönköping, Kristianstad, Linköping, Lund, Norrköping, Örebro, Umeå, Uppsala, Västerås, and Växjö). Since the main purpose of gathering the data was to estimate heritability of behaviour in experiments designed to elicit social preferences[3], subjects were recruited from the population of twins in Sweden. Contact information on twins in the region of a particular experimental session was obtained from the Swedish Twin Registry, which contains all twins born in Sweden from 1926 and onwards[4]. Twins born between 1960 and 1985 were contacted by email and mail. Only pairs where both twins could participate in the same session were ultimately included in the experiment, to ensure that conditions for both twins would be as similar as possible, and, specifically, to avoid communication between sessions.

2.3.2 Ultimatum game test procedure

The ultimatum game was implemented using the strategy method: In the first stage, all subjects were assigned the role of proposer, were matched randomly to an anonymous responder from a different experimental session, and were asked

[2] David Cesarini, Magnus Johannesson, Björn Wallace

[3] For other studies using data from these experimental sessions, see Burnham et al. (2008), Cesarini et al. (forthcoming), Cesarini et al. (2008), Wallace et al. (2007).

[4] See Lichtenstein et al. (2006) for a description of the Swedish Twin Registry.

to indicate in writing what share, in increments of 10, out of a total amount of 100 Swedish krona[5] they would offer to the respondent. In the second stage, all subjects were assigned the role of responder, were matched randomly with an anonymous proposer from a different experimental session and different from the one in the first stage, and were asked what offers they would and would not accept. From this strategy set, I use the minimum acceptable offer (MAO) in the range 0-50% as my measure of responder behaviour. In the range 0-50%, all subjects except one (who was therefore excluded) indicated a monotonic strategy, *i.e.* accepting all offers greater than the MAO. The relevant experimental data thus contains, for each subject, one offer and one MAO as measures of behaviour in the ultimatum bargaining game.

2.3.3 Cognitive ability

As part of the experiment, the subjects were asked to take a short standard IQ test developed by Psykologiförlaget (Sjöberg, Sjöberg, and Forssen (2006)). This test consists of three sections: logical problems, verbal analogies, and mathematical series, and the maximum time for completion is 20 minutes. The population average and standard deviation, as provided by the test developers, were then used to standardise each participant's score. For simplicity, the unit of measurement used in this chapter is simply the number of population standard deviations away from the population mean, rather than linearly transforming the scores into an IQ value.

[5] 6 Swedish krona is approximately 1 US dollar.

2.3.4 Representativeness

To assess the representativeness of the sample, the experimental data was matched to data on taxable income, highest educational level attained, and marital status from Statistics Sweden. The educational data was then converted into years of schooling using estimates from a representative random sample reported in Isacsson (2004).

Table 2.1 contains a comparison of sample income, education, and marital status with the population average for the cohorts aged 25-44.

As is common in laboratory experiments, participants are more educated than the average in the population, in this case 14.0 years compared to 12.8. IQ test results are higher as well, with the average participant in the experiment having an IQ about half a standard deviation higher than the population mean.

Average income is similar to the population average for women, but about 10% lower for men, and marital status is lower for men as well as for women, with 27% married in the sample, as opposed to 37% in the population as a whole. The average age in the sample is 34.8 years.

2.4 Results

2.4.1 Observed behaviour

Figure 1 shows the distribution of offers. As is generally observed in the ultimatum bargaining game, almost all offers are within the 30-50[6] range. The mean is 48.6, and the median and mode are 50. Only 2.7% of the subjects offer 0 or 10 (the only sub-game perfect Nash equilibria), but more than 80% offer a fifty-fifty split. Figure 2 shows the distribution of recorded minimum offers that an individual would accept. The mean MAO is 36.8, with median 40 and mode 50. These results conform well to those found in the literature. (see, *e.g.*, the summary in Falk and Fischbacher (2006)). A complete set of summary statistics on the observed experimental behaviour is given in Table 2.2.

The expected payoff from any particular offer or minimum acceptable offer depends on the distribution of strategies pursued in the relevant population. Using the sample distributions of MAOs and offers, the expected payoff from each possible offer and MAO, respectively, is calculated.

Table 2.3 contains expected monetary payoffs in the experiment for each value of minimum acceptable offer. By mathematical necessity, the expected payoffs are weakly decreasing in MAO, although the expected payoff from accepting no less than 0, *i.e.* accepting any offer, and accepting no less than 40 are virtually the same. In other words, in this setting, the expected cost from devations from the simple sub-game equilibrium are very small. The bottom row of Table 2.3

[6]For expositional clarity, the outcome variables will be referred to without a % sign or unit, rather than in percent of the pie or Swedish krona.

displays the number of subjects choosing each particular MAO.

Table 2.4 presents the expected payoff from an offer, conditional on the observed distribution of minimum acceptable offers. An offer of 50 has the highest expected payoff, and offering 60 has a higher expected payoff than offering 40. As offers get closer to zero, expected payoffs to the proposer get smaller, although the expected payoff from offering 10 and 20, respectively, are virtually the same. The bottom row of Table 2.4 displays the number of subjects choosing each particular offer.

2.4.2 Regression analysis of the role of IQ

Responder behaviour

In this section, I analyse the role of cognitive abilities in explaining behaviour where the income-maximizing choice is not affected by the behaviour of any other player. In the ultimatum game, this type of weakly monotonic situation arises with respect to the minimum acceptable offer (MAO). The extent to which cognitive abilities can explain the expected payoff of the minimum acceptable offer chosen is also examined.

In Table 2.5, results are presented from a regression analysis of minimum acceptable offer (MAO) and the expected value of MAO on IQ, and on IQ and two sets of covariates.

Columns 1-3 contain coefficients estimated when using offer as the dependent variable. Five results emerge: First, IQ matters for minimum acceptable offer, but although this effect is statistically significant, it is small in magnitude. A difference in IQ of one standard deviation is associated with a lower MAO of approximately 3-3.5, or roughly one tenth of the average MAO. Second, this effect is only marginally affected by the inclusion of covariates, dropping to 2.7 (negative) when including the full set of covariates - age, sex, yearly income, years of education, and marital status. Third, age matters positively, but weakly, with 10 years corresponding to an increase in MAO of about 3. Fourth, the share of variance explained is small, about 6%, even when including the full set of covariates. Finally, the inclusion of income, education, and marital status does not significantly increase the predictive fit ($p = 0.44$) compared to the model with only age and sex as covariates.

Columns 4-6 contain the corresponding results using expected value of MAO as the dependent variable. All results are similar in terms of significance and interpretation, but far removed in terms of magnitude. An increase in IQ of one standard deviation corresponds to an increase in the expected payoff of only 0.2. This discrepancy is not surprising given the compressed distribution of the expected payoffs. As shown in Table 2.3, the expected payoff varies very little even between the extremes of 0 and 50.

Proposer behaviour

In this section, I examine whether cognitive ability predicts behaviour in cases where the income-maximising choice is a function of other players' actions, in this game corresponding to the choice of offer. As in the preceeding section, the expected payoff of the behaviour is employed as an alternative dependent variable.

It should be noted that in the case of offers, this transformation is not monotonous, since expected payoff is maximised for a value of 50 (as seen in Table 2.4), rather than 0 or 100.

In Table 2.6, results are presented from a regression analysis of offer and the expected value of offer on IQ, and on IQ and the same two sets of covariates as in the preceeding section. Four results can be noted: First, across all specifications, the estimated coefficient on IQ is small and statistically insignificant. Second, of all covariates only sex explains behaviour, with women offering about 1.9 less then men. Third, again the inclusion of income, education, and marital status does not significantly increase the predictive fit for either of the two dependent variables ($p = 0.38$ and $p = 0.97$) compared to the models with only age and sex as covariates. Finally, the total share of variance explained is negligible, less than 1.2%, in all specifications.

Exclusion of potential outliers

In the proposer stage, seven individuals offered 100 to the responder. Although such preferences in this particular game are surely feasible, it is still noteworthy that noone offered either 90 or 80, and only one subject offered 70. Based on this curious observed distribution of offers, it is possible that the offers of 100 were the results of misunderstanding the answer sheets.

For this reason, I re-estimate the effect on MAO, offer, and the expected payoffs from MAO and offer, excluding these seven observations. The results of this robustness test are presented in Table 2.7. Excluding the seven potential outliers does not appreciably affect the estimated association between IQ and behaviour, which remains statistically significant only for MAO and expected payoff from MAO, and then only with limited explanatory power.

2.4.3 Group comparisons

High-IQ individuals

In this subsection, I examine whether the assumption of rationality may be properly viewed as applying fully only to the top percentiles of the IQ distribution. In other words, rather than taking a continuous approach to the relationship between cognitive ability and behaviour, as in the previous sections, I simply examine whether the observed behaviour differs between individuals whose measured IQ corresponds to the the top 5% of the population (approximately the top 15% of the sample), and the rest of the sample. These results are presented in Table 2.8. I also display results from a corresponding comparison using individuals from the top 1% of the population IQ distribution.

Three results can be noted from this group comparison. First, although the difference between either of the two high-IQ groups and the rest of the sample is statistically significant for both MAO and the expected payoff from MAO, the bulk of observed deviations from the weakly dominant strategies (to accept offers as low as 0 or 10, respectively) remains. The average subject in the top 1% of the population distribution still rejects offers of less than about 30.

Second, the general pattern observed is closely in line with the results from the regression analysis above, in that IQ appears to matter somewhat for the responder stage, where there is a weakly dominant strategy (accept or reject), whereas it appears to play no role for the proposer stage, where payoffs depend on the successive play of the opponent (the offer stage)

Third, even for individuals in the top 1% of the population IQ distribution, corresponding to approximately the top 5% of the sample distribution, the expected payoff from the MAO is only a meagre 0.6 higher on average than for the average individual in the sample. In the actual experimental situation, the expected gains from higher IQ are negligible in magnitude.

"Narrowly rational" individuals

To examine whether there is anything which distinguishes players who adhere to the subgame perfect equilibrium generally associated with the assumption of selfish income-maximising agents, I partition the sample into those who adhere to the corresponding SPNE strategies, and those who do not. There are 18 individuals who both (i) offered 0 or 10, and (ii) accepted all offers higher or equal to 0 or 10. I label these as "narrowly rational", to indicate that although their behaviour is consistent with play in an environment where all agents are selfishly rational, it is not necessary consistent with income-maximising play when other players are believed, with positive probability, to play strategies other than those of the SPNE referred to above. In other words, these are the individuals exhibiting the behaviour predicted, through backward induction, for homo oeconomicus *in an environment of likes.*

Table 2.9 presents a comparison of means. The narrowly rational individuals tend to have slightly lower income, be slightly less likely to be married, have slightly higher IQ, and are slightly more likely to be male, but none of these differences are statistically significant.

As some of the observed differences appear to be negatively correlated, in particular lower income and higher male ratio, I examine the probability of being narrowly rational in a probit model. The results are presented in Table 2.10. Across a range of specifications, none of the explanatory variables have a statistically significant coefficient, including IQ.

2.5 Discussion

The results above give weak support for the idea that a failure of rationality is a contributing factor in explaining the commonly observed behaviour in ultimatum bargaining games, that players depart substantially from the self-interested income-maximising subgame perfect equilibrium. Responders have a weakly dominant strategy of always accepting positive (or non-negative) offers. Yet, even in this unconditional setting, IQ alone explains less than five percent of variation, and even among individuals from the top 5% of the population in terms of IQ, the bulk of the deviation from the weakly dominant strategy persists. A one standard deviation increase in IQ is associated with a drop of 3-3.5 in minimum acceptable offer (MAO).

In cases where the optimality of a strategy depends on the actions of the other player(s), I find no association between IQ and behaviour. Neither the actual offers, nor their expected payoff based on the ex post realised distribution of

MAOs, are even weakly associated with IQ. It should also be noted that more than 80% of all subjects play the optimal strategy - to offer 50 - given the ex post choices of the respondents

Yet, as is generally found, there is substantial heterogeneity with respect to responder behaviour. A failure of rationality to be the culprit naturally lends support for preference-based explanations holding the key to understanding these heterogeneities. But what might account for this heterogeneity in preferences?

In this chapter, I provide some explorative empirical analysis of the role of income, education, marital status, age, and sex in explaining heterogeneities in observed behaviour. Admittedly, when including demographic covariates, there is potentially important endogeneity between both stipulated dependent variable (behaviour in the ultimatum game), stipulated treatment variable (IQ), and stipulated controls (*e.g.* income and education). The interpretation of the estimated coefficients is therefore not straightforward. It can nevertheless be noted that the effects of IQ across dependent variables are fairly robust to the inclusion of these demographic variables in the estimating equations.

Whereas some variation in MAO can be explained by age and sex, the inclusion, in addition to IQ, sex, and age, of income, education, and marital status in a linear regression model does not contribute significantly to its explanatory power. Furthermore, the joint addition to R^2 from inclusion of all these five variables is never more than 0.02, which still in all cases leaves more than 90% of the variation unexplained by IQ or these five conventional demographic variables.

The potential role of genetic differences as an underlying cause of observed heterogeneities in economic preferences has been explored by Burnham *et al.* (2008), Cesarini *et al.* (forth), Cesarini *et al.* (2008), Wallace *et al.* (2007).[7] Other recent research along the same lines includes Fowler *et al.* (2007), who examine the role of genetics in explaining variation in political preferences. In all the above five papers, genetic variation is robustly found to explain a moderate share of observed variation, in the range of 25-40%. The results in this chapter suggest that the mediating channel is not the genetic transmission of general cognitive abilities.

Obviously, IQ and rationality are not immediately corresponding notions. The findings of this chapter however complement those of Oechssler *et al.* (2008), as reported above, who find that more reflective decision-makers (as measured by a three-item "cognitive reflection test", due to Frederick (2005)) are equally likely as impulsive decision-makers to reject an unequal split in the ultimatum bargaining game. Furthermore, Oechssler *et al.* (2008) find that in the standard implementation where subjects are paid in money, rejection rates do not significantly decrease even when responders are given, 24 hours later, a chance to revise their action.

Taken together, these findings suggest that neither lack of sufficient cognitive ability, nor failure to apply it properly, is responsible for the observed rejections in the ultimatum bargaining game. Instead, preference-based explanations allowing for variations of the assumptions of self-interest are more likely to hold the answer.

[7]These four papers use data from the same experimental sessions as is this paper.

2.6 Conclusion

Using a large and representative sample of 895 individuals, I examine the effect of IQ on proposer and responder behaviour in the ultimatum game. I find no effects on proposer behaviour and statistically significant, but small, effects on responder behaviour. These findings remain when controlling for a range of demographic variables, and when juxtaposing individuals from the top 5% and 1% of the population with the rest of the sample. Furthermore, the 2% of subjects who played the simple subgame perfect Nash equilibrium did not differ significantly on neither IQ nor any of the demographic variables examined. No statistically or economically significant additional explanatory power was obtained from income, education, and marital status.

2.7 Tables

Table 2.1: Comparison of demographic characteristics

	Main Sample			Population		
	All	Women	Men	All	Women	Men
Education	14.0	14.0	14.0	12.8	13.1	12.5
	(2.3)	(2.4)	(2.2)			
IQ test	0.53	0.43	0.88	0.0	-	-
	(1.03)	(1.00)	(1.04)	1.0	-	-
Income	207	198	242	235	197	271
	(144)	(132)	(177)			
Marital status	0.27	0.28	0.22	0.37	0.40	0.33
	(0.44)	(0.45)	(0.42)			
Age	34.8	35.3	33.1	25-44	25-44	25-44
	(7.5)	(7.5)	(7.2)			
N	895	700	195	2.4 M	1.2 M	1.2 M

Notes: Standard deviations within parentheses.

Income in 1000 SEK,

IQ test results standardised to the entire population

Marital status is 1 for "married", where co-habitation counted as "not married"

In population data, education data is missing for about 0.5%

For 4 female observations education is missing, and for 2 observations income and marital status are missing

Table 2.2: Summary statistics of behaviour

	All		Women		Men	
	Mean	S.D	Mean	S.D	Mean	S.D
MAO	36.8	17.3	37.0	17.3	36.1	17.0
Expected payoff from MAO	47.4	0.9	47.4	0.9	47.5	0.9
Offer	48.6	9.4	48.2	9.0	49.8	10.7
Expected payoff from offer	47.2	9.0	47.5	8.4	46.1	10.7
N	895		700		195	

Table 2.3: Expected payoff from minimum acceptable offer (MAO)

Share	50	40	30	20	10	0
E(payoff)	46.5	48.0	48.4	48.5	48.6	48.6
N	432	203	62	25	87	86

Notes: The expected payoff is calculated as the expectation over offers, where offers lower than the minimum acceptable offer are set to zero.

Table 2.4: Expected payoff from offer

Share	0	10	20	30	40	50	60	70	80	90	100
E(payoff)	9.6	17.4	17.7	20.3	31.0	50.0	40.0	30.0	20.0	10.0	0.0
N	15	8	5	10	35	801	13	1	0	0	7

Notes: The offer is the share proposed to the responder, whereas the expected payoff is to the proposer.

The expected payoff is calculated as the share left to the proposer, times the ex post probability that the offer is accepted by the responder.

Table 2.5: Responder behaviour - results from regressing MAO and expected payoff from MAO on IQ and covariates

	MAO			Expected payoff from MAO		
IQ	-3.51**	-3.05**	-2.75**	0.24**	0.21**	0.19**
	(0.55)	(0.57)	(0.61)	(0.03)	(0.03)	(0.03)
Age		0.28**	0.26**		-0.02**	-0.02**
		(0.08)	(0.10)		(0.00)	(0.01)
Sex		-1.05	-0.74		0.05	0.03
		(1.39)	(1.43)		(0.07)	(0.08)
Income			0.28			-0.01
			(0.50)			(0.03)
Education			-0.43			0.02
			(0.26)			(0.01)
Married			-0.52			0.04
			(1.39)			(0.07)
Constant	38.67**	29.41**	35.15**	47.30**	47.95**	47.72**
	(0.63)	(2.98)	(4.73)	(0.03)	(0.16)	(0.25)
Observations	895	895	891	895	895	891
R^2	0.044	0.058	0.061	0.070	0.093	0.093

Notes: Standard errors within parentheses

** p<0.01; * p<0.05

Income in 100.000 SEK

Education in years of schooling, based on registry data from Statistics Sweden

Co-habitation counted as "not married"

Table 2.6: Proposer behaviour - results from regressing offer and expected payoff from offer on IQ and covariates

	Offer			Expected payoff from offer		
IQ	-0.30	-0.36	-0.44	-0.40	-0.16	-0.15
	(0.31)	(0.32)	(0.34)	(0.29)	(0.30)	(0.33)
Age		0.04	-0.01		0.08	0.09
		(0.04)	(0.06)		(0.04)	(0.05)
Sex		-1.87*	-1.80*		1.20	1.20
		(0.78)	(0.80)		(0.74)	(0.76)
Income			0.17			0.00
			(0.28)			(0.27)
Education			0.03			-0.02
			(0.15)			(0.14)
Married			1.24			-0.37
			(0.78)			(0.74)
Constant	48.74**	48.87**	49.54**	47.43**	43.62**	43.58**
	(0.35)	(1.67)	(2.66)	(0.34)	(1.59)	(2.53)
Observations	895	895	891	895	895	891
R^2	0.001	0.008	0.012	0.002	0.010	0.010

Notes: Standard errors within parentheses

** p<0.01; * p<0.05

Income in 100.000 SEK

Education in years of schooling, based on registry data from Statistics Sweden

Co-habitation counted as "not married"

Table 2.7: Results from regressing behaviour on IQ and age - without possible outliers w r t behaviour variables

	MAO	E(MAO)	Offer	E(offer)
IQ	-3.14**	0.21**	-0.14	-0.37
	(-5.55)	(6.92)	(-0.51)	(-1.36)
Age	0.30**	-0.02**	0.05	0.07*
	(3.86)	(-4.85)	(1.21)	(1.98)
Sex	-1.83	0.07	-0.62	0.04
	(-1.32)	(0.94)	(-0.89)	(0.06)
Constant	29.75**	47.95**	47.11**	45.22**
	(10.03)	(301.56)	(31.70)	(31.82)
Observations	888	888	888	888
R^2	0.063	0.096	0.003	0.009

Notes: Excluding the seven observations who as proposers offered 100% to the respondent

Standard errors within parentheses

** p<0.01; * p<0.05

Table 2.8: High-IQ behaviour

	All		Top 5%			Top 1%		
	Mean	S.D	Mean	S.D	p-value	Mean	S.D	p-value
MAO	36.8	17.3	31.6	17.8	0.000	27.7	18.0	0.001
E(payoff) of MAO	47.4	0.9	47.8	0.9	0.000	48.0	0.8	0.000
Offer	48.6	9.4	48.0	7.1	0.378	47.3	7.9	0.270
E(payoff) of offer	47.2	9.0	46.8	8.4	0.525	46.4	9.5	0.545
N	895		133			44		

Notes: p-value refers to t-test of equal means between high-IQ group and its complement (i.e., not against the full sample)

Table 2.9: Narrowly "selfish rational" behaviour

	Not SPNE		SPNE		p-value
	Mean	S.D	Mean	S.D	
Education	14.0	2.3	14.6	2.0	0.25
IQ	0.525	1.03	0.687	1.02	0.51
Income	208	144	180	175	0.51
Marital status	0.27	0.44	0.17	0.38	0.27
Age	34.9	7.5	33.6	8.4	0.54
Sex	0.78	0.41	0.89	0.32	0.18
N	877		18		

Note: Subgame perfect Nash Equilibrium play is when offer is 0 or 10, and minimum acceptable offer is 0 or 10.

For 4 observations education is missing, and for 2 observations income and marital status are missing

Table 2.10: Probit analysis of "narrowly rational" (SPNE) behaviour

Dep. variable: 1 if individual is "narrowly rational"						
IQ	0.003		0.003		0.001	0.002
	(0.66)		(0.67)		(0.15)	(0.39)
Income				-0.003	-0.003	-0.002
				(-0.81)	(-0.79)	(-0.51)
Education				0.003	0.002	0.002
				(1.23)	(1.12)	(1.02)
Married				-0.009	-0.009	-0.010
				(-0.83)	(-0.82)	(-0.91)
Age		-0.000	-0.000			0.000
		(-0.79)	(-0.58)			(0.08)
Sex		0.013	0.014			0.012
		(1.18)	(1.25)			(1.14)
Observations	895	895	895	891	891	891

Notes: Regression coefficients are marginal effects

z-statistics within parentheses

Income in 100.000 SEK

Education in years of schooling, based on registry data from Statistics Sweden

Co-habitation counted as "not married"

** p<0.01; * p<0.05

2.8 Figures

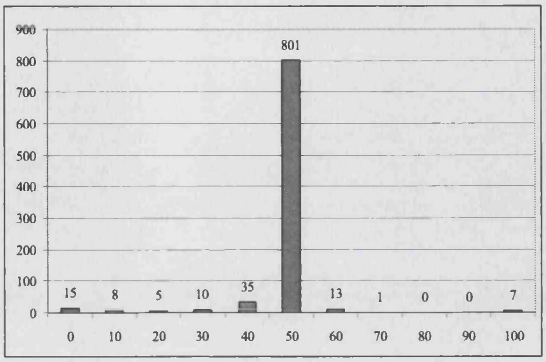

Figure 2.1: Distribution of offers

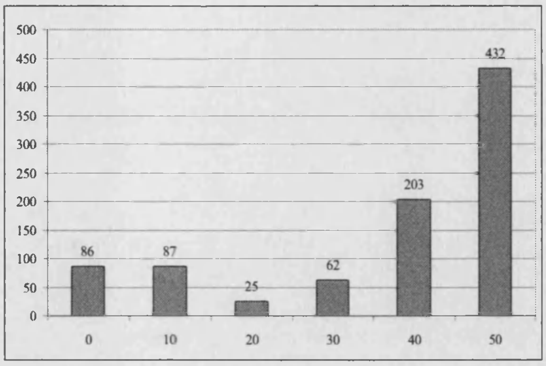

Figure 2.2: Distribution of minimum acceptable offers

Chapter 3

Is financial risk-taking behaviour genetically transmitted?

(Co-authored with David Cesarini, Magnus Johannesson, and Björn Wallace)

3.1 Introduction

Parents and their children exhibit considerable similarity in self-reported attitudes toward risk (Charles and Hurst (2003), Dohmen *et al.* (2006), Hryshko *et al.* (2007)), as well as in their choice of what assets to hold (Chiteji and Stafford (1999)). Yet, little is known about the mechanisms generating these correlations. Do they arise because parents pass on genes for certain traits associated with risk preferences to their children, or is it, as often postulated, merely a reflection of parental socializing influences? In a recent string of papers (Wallace *et al.* (2007), Cesarini *et al.* (2008, forthcoming), laboratory experiments designed to elicit preferences were run on a sample of twins. Comparing the behaviour of monozygotic (MZ) twins to that of dizygotic (DZ) twins is a form of quasi-experiment. MZ twins reared together share the same environment and the same genes, and while DZ twins reared together also share the same environment, their degree of genetic relatedness is no greater than that of ordinary siblings. A significantly higher observed correlation for MZ twins than DZ twins is therefore usually taken as evidence that a trait is under genetic influence. Estimating fairly standard behaviour genetic models, the results in Wallace *et al.* (2007) and Cesarini *et al.* (2008, forthcoming) suggest that heritability - the share of individual variation that can be explained by genetic influences - for a number of economic preferences, including risk preferences, is typically somewhere between 20 and 40 %.

Yet, eliciting preferences experimentally has at least two distinct disadvantages. First, there is genuine uncertainty about the extent to which laboratory behaviour generalizes to the field (Levitt and List (2007)). Second, the sample sizes in the above cited studies, though very large by behavioural economic standards, still do not allow for precise inference. In this chapter, we use microdata from the Swedish individualized pension savings account introduced in 2000 to extend the previous literature from the laboratory to the field. As part of the transition to a new pension system, virtually all adult Swedes born after 1938 had to make simultaneous investment decisions with potentially far-reaching effects on their post-retirement wealth. In particular, they had to compose an investment portfolio from a menu of more than six hundred funds. We take this event, which is known as the "Big Bang" of the Swedish financial sector, as a field experiment to infer risk preferences. Matching individual portfolio data to the Swedish Twin Registry, we then employ standard methods from behaviour genetics and estimate the heritability of preferences for financial risk-taking. Unlike small stake gambles in the laboratory, or attitudinal risk questions, the investment decisions made in the pension savings accounts are real financial decisions that have real economic consequences. Moreover, our dataset is very large, allowing us to estimate parameters with much greater precision than previous experimental studies.

To our knowledge, this study is the first to use behaviour genetic techniques to document the heritability of risk-taking in the financial market, as well as outside the laboratory. There is, however, a related literature in economics which has considered economic outcome variables such as educational attainment, income and socioeconomic status (Taubman (1976), Behrman and Taubman (1989), Lichtenstein *et al.* (1992), Plug and Vijverberg (2003), Björklund, Lindahl and Plug (2006), Björklund, Jäntti, and Solon (2007), Sacerdote (2002, 2007)). The general idea behind these papers is to make reasonable assumptions about the genetic relationships of relatives to separate the effects of genetic and environmental variation (Sacerdote (forthcoming)). Behavior genetic techniques are in no way restricted to twins, and many of the above studies also include adoptees, as well as other sibling types. Taken together, adoption, sibling and twin studies point to a role for both genetic and cultural transmission of economic outcomes.

The estimates of heritability that we obtain match the laboratory evidence in Cesarini *et al.* (forthcoming) very closely, and suggest that approximately 25 % of the individual variation in financial risk-taking is due to genetic influences.

This implies that a significant portion of the previously observed parent-child resemblance in risk attitudes is due to genetic transmission. Furthermore, besides establishing that a key economic preference is heritable, an important result in and of itself, we believe that our findings have broader implications. For instance, the share of individual variation explained by genes is much higher than the R^2s typically obtained in standard empirical models of financial risk-taking (Cohn *et al.* (1975), Pålsson (1996), Palme *et al.* (2007)). In an early microdata study of portfolio choice Cohn *et al.* (1975) found R^2s of approximately 0.10 using a set of demographic and socio-economic controls, while Pålsson (1996) report substantially lower estimates for Swedish registry based data. Importantly, these results for individual portfolios of total asset holdings are close to the R^2's found when considering portfolio choice for the Swedish individualized pension savings accounts in isolation (Säve-Söderbergh (2005), Palme *et al.* (2007)).

Our findings may also be relevant for research on the intergenerational transmission of economic status. Reviewing the literature, Bowles and Gintis (2002) suggest that further study of non-cognitive behavioural traits and preferences may help explain the fact that even though income is heritable (Taubman (1976)), simple calibration exercises show that the genetic transmission of intelligence can account for at most a moderate share of the parent-child correlation. Poterba *et al.* (2000) show the substantial effects risk-preferences can have on accumulation of post-retirement wealth and thus potentially on intergenerational transmission of economic status.[1]

[1] However, preferences are notoriously diffucult to measure, and attitudes toward risk is only one of many dimensions of preferences, whose individual effects may be small, but whose com-

The chapter is structured as follows. In Section 3.2, we describe the Swedish Pension reform and our dataset. In Section 3.3, we describe twin methodology. In Section 3.4, we present our results, and relate them to previous findings. In section 3.5 we investigate and discuss the robustness and generalizability of our findings, and in section 3.6 we relate them to the literature on behaviour genetics and the biological basis of risk preferences. Section 3.7 concludes.

3.2 Data

3.2.1 The Swedish Pension Reform

In 1994, legislation gradually introducing a new pension system was passed by the Swedish parliament in response to demographic challenges and underfinancing of the pay-as-you-go system that had been in place since the 1960s.[2] The new system is based on a contribution rate of 18.5 percent on earnings, whereof 2.5 percentage points accrue to mandatory individual self-directed accounts, one of the system's key features.

As part of the introduction of the new system, a government body - the Premium Pension Agency - was set up and assigned the responsibility of handling the individual investment accounts. Most adult Swedes born after 1938 were invited to decide how to invest the balance on their individualized pension savings ac-

bined effect might be substantial. Moreover, attitudes towards risk may well interact with other variables and form a non-linear relationship with socio-economic status. Such non-linearities have been documented by Turner and Martinez (1977) in the context of scores on the Mach V scale, which measures the degree of Macchiavellian personality traits. They provide evidence that individual scores on this personality test have differential effects on income depending on social stratum.

[2] See Palmer (2000) for a detailed exposition of the new system.

counts, but the system only fully applied to individuals born 1954 and onwards.[3] The "Big Bang" occurred toward the end of 2000, when all participants in the new system had to simultaneously decide how to invest their balances. Some 68 percent of the eligible population made an active decision. Individuals who did not make an active choice had their money invested in a default fund.

Participants could compose a portfolio consisting of no more than 5 funds from a very large menu of options comprising more than 600 different funds.[4] All eligible Swedes were sent a catalogue in which available funds were listed with information on management fees and the investment strategy of each fund. For the approximately 400 funds that had a historical record, returns and standard deviation of returns for the preceding three years were also given. These funds were also color-coded by risk level: from red (high risk) to green (low risk). The circumstances under which these investment decisions were made make the experiment uniquely suitable for inferring risk preferences among individuals with little or no financial fluency.

3.2.2 Portfolio Risk Data

Our primary measure of portfolio risk, which we denote Risk 1, is the average risk level of the funds invested in by the individual, with the risk of each fund measured as the standard deviation of the rate of return over the previous three

[3]Only Swedes whose income exceeded SEK 36000 ($1 is roughly 6 SEK) in 1995, 36800 in 1996, 37000 in 1997 and 37100 in 1998 were eligible for fund selection in the year 2000.

[4]The official justification for this policy was that individuals should be able to select a portfolio that suited their preferences. For a criticism of this feature of the system, see Cronqvist and Thaler (2004) and Palme and Sundén (2004).

years. In cases where historical returns were not available, these values were imputed by assigning the average value of risk for similar types of funds in the sample.[5] This measure is similar to that employed in Säve-Söderbergh (2008) and Palme *et al.* (2007), with the one notable exception that we also include twins whose money was invested in the default fund.[6] As a robustness check, we also calculated a second risk measure, Risk 2, as the weighted share of high-risk funds in an individual's portfolio.[7]

3.2.3 The Swedish Twin Registry

The Swedish Twin Registry, the largest in the world, contains information on nearly all twin births in Sweden since 1886, and has been described in detail elsewhere (Lichtenstein *et al.* (2006)). Our sample includes individuals who have participated in at least one of the Twin Registry's surveys. For these respondents, we can establish zygosity with reasonable confidence based on survey questions with proven reliability (Lichtenstein *et al.* (2006)). In practice, roughly 90 % of the twins in our dataset come from one of two sources. The primary source is the web-based survey STAGE (The Study of Twin Adults: Genes and Environment).

[5] The classification of funds was made by the Premium Pension Agency. Examples of types are "New Markets", "IT and Communication", and "Europe Small Enterprises". Our method of imputing missing values has no interesting effects on the estimates we report in this paper.

[6] Säve Söderbergh (2008) excludes individuals with the default portfolio on the grounds that its investment profile was not fully known when investment decisions were made in the fall of the year 2000. The reason its risk profile was not known is that it was constructed to reflect the profile of an average investor. On the other hand, it seems reasonable to assume that people had some expectation about the future level of risk in the default fund. In practice, none of the results reported in this paper are sensitive to this inclusion. This supports the notion that individuals not actively choosing a portfolio nevertheless conveyed some information about their risk preferences.

[7] A high risk fund was defined by the Premium Pension Agency as one holding at least 75% equity investments.

This survey was administered between November 2005 and March 2006 to all twins born in Sweden between 1959 and 1985, and it attained a response rate of 60 %. Data on individuals born between 1938 and 1958 were obtained from SALT (Screening Across the Lifespan Twin study), a survey conducted by telephone in 1998. SALT attained a response rate of 74 % (Lichtenstein *et al.* (2006)).[8] Though these response rates are not alarmingly low, we acknowledge that our sample is not fully representative of the population of twins. Considering all complete same sex twin pairs born after 1938 gives a total of 7224 female pairs, of which 3346 are monozygotic, and 6338 male pairs, of which 2747 are monozygotic.

3.3 Method

Our analysis uses a measure of the portfolio risks chosen by twins to estimate the degree to which variation is influenced by additive genetic factors (A), environmental factors shared or common to the two twins in a pair (C), and unshared environmental (E) factors which are specific to each twin. Additive genetic effects are defined as the sum of the effects of individual alleles influencing a trait. Common environment effects are those environmental influences shared by both twins. Examples include childhood diet, schooling, parental socialization and shared peer influences. Unshared environmental effects include influences not shared by the co-twins as well as measurement and response error.

The basic idea behind a behaviour genetic decomposition is simple. MZ and

[8] Additionally, a small number of individuals in our sample responded to a survey sent out in 1973 (See, again, Lichtenstein *et al.*, 2002). These are also included.

DZ twins differ in their genetic relatedness. If one is willing to assume that the common environment does not exert greater influence in MZ twins, then a greater similarity between MZ twins can be taken as evidence that the trait is under genetic influence.

Several authors, most recently Sacerdote (forthcoming), have noted that moving from a crude comparison of correlations to a full-fledged variance decomposition requires making strong independence and functional form assumptions. Therefore, our empirical analysis proceeds in two steps. We first abstain from imposing any structural assumptions, and simply compare the correlations of MZ and DZ twins using the bootstrap. Letting N_{MZ} be the number of MZ pairs without missing data, we draw N_{MZ} pairs with replacement 1000 times and calculate the non-parametric correlation each time. We proceed analogously for DZ twins, and then create a 1000 by 1 vector where the DZ correlation is subtracted from the MZ correlation for each draw. This gives a distribution for the difference in correlation between the two samples. The p-value for the test of the hypothesis that the two correlations are equal is then easily computed by counting the number of instances where the vector of differences takes a negative value and dividing by ten.

We then proceed to a standard behaviour genetic variance decomposition. The workhorse model in the behaviour genetics literature, known as the ACE model, posits that additive genetic factors (A), common environmental factors (C), and specific environmental factors (E) account for all individual differences in the trait of interest. Start with the case of MZ twins. Let all variables, including the trait, be expressed as deviations from zero and standardize them to have unit variance. Consider a pair of MZ twins and suppose first that the outcome variable can be written as the sum of two independent influences: additive genetic effects, A, and environmental influences, U. We then have that,

$$P = aA + uU,$$

and, using a superscript to denote the variables for twin 2 in a pair:

$$P' = aA' + uU'.$$

Since for MZ twins $A = A'$, the covariance (which, due to our normalization, is also a correlation) between the outcome variables of the two twins is given by,

$$\rho_{MZ} = a^2 + u^2 COV(U, U')_{MZ}.$$

Now consider a DZ pair. Under the assumptions of random-assortative mating with respect to the trait of interest, it will be the case that $COV(A, A') = 0.5$.[9] We then have that,

$$\rho_{DZ} = \frac{1}{2}a^2 + u^2 COV(U, U')_{DZ}.$$

Finally, we impose the equal environment assumption, namely that,

$$COV(U, U')_{MZ} = COV(U, U')_{DZ}.$$

[9] A full derivation of the latter result can be found in any text on quantitative genetics, for instance Falconer (1996) or Mather and Jinds (1982).

Under these, admittedly strong, assumptions it is easy to see that heritability, the fraction of variance explained by genetic factors, is identified as $a^2 = 2(\rho_{MZ} - \rho_{DZ})$. In the standard behaviour genetics framework, environmental influences are generally written as the sum of a common environmental component (C) and a non-shared environmental component (E) such that,

$$P = aA + cC + eE.$$

With this terminology, the environmental covariance component of the trait correlation, $u^2 COV(U, U')$, can be written as c^2, since by definition any covariance must derive only from the common component. This allows us to write the individual variation as the sum of three components a^2, c^2, and e^2; a^2 is the share of variance explained by genetic differences, c^2 is the share of variance explained by common environmental influences, and e^2 the share of variance explained by non-shared environmental influences. There are a number of ways in which the parameters of this model can be estimated. We follow standard practice in using maximum likelihood under the assumption that the outcome variables come from a bivariate normal distribution.[10]. In particular, following directly from the above

[10]Estimation of variance and covariances by maximum likelihood is consistent even if the normality assumption does not hold. However, standard errors will be biased, even though simulations show that small departures from normality are not too great a concern (Enders, 2001). As a robustness check, we have also estimated the model using the estimator of DeFries and Fulker (1985). They propose regressing twin 1's phenotype on: a constant, twin 2's phenotype and twin 2's phenotype interacted with the coefficient of genetic relatedness for the pair in question. DeFries and Fulker (1985) demonstrate that, under the additive genetic model, this produces unbiased estimates of the variance components. Kohler and Rodgers (2001) establish the asymptotic properties of this least square estimator with double entry. Computing standard errors using their method, we obtain heritability estimates that are extremely similar to those reported in the main body of the text.

derivation, the likelihood is maximized under the restriction that the variance-covariance matrix is of the form,

$$\sum = \begin{bmatrix} a^2 + c^2 + e^2 & R_i a^2 + c^2 \\ R_i a^2 + c^2 & a^2 + c^2 + e^2 \end{bmatrix},$$

where R_i takes the value 1 if the observation is of an MZ pair, and 0.5 otherwise. The analyses are run in MPLUS (Muthén and Muthén (2006)), a numerical optimizer often used in behaviour genetics.

3.4 Results

A first diagnostic of genetic influences comes from examining the MZ and DZ correlations. These are reported in Table 3.1. Interestingly, there are no major differences between men and women in the patterns of correlations, with MZ correlations being consistently higher than the DZ correlations. In women the correlations are 0.27 and 0.16. In men, the correlations are 0.29 and 0.13. An MZ correlation, as we have noted, captures all determinants of financial risktaking that identical twins share; that is, genotype and shared environmental influences. In other words, the joint influence of genes and shared environment explains nearly 30 percent of the variation portfolio risk. The correlations for our second risk measure, Risk 2, are very similar, which demonstrates that most variation in risk is driven by differences in the share of equity in the portfolio. Some summary statistics are reported in Table 3.2.

In the two columns of Table 3.3 we report results from the basic model, without age controls (Column 1) and with age controls (Column 2).[11] In the top panel, we report results from a model where variance components are allowed to differ by gender. Similar patterns hold for men and women. Consider for example the results from Model 1. In women, heritability is estimated at 0.22 (99% CI, 0.07-0.31) and in men, heritability is estimated at 0.28 (99 % CI, 0.15-0.32). In both cases, most of the remaining variation comes from non-shared environment.

The lower panel reports results from a model where the restriction that variance components are the same in men and women has been imposed. Whether this restriction entails a significant deterioration in fit can be tested using a likelihood-ratio test. We can reject the hypothesis that the variance components are the same in men and women ($\Delta\chi^2 = 10.14$, $df = 3$, $p < 0.05$) but this is probably a consequence of the large sample size rather than of economically interesting differences. In the pooled model, heritability is estimated at 0.26 (99 % CI, 0.15-0.31) and common environment is estimated at 0.01 (99 % CI, 0.00-0.10). The estimates are very similar when risk residualized on age is used as the dependent variable. Heritability is estimated at 0.22 in women (99 % CI, 0.07-0.31), 0.28 in men (99 % CI, 0.15-0.32) and the pooled model produces a heritability estimate of 0.23 (0.17-0.23).

[11] It is common in behavior genetics studies to residualize the phenotype on age, but interpretational issues arise. For example, age is obviously confounded with cohort effects, so removing age-related variation might actually remove environmental variation inadvertently. Or, gene expression might vary with age, in which case purging the outcome variable from age-related variation might actually have the unintended consequence of removing genetic variation.

3.4.1 Relationship to Previous Findings

As noted above, previous studies have shown that there is moderate parent-child correlation both in attitudes toward risk (Charles and Hurst (2003), Dohmen *et al.* (2006), Hryshko *et al.* (2007)), and in choice of asset holdings (Chiteji and Stafford (1999)), but a parent-child correlation in isolation cannot inform us about the relative importance of genetic and environmental influences. The magnitude of our estimates can easily be reconciled with this existing literature on intergenerational transmission. For instance, the parent-child correlation found in Dohmen *et al.*'s (2006) representative German sample imply upper bounds on heritability of approximately 0.35, and the point estimates of heritability in Cesarini *et al.* (forthcoming) range from 0.14 to 0.35.[12] This convergence of results across different methodologies is reassuring because it suggests that the findings are not driven by confounding factors particular to our study. Such include the fact that our sample is not fully representative (unlike the sample in Dohmen *et al.* (2006)), or the fact that we cannot rule out that twins have communicated about their choice of portfolio (unlike the experimental evidence in Cesarini *et al.* (forthcoming) where twins always participated in the same session).

Furthermore, it is interesting to note that the share of individual variation that is explained by genes as reported above is much higher than the R^2 typically found in standard empirical models of financial risk-taking (Cohn *et al.* (1975), Pålsson (1996), Palme *et al.* (2007)). In an early microdata study of portfolio choice using

[12]If the coefficient of genetic relatedness is 0.5, and only genes explain parent-child resemblance, then doubling the correlation will produce an estimate of heritability. If there are other, non-genetic, forces that can account for the correlation, then heritability estimated from parent offspring correlations will be upward biased.

a non-representative sample, Cohn *et al.* (1975) obtained R^2's of approximately 0.10 using a set of demographic and socio-economic controls, while Pålsson (1996) report substantially lower estimates for Swedish registry based data. Perhaps most strikingly, our single variable A typically explains a substantially larger fraction of individual variation in risk-taking for the Swedish individualized pension accounts than the up to 8 controls in Palme *et al.* (2007), $R^2 \leq 0.042$, and the approximately 20 controls in Säve-Söderbergh (2005), $R^2 \leq 0.112$. A fairly robust finding is that there are differences between men and women in their average propensity to take financial risk (Sunden and Surette (1998), Save-Soderbergh (2008)). In this context it is interesting to note that these sex differences are small relative to the genetic differences within-sex suggested by our estimates.

3.5 Robustness and Generalizeability

To establish how sensitive our results are to variations in the underlying assumptions, we now turn to an examination of the numerous potential sources of bias, their direction, and the extent to which they might be expected to impact our findings.

3.5.1 Representativeness and Generalizeability

In order to ascertain how representative our sample is of the population at large, we compare it disaggregated on zygosity and sex to the Swedish population born between 1938 and 1978 on a number of demographic background variables. The results are reported in Table 3.4.[13] Respondents tend to have slightly higher incomes than the population average, but unlike other studies (Behrman *et al.* (1994), Ashenfelter and Krueger (1994), Rouse (1999)), we do not find any economically interesting attrition with respect to education. There is however a slight tendency for participants to have higher marriage rates than the population as a whole. Finally, STAGE and SALT respondents are also somewhat older than the average for the 1938-1978 cohorts.

Obviously, it is impossible to fully establish the "selectivity" of our sample. The propensity to respond to a survey is likely associated with a number of background characteristics which are not readily measurable but which may nevertheless be influencing our findings, such as general motivational factors. If people with certain background characteristics are overrepresented, and if heritability is associated with these background characteristics, then the heritability estimate will be biased in the direction of this association.

In addition to asking how representative our sample of twins is, it is also important to consider whether twins as a group differ from the population as a whole with respect to unobservables. Few variables have been found to differ between twins and non-twins (Kendler *et al.* (1996)) and we can think of no good reason why the experience of having grown up with a twin should have idiosyncratically affected financial decisionmaking in adult life.

[13]As is common in twin studies, women are slightly overrrepresented (McGue and Tellegen, 1980) in both STAGE and SALT, comprising 53 % of our sample.

3.5.2 Equal Environment Assumption

Critics of the classical twin design cite a number of alleged failures of the equal environment assumption which states that shared environmental influences are not more important for monozygotic twins than for dizygotic twins. A number of objections have been raised, including that parents, on average, give MZ twins more similar treatment.[14] It is important to emphasize that even if MZ twins receive more similar treatment from their parents, this does not in and of itself constitute a violation of the assumption; greater similarity in environment may be caused by the greater similarity in genotypes (Plomin *et al.* (2001)). In the context of research on personality and IQ, where the equal environment assumption has been tested most rigorously, the evidence is fairly convincing that any bias that arises from this restriction is not of first order (Bouchard (1998)).

Most importantly, for measures of personality and cognitive ability, studies of MZ and DZ twins who were reared apart tend to produce estimates of heritability similar to those using twins reared together (Bouchard (1998)). Since studies of twins reared apart do not rely on the equal environments assumption, findings from such studies seem to validate the basic model. Also, in the relatively rare cases where parents miscategorize their twins as MZ instead of DZ (or the converse), differences in correlations of cognitive ability and personality persist (Bouchard and McGue (2003)).

[14] For further criticisms of the equal environment assumption, see Joseph (2002) and Pam et al., (1996), and the references therein.

3.5.3 Reciprocal Influences

The basic model assumes an absence of reciprocal influences between twins. If twins influence each other's choices positively, their degree of similarity will be inflated. Moreover, if this effect is stronger in MZ twins than in DZ twins, it will bias upward the estimate of heritability. The STAGE and SALT datasets both contain information on the frequency of contact between twins. As is commonly found in twin studies, monozygotic twins do interact more than dizygotic twins. On average, MZ pairs reported 3.3 interactions per week at the time of the survey, whereas DZ pairs reported an average of 1.8 interactions per week.[15]

Running separate regressions by gender, where the dependent variable is the squared within-pair difference in portfolio risk, and the independent variables are frequency of contact and zygosity, frequency of contact is a significant predictor of within twin-pair squared difference in portfolio risk, for both men and women. The presence of a statistically significant effect does not, however, prove that the frequency of contact is causing increased similarity. Much research has been devoted to establishing the direction of causality. Lykken *et al.* (1990) and Posner *et al.* (1996) offer some evidence suggesting that twins similar in personality tend to stay in contact with one another, and not the other way round.

One crude way of examining whether twins have communicated about their choice of funds is to ask how common it is for both twins to choose the same portfolio. Excluding pairs where both twins selected the default portfolio, of the

[15]We construct the frequency of contact variable as follows. Subjects who report seven or more interactions (by e-mail, telephone or letter) per week are assigned a value of 7. All other subjects are assigned the number of interactions per week that they report. If we have data on both twins, we use the mean of the two reports.

remaining MZ twins, 8 % choose the same portfolio as their co-twin. In DZ twins the corresponding figure is 3 %. To further examine the sensitivity of our results to this source of bias, we conduct two robustness checks, the results of which are reported in Table 3.5.

First, we drop all pairs in which both individuals chose the same portfolio, and rerun the analyses. Obviously, by discarding these observations, both MZ and DZ correlations will drop. Furthermore, these adjusted correlations will be downward biased if twins choosing identical portfolios are more similar than average with respect to risk-preferences. This sample restriction produces a pooled heritability estimate of 0.20 (99% CI, 0.11-0.23) which, under the assumption that communication only affects choices through identical portfolios, can serve as a lower bound to our heritability estimate in the presence of reciprocal action.

Second, we make use of our frequency of contact variable. Specifically, we stratify frequency of contact into 15 groups, and for each sex and level of contact we then randomly drop the required number of either MZ or DZ pairs to make the number of MZ and DZ pairs equal. In this restricted sample, the distribution of frequency of contact is, by construction, virtually the same in the MZ and DZ groups. Rerunning the analyses on this subset of the data, the heritability estimate in the pooled model falls to 0.19 (99% CI, 0.07-0.28). The finding that the heritability estimates only fall marginally is reassuring since it demonstrates that frequency of contact is not a major influence on our main result.[16]

Our interpretation of these results is that the twins who opted for the same

[16] A significant drop in estimated heritability is, however, a necessary but not sufficient condition for frequency of contact to be the cause of greater similarity.

retirement fund would generally have chosen portfolios with similar levels of risk even without the opportunity to consult each other.

3.5.4 Misclassification and Measurement Error

We use the Swedish Twin Registry's standard algorithm to establish zygosity. The algorithm has been validated against DNA-based evidence, and studies show that misclassification is typically of the order 2-5 % (Lichtenstein *et al.* (2006)). Purely random assignment error would bias heritability downward, since the difference in genetic relatedness between pairs assigned as MZ or DZ would diminish to less than one half. However, misclassification is probably not random, but related to physical similarity (notice that the questions we use to establish zygosity are solely based on assessments of physical similarity). The relevant question is then if physical similarity is somehow related to the similarity with respect to behaviour. The classical reference on this topic is Matheny *et al.* (1976), who administered two intelligence tests, two perceptual tests, one reading test, one test of speech articulation, and one personality inventory to twins and found that "correlations revealed no systematic relation between the similarity of appearance and the similarity of behaviours for either the identical twin pairs or the same-sex fraternal twin pairs."[17] We conclude that the bias which arises due to misclassification is likely small and leads to an understatement of heritability.

As in the case of misclassification, measurement errors tend bias a^2 and c^2

[17] More recently, Hettema, Neale and Kendler (1995) report no significant associations between physical similarity and phenotypic resemblance in four out of the five psychological disorders they consider (the one exception is bulimia.)

downwards since any such error will be subsumed under the estimate of e^2. In the simplest case where the preference is observed with a mean zero random error with variance σ_ϵ^2, it is easy to show that the estimates of a^2 and c^2 need to be scaled up by a factor of $\frac{1}{1-\sigma_\epsilon^2}$. But, whereas measurement error is easy to conceptualize in psychometric research as the test-retest reliability of some instrument designed to measure a personality trait, it is less clear in the present case where it presumably would involve the choice of actual portfolio risk to be related to factors other than risk preferences. While this is certainly likely to be the case, it is far from obvious how the reliability of actual observed risk-taking in the field convincingly could be tested.

3.6 Heritability and the Biological Basis of Risk-Taking

It is important to emphasize two features of this behaviour genetic model when interpreting our findings. First, the model produces estimates of the proportion of variance explained. Thus, it does not shed any direct light on the determinants of the average proclivity to take risks. This distinction is important. For instance, if genetic transmission in a studied population is uniform, then a trait that is primarily acquired through genes might actually show a low, or zero, estimate of heritability. Alternatively, consider a culturally homogenous environment with little variation in how parents, whether consciously or not, instill certain beliefs and values in their children. In such an environment, it is quite possible that common environmental influences are important determinants of the average propensity to take financial risks, but that differences in common environmental influences are not an important source of variation. Second, the model is based on strong functional form and independence assumptions.

Reassuringly, our results are in line with the very voluminous and closely related behaviour genetic literature on personality and attitudes (Bouchard (1998), Bouchard and McGue (2003), Plomin *et al.* (2006)), much of which has employed other types of sibling relations. While not measuring risk preferences *per se*, several suggested dimensions of personality are thought to be correlated with actual risk-taking. In a recent metastudy of parent-child resemblance in personality, Loehlin (2005) report average correlations of 0.13 for personality and 0.26 for attitudes in families with children reared by their biological parents.[18] As with our findings, twin and adoption studies strongly suggest that the primary explanation for these correlations is genetic transmission (Bouchard and McGue (2003), Loehlin (2005)). For instance, the correlations for personality and attitudes are 0.04 and 0.07 respectively between adopted children and their non-biological parents, but 0.13 and 0.20 between adopted children and their biological parents (Loehlin (2005)). Thus, seen in the context of the behaviour genetic literature there is nothing anomalous about the finding of moderate heritability, a low effect of shared environment, and a large effect of non-shared environment for finan-

[18]Loehlin (2005) distinguishes young children from other children. When he only considers young children, the association between non-biological, but rearing, parents and their children is stronger. This finding that is consistent with the literature documenting increasing heratibility in adolescence (Bouchard and McGue 2003). Notice also that Loehlin's parent-offspring correlations yield considerably lower estimates of heritability than estimates based on samples of twins. He suggests that the difference is accounted for by non-additivity.

cial risk-taking. Indeed, these findings match Turkheimer's (2000) three laws of behaviour genetics perfectly.[19]

The fact that a trait is heritable does not imply that there are genes with a direct effect on the trait. However, sensation and novelty seeking are both heritable and presumably correlated with risk-taking, and molecular genetic studies have implicated a number of particular genes associated with these traits (Koopmans *et al.* (1995), Zuckerman and Kuhlman (2000), Munafò *et al.* (2002), Kreek *et al.* (2005), Boyer (2006)). In addition to particular genes, several studies have found significant relationships between risk-taking and other biological factors such as patterns of brain activation and testosterone levels (Kuhnen and Knutson (2005), Cardinal (2006), Apicella *et al.* (2008)). It is worth noting that hormone levels (Harris *et al.* (1998)) and brain structure (Toga and Thompson (2005)) are both heritable, providing some indirect support for our hypothesis.

Yet, it seems very likely that some of the genetic effects may operate through genome-wide influences on variables which in turn affect risk-taking. For instance, one early paper found that participants' education and income levels were related to asset allocation decisions in mandatory private savings accounts, with less educated and lower income participants being less inclined to invest in equity securities (Poterba and Wise (1996)), although this finding is not supported by Palme *et al.* (2007). Differences in financial fluency is another candidate variable (Bhandari and Deaves (2007)).[20].

[19] Turkheimer's three laws are the following. First, all human behavioral traits are heritable. Second, the effect of being raised in the same family is smaller than the effect of genes. Third, a substantial portion of variation in complex human behavioral traits is not accounted for by the effects of genes and family.

[20] See Benartzi and Thaler (2001) for some evidence suggesting that individuals apply a diver-

3.7 Conclusion

In this chapter, we have matched data on the mandatory pension investment decisions made in the fall 2000 to the Swedish Twin Registry in an attempt to estimate the genetic influence on variation in financial risk-taking. Relative to the experimental and survey evidence reported in Cesarini *et al.* (forthcoming), a distinct advantage of our approach is that we examine risk-taking behaviour in a field setting with large financial incentives attached to performance. Furthermore, relative to Dohmen *et al.* (2006), a second advantage of our approach is that the use of twin data allows us to shed light on the relative importance of environmental and genetic differences as sources of variation. Our finding that approximately 25% of variation in portfolio risk is due to genetic influences is in line with this previous, but small, experimental and intergenerational literature as well as the behaviour genetics literature in general. The explanatory power of the genetic effect that we find is also typically at least twice as large as the R^2s found in previous non-twin studies using the same data and up to as many as 20 controls. In short, this study is the first to use behaviour genetic techniques to document the heritability of risk-taking in financial markets, as well as outside the laboratory, and the results strongly suggest that genetic variation is an important source of individual heterogeneity.

In addition to exploring specific mechanisms, we can think of a number of avenues for further work along the lines of this study. Constructing a dataset

sification heuristic which is inconsistent with mean-variance optimizing behavior. In particular, individuals overinvest in asset types that are overrrepresented in the menu of funds.

similar to ours but with adoptees instead of twins would provide more precise estimates of the relative importance of common environmental influences. Also, augmenting the twin dataset with other sibling types of varying degrees of genetic relatedness, and, ideally, rearing environments, would allow researchers to explore the possibility of non-additivity or to test the equal environment assumption (See, *e.g.*, Björklund *et al.* (2005)) . If other work on attitudes and personality provides any guidance, we would expect that some of the genetic influences reported in this study are in fact non-additive.[21] Regardless of what evolutionary dynamics led to the genetic variation that we observe for preferences in financial risk-taking[22], the fact is that genetic differences explain a large share of individual variation in risk-taking. In light of these findings, we suggest that the further study of the biological and genetic basis of human risk-taking behaviours will lead to a more comprehensive theory of financial decision-making.

[21]The basic ACE-model - like most behavior genetic models - assumes that genes influence the trait in an additive manner. That is to say, the genetic effect is simply the sum of all individual effects. This rules out epistasis (interaction between alleles at different loci) and dominance (interaction between alleles at a locus). A possible way to test for this would be to extend the dataset to include also sibling, parent-child, or even cousin data. The correlation between siblings, under an additive model, ought to be at least half the heritability obtained from a twin study. Were this assumption to fail, it would be diagnostic of some non-additivity being present in the data. This issue is explored in Loehlin (2005) in the context of the heritability of personality.

[22]Dall *et al.* (2004) and Penke *et al.* (2007) are two recent papers exploring the issue of how genetic variation can be maintained.

3.8 Tables

Table 3.1: Within-pair correlations

		Women		p-value of diff	Men		p-value of diff
		MZ	DZ		MZ	DZ	
Risk 1	Pearson	0.27	0.16	<0.01	0.29	0.13	<0.01
	Spearman	0.28	0.16	<0.01	0.30	0.13	<0.01
	# pairs	3346	3878		2747	3591	
Risk 2	Pearson	0.26	0.13	<0.01	0.24	0.11	<0.01
	Spearman	0.26	0.14	<0.01	0.23	0.10	<0.01
	# pairs	3346	3878		2747	3591	

Note. One sided p-values testing the equality of MZ and DZ correlations are reported.

Table 3.2: Summary statistics for risk measures

	Women		Men		Total	
	MZ	DZ	MZ	DZ	MZ	DZ
Risk 1	19.0	18.7	19.3	19.1	19.2	18.9
S.D.	(4.2)	(4.4)	(4.4)	(4.6)	(4.3)	(4.5)
Risk 2	0.77	0.77	0.81	0.80	0.79	0.78
S.D.	(0.34)	(0.34)	(0.32)	(0.33)	(0.33)	(0.34)
Active Choice	0.72	0.69	0.71	0.67	0.71	0.68
S.D.	(0.36)	(0.36)	(0.36)	(0.35)	(0.36)	(0.36)
# observations	6692	7756	5494	7182	12186	14938

Note. Standard deviations in parenthesis.

Active Choice is a binary variable taking the value 1 if individual made an active portfolio investment decision and 0 otherwise.

Table 3.3: Results of the ACE model

			Model 1	Model 2
Separate				
		a^2	0.22** (0.07-0.31)	0.22** (0.08-0.28)
	Women	c^2	0.04 (0.00-0.17)	0.00 (0.00-0.10)
		e^2	0.73** (0.68-0-78)	0.78** (0.72-0.83)
		a^2	0.28** (0.15-0.32)	0.24** (0.17-0.29)
	Men	c^2	0.00 (0.00-0.08)	0.00 (0.00-0.02)
		e^2	0.72** (0.68-0.78)	0.76** (0.71-0.82)
	$ln(L)$		-180051.48	-121625.98
Pooled		a^2	0.26** (0.15-0.31)	0.23** (0.17-0.27)
		c^2	0.01 (0.00-0.10)	0.00 (0.00-0.04)
		e^2	0.73**(0.69-0.76)	0.77** (0.74-0.80)
	$ln(L)$		-180056.55	-121636.13

* 99% confidence intervals within parentheses; A is the genetic contribution; C is the common environment contribution; E is the unique environment contribution. Two stars denote statistical significance at the 1 % level, and one star denotes statistical significance at the 5 % level. Dependent variable is Risk 1. The top panel contains results from a model where separate variance components are estimated for women (subscript w) and men (subscript m). The lower panel reports a restricted model where $a_w^2 = a_m^2$, $c_w^2 = c_m^2$ and $e_w^2 = e_m^2$. All models are estimated allowing the mean and the variance to differ by gender. Confidence intervals are constructed using the bootstrap with 1000 draws. Model 1 is the baseline model without age moderation. Model 2 is the baseline model where the risk measure has been residualized on age.

Table 3.4: Background variables

	Women		Men		Population		
	MZ	DZ	MZ	DZ	Women	Men	Total
Income	234	231	326	325	210	288	251
S.D.	111	108	216	293	-	-	-
Education (years)	12.3	11.9	12.0	11.6	12.2	11.9	12.11
S.D.	2.6	2.7	2.8	2.9	-	-	-
Marital Status	0.52	0.55	0.55	0.56	0.52	0.48	0.50
S.D.	0.50	0.50	0.50	0.50	-	-	-
Age	48.7	51.8	50.1	52.8	46.6	46.5	46.6
S.D.	11.3	10.0	10.9	9.2	-	-	-

Note. Income is in 1000 SEK (1 USD = 6 SEK). Population mean is defined as the average for individuals born 1938 to 1978. Education refers to years of education. Marital status is a variable taking the value 1 if the individual is married. All data is for the year 2005 and population means were computed using data from Statistics Sweden.

Table 3.5: Robustness checks of the ACE model

			Dropped	Matched
Separate				
		a^2	0.16** (0.01-0.22	0.15* (0.00-0.29)
	Women	c^2	0.01 (0.00-0.11)	0.10 (0.00-0.23)
		e^2	0.83** (0.78-0.88)	0.75 (0.69-0.82)
		a^2	0.23** (0.13-0.28)	0.23* (0.00 -0.32)
	Men	c^2	0.00 (0.00-0.06)	0.03 (0.00-0.17)
		e^2	0.77 (0.72-0.83)	0.74 (0.68-.0.82)
	$ln(L)$		-147933.61	-91029.09
Pooled		a^2	0.20** (0.11-0.23)	0.19** (0.07-0.28)
		c^2	0.00 (0.00-0.06)	0.07 (0.00-0.17)
		e^2	0.80** (0.78-0.84)	0.75** (0.70-0.79)
	$ln(L)$		-147937.36	-91033.66

* 99% confidence intervals within parentheses; A is the genetic contribution; C is the common environment contribution; E is the unique environment contribution. Two stars denote statistical significance at the 1 % level, and one star denotes statistical significance at the 5 % level. In the "Dropped" column, pairs where both twins selected identical portfolios are excluded. In the "Matched" column, we stratified the data by frequency of contact into 15 groups, and for each sex and level of contact we then randomly dropped the required number of either MZ or DZ pairs to make the number of MZ and DZ pairs equal. In this restricted sample, the distribution of frequency of contact is, by construction, virtually the same in the MZ and DZ groups. The top panel contains results from a model where separate variance components are estimated for women (subscript w) and men (subscript m). The lower panel reports a restricted model where $a_w^2 = a_m^2$, $c_w^2 = c_m^2$ and $e_w^2 = e_m^2$. All models are estimated allowing the

mean and the variance to differ by gender. Confidence intervals are constructed using the bootstrap with 1000 draws.

Chapter 4

The heritability of entrepreneurship - men are not like women

4.1 Introduction

In entrepreneurship research, there is by now a rich body of evidence which documents a strong connection between parental self-employment and the self-employment of their children (de Wit and Van Winden (1989), Fairlie (1999), Taylor (1996), Uusitalo (2001)). Yet, the mechanism of intergenerational transmission is frequently left unspecified, and is usually assumed to be environmental. For instance, Falter (2007) highlights that children acquire "specific skills learned from self-employed parents during their youth" and "market imperfections" (p. 132), while Lentz and Laband (1990) assert that the father "passes on to his son valuable human capital about running a business operation; the son acquires this integrated, managerial human capital as a by-product of growing up" (Lentz and Laband (1990), p. 564). In a similar spirit, Krueger (1993) and Sorenson (2007) interpret the higher probability of self-employment solely in terms of information acquisition during childhood and other environmental variables. Another environmental source of intergenerational transmission of owner-manager status, the inheritance of family firms, is analysed in Burkart *et al.* (2003), Bertrand and Schoar (2006), and Bennedsen *et al.* (2007)[1].

In a recent string of papers, however, this environmental approach is complemented by investigation and discussion of the biological basis of entrepreneurial behaviour (Nicolaou *et al.* (2008), Nicolaou and Shane (2008), White *et al.* (2006)). Many behavioural traits have been shown to be under considerable genetic influence (Plomin et al. (2001), and a priori, there are no reasons to expect the contrary for entrepreneurial behaviour. In fact in behaviour genetics, the ubiquity of genetic influences on most behavioural traits has become so well established that it is now elevated to the status of a law (Turkheimer (2000).

Nicolaou *et al.* (2008) examine survey data from a sample of UK twins. Their primary proxy variable for entrepreneurship is self-reported self-employment status. Applying the classical twin design, which relies on comparing the within-pair correlations of a trait among monozygotic and dizygotic twins, they estimate the heritability of entrepreneurship, *i.e.* the share of variation which can be explained by genetic differences. Nicolaou *et al.* (2008) find that this share is moderately

[1]Bennedsen (2007) examine CEOs in listed companies, who arguably should not be labelled self-employed even when originating in a family of owners.

high, around 48% for their primary measure, and in a companion paper, Nicolaou and Shane (2008) pontificate about the possible biological pathways that this correlation might take.

The results in Nicolaou *et al.* (2008) are however based on a sample including virtually only women (92-94%). Several authors have noted that female entrepreneurship is structurally different from that of men in a number of ways (Du Rietz and Henrekson (2000), Anna *et al.* (1999)). A major difference is that there are large differences in the shares of women and men who become entrepreneurs. Furthermore, the sectoral composition differs by sex, with female entrepreneurs more likely to operate in service industries such as health and catering. Finally, on average, the companies operated by female entrepreneurs are smaller in terms of revenue and number of employees (Holmquist and Sundin (2002)). These empirical patterns are not unique to Sweden; similar patterns have been observed in other countries, *e.g.* the US (Anna *et al.* (1999)). Such differences thus suggest that the role of genetic differences in explaining individual variation in entrepreneurial behaviour may well vary by sex. Not only do the mean levels of entrepreneurial activity differ quite dramatically by sex, with much fewer women than men being entrepreneurs; the average female entrepreneur is quite different than the average male entrepreneur.

To investigate if there are differences in heritability, I replicate the Nicolaou *et al.* (2008) study using a large sample of same-sexed Swedish twin pairs. I find large, but not statistically significant, differences in the heritability for men and for women. For women, my results are broadly consistent with those reported in

Nicolaou *et al.* (2008), with point estimates in the range 0.35 to 0.62, depending on how entrepreneurship is operationalised. For men, I find lower, and statistically insignificant, estimates of heritability, ranging from 0.00, for my main measure of entrepreneurship, to 0.16 for one of the alternative measures.

These results suggest that it may not be appropriate to extrapolate heritability estimates based on samples of women to men. At a more general level, my findings indicate that, at least in Sweden, differences in common environment are a much more important source of variation in the entrepreneurial activities of men than in the entrepreneurial activities of women.

In the next section, I present the model used to estimate heritability from data on twins reared together. In Section 4.3, I describe the data, followed by the results in Section 4.4, and robustness checks in Section 4.5. Section 4.6 contains a discussion of how to interpret these findings, and Section 4.7 concludes.

4.2 Theoretical framework

Heritability refers to the share of variance in a trait which can be explained by genetic differences in a given population. It is thus important to note that heritability says nothing of the average occurrence of a trait in an environment. For example, the number of digits on a human hand is clearly under strong genetic influence, yet the share of variance explained by genetic variation is probably low, as virtually all people have genes coding for exactly five digits.

The standard ACE model assumes that genetic determinants are additive, so that there are no interaction effects neither between alleles at the same locus (dominance) nor at different loci (epistasis)[2]. Consider an indicator variable P^* taking the value 1 if an individual is self-employed, and let P be a normally distributed latent propensity to become self-employed, such that:

$$P^* = 1, \qquad \text{iff } P \geqslant T,$$

Where T is the threshold level of P for becoming self-employed.

We then apply a standard behaviour genetic variance decomposition of the latent trait P. The workhorse model in the behaviour genetics literature, known as the ACE model, posits that additive genetic factors (A), common environmental factors (C), and specific environmental factors (E) account for all individual differences in the trait of interest. Start with the case of MZ twins.

Consider a pair of MZ twins and suppose first that the latent trait, P, can be written as the sum of two independent influences: additive genetic effects, A, and environmental influences, U. We then have that,

$$P = A + U,$$

and, using a superscript to denote the variables for twin 2 in a pair,

$$P' = A' + U'.$$

Since for MZ twins $A = A'$, the covariance between the traits of the two twins

[2]See Mather and Jinks (1982) for a careful explanation of the biological underpinnings of behavioural genetics models.

is given by,

$$COV(P,P')_{MZ} = \sigma_A^2 + COV(U,U')_{MZ.}$$

Now consider a DZ pair. Under the assumptions of random-assortative mating with respect to the trait of interest, it will be the case that $COV(A,A')_{DZ} = \frac{1}{2}COV(A,A')_{MZ}$.[3] We then have that,

$$COV(P,P')_{DZ} = \frac{1}{2}\sigma_A^2 + COV(U,U')_{DZ}.$$

Finally, impose the equal environment assumption, namely that,

$$COV(U,U')_{MZ} = COV(U,U')_{DZ}.$$

Under these, admittedly strong, assumptions it is easy to see that heritability, the share of variance explained by genetic factors, is identified as $2(\rho_{MZ} - \rho_{DZ})$.[4] In the standard behaviour genetics framework, environmental influences are commonly written as the sum of a common environmental component and a non-shared environmental component such that,

$$P = A + C + E.$$

With this terminology, the environmental covariance component of the phenotypic correlation, $COV(U,U')$, can be expressed as the variance of C, σ_C^2, since

[3]A full derivation of the latter result can be found in any text on quantitative genetics, for instance Falconer and Mackay (1996) or Mather and Jinks (1982).

[4]Dividing by the variance of P, σ_P^2, and solving for σ_A^2/σ_P^2.

by definition any covariance must derive only from the common component. The model thus implies a variance-covariance matrix of the form,

$$\sum = \begin{bmatrix} \sigma_A^2 + \sigma_C^2 + \sigma_E^2 & R_i\sigma_A^2 + \sigma_C^2 \\ R_i\sigma_A^2 + \sigma_C^2 & \sigma_A^2 + \sigma_C^2 + \sigma_E^2 \end{bmatrix},$$

where R_i takes the value 1 if the observation is of an MZ pair (since MZ twins have identical genes), and 0.5 otherwise (since under the assumption of additivity and random assortative mating, the genotypic correlation in DZ twins is 0.5).

I set the variance of P to 1, and use a robust weighted least-squares mean- and variance adjusted (WLSMV) estimator (Muthen and Satorra (1995), Muthen *et al.* (1997))[5], to estimate the variances of A, C, and E, and the value of the threshold, T, in units of σ_P^2. With this normalisation, σ_A^2 is our main object of interest, the share of variance explained by genetic factors, and σ_C^2 is the share of variance explained by environmental factors common to both twins in a pair. Finally, σ_E^2 is the share of variance explained by non-shared environmental influences. Confidence intervals are estimated using the bootstrap with 5000 draws with replacement.

4.3 Data

My sample of twin pairs is drawn from the Swedish Twin Registry, which is the largest registry of its type in the world. This registry was assembled for the purpose of conducting epidemiological studies, and contains virtually all twins

[5]See Flora and Curran (2004), for an evaluation of this and similar estimation methods.

born in Sweden from 1926 onwards. In late 2005, a survey was conducted aimed at twins born in 1959-1985, the STAGE survey (The Study of Twin Adults: Genes and Environment). All twins in the Registry were contacted by mail and invited to participate in a web-based survey. To non-respondents, several reminders were sent out by mail and email, and the possibility of submitting answers by phone was offered. The final response rate was 60%, out of the 42,582 twins in the registry in this age group. Zygosity was determined by the Swedish Twin Registry by aggregating answers to a battery of questions concerning intrapair physical similarity in childhood. The validity of this method of determining zygosity has been repeatedly estimated to be 95-98% (Lichtenstein *et al.* (2002)).[6]

My main measure of entrepreneurship is very similar to the one used by Nicolaou *et al.* (2008), who take the answer to the survey question "Are you self-employed" as their main measure. My primary measure of entrepreneurship is based on an individual's answer to the following multiple choice question "Describe your work situation by entering the option(s) which fit your current situation". Individuals who ticked the box "Full-time managing own business" are coded for the purposes of this study as entrepreneurs. To analyze whether results are robust across different operationalisations of entrepreneurship, I also use answers to two other questions that were administered in the survey. The first question asks the individual about his or her main mode of employment for the last three years. Individuals who respond "Self-employed/Co-owner of company" are coded as entrepreneurs. The second question asks an individual who his or

[6]For a description of the STAGE survey, see Lichtenstein et al. (2006).

her main employer was during the last three years. Individuals who answer "own company" are coded as entrepreneurs. To preview our results, the three measures yield qualitatively very similar estimates of heritability, and the polychoric correlation between an individual's answer to any two of the three questions always exceeds 0.91. Ideally, we would like some measure which more closely captures the essence of entrepreneurial behaviour, as in some of Nicolaou *et al.*'s (2008) auxiliary measures, rather than merely measuring self-employment. However, it should be noted that the estimated heritabilities reported in Nicolaou *et al.* (2008) do not vary substantially across operationalisations.

In Table 4.1, summary statistics for the three questions are reported. Approximately 5% of women and 11.5% of men are entrepreneurs under our primary measure. These figures are roughly consistent with the overall population share of self-employed individuals in 2005 among individuals aged 25-44. According to Statistics Sweden, the share of men who were self-employed in this cohort was 10.1%, and the share for women was 3.6%.

Table 4.2 contains a comparison of demographic characteristics between the full sample - pairs where for any of the three questions regarding self-employment, as described above, both twins responded - and the total population aged 25-44. As is common in survey data, the respondents are higher than average on income, education, and marital status, although differences are not large. In the case of income, the full sample mean is approximately 10% above the population mean, and the average years of education is about 0.5 years longer.

STAGE contains data on 12124 individuals from complete same-sex pairs. Of these, 2915 individuals from 2170 pairs did not answer any of the three questions on self-employment, despite answering other question in STAGE. It should be noted that the list of response options for my primary measure ("Describe your work situation by entering the option(s) which fit your current situation") is meant to be exhaustive, and includes response options such as "unemployed", "parental leave", and "disability pension".

4.4 Results

In Table 4.3, I present simple within-pair polychoric correlations in self-employment for MZ and DZ twins, by gender. MZ twins have similar correlations for men and women and across my two measures of self-employment, in the range 47-64%. DZ twins however vary substantially between men and women. Whereas for men correlations are roughly the same as for male MZ twins, female DZ twins tend to have markedly lower correlations than female MZ twins.

In Tables 4.4 and 4.5, I report estimates of the ACE model for my two measures of self-employment, current and last three years, for women and men separately. My estimates for females are similar to those of Nicolaou and Shane (2008), who estimate heritability for current self-employment, their main measure, at 48%. My estimate for current self-employment is 35%, and for self-employment in the last three years 61-62%, both of which are within the 95% confidence interval reported by Nicolaou *et al.* (2008).

For males, however, I find that heritability is very low, 0 % for current self-employment, and 0.00 and 0.16 for the measures based on the last three years.

None of these three estimates is statistically significantly different from zero. Instead, common environment (C) accounts for about 30-59% of the variation in male self-employment, depending on measure. As is typical in behaviour genetics decompositions, the role of unique environment is large, both for men and for women, ranging from 38% to 55%.

Table 4.6 provides confidence intervals for the difference in estimated heritability between men and women in my sample. As with the confidence intervals of the model parameters as provided above, we estimate this using a bootstrap estimator with 5000 draws. Applying a two-sided test, we cannot reject the hypothesis of equal heritability in men and women, at the 5% level, despite the large differences in estimated heritability, although for both main employer and main mode of employer the 2.5 percentile is close to zero.

4.5 Assortative mating

To assess the overall sensitivity of our results to the assumption of no assortative mating, we re-estimate the model assuming various nonzero parent correlations in latent entrepreneurial propensity. Tables 4.7 and 4.8 contain estimates based on parental correlations, λ, of 0.1, 0.2, and 0.3, for men and women separately. In the variance-covariance matrix above, these correlations correspond to values of R_i for dizygotic twins of 0.55, 0.6, and 0.65, respectively. Estimated heritabilities increase with increasing degrees of assortative mating, although for all three measures, and for both women and men, this effect is small. The largest change is for the heritability of current self-employment for women, where the estimate increases from 0.39 to 0.47 when changing from the assumption of no assortative mating ($\lambda = 0$) to a degree of assortativity of 0.3 ($\lambda = 0.3$). The estimated magnitude of the difference is very large in all cases, ranging between 0.39 and 0.59.

Table 4.9 contains confidence intervals for the difference in estimated heritability between men and women under the above alternative assumptions. The degree of assortative mating has virtually no impact on the confidence intervals for the heritability differences, which again remain statistically insignificant for current self-employment, and only bordering to significant at the 5% level for main employer and mode of employer.

4.6 Discussion

The most notable aspect of these results is the sharp contrast between finding, as in Nicolaou *et al.* (2008), moderate heritability (35% in our main specification) among women yet virtually no heritability among men. Although the difference in point estimates is not statistically significant at conventional levels, these results nevertheless indicate that caution must be exercised before claiming moderate and robust heritability in entrepreneurship. In essence, our results suggest that whereas genetic differences may be an important source of individual variation in the entrepreneurial activity of women, there is no support for this statement for men. This finding highlights the population-specificity of heritability estimates, and the importance of exercising care when extrapolating findings from one sample to another.

What, apart from sampling variation, might account for such sharply different heritability estimates? One interpretation is that our primary measure of entrepreneurship captures different phenomena in women than it does in men. As we have noted, the share of women who are entrepreneurs is much lower than the corresponding figure for men, and attempts to characterize female entrepreneurship have revealed that there are large structural differences with respect to firm size, sector and managerial strategies. Such differences obviously render extrapolations from estimates based on a sample of women to men difficult. Our findings of very different heritabilities reinforce the point that male and female entrepreneurship differs along important dimensions and that distinct mechanisms may explain within group variation.

Following the work of Hisrich and Brush (1983), a number of papers have sought to characterize "male" and "female" entrepreneurship. The following stylized facts emerge from these papers. Businesses run by women tend to grow slower, have fewer employees and lower revenue (Brusch (1992), Chaganti and Parasuraman (1996), Rosa, Carter, and Hamilton (1996), Fischer *et al.* (1993)). To a significant extent, it is quite possible that these differences are driven by differences in preferences and motivations for undertaking entrepreneurial activity (Du Rietz and Henrekson (2000)) and industry selection (Holmquist and Sundin (1988)). For instance, Stevenson (1986) reports that women value occupational flexibility more than men and that the desire to achieve a work life balance is a more commonly stated reason for starting a business (Carter *et al.*, 2003), Boden (1999)). Some women also pursue self-employment because it allows them to work at home. Welter (2001) reports evidence suggesting women are less interested in developing their own businesses. Female entrepreneurs are also much less likely to state that a desire to build a company or make money motivated the decision to start a business (Georgellis and Wall (2004)). Research has also shown that women are less likely to seek equity capital (Orser *et al.* (2006)). Finally, Burke *et al.* (2002) found that the desire to be one's own boss had a three times greater effect on the probability of self-employment in men compared to in women. Interestingly, Burke *et al.* (2002) also found that it was only for men that such a desire was significantly associated (positively) with firm value. I conclude that female entrepreneurship is a distinct phenomenon from that of men. There are fewer female entrepreneurs, and, on average, their companies differ markedly from the companies operated by men.

Yet, it seems plausible that there is at least some overlap in the causes of male and female self-employment. If so, any large differences in estimates beg for an explanation. The considerably higher estimates of c^2 in men suggest that differences in family environment are a more important source of variation in self-employment of men. One possible explanation for this would be if family firms get passed on and divided among male twins more often than among female twins. Of course, this begs the question why correlations among monozygotic twins are no higher for males than for females. As I do not have data on family firms getting passed on, nor any information on the employment status of the previous generations, this is an avenue of inquiry which I will have to leave for further research. In this context, it deserves mentioning that although my survey data is based on *self*-employment and *own* business, it is possible that some twins have also included businesses managed in cooperation with their co-twin. This would bias our estimates, unless twin-twin interaction was assumed negligible so that twins only formed business relationships in cases where either would have been self-employed even absent their co-twin.

My results shed some light on the plausibility of some of the genetic pathways presented by Nicolaou and Shane (2008) as potential mediators of heritability in entrepreneurship. Nicolaou and Shane (2008) propose three personality traits which have been shown both to be under genetic influence and to be associated with entrepreneurial behaviour: achievement motivation, extraversion, and locus of control[7]

Finkel and McGue (1997) examine about 4000 twins and other relatives and find no evidence for sex differences in achievement motivation. Finkel and McGue (1997) also provide a survey of results on such differences on various other personality measures. Although they do conclude that "there is some suggestion of higher heritability in women for some personality traits", this is not based on any of the three traits proposed by Nicolaou and Shane (2008). Eaves *et al.* (1998) examine the heritability of extraversion using on a sample of more than 12000 pairs of twins from Australia, Finland, and USA. They find no consistent differences in the heritability of extraversion between men and women, although they do find that epistatic (non-additive) effects are appreciable. In an earlier study

[7]Evidence on the link between these personality traits and entrepreneurial behaviour can be found in Rauch and Frese (2000) (locus of control), Burke et al. (2000) (extraversion), and Collins et al. (2004) and Stewart and Roth (2007) (acheivement motivation)

using Swedish data, Eaves *et al.* (1989) find slightly higher heritability of extraversion for women than for men, whereas Macaskill *et al.* (1994) report higher male heritability for extraversion based on an Australian sample. In summary, the possibly differential heritability in entrepreneurship reported in this chapter is unlikely to be rationalised by sex differences in the heritability of extraversion or acheivement motivation. [8]

Large differences between heritability in men and women, respectively, would indicate a complex gene-to-outcome mechanism. Finding a candidate set of genes is therefore likely to be even harder than for example for other complex traits such as risk-taking. This is in itself not surprising, given the widely differing sets of skills and abilities which have been associated with the typical entrepreneur. In this context, it deserves mentioning that high heritability for a trait does not guarantee that it will be possible to identify specific genes that contribute to it. For example, molecular genetics analysis has had more success explaining conditions with low estimated heritability - such as breast cancer and colon cancer - than it has for diseases such as multiple sclerosis and schizophrenia, with considerably higher estimated heritabilities (Risch (2001)).

The results in this chapter are estimated under some potentially quite restrictive assumptions. To what extent are these conclusions robust to relaxing these assumptions? First, the assumption of no non-additive genetic effects is necessary to identify the model when using only non-adopted twin data. In general, when the data contains a richer set of sibling types, the validity of this assumption

[8] I am not aware of any study on sex differences in the heritability of locus of control to date.

can be verified by comparing the estimates obtained from only using non-adopted twins, to those were the full dataset – and hence less restrictive assumptions – have been employed. Coventry and Keller (2005) survey behavioural studies on 17 traits where this type of comparison has made and conclude that: (i) twin studies tend to have a small upward bias of on average roughly one sixth (in their study, an average of 0.39 compared to the true average of 0.33), in their estimated total genetic effect, and (ii) twin studies tend to underestimate the role of common environmental factors. As our main result is that heritability in men is vastly lower, and in no case statistically significantly different from zero at the 5%-level, than heritability in women, this would appear to be robust to the restrictiveness of the twin model. Our estimates, however, have large standard errors and the hypothesis of equal heritability among men and women could not be rejected at the 5% level. The findings of Keller and Coventry (2005) imply that the point estimates must be interpreted with further caution.

4.7 Conclusion

This chapter extends work by Nicolaou *et al.* (2008) on the heritability of entrepreneurship. Nicolaou *et al.* (2008) tested the hypothesis that entrepreneurship, as proxied by answers to survey questions about self-employment, is heritable and found support for this hypothesis in a highly selected sample of UK twins including virtually only women. Using a much larger and more representative sample, I estimate heritabilities for both men and women separately. The estimated heritabilities for women are consistent with those found by Nicolaou *et al.* (2008), but for men, I find heritabilities which are substantially lower and not statistically significantly different from zero. The difference in my sample between the estimated heritabilities of women and men is however not statistically significant at the 5% level with a two-sided test. Apart from sampling variation, I hypothesize that this discrepancy may be due partly to differences between women's and men's self-employment. The findings of this study reinforce the notion that caution be exerted when extrapolating findings between samples of men and women.

4.8 Tables

Table 4.1: Self-employment in sample and in population

		Women		Men	
		MZ	DZ	MZ	DZ
Current		4.4%	5.0%	11.5%	11.6%
	# obs	1902	1352	1132	764
Last 3 yrs - employer		4.1%	4.3%	6.9%	7.8%
	# obs	2394	1872	1690	1262
Last 3 yrs - mode of employment		3.1%	3.9%	5.6%	7.3%
	# obs	2636	2052	1774	1320
Last 3 yrs - either		4.2%	4.5%	7.2%	8.8%
	# obs	2636	2052	1774	1320
Statistics Sweden		3.6%		10.1%	

Note: Figures for population is from 2005 for the age group 25-44

Table 4.2: Comparison of demographic characteristics

	Main Sample				STAGE		Population	
	Women		Men		W	M	W	M
	MZ	DZ	MZ	DZ				
Income	219	224	309	312	195	269	197	271
S.D.	115	106	182	157	113	191		
Education	13.5	13.3	13.2	13.0	13.1	12.8	13.1	12.5
S.D.	2.5	2.6	2.6	2.6	2.4	2.4		
Marital status	0.38	0.43	0.36	0.40	0.33	0.29	0.40	0.33
S.D.	0.49	0.50	0.48	0.49	0.47	0.45		
Age	35.0	37.0	34.8	36.8	33.5	33.6	25-44	25-44
S.D.	6.6	6.5	6.6	6.4	7.7	7.7		
# obs	2638	2052	1774	1320	14113	11264	1.2M	1.2M

Notes: Main sample is total sample used for estimation on any self-employment measure (all same-sex pair where for any of the three questions on self-employment, both twins answered)

All population data are from Statistics Sweden, and is for the age group 25-44, corresponding roughly to the average age of the sample

Education data is for 2007; Income and marital status are for 2005

Table 4.3: Within-pair polychoric correlations

		Women		Men	
		MZ	DZ	MZ	DZ
Current		0.56	0.39	0.59	0.59
	# pairs	951	676	566	382
Last 3 years:					
main employer		0.64	0.22	0.47	0.38
	# pairs	1197	936	845	631
main mode of employment		0.63	0.30	0.52	0.51
	# pairs	1318	1026	887	660

Table 4.4: Parameter estimates from ACE model - Women

	Genetic	Shared env.	Non-shared env.
Current	0.35 (0.00-0.68)	0.21 (0.00-0.55)	0.44 (0.29-0.64)
Last 3 years:			
main employer	0.62 (0.28-0.74)	0.00 (0.00-0.28)	0.38 (0.26-0.55)
main mode of empl.	0.61 (0.08-0.73)	0.00 (0.00-0.41)	0.39 (0.26-0.59)

Notes: WLSMV (weighted least squares, mean- and variance-adjusted) using MPlus (Muthen and Muthen, 2007)

95% Confidence intervals using the bootstrap from 5000 draws with replacement

No assortative mating

Table 4.5: Parameter estimates from ACE model - Men

	Genetic	Shared env.	Non-shared env.
Current	0.00 (0.00-0.45)	0.59 (0.18-0.68)	0.41 (0.28-0.52)
Last 3 years:			
main employer	0.16 (0.00-0.55)	0.30 (0.00-0.51)	0.55 (0.40-0.72)
main mode of empl.	0.00 (0.00-0.50)	0.50 (0.08-0.62)	0.49 (0.34-0.64)

Notes: WLSMV (weighted least squares, mean- and variance-adjusted) using MPlus (Muthen and Muthen, 2007)

95% Confidence intervals using the bootstrap from 5000 draws with replacement

Table 4.6: Confidence intervals of difference in heritability between men and women

	2.5%	97.5%
Current	-0.28	0.67
Last 3 years:		
main employer	-0.05	0.71
main mode of employment	-0.08	0.72

Notes: WLSMV (weighted least squares, mean- and variance-adjusted) using MPlus (Muthen and Muthen, 2007)

Confidence intervals using the bootstrap from 5000 draws with replacement

No assortative mating

Table 4.7: Parameter estimates from ACE model - Women. Alternative assumptions on assortative mating

	Genetic	Shared env.	Non-shared env.
$\lambda = 0.1$			
Current	0.39 (0.00-0.68)	0.17 (0.00-0.55)	0.44 (0.29-0.64)
Last 3 years:			
main employer	0.62 (0.31-0.73)	0.00 (0.00-0.24)	0.38 (0.26-0.55)
main mode of empl.	0.60 (0.10-0.72)	0.00 (0.00-0.40)	0.40 (0.27-0.59)
$\lambda = 0.2$			
Current	0.43 (0.00-0.68)	0.13 (0.00-0.55)	0.44 (0.30-0.64)
Last 3 years:			
main employer	0.61 (0.34-0.73)	0.00 (0.00-0.20)	0.39 (0.27-0.55)
main mode of empl.	0.59 (0.11-0.72)	0.00 (0.00-0.38)	0.40 (0.28-0.60)
$\lambda = 0.3$			
Current	0.47 (0.00-0.68)	0.08 (0.00-0.55)	0.44 (0.30-0.65)
Last 3 years:			
main employer	0.60 (0.38-0.72)	0.00 (0.00-0.15)	0.40 (0.28-0.56)
main mode of empl.	0.58 (0.13-0.71)	0.00 (0.00-0.36)	0.42 (0.28-0.60)

Notes: WLSMV (weighted least squares, mean- and variance-adjusted) using MPlus (Muthen and Muthen, 2007)

95% Confidence intervals using the bootstrap from 5000 draws with replacement

Table 4.8: Parameter estimates from ACE model - Men. Alternative assumptions on assortative mating

	Genetic	Shared env.	Non-shared env.
$\lambda = 0.1$			
Current	0.00 (0.00-0.50)	0.59 (0.13-0.68)	0.41 (0.28-0.52)
Last 3 years:			
main employer	0.18 (0.00-0.56)	0.28 (0.00-0.51)	0.55 (0.40-0.72)
main mode of empl.	0.01 (0.00-0.53)	0.50 (0.02-0.61)	0.49 (0.34-0.63)
$\lambda = 0.2$			
Current	0.00 (0.00-0.55)	0.59 (0.07-0.68)	0.41 (0.28-0.52)
Last 3 years:			
main employer	0.20 (0.00-0.56)	0.25 (0.00-0.51)	0.55 (0.40-0.72)
main mode of empl.	0.02 (0.00-0.56)	0.49 (0.00-0.61)	0.49 (0.34-0.63)
$\lambda = 0.3$			
Current	0.00 (0.00-0.57)	0.59 (0.00-0.68)	0.41 (0.28-0.52)
Last 3 years:			
main employer	0.22 (0.00-0.57)	0.23 (0.00-0.50)	0.55 (0.40-0.72)
main mode of empl.	0.00 (0.00-0.58)	0.51 (0.00-0.61)	0.49 (0.34-0.63)

Notes: WLSMV (weighted least squares, mean- and variance-adjusted) using MPlus (Muthen and Muthen, 2007)

95% Confidence intervals using the bootstrap from 5000 draws with replacement

Table 4.9: Confidence intervals of difference in heritability between men and women. Alternative assumptions on assortative mating

	2.5%	97.5%
$\lambda = 0.1$		
Current	-0.31	0.66
Last 3 years:		
main employer	-0.05	0.70
main mode of employment	-0.09	0.71
$\lambda = 0.2$		
Current	-0.32	0.66
Last 3 years:		
main employer	-0.04	0.70
main mode of employment	-0.09	0.70
$\lambda = 0.3$		
Current	-0.35	0.65
Last 3 years:		
main employer	-0.04	0.69
main mode of employment	-0.11	0.69

Notes: WLSMV (weighted least squares, mean- and variance-adjusted) using MPlus (Muthen and Muthen, 2007)

Conclusion

This thesis provides four separate studies concerning generally unobservable skills and preferences of economic agents. In the first study, monozygotic twins were shown to be substantially different with respect to latent "ability", in the sense commonly used to describe the latent skill which both correlates with schooling and which drives income in itself. These results were estimated under necessary assumptions on the interrelationship of IQ with either (i) the exogenous preferences for schooling, or (ii) the exogenous determinants of income.

In the second study, the relationship between IQ and ostensibly social behaviour, as expressed in the ultimatum bargaining game, was analysed. It was found that IQ has statistically significant explanatory power of the commonly reported "anomalies", i.e behaviour not.conforming to standard self-interested rationality, with respect to responder behaviour, but that its effect is small in magnitude. Admittedly, the extent to which this study shows that a failure of rationality is not the "culprit" hinges on how anomalous we consider this experimental situation to be. An experiment specifically designed to elicit anomalous responses may require truly anomalous skills on the part of the experimental subject in order to produce a non-anomalous outcome. In contrast, the requirements on rationality may be more modest when making economic decisions in the field.

In the third and fourth chapters, a moderately large genetic influence on the variation in risk preferences was found, but for the propensity to become self-employed, the imprecisely estimated genetic components differed widely between men and women. The study of risk preferences is can serve to increase our understanding of both heterogeneity in economic outcomes within populations, and aggregate economic phenomena based on the specialisation of labour. This is potentially true as well for the propensity to become self-employed, although this latter notion is of course far less analysed from a theoretical perspective in the literature, and as such is much less well-defined. Both these studies rely on strong assumptions with regard to the functional form, and the second study suffers from a sample size which, although large by comparison to previous studies on the subject, is still too small to allow for precise statistical inference.

In the respective chapters, the implications and possible future directions of research based on these findings have been outlined. Hopefully, these studies together also inspire future study in general into the fascinating world of heterogeneities among economic agents at the individual level.

Bibliography

Anna, A. L., G. N. Chandler, E. Jansen and N. P. Merod, 1999, "Women business owners in traditional and non-traditional industries", *Journal of Business Venturing*, 15: 279-303.

Apicella, C. L., A. Dreber Almenberg, B. Campbell, P. Gray, M. Hoffman and A. C. Little, 2008, "Testosterone and Financial Risk-Taking", *Evolution and Human Behavior*, 29: 385-390.

Ashenfelter, O. and A. Krueger, 1994, "Estimating the Returns to Schooling Using a New Sample of Twins", *American Economic Review*, 84: 1157-1173.

Ashenfelter, O. and C. Rouse, 1998, "Income, Schooling and Ability: Evidence from a New Sample of Identical Twins", *Quarterly Journal of Economics*, 113: 253-84.

Behrman, J. R. and M. Rosenzweig, 1999, "'Ability' Biases in Schooling Returns and Twins: A Test and New Estimates", *Economics of Education Review*, 18: 159-67.

Behrman, J. R. and P. Taubman, 1976, "Intergenerational Transmission of Income and Wealth", *American Economic Review*, 66: 436-40.

Behrman, J. R. and P. Taubman, 1989, "Is Schooling Mostly in the Genes?

Nature-Nurture Decomposition Using Data on Relatives", *Journal of Political Economy*, 97: 1425-1446.

Behrman, J. R., M. R. Rosenzweig and P. Taubman, 1994, "Endowments and the Allocation of Schooling in the Family and in the Marriage Market: The Twins Experiment", *Journal of Poltical Economy*, 102: 1131-1174.

Benartzi, S. and R. H. Thaler, 2001, "Naive Diversification Strategies in Defined Contribution Saving Plans", *American Economic Review*, 91: 79-98.

Benjamin, D., S. Brown and J. Shapiro, 2006, "Who is 'behavioural'? Cognitive ability and anomalous preferences", *mimeo*, Harvard University.

Benjamin, J., L. Li, C. Patterson, B. D. Greenberg, D. L. Murphy and D. H. Hamer, 1996, "Population and familial association between the D4 dopamine receptor gene and measures of Novelty Seeking", *Nature Genetics*, 12: 81-84.

Bennedsen, M., K. M. Nielsen, F. Pérez-González and D. Wolfenzon, 2007, "Inside the Family Firm: the Role of Families in Succession Decisions and Performance", *Quarterly Journal of Economics*, 122: 647-691.

Bertrand, M. and A. Schoar, 2006, "The Role of Family in Family Firms", *Journal of Economic Perspectives*, 20: 73-96.

Bhandari, G. and R. Deaves, 2008, "Misinformed and informed asset allocation decisions of self-directed retirement plan members", *Journal of Economic Psychology*, 29: 473-490

Björklund, A., M. Jäntti and G. Solon, 2005, "Influences of Nature and Nurture on Earnings Variation: A Report on a Study of Various Sibling Types in Sweden", in S. Bowles, H. Gintis, and M. Osborne Groves, eds.: *Unequal Chances: Family*

Background and Economic Success, Princeton University Press, Princeton, NJ.

Björklund, A., M. Jäntti and G. Solon, 2007, "Nature and Nurture in the Intergenerational Transmission of Socioeconomic Status: Evidence from Swedish Children and Their Biological and Rearing Parents", *B.E. Journal of Economic Analysis & Policy* 7 (2), art. 4.

Björklund, A., M. Lindahl and E. Plug, 2006, "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data", *Quarterly Journal of Economics*, 121: 999-1028.

Boden, R., 1999, "Flexible Working Hours, Family Responsibilities, and Female Self-Employment", *American Journal of Economics and Sociology*, 58: 71-83.

Bonjour, D., L. Cherkas, J. Haskel, D. Hawkes and T. Spector, 2003, "Returns to education: evidence from UK twins", *American Economic Review*, 93: 1799-1812.

Bouchard, T. J. Jr, 1998, "Genetic and environmental influences on adult intelligence and special mental abilities", *Human Biology*, 70: 257-279.

Bouchard, T. J. Jr. and M. McGue, 1981, "Familial studies of intelligence: a review", *Science*, 212: 1055-1059.

Bouchard, T. J. Jr. and M. McGue, 2003, "Genetic and environmental influences on human psychological differences", *Journal of Neurobiology*, 54: 4-45.

Bound, J. and G. Solon, 1999, "Double Trouble: On the Value of Twins-Based Estimation of the Return to Schooling", *Economics of Education Review*, 18: 169-82.

Brandstatter, H. and W. Güth, 2002, "Personality in dictator and ultimatum games", *Central European Journal of Operations Research*, 3: 191-215.

Brush, C. G., 1992, "Research on Women Business Owners: Past Trends, a New Perspective and Future Directions", *Entrepreneurship Theory and Practice*, 16: 5-30.

Bowles, S. and H. Gintis, 2002, "The Inheritance of Inequality", *Journal of Economic Perspectives*, 16: 3-30.

Bowles, S. and H. Gintis, "Homo Reciprocans", *Nature*, 415: 125-126.

Boyer, T. W., 2006, "The Development of risk-taking: A multi-perspective review", *Developmental Review*, 26: 291-345.

Burkart, M., F. Panunzi and A. Shleifer, 2003, "Family firms", *Journal of Finance*, 58: 2167-2201.

Burke, A. E., F. R. FitzRoy and M. A. Nolan, 2000, "When less is more: distinguishing between entrepreneurial choice and performance", *Oxford Bulletin of Economics and Statistics*, 62: 567–587.

Burke, A. E., F. R. Fitzroy and M. A. Nolan, 2002, "Self-employment Wealth and Job Creation: The Roles of Gender, Non-pecuniary Motivation and Entrepreneurial Ability", *Small Business Economics*, 19: 255-270.

Burks, S., J. Carpenter, L. Gotte, A. Rustichini, 2008, "Cognitive skills explain economic preferences, strategic behaviour, and job attachment", *IZA Discussion Paper No. 3609.*

Burnham T., D. Cesarini, P. Lichtenstein, M. Johannesson and B. Wallace, "Cognitive Ability Correlates with Nash Equilibrium Behavior in a p-Beauty Contest", *Journal of Economic Behavior and Organization*, forthcoming.

Burnham, T. and D. Johnson, 2005, "The biological and evolutionary logic of human cooperation", *Analyse and Kritike,* 27: 113-135.

Camerer, C., 2003, *Behavioral game theory: experiments in strategic interaction,* Princeton University Press, Princeton, NJ.

Card, D., 1999, "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card, eds.: *Handbook of labor economics,* Vol. 3A, North Holland, Amsterdam, NL.

Cardinal, R. N., 2006, "Neural systems implicated in delayed and probabalistic reinforcement", *Neural Networks,* 19: 1277-1301.

Carlstedt, B., 2000, *Cognitive Abilities - aspects of structure, process and measurement,* PhD dissertation, Acta Universitatis Gothoburgensis, University of Gothenburg, Sweden.

Carter, N. M., W. B. Gartner, K. G. Shaver and E. J. Gatewood, 2003, "The Career Reasons of Nascent Entrepreneurs", *Journal of Business Venturing,* 18: 13-39.

Cesarini D., C. T. Dawes, M. Johannesson, P. Lichtenstein, B. Wallace, "Genetic Variation in Preferences for Giving and Risk-Taking", *Quarterly Journal of Economics,* forthcoming.

Cesarini, D., C. T. Dawes, J. H. Fowler, M. Johannesson, P. Lichtenstein and B. Wallace, 2008, "Heritability of Cooperative Behavior in the Trust Game", *Proceedings of the National Academy of Sciences,* 104: 3721–3726.

Chaganti, R. and S. Parasuraman, 1996, "A study of the impacts of gender on business performance and management patterns in small businesses", *Entre-*

preneurship Theory and Practice, 21: 73-75.

Charles, K. K. and E. Hurst, 2003, "The Correlation of Wealth across Generations", *Journal of Political Economy*, 111: 1155-1182.

Chiteji, N. S. and F. P. Stafford, 1999, "Portfolio Choices of Parents and Their Children as Young Adults: Asset Accumulation by African-American Families", *American Economic Review*, 89: 377-380.

Cohn, R. A., W. G. Lewellen, R. C. Lease and G. G. Schlarbaum, 1975, "Individual Inverstor Risk Aversion and Investment Portfolio Composition", *The Journal of Finance*, 30: 605-620.

Collins, C., P. Hanges, E. Locke, 2004, "The relationship of achievement motivation to entrepreneurial behaviour: a meta-analysis", *Human Performance*, 17: 95-117.

Coventry, W. L. and M. C. Keller, 2005, "Estimating the Extent of Parameter Bias in the Classical Twin Design: A Comparison of Parameter Estimates From Extended Twin-Family and Classical Twin Designs", *Twin Research and Human Genetics*, 8: 214-223.

Cronquist, H. and R. H. Thaler, 2004, "Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience", *American Economic Review*, 94: 424-428.

Dall, S. R. X., A. I. Houston and J. M. McNamara, 2004, "The behavioural ecology of personality: consistent individual differences from an adaptive perspective", *Ecology Letters*, 7: 734-739.

DeFries, J. C. and D. W. Fulker, 1985, "Multiple regression analysis of twin data", *Behavior Genetics*,15: 467-473.

de Wit, G. and F. A. A. M. van Winden, 1989, "An empirical analysis of self-employment in the Netherlands", *Small Business Economics*, 1: 263-272.

Dohmen, T., A. Falk, D. Huffman and U. Sunde, 2006, "The Intergenerational Transmission of Risk and Trust Attitudes", *IZA Discussion Paper No. 2380.*

Dohmen, T., A. Falk, D. Huffman and U. Sunde, 2007, "Are risk aversion and impatience related to cognitive ability?", *IZA Discussion Paper No. 2735.*

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner, 2005, "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey", *IZA Discussion Paper No. 1730.*

Du Rietz, A. and M. Henrekson, 2000, "Testing the female underperformance hypothesis", *Small Business Economics*, 14: 1-10.

Eaves, L., A. Heath, M. Neale, J. Hewitt and G. Martin, 1989, "Sex differences and non-additivity in the effects of genes on personality", *Twin Research*, 1: 131-137.

Ebstein, R. P., O. Novick, R. Umansky, B. Priel, Y. Osher, D. Blaine, E. R. Bennett, L. Nemanow, M. Katz and R. H. Belmaker, 1996, "Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of Novelty Seeking", *Nature Genetics*, 12: 78–80.

Enders, C.K., 2001, "The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data", *Psychological Methods*, 6: 352-370.

Fairlie, R. W., 1999, "The Absence of African-American Owned Businesses:

An analysis of the Dynamics of Self-employment", *Journal of Labor Economics*, 17: 80-108.

Falconer, D. S. and T. F. C. Mackay, 1996, *Introduction to Quantitative Genetics,* Benjamin Cummings, UK.

Falk, A. and U. Fischbacher, 2006, "A theory of reciprocity", *Games and Economic Behavior*, 54: 293-315.

Falter, J-M., 2007, "Entrepreneurship and Intergenerational Links In Switzerland", *Labour: Review of Labour Economics and Industrial Relations*, 21: 121-134.

Fehr, E. and S. Gachter, 2002, "Altruistic punishment in humans", *Nature*, 415:137-140.

Fehr, E. and U. Fischbacher, 2003, "The nature of altruism", *Nature*, 425: 785-791.

Fehr, E., S. Gachter and G. Kirchsteiger, 1997, "Reciprocity as a contract enforcement device: experimental evidence", *Econometrica*, 65: 833-860.

Fehr, E., U. Fischbacher, and S. Gachter, 2002, "Strong reciprocity, human cooperation, and the enforcement of social norms", *Human Nature*, 13: 1-25.

Finkel, D. and M. McGue, 1997, "Sex differences and nonadditivity in heritability of the multidimensional personality questionnaire", *Journal of Personality and Social Psychology*, 72: 929-938.

Fischer, E. M., A. R. Reuber and L. S. Dyke, 1993, "A Theoretical Overview and Extension of Research on Sex, Gender, and Entrepreneurship", *Journal of Business Venturing*, 8: 151-168.

Flora, D. and P. Curran, 2004, "An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data", *Psychological Methods*, 9: 466-491.

Fowler, J. H., C. T. Dawes and L. Baker, 2008, "The Genetic Basis of Political Participation", *American Political Science Review*, 102: 233-248.

Frederick, S., 2005, "Cognitive reflection and decision making", *Journal of Economic Perspectives*, 19: 25-42.

Georgellis, Y. and H. J. Wall, 2005, "Gender Differences in Self-Employment", *International Review of Applied Economics*, 19: 321-342.

Gintis, H., 2000, "Strong reciprocity and human sociality", *Journal of Theoretical Biology*, 206:169-179.

Gneezy, U. and J. List, 2006, "Putting behavioural economics to work: testing for gift exchange in labor markets using field experiments", *Econometrica*, 74: 1365-1384.

Gottschalk, P. and T. Smeeding, 1997, "Cross-national comparisons of earnings and income inequality", *Journal of Economic Literature*, 35: 633-687.

Griliches, Z., 1977, "Estimating the Returns to Schooling: Some Econometric Problems", *Econometrica*, 45: 1-22.

Griliches, Z., 1979, "Sibling Models and Data in Economics: Beginnings of a Survey", *Journal of Political Economy*, 87: S37-S65.

Güth, W., R. Schmittberger, B. Schwarze, 1982, "An experimental study of ultimatum bargaining", *Journal of Economic Behavior and Organization*, 3: 367-388.

Harmon, C., H. Oosterbeek and I. Walker, 2003, "The Returns to Education:

Microeconomics", *Journal of Economic Surveys*, 17: 115–156.

Harris, J. A., P. A. Vernon and D. I. Boomsma, 1998, "The Heritability of Testosterone: A Study of Dutch Adolescent Twins and Their Parents", *Behavior Genetics*, 28: 165-171.

Hausman, J., 1978, "Specification tests in econometrics", *Econometrica*, 46: 1251-1271.

Hettema, J. M., M. C. Neale and K. S. Kendler, 1995, "Physical similarity and the equal-environment assumption in twin studies of psychiatric disorders", *Behavior Genetics*, 25: 327-335.

Hisrich, R. D. and C. G. Brush, 1983, "The woman entrepreneur: implications of family, educational and occupational experience", in J. A. Hornaday, J. A. Timmons, K. H. Vesper, eds.: *Frontiers in Entrepreneurship Research: Proceedings of the 1983 Conference on Entrepreneurship*, Babson College, Wellesley, MA.

Holmquist, C. and E. Sundin, 1988, "Women as entrepreneurs in Sweden - conclusions from a survey", in Kirchhoff B. A. *et al.*, eds.: *Frontiers of entrepreneurship research 1988*, Babson College, Wellesley, MA.

Holmquist, C. and E. Sundin, 2002, "Kvinnors företagande, siffror och synliggörande", in C. Holmquist and E. Sundin, eds.: *Företagerskan – om kvinnor och enterprenörskap*, SNS, Stockholm, SE.

Hryshko, D., M. J. Luengo-Prado and B. E. Sorensen, 2007, "Childhood Determinants of Risk Aversion: The Long Shadow of Compulsory Education", *mimeo*, Northeastern University.

Isacsson, G., 1999, "Estimates of the Return to Schooling in Sweden from a

Large Sample of Twins", *Labour Economics*, 6: 471-89.

Isacsson , G., 2004, "Estimating the economic return to educational levels using data on twins", *Journal of Applied Econometrics*, 2004, 19: 99-119.

Joseph, J., 2002, "Twin Studies in Psychiatry and Psychology: Science or Pseudoscience?", *Psychiatric Quarterly*, 73: 71-82.

Kanazawa, S., and L. Fontaine, 2007, "Intelligent people defect more in a one-shot prisoner's dilemma game", *mimeo.*

Kendler, K.S., N. G. Martin, A.C. Heath and L.J. Eaves, 1995, "Self-report psychiatric symptoms in twins and their nontwin relatives: are twins different?", *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, 60: 588-591.

Kohler, H-P. and J. L. Rodgers, 2001, "DF-Analyses of Heritability with Double-Entry Twin Data: Asymptotic Standard Errors and Efficient Estimation", *Behavior Genetics*, 31: 179-191.

Koopmans, J. R., D. I. Boomsma, A. C. Heath and L. J. P. van Doornen, 1995, "A Multivariate Genetic Analysis of Sensation Seeking", *Behavior Genetics*, 25: 349-356.

Krain, A. L., A. M. Wilson, R. Arbuckle, F. Xavier Castellanos and M. P. Milham, 2006, "Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making", *NeuroImage*, 32: 477-484.

Kreek, M. J., D. A. Nielsen, E. R. Butelman and S. Laforge, 2005, "Genetic Influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction", *Nature Neuroscience*, 8: 1450-1457.

Krueger, N. F., 1993, "Growing up entrepreneurial", *Proceedings of the Acad-*

emy of Management, GA.

Kuhnen, C. M. and B. Knutson, 2005, "The Neural Basis of Financial Risk Taking", *Neuron*, 47: 763-770.

Levitt, S. and J. List, 2007, "What do laboratory experiments measuring social preferences reveal about the real world", *Journal of Economic Perspectives*, 21: 53-174.

Lentz, B. F. and D. N. Laband, 1990, "Entrepreneurial success and occupational inheritance among proprietors", *Canadian Journal of Economics*, 23: 563-579.

Lichtenstein, P., N. L. Pedersen and G. E. McClearn, 1992, "The origins of individual differences in occupational status and educational level: A study of twins reared apart and together", *Acta Sociologica*, 35: 13-31.

Lichtenstein, P., U. deFaire, B. Floderus, M. Svartengren, P. Svedberg and N. Pedersen, 2002, "The Swedish Twin Registry: a unique resource for clinical, epidemiological and genetic studies", *Journal of Internal Medicine*, 252: 184-205.

Lichtenstein, P., P. Sullivan, S. Cnattingius, M. Gatz, S. Johansson, E. Carlström, C. Björk, M. Svartengren, A. Volk, L. Klareskog, U. deFaire, M. Schalling, J. Palmgren and N. Pedersen, 2006, "The Swedish Twin Registry in the Third Millennium : An Update", *Twin Research and Human Genetics*, 9: 875-882.

List, J., 2006, "The behaviouralist meets the market: measuring social preferences and reputation effects in actual transactions", *Journal of Political Economy*, 114: 1-37.

Loehlin, J. C., 2005, "Resemblance in Personality and Attitudes between Parents and Their Children: Genetic and Environmental Contributions", in S. Bowles, H. Gintis and M. Osborne Groves, eds.: *Unequal Chances: Family Background and Economic Success*, Princeton University Press, Princeton, NJ.

Lykken, D. T., M. McGue and A. Tellegen, 1987, "Recruitment bias in twin research: The rule of two-thirds reconsidered", *Behavior Genetics*, 17: 343-362.

Lykken, D. T., M. McGue, T. J. Bouchard Jr. and A. Tellegen, 1990, "Does contact lead to similarity or similarity to contact?", *Behavior Genetics*, 20: 547-561.

Macaskill, G., J. Hopper, V. White, D. Hill, 1994, "Genetic and environmental variation in eysenck personality qustionnaire scaltes measured on Australian adolescent twins", *Behavior Genetics*, 24: 481-491.

Matheny, A. P., R. S. Wilson and A. Brown Dolan, 1976, "Relations between twins' similarity of appearance and behavioural similarity: Testing an assumption", *Behavior Genetics*, 6: 343-351.

Mather, K. and J. L. Jinks, 1977, *An introduction to biometrical genetics*, Chapman and Hall, London, UK.

Mather, K. and J. L. Jinks, 1982, *Biometrical Genetics*, 3rd ed., Chapman and Hall, London, UK.

Miller, P., C. Mulvey and N. Martin, 1995, "What Do Twins Studies Reveal About the Economic Returns to Education? A Comparison of Australian and U.S. Findings", *American Economic Review*, 85: 586-99.

Munafò, M. R., T. G. Clark, L. R. Moore, E. Payne, R. Walton and J. Flint, 2003, "Genetic Polymorphisms and Personality in Healthy Adults: A systematic review and meta-analysis", *Molecular Psychiatry*, 8: 471-484.

Muthén, Linda. K. and Bengt O. Muthén, 2006, *Mplus. Statistical Analysis With Latent Variables. User's Guide, Version 4.1*, Los Angeles, CA.

Muthén, B and A. Satorra, 1995, "Technical aspects of Muthen's LISCOMP approach to estimation of latent variable relations with a comprehensive measurement model", *Psychometrica*, 60: 489-503.

Muthén, B., S. du Toit and D. Spisic, 1997, "Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes", *mimeo*.

Neale, M. C. and H. H. M. Maes, 2004, *Methodology for Genetic Studies of Twins and Families*, Kluwer Academic, Dordrecht, NL.

Neisser, U., G. Boodoo, T. J Bouchard. Jr., A. Wade Boykin, N. Brody, S. J. Ceci, D. F. Halpern, J. C. Loehlin, R. Perloff, R. J. Sternberg and S. Urbina, 1996, "Intelligence: Knowns and Unknowns", *American Psychologist*, 51: 77-101.

Nicolaou, N. and S. Shane, 2009, "Can genetic factors influence the likelihood of engaging in entrepreneurial activity?", *Journal of Business Venturing*, 24: 1-22.

Nicolaou, N., S. Shane, L. Cherkas, J. Hunkin and T. D. Spector, 2008, "Is the tendency to engage in entrepreneurship genetic?", *Management Science*, 54: 167-179.

Oechssler, J., A. Roider, P. Schmitz, 2008,"Cooling-off in negotiations - does it work?", *CEPR Discussion Paper No. 6807*.

Orser, B. J. , A. L. Riding and K. Manley, 2006, "Women Entrepreneurs and Financial Capital", *Entrepreneurship Theory and Practice*, 30: 643 - 665.

Palme, M. and A. Sundén, 2004, "Premiepensionen i det reformerade pensionssystemet – är det önskvärt att kunna välja mellan 663 fonder", *Ekonomisk debatt*, 3: 6-15.

Palme, M., A. Sundén and P. Söderlind, 2007, "Investment Choice in the Swedish Premium Pension Plan", *Journal of the European Economic Association Papers and Proceedings*, 5: 636-646.

Palmer, E., 2000, "The Swedish Pension Reform Model: Framework and Issues", *The World Bank SP Discussion Paper No. 0012.*

Pam, A., S. S. Kemker, C. A. Ross and R. Golden, 1996, "The "equal environments assumption" in MZ-DZ twin comparisons: an untenable premise of psychiatric genetics?", *Acta Genetеticae Medicae Gemellologiae (Roma)*, 45: 349-360.

Penke, L., J. J. A. Denissen, and G. F. Miller, "The Evolutionary Genetics of Personality", *European Journal of Personality*, 21: 549-587.

Pinker, S., 2004, "Why nature & nurture won't go away", *Dœdalus* 133: 5-17.

Plomin, R. D., J. C. DeFries, G. E. McClearn and P. McGuffin, 2001, *Behavioral genetics*, 4th ed., Freeman, New York, NY.

Plug, E. and W. Vijverberg, 2003, "Schooling, Family Background, and Adoption: Is It Nature or Is It Nurture?", *Journal of Political Economy*, 111: 611-641.

Posner, S. F., L. Baker, A. Heath and N. G. Martin, 1996, "Social contact and attitude similarity in Australian twins",. *Behavior Genetics*, 26: 123-133.

Poterba, J. M. and D. Wise, 1998, "Individual Financial Decisions in Retirement Saving Plans and the Provision of Resources for Retirement", in M.

Feldstein, ed.: *Privatizing Social Security*, University of Chicago Press, Chicago, IL.

Poterba, J. M., S. F. Venti and D. A. Wise, 2000, "Saver Behavior and 401(k) Retirement Wealth", *American Economic Review*, 90: 297-302.

Pålsson, A-M., 1996, "Does the degree of relative risk aversion vary with household characteristics?", *Journal of Economic Psychology*, 17, 771-787.

Rauch, A. and Frese, M., 2000, "Psychological approaches to entrepreneurial success. A general model and an overview of findings" in C. L. Cooper and I. T. Robertson, eds.: *International Review of Industrial and Organizational Psychology*, Wiley, Chichester, UK.

Risch, N., 2001, "The Genetic Epidemiology of Cancer: Interpreting Family and Twin Studies and Their Implications for Molecular Genetic Approaches", *Cancer Epidemiology: Biomarkers and Prevention*, 10: 733–41.

Rivkin, S., E. Hanushek, and J. Kain, 2005, "Teachers, schools, and academic achievement", *Econometrica*, 73: 417-458.

Rosa, P., S. Carter and D. Hamilton, 1996, "Gender as a Determinant of Small Business Performance: Insights from a British Study", *Small Business Economics*, 8: 463-478.

Rouse, C., 1999, "Further Estimates of the Economic Return to Schooling from a New Sample of Twins", *Economics of Education Review*, 18: 149-57.

Rowe, D. C., W. J. Vesterdal and J. L. Rodgers, 1998, "Herrnstein's syllogism: genetic and shared environmental influences on IQ, education, and income", *Intelligence*, 26: 405-423.

Sacerdote, B., 2001, "Peer effects with random assignment: results for Dartmouth roommates", *Quarterly Journal of Economics*, 116: 681-704.

Sacerdote, B., 2002, "The Nature and Nurture of Economic Outcomes", *American Economic Review*, 92: 344-348.

Sacerdote, B., 2007, "How Large Are the Effects from Changes in Family Environment? A Study of Korean American Adoptees", *Quarterly Journal of Economics*, 122: 119-157.

Sacerdote, B., 2009, "Nature And Nurture Effects On Children's Outcomes: What Have We Learned From Studies Of Twins And Adoptees?", in J. Benhabib, A. Bisin and M. O. Jackson, eds.: *Handbook of Social Economics*, Elsevier, North Holland, Amsterdam, NL.

Sjöberg, A., S. Sjöberg and K. Forssén, 2006, "Pjp: Predicting Job Performance, Manual", Assesio International B, Stockholm, SE.

Sobel, M., 1982, "Asymptotic confidence intervals for indirect effects in structural equation models", *Sociological Methodology*, 13: 290-312.

Sorenson, J. B., 2007, "Closure versus exposure: Mechanisms in the intergenerational transition of self-employment", in M. Ruef and M. Lounsbury eds.: *Research in the Sociology of Organizations*, Elsevier, North Holland, Amsterdam, NL.

Stevenson, L. A., 1986, "Against all odds: The entrepreneurship of women", *Journal of Small Business Management*, 24: 30–36.

Stewart, W. and P. Roth, 2007, "A meta-analysis of achievement motivation differences between entrepreneurs and managers", *Journal of Small Business*

Management, 45: 401-421.

Sundén, A. E. and B. J. Surette, 1998, "Gender Differences in the Allocation of Assets in Retirement Savings Plans", *American Economic Review Papers and Proceedings*, 88: 207-211.

Säve-Söderbergh, J, 2008, "Self-Directed Pensions: Gender, Risk & Portfolio Choices", *mimeo*, Stockholm University.

Taubman, P., 1976, "Earnings, education, genetics, and environment", *Journal of Human Resources*, 11: 447–461.

Taubman, P., 1976, "The Determinants of Earnings: Genetics, Family, and Other Environments: A Study of White Male Twins", *American Economic Review*, 66: 858-870.

Taylor, M. P., 1996, "Earnings, independence or unemployment:why become self-employed?", *Oxford Bulletin of Economics and Statistics*, 58: 253–266.

Toga, A. W. and P. M. Thompson, 2005, "Genetics of brain structure and intelligence", *Annual Review of Neuroscience*, 28: 1-23.

Turkheimer, E., 2000, "Three Laws of Behavior Genetics and What They Mean", *Current Directions in Psychological Science*, 9: 160-164.

Turner, C. F. and D. C. Martinez, 1977, "Socioeconomic Achievement and the Machiavellian Personality", *Sociometry*, 40: 325-336.

Uusitalo, R., 2001, "Homo Entreprenaurus?", *Applied Economics*, 33: 1631–1638.

Wallace B, D. Cesarini, P. Lichtenstein, M. Johannesson, 2007, "Heritability of ultimatum game responder behaviour", *Proceedings of the National Academy* *of Sciences*, 104: 15631-15634.

Welter, F., 2001, "Who wants to grow? Growth intentions and growth profiles of (nascent) entrepreneurs in Germany", in W. D. Bygrave, E. Autio, C. G. Brush, P. Davidsson, P. G. Greene, P. D. Reynolds and H. J. Sapienza, eds.: *Frontiers of entrepreneurship research*, Babson College, Wellesley, MA.

White, R. E., S. Thornhill and E. Hampson, 2006, "Entrepreneurs and evolutionary biology: the relationship between testosterone and new venture-creation", *Organizational Behavior and Human Decision Processes*, 100: 21–34.

Zhang, J., P-W. Liu and L. Yung, 2007, "The Cultural Revolution and returns to schooling in China: Estimates based on twins", *Journal of Development Economics*, 84: 631-639.

Zuckerman, M. and D. M. Kuhlman, 2000, "Personality and Risk-Taking: Common Biosocial Factors", *Journal of Personality*, 68: 999-1028.