De Neve, Jan-Emmanuel; Fowler, James H.; Frey, Bruno S.

Working Paper
Genes, economics, and happiness

CESifo Working Paper, No. 2946

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: De Neve, Jan-Emmanuel; Fowler, James H.; Frey, Bruno S. (2010) : Genes, economics, and happiness, CESifo Working Paper, No. 2946, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:
http://hdl.handle.net/10419/30746

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Genes, Economics, and Happiness

Jan-Emmanuel De Neve
James H. Fowler
Bruno S. Frey

CESifo Working Paper No. 2946
Category 12: Empirical and Theoretical Methods
February 2010

An electronic version of the paper may be downloaded
• from the SSRN website: www.SSRN.com
• from the RePEc website: www.RePEc.org
• from the CESifo website: www.CESifo-group.org/wp
Genes, Economics, and Happiness

Abstract

Research on happiness has produced valuable insights into the sources of subjective well-being. A major finding from this literature is that people exhibit a “baseline” happiness that shows persistent strength over time, and twin studies have shown that genes play a significant role in explaining the variance of baseline happiness between individuals. However, these studies have not identified which genes might be involved. This article presents evidence of a specific gene that predicts subjective well-being. Using data from the National Longitudinal Study of Adolescent Health, we show that individuals with a transcriptionally more efficient version of the serotonin transporter gene (5HTT) are significantly more likely to report higher levels of life satisfaction. Having one or two alleles of the more efficient type raises the average likelihood of being very satisfied with one’s life by 8.5% and 17.3%, respectively. This result may help to explain the stable component of happiness and suggests that genetic association studies can help us to better understand individual heterogeneity in subjective well-being.

JEL-Code: A12, Z00.

Keywords: happiness, subjective well-being, genetics.
So our happiness depends on our genes and our experience (past and present). Any social reformer has to be mainly interested in the role of experience since that is all that we can change. But we will never understand that bit unless we understand the complete reality, and the complete reality includes a strong role for the genes.


1 Introduction

Happiness research has become one of the liveliest subjects in economics in recent years. Its main goal is to explain the determinants of individual life satisfaction or subjective well-being (often loosely called happiness). Economists have mainly dealt with economic influences, in particular, income and its distribution, unemployment and inflation. Systematic influences on life satisfaction have also been found for socio-demographic factors (age, race, marital status and children) as well as for political and cultural factors (such as democracy, decentralization, and ethnicity). While all studies find significant influences by these variables, it has also become clear that there remains an important part unexplained and hidden behind the constant in the estimates. Thus, the “Easterlin Paradox” (Easterlin 1974, 2004, Clark, Frijters and Shields 2008) suggests that due to adaptation processes, increases in real income either have no lasting effect on happiness, or only a quite small one (Stevenson and Wolfers 2008). Psychologists speak of a “set point” or “baseline” of happiness that is influenced by personality and genetic predispositions (Kahneman et al. 1999). So far, however, no genetic association with subjective well-being has been established. This article presents the first such evidence of a specific gene that is associated with life satisfaction.

Studies using twin design research techniques have been able to gauge the relative explanatory power of genes and environment, and they show that genes play a significant role in explaining happiness levels. In a seminal study of the Minnesota twin registry, Lykken and Tellegen (1996) report that between 44% and 52% of the variance in subjective well-being is due to genetic variation. None of the environmental variables such as education, socio-economic status, income, marriage, or religiosity could account for more than 3% of the variance. When considering a subsample of the twins for which time series was available, Lykken and Tellegen (1996: 186) estimated that “the heritability of the stable component of subjective well-being approaches 80%.” Similarly, Nes, Roysamb, Tambs, Harris and Reichborn-Kjennerud (2006) studied genetic and environmental influences on stability and change in subjective well-being over time in a sample of Norwegian twins. They also found that for both males and females, long-term stability of well-being is mainly attributable to stable additive genetic factors. In fact, the authors also report that additive genetic effects explain approximately 80% of the cross-time correlation.

Although twin studies are an important step in establishing the influence of genes in subjective well-being, they do not identify the specific genes involved. The increasing availability of genotypic information now allows us to test hypotheses about targeted genes and their effects. The natural place to start the search for such genes is among those that have already been shown to account for variation in emotional states. Among these, 5HTT is a prime candidate:\footnote{The 5HTT gene has several other names, including HTT, SLC6A4, and SERT.} The 5HTT gene encodes a transporter in the cell wall that absorbs excess serotonin into the pre-synaptic neuron in parts of the brain that influence mental states (Hariri, Mattay, Tessitore, Kolachana, Fera and Goldman 2002, Bertolino, Arciero, Rubino, Latorre, Candia and Mazzola 2005, Heinz, Braus, Smolka, Wrase, Puls, Hermann and et al. 2005, Canli and Lesch 2007). 5HTT has been studied for more than twenty years and much is known about the way different versions of this gene influence transcription, metabolism, and signal transfers between neurons, all of which may influence personality. In
particular, the less transcriptionally efficient alleles of this gene have been shown to moderate
the influence of life stress on depression (Caspi, Sugden, Moffitt, Taylor, Craig, Harrington,
McClay, Mill, Martin, Braithwaite and Poulton 2003); and the more transcriptionally effi-
cient alleles have been linked to optimism (Fox, Ridgewell and Ashwin 2009). As a result,
economists have specifically identified 5HTT as a candidate gene for further study (Benjamin,
Chabris, Glaeser, Gudnason, Harris, Laibson, Launer and Purcell 2007).

It is crucial to point out at the outset that this association study does not establish that
5HTT “causes” happiness, nor does it exclude the possibility that several other genes may
play a role, but it does suggest at least one possible causal pathway able to account for
the influence of genes on happiness. Using data from the National Longitudinal Study of
Adolescent Health (Add Health), we analyze the relationship between 5HTT and life satis-
faction. The results indicate that the 5HTT gene is significantly associated with subjective
well-being. The direct association of a specific gene variant with life satisfaction has, to our
knowledge, so far not been shown.

2 Basic Concepts in Genetics

The human DNA is composed of an estimated 25,000 genes that form the blueprint for
molecules that regulate the development and function of the human body. Genes are distinct
regions of human DNA that are placed in the 23 pairs of chains, or chromosomes, that make
up all human DNA. Almost all human cells contain the same inherited DNA chains that are
fixed from the moment of conception.

Individuals inherit one half of their DNA from each parent, with one copy of each gene
coming from the mother and one copy from the father. Some genes come in different versions,
known as “alleles”—for example, sickle cell disease results from a particular allele coding for
abnormal rather than normal hemoglobin. Each parent has two separate copies of an allele
at each “locus”, or location, on the chromosome, but each sperm or egg cell contains only
one of these alleles. Thus a child has a 50% chance of receiving a particular allele from a
particular parent. For example, suppose that at a given locus there are two possible alleles, A and B. If both parents are “heterozygous” at that locus, meaning they each have an A and a B allele (AB or BA—order is irrelevant), then a given offspring has a 25% chance of being “homozygous” for A (AA), a 25% chance of being homozygous for B (BB) and a 50% chance of being heterozygous (AB or BA). If an individual is heterozygous at a locus, a “dominant” allele may impose itself on the “recessive” allele and the expression of the latter allele will not be observed.

Genes transcribe proteins that begin a cascade of interactions that regulate bodily structure and function. Many of the observable traits and behaviors of interest, referred to as “phenotypes”, are far downstream from the original “genotypes” present in the DNA. While in some cases one allele can single-handedly lead to a disease (such as Sickle Cell Anemia, Huntingtons disease, and cystic fibrosis), the vast majority of phenotypes are “polygenic”, meaning they are influenced by multiple genes (Mackay 2001, Plomin 2008), and are shaped by a multitude of environmental forces. As a result, simple association models between genotypes and phenotypes are an important first step, but they are not the end of the story. It is also important to investigate the extent to which genetic associations are moderated by environmental factors and other genes.

3 5HTT, Serotonin, and Happiness

Research in behavioral genetics typically starts with “candidate” genes that are known to influence behaviors or processes in the body that are related to the phenotype of interest. For subjective well-being, this means focusing on genes that affect brain development, neurotransmitter synthesis and reception, hormone regulation, and transcriptional factors (Damberg, Garpenstrand, Hallman and Oreland 2001, Benjamin et al. 2007).

To study whether genes affect happiness we choose a candidate gene that has already received a great deal of attention for its association with mental states. The 5HTT gene is critical to the metabolism of serotonin in the brain. As shown in Figure [1], serotonin is
a chemical that is released by a neuron and sensed by a receptor on the receiving neuron, passing an electric potential across a gap called a nerve synapse (the nerve that emits the serotonin is on the “pre-synaptic” side of the gap). Signals are carried throughout the body by the sequential release of a neurotransmitter by one neuron after another across these synapses. When an individual experiences stress, it causes increased neuron activity, stimulating the release of excess serotonin into the gaps between the synapses (Chaouloff, Berton and Mormede 1999). If serotonin remains outside the cells, it can oxidize into a toxin that kills both the pre-synaptic and post-synaptic neurons. The body’s homeostatic response to this excess serotonin is to reabsorb it into the pre-synaptic neuron via a transporter in the cell wall, called 5HTT.

The 5HTT gene contains a 44 base-pair variable-number tandem repeat (VNTR) polymorphism\(^3\) in the promoter region\(^4\) that is believed to be responsible for variation in transcriptional activity. The “long” (5-HTTLPR) and “short” (5-HTTSPR) polymorphism produce the same protein but the transcriptional efficiency of the “long” version of this allele is associated with a much higher basal activity than the shorter allele (Lesch, Bengel, Heils, Sabol, Greenberg, Petri and et al. 1996, Little, McLaughlin, Zhang, Livermore, Dalack and McFinton 1998, Glatz, Mössner, Heils and Lesch 2003, Canli and Lesch 2007). The “short” variant produces significantly less 5HTT mRNA\(^5\) and protein. The “short” polymorphism thus results in decreased gene expression and less serotonin transporters in the cell membrane. As a consequence, less serotonin is reintroduced into the pre-synaptic cell and higher concentrations of serotonin remain in the synaptic cleft.


\(^{3}\)A VNTR polymorphism is a repeated segment of DNA that varies among individuals in a population.  
\(^{4}\)A promoter region is the regulatory region of DNA that tells transcription enzymes where to begin. These promoter regions typically lie upstream from the genes they control.  
\(^{5}\)Messenger ribonucleic acid (mRNA) is a type of RNA that carries information from DNA to ribosomes. In turn, these ribosomes “read” messenger RNAs and translate their information into proteins.
Figure 1: Representation of the long/short variant of the 5HTT gene and the release, reception, and recycling of serotonin in neurons. Adapted from Canli & Lesch (2007), with permission from the Nature Publishing Group.

studies on rhesus macaque monkeys report that impaired serotonin metabolisms exaggerate the neurobiological response to stress (Barr, Newman, Schwandt, Shannon, Dvoskin, Lindell, Taubman, Thompson, Champoux, Lesch, Goldman, Suomi and Higley 2004). A number of human studies show different reactivity in the amygdala for variation in the 5HTT gene and the serotonin system. Because the amygdala is a central part of the human neural system this has important implications for emotion processing and, thus, mental health and well-being (Phelps and LeDoux 2005).

A recent study by Fox et al. (2009) further suggests that 5HTT may influence optimism. The authors obtained DNA samples from about 100 participants and compared reaction
times to pictures with positive, negative, and neutral valence, in line with a common experiment in psychopathology research. The results show that individuals with both “long” 5HTT alleles display a significant bias towards processing positive information and selectively avoiding negative information. This emotionally self-protective pattern does not obtain in individuals carrying one or both “short” alleles.

Not all studies show a direct relationship between a gene variant and a phenotype. Instead, developmental or concurrent environments may moderate an association between genes and phenotypes. For example, a well-known study by Caspi et al. (2003) identifies a gene-environment (GxE) interaction for the influence of life stress on depression. The authors find that individuals with a functional polymorphism in the 5HTT gene are more vulnerable to stress-induced depression. Among those individuals that had experienced a relatively large number of stressful life events, about 33% of the carriers of the less efficient “short” allele were cases of diagnosed depression as compared to only 17% of the individuals that carried both long alleles. Thus, in the Caspi et al. (2003) study the gene itself is not associated with depression. Rather, it is the combination of both gene and environment that yields a significant association. In this study we do not find evidence of a gene-environment interaction, but the direct association between the number of long 5HTT alleles and subjective well-being that we establish suggests that future research could benefit from exploring how environmental factors moderate this association.

4 The Add Health Data

This research is based on genetic and survey data collected as part of The National Longitudinal Study of Adolescent Health (Add Health). Add Health explores the causes of health-related behavior of adolescents in grades 7 through 12 and their outcomes in young adulthood. It has been employed widely across disciplines and recently it has produced important contributions in economics (Echenique, Fryer and Kaufman 2006, Echenique and Fryer 2007, Alcott, Karlan, Mobius, Rosenblat and Szeidl 2007, Norton and Han 2009). The
first wave of the Add Health study (1994-1995) selected 80 high schools from a sampling frame of 26,666. The schools were selected based on their size, school type, census region, level of urbanization, and percent of the population that was white. Participating high schools were asked to identify junior high or middle schools that served as feeder schools to their school. This resulted in the participation of 145 middle, junior high, and high schools. From those schools, 90,118 students completed a 45-minute questionnaire and each school was asked to complete at least one School Administrator questionnaire. This process generated descriptive information about each student, the educational setting, and the environment of the school. From these respondents, a core random sample of 12,105 adolescents in grades 7-12 were drawn plus several over-samples, totaling more than 27,000 adolescents. These students and their parents were administered in-home surveys in the first wave.

Wave II (1996) was comprised of another set of in-home interviews of more than 15,000 students from the Wave I sample and a follow-up telephone survey of the school administrators. Finally, Wave III (2001-2002) consisted of an in-home interview, six years later, of 15,170 Wave I participants. The result of this sampling design is that Add Health is a nationally representative study. Women make up 49% of the study’s participants, Hispanics 12.2%, Blacks 16.0%, Asians 3.3%, and Native Americans 2.2%. Participants in Add Health also represent all regions of the country: the Northeast makes up 17% of the sample, the South 27%, the Midwest 19%, and the West 17%.

Allelic information for six genetic markers are available for 2,574 individuals as part of Wave III, including markers that identify alleles of 5HTT. Details of the DNA collection and genotyping process are available at the Add Health website (Add Health Biomarker Team 2007). 5HTT alleles vary in their transcriptional activity, with a “long” 528 base-pair allele associated with a much higher basal activity than the shorter 484 base-pair allele. Allele frequency for the short allele is 42% and for the long allele is 58%.

In Wave III, subjects were asked “How satisfied are you with your life as a whole?”

A breakdown for those providing DNA samples is presented in the Appendix.
Answer categories ranged from very dissatisfied, dissatisfied, neither satisfied nor dissatisfied, satisfied, to very satisfied. Alternative answers were “refused” or “don’t know” and these were discarded for the purpose of this study (less than 1% of interviewees gave such a response). This question and answer formulation is standard in the economics of happiness literature (Di Tella et al. 2001, Di Tella et al. 2003, Kahneman and Krueger 2006, Frey 2008). The distribution of answers to the life satisfaction question is shown in Appendix. In line with the happiness literature, a large majority of respondents report being satisfied or very satisfied (Frey and Stutzer 2002a). That most people, in fact, report a positive level of subjective well-being is the object of a paper by Diener and Diener (1996), where the authors find this distribution to be representative in a wide cross-national analysis.

While this question gives us a valuable opportunity to explore the genetic basis of subjective well-being, we want to make clear two limitations of the data. First, it would be preferable to not only have self-reported life satisfaction answers but also alternative measures of well-being. Second, it would also be preferable to have information about the happiness levels of older adults. The Add Health sample is restricted to individuals who are 18-26 years old during Wave III, so it is possible that our results apply only to subjective well-being by young adults and not to people in different age categories. However, the strong similarity in the distribution of answers in the Add Health data as compared to other life satisfaction surveys used in the happiness literature suggests that the age limits are not likely to gravely distort our results.

5 Genetic Association

Genetic association studies test whether an allele or genotype occurs more frequently within a group exhibiting a particular phenotype than those without the phenotype. However, a significant association can mean one of three things: (1) The allele itself influences subjective well-being; (2) the allele is in “linkage disequilibrium” with an allele at another locus that influences subjective well-being; or (3) the observed association is a false positive signal due
Population stratification occurs because groups may have different allele frequencies due to their genetic ancestry. Subjective well-being in these groups may be the product of their environments, alleles other than the one of interest, or some unobserved reason. For example, two groups may not have mixed in the past for cultural reasons. Through the process of natural selection or genetic drift these groups may develop different frequencies of a particular allele. At the same time, the two groups may also develop divergent behaviors that are not influenced by the allele but solely by the environment in which they live. Once these two groups mix in a larger population, simply comparing the frequency of the allele to the observed behavior would lead to a spurious association.

There are two main research designs employed in association studies, case-control designs and family-based designs. Case-control designs compare the frequency of alleles or genotypes among subjects that exhibit a phenotype of interest to subjects who do not. As a result, case-control designs are vulnerable to population stratification if either group is especially prone to selection effects. A typical way to control for this problem is to include controls for the race or ethnicity of the subject or to limit the analysis to a specific racial or ethnic group. Family-based designs eliminate the problem of population stratification by using family members, such as parents or siblings, as controls. Tests using family data compare whether offspring exhibiting the trait receive a risk allele from their parents more often than would be expected by chance. This design is very powerful in minimizing type I error but also suffers from much lower power in detecting a true association. Xu and Shete (2006) show, based on extensive simulation work, that a case-control association study using mixed-effects regression analysis outperforms family-based designs in detecting an association while at the same time effectively limiting type I error.

Given our data, we cannot differentiate between 1 and 2. In order to do so we would need additional genetic information about loci in close proximity to the locus of interest. Thus, a significant association means that either a particular allele, or one likely near it on the same gene, significantly influences subjective well-being.
To test for genetic association we employ a mixed-effects OLS regression model:

\[ Y_{ij} = \beta_0 + \beta_G G_{ij} + \beta_k Z_{kij} + U_j + \epsilon_{ij} \]

where \(i\) and \(j\) index subject and family respectively. For the 5HTT gene, \(G = 2\) if the subject’s genotype is LL, \(G = 1\) for genotypes LS or SL, and \(G = 0\) if the subject’s genotype is SS (where L represents having a copy of a 528 base-pair “long” allele, and S represents having a copy of a 484 base-pair “short” allele). \(Z\) is a matrix of variables to control for the underlying population structure of the Add Health sample as well as potentially mediating factors such as age, gender, education, religiosity, marriage, job, welfare, or medication that may all influence subjective well-being. Finally, the variable \(U\) is a family random effect that controls for potential genetic and environmental correlation among family members, and \(\epsilon\) is an individual-specific error.

To control for the effects of the underlying population structure, we include indicator variables for whether a subject self-reported as Black, Hispanic, or Asian (base category is White). Following the policy of the United States Census, Add Health allows respondents to mark more than one race. Since this complicates the ability to control for stratification, we exclude these individuals \((N = 117)\), but a supplementary analysis including them yields substantively equal results.

6 Results

Table 1 shows the results of several specifications of the models to test the hypothesis that the 5HTT gene is associated with subjective well-being. Each of these specifications includes

---

8The choice between OLS and ordered probit regression analysis rests on whether the categories of the life satisfaction are considered cardinal or ordinal. Economists typically consider these happiness scores as ordinal and have mainly opted for the ordered type of analysis. Psychologists and sociologists interpret happiness categories as cardinal and therefore use OLS. Ferrer-i-Carbonell and Frijters (2004) survey and test both empirical literatures to conclude that assuming cardinality or ordinality of happiness surveys makes little difference in studies where the dependent variable is measured at a single point in time. We opted for OLS, but other analyses using ordered probit reveal no meaningful differences in coefficients or significance.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Coeff.</th>
<th>SE</th>
<th>P-value</th>
<th>Coeff.</th>
<th>SE</th>
<th>P-value</th>
<th>Coeff.</th>
<th>SE</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>5HTT long</td>
<td>0.059</td>
<td>0.023</td>
<td>0.012</td>
<td>0.065</td>
<td>0.023</td>
<td>0.005</td>
<td>0.070</td>
<td>0.029</td>
<td>0.017</td>
</tr>
<tr>
<td>Black</td>
<td>-0.111</td>
<td>0.048</td>
<td>0.021</td>
<td>-0.114</td>
<td>0.049</td>
<td>0.020</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.198</td>
<td>0.117</td>
<td>0.092</td>
<td>0.216</td>
<td>0.118</td>
<td>0.067</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>-0.196</td>
<td>0.073</td>
<td>0.007</td>
<td>-0.221</td>
<td>0.072</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.004</td>
<td>0.009</td>
<td>0.705</td>
<td>-0.011</td>
<td>0.009</td>
<td>0.262</td>
<td>-0.031</td>
<td>0.012</td>
<td>0.008</td>
</tr>
<tr>
<td>Male</td>
<td>0.014</td>
<td>0.033</td>
<td>0.682</td>
<td>0.028</td>
<td>0.033</td>
<td>0.406</td>
<td>0.039</td>
<td>0.041</td>
<td>0.341</td>
</tr>
<tr>
<td>Job</td>
<td>0.093</td>
<td>0.041</td>
<td>0.024</td>
<td>0.104</td>
<td>0.057</td>
<td>0.071</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College</td>
<td>0.115</td>
<td>0.033</td>
<td>0.001</td>
<td>0.238</td>
<td>0.042</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.232</td>
<td>0.041</td>
<td>0.000</td>
<td>0.318</td>
<td>0.050</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Divorced</td>
<td>-0.313</td>
<td>0.153</td>
<td>0.041</td>
<td>-0.310</td>
<td>0.155</td>
<td>0.047</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Religiosity</td>
<td>0.103</td>
<td>0.017</td>
<td>0.000</td>
<td>0.082</td>
<td>0.023</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Welfare</td>
<td>-0.236</td>
<td>0.098</td>
<td>0.017</td>
<td>-0.121</td>
<td>0.153</td>
<td>0.432</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medication</td>
<td>-0.045</td>
<td>0.032</td>
<td>0.162</td>
<td>-0.095</td>
<td>0.041</td>
<td>0.021</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>4.078</td>
<td>0.208</td>
<td>0.000</td>
<td>4.096</td>
<td>0.210</td>
<td>0.000</td>
<td>4.514</td>
<td>0.262</td>
<td>0.000</td>
</tr>
</tbody>
</table>

| N     | 2545   | 2528  | 1446   |
| R²    | 0.01   | 0.06  | 0.08   |

Table 1: OLS models of association between 5HTT and life satisfaction. Variable definitions are in the Appendix. Standard errors (SE) and P-values are also presented.

Model 1 shows that the long allele of 5HTT is significantly associated with increased life satisfaction ($p = 0.012$). In Figure 2, we summarize the results for 5HTT by simulating first differences from the coefficient covariance matrix of Model 1. Holding all else constant and changing the 5HTT gene of all subjects from zero to one long allele would increase the reporting of being very satisfied with one’s life in this population by about 8.5%. Similarly, changing the 5HTT gene from zero to two long alleles would increase the reporting of being very satisfied by about 17.3%.

Model 2 includes a number of socio-economic factors that are known to influence subjective well-being. In particular, having a job, education, marriage, divorce, religiosity, welfare assistance, and being on medication. This model also suggests that there is a statistically significant association ($p = 0.005$) between additional long alleles of the 5HTT gene and the
Figure 2: Increasing the number of “long,” more efficient 5HTT alleles yields significantly higher life satisfaction. First differences, based on simulations of Model 1 parameters, are presented along with 95% confidence intervals. All other variables are held at their means.
reporting of life satisfaction. These variables may in fact mediate the relationship between the gene we have identified and life satisfaction.\(^9\) We might also expect genes to contribute to variation in socio-economic factors (Bowles and Gintis 2002), which in turn would impact subjective well-being. In order to test whether these variables are potentially mediators, we regress each of them separately on 5HTT long along with race, age, and gender controls. Since 5HTT long is not significantly associated with any of these variables, we can rule them out as mediators.\(^9\)

Following Xu and Shete (2006), as a robustness test for population stratification, we also include Model 3 that is a case-control association model for those subjects that uniquely identified themselves as being white. The coefficient on 5HTT and its p-value \((p = 0.017)\) suggest that population stratification between self-reported racial categories is not driving the association between 5HTT and life satisfaction.

7 Discussion

The goal of this paper is to provide empirical evidence that individuals having a specific gene—a polymorphism of the serotonin transporter (5HTT) gene—report significantly higher levels of life satisfaction. We thus address the “baseline” or “set point” of subjective well-being which previous research has identified to be of great importance (Kahneman et al. 1999, Graham 2008). The existence of a stable baseline does not mean that the socio-economic influences on happiness so far identified by researchers are unimportant. Rather, our results complement these studies and suggest a new direction for research. As shown by the \(R^2\) value in Table 1, the 5HTT gene explains less than one percent of the variation in life satisfaction, but past twin studies show that genes account for about half the total variance (Lykken

\(^9\)A variable \(M\) mediates the relationship between an independent variable \(X\), in our case a genotype, and a dependent variable \(Y\), in our case life satisfaction, if (1) \(X\) significantly predicts \(Y\), (2) \(X\) significantly predicts \(M\), and (3) \(M\) significantly predicts \(Y\) controlling for \(X\) (Baron and Kenny 1986).

\(^{10}\)The \(p\) values for 5HTT long in each regression are presented in the appendix. Medication has a \(p = 0.08\) but looses its significance \((p = 0.17)\) when run on life satisfaction while controlling for 5HTT long along with age, gender, and race. Hence, medication cannot be considered a mediating variable (Baron and Kenny 1986).
and Tellegen 1996). Therefore, there are probably many other genes which, in conjunction
with environmental factors, help to explain how baseline happiness varies from one person
to another. The association with 5HTT is just the first of many such associations that will
likely be identified over the course of the next few years.

While the Add Health data set we utilize refers to young adults living in the United States,
the distribution of responses to the life satisfaction question mirrors the results reported by
other scholars in earlier research (Frey and Stutzer 2002a). Furthermore, the estimates
of the influence of socio-demographic, economic, and cultural covariates on subjective life
satisfaction in Models 2 and 3 corroborate the generally identified systematic effects of these
variables in the literature (for a survey, see Dolan et al. 2008). In particular, gender does not
systematically affect happiness. Higher age has a negative, though not statistically significant
effect (this is not surprising considering that our sample refers to young adults). Blacks and
Asians are systematically less happy than are Whites, while Latinos are somewhat happier,
but not in a statistically significant way. Better educated and married individuals report to
have significantly higher life satisfaction, while divorced persons a more unhappy. Having a
job strongly raises life satisfaction. This reflects the psychic benefits of being occupied and
integrated in society. At the same time it shows that having an income raises life satisfaction.
In contrast, persons on welfare are much less happy than those employed which reflects the
psychic costs of unemployment. Religious individuals are significantly more happy than
those without religious beliefs. Persons with less good health, as measured by the need to
be on medication, are also less happy. As is the case with most research on happiness, these
estimates identify correlations, not causality, given the difficulty in disentangling endogeneity.
But consistency with previous studies suggests that results using the Add Health data may
generalize to other populations.

We believe that genetic association studies such as ours may complement neurosience-
tific research. For example, the work by (Urry, Nitschke, Dolski, Jackson, Dalton, Mueller,
Rosenkranz, Ryff, Singer and Davidson 2004) presents neural correlates of subjective well-
being. Some of the neurological variation they observe may result from differences in genotype. Genes are upstream from the neurological processes that Urry et al. (2004) identify and may thus bring us closer to understanding the sources of subjective well-being.

8 Conclusion

Our results suggest that genetic factors significantly influence individuals' subjective well-being. We have identified one particular gene variant—5HTT long—as having a sizeable positive association with self-reported life satisfaction. More broadly, these results suggest a tentative new step forward for economics as a discipline. They show that incorporating genetic information into our theories and analysis may contribute to a greater understanding of the sources of subjective well-being. Integrating the unique biology of each individual, in addition to studying experience and environment may usefully complement existing models (Caplin and Dean 2008).

The empirical approach we employ in this paper improves on twin study designs in an important way. Twin studies are valuable for measuring the total influence of all genes, but they are agnostic about which genes contribute to variation (Cesarini, Dawes, Johannesson, Lichtenstein and Wallace 2009). By focusing on specific genes, our analysis is able to suggest potential causal pathways through which genes influence happiness levels. A significant body of research has found that the 5HTT gene influences human psyche via its impact on neurological processes, thereby establishing a potential causal chain leading from this gene to self-reported life satisfaction. Given prior research linking the “short,” less efficient alleles of the 5HTT gene to depression, and the “long,” more efficient alleles to optimism bias, we hypothesized that carriers of the “long” alleles would be more likely to report being happy, and this intuition is supported by the data.

The causal structure must be further studied once additional data reporting the genetic endowment of individuals coupled with data on their subjective well-being become available. We have stressed that genetic factors complement, rather than substitute for, the existing
studies showing the influence of socio-demographic, economic and cultural variables on life satisfaction. This study seeks to show the combined influences of biological and environmental influences on subjective well-being; but it is only a beginning and may open up a fascinating new research frontier.

Future work should use genetic association studies to identify other specific genes that are implicated in subjective well-being. Finding out which genes they are and what physical function they have will improve our understanding of the biological processes that underlie well-being and may also shed light on their evolutionary origin (Fitzpatrick, Ben-Shahar, Smid, Vet, Robinson and Sokolowski 2005). While the 5HTT gene may explain a significant portion of the variation in happiness, it is important to re-emphasize that there is no single “happiness gene.” Instead, there is likely to be a set of genes whose expression, in combination with environmental factors, influences subjective well-being.
Appendix

Variable Definitions

5HTT long is a variable for having 0, 1, or 2 of the 528 base-pair alleles of the 5HTT gene (as opposed to the 484 base-pair version). The race/ethnicity indicator variables are based on the questions “Are you of Hispanic or Latino origin?” and “What is your race? [white/black or African American/American Indian or Native American/Asian or Pacific Islander]”. Age is self-reported age and Male is an indicator taking the value of 1 if the respondent is a male and 0 for a female. Job is the response to the question “Do you currently have a job?” College is an indicator variable taking the value 1 if the respondent completed at least one year of college and 0 for no college. It is based on the question “What is the highest grade or year of regular school you completed?” Married and Divorced are dummies derived from the population subset that have married and answered “Are you still married?” Religiosity relies on “To what extent are you a religious person?” and takes a value between 0 and 3 for very religious. Welfare is a dummy for “Are you receiving welfare?” Medication is a dummy for “In the past 12 months, have you taken any prescription medication—that is, a medicine that must be prescribed by a doctor or nurse?”
Summary Statistics

Table 2: Sample means.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std Dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life satisfaction</td>
<td>4.20</td>
<td>0.79</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>5HTT long</td>
<td>1.14</td>
<td>0.72</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Age</td>
<td>21.9</td>
<td>1.7</td>
<td>18</td>
<td>26</td>
</tr>
<tr>
<td>Religiosity</td>
<td>1.43</td>
<td>0.92</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3: Percentage of subjects exhibiting these characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>70.9</td>
</tr>
<tr>
<td>Black</td>
<td>19.0</td>
</tr>
<tr>
<td>Hispanic</td>
<td>14.7</td>
</tr>
<tr>
<td>Asian</td>
<td>8.2</td>
</tr>
<tr>
<td>Male</td>
<td>47.8</td>
</tr>
<tr>
<td>College</td>
<td>54.9</td>
</tr>
<tr>
<td>Married</td>
<td>17.3</td>
</tr>
<tr>
<td>Divorced</td>
<td>1.4</td>
</tr>
<tr>
<td>Welfare</td>
<td>4.2</td>
</tr>
<tr>
<td>Medication</td>
<td>61.2</td>
</tr>
</tbody>
</table>
Figure 3: Distribution of life satisfaction.

How satisfied are you with your life as a whole?
### Table 4: Cross-tabs

<table>
<thead>
<tr>
<th>Life satisfaction</th>
<th>5HTT long</th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very dissatisfied</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>(0.8%)</td>
<td>(0.3)%</td>
<td>(0.6)%</td>
<td>(0.5)%</td>
</tr>
<tr>
<td>Dissatisfied</td>
<td>17</td>
<td>35</td>
<td>13</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>3.3%</td>
<td>2.9%</td>
<td>1.6%</td>
<td>2.6%</td>
</tr>
<tr>
<td>Neither</td>
<td>72</td>
<td>149</td>
<td>97</td>
<td>318</td>
</tr>
<tr>
<td></td>
<td>14.2%</td>
<td>12.6%</td>
<td>11.3%</td>
<td>12.4%</td>
</tr>
<tr>
<td>Satisfied</td>
<td>226</td>
<td>544</td>
<td>394</td>
<td>1,164</td>
</tr>
<tr>
<td></td>
<td>44.4%</td>
<td>45.9%</td>
<td>45.7%</td>
<td>45.6%</td>
</tr>
<tr>
<td>Very satisfied</td>
<td>190</td>
<td>453</td>
<td>353</td>
<td>996</td>
</tr>
<tr>
<td></td>
<td>37.3%</td>
<td>38.2%</td>
<td>41.0%</td>
<td>39.0%</td>
</tr>
<tr>
<td>Total</td>
<td>509</td>
<td>1,185</td>
<td>862</td>
<td>2,556</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

### Table 5: Life satisfaction and genotype, by race.

<table>
<thead>
<tr>
<th>Race</th>
<th>Life satisfaction Mean</th>
<th>5HTT long Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>4.24</td>
<td>1.12</td>
</tr>
<tr>
<td>Black</td>
<td>4.13</td>
<td>1.47</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4.22</td>
<td>0.93</td>
</tr>
<tr>
<td>Asian</td>
<td>4.01</td>
<td>0.69</td>
</tr>
</tbody>
</table>
Potential Mediators

<table>
<thead>
<tr>
<th></th>
<th>5HTT long</th>
</tr>
</thead>
<tbody>
<tr>
<td>DV</td>
<td>p – value</td>
</tr>
<tr>
<td>Job</td>
<td>0.17</td>
</tr>
<tr>
<td>College</td>
<td>0.99</td>
</tr>
<tr>
<td>Married</td>
<td>0.33</td>
</tr>
<tr>
<td>Divorced</td>
<td>0.16</td>
</tr>
<tr>
<td>Religiosity</td>
<td>0.48</td>
</tr>
<tr>
<td>Welfare</td>
<td>0.25</td>
</tr>
<tr>
<td>Medication</td>
<td>0.08</td>
</tr>
</tbody>
</table>

Table 6: Table presents p values for 5HTT long in models with job, college attendance, married, divorced, religious, welfare, and medication as dependent variables. Regressions also include race, age, and gender controls.

References


Thomas Aronsson and Erkki Koskela, Outsourcing, Public Input Provision and Policy Cooperation, December 2009

Andreas Ortmann, “The Way in which an Experiment is Conducted is Unbelievably Important”: On the Experimentation Practices of Economists and Psychologists, December 2009

Andreas Irmen, Population Aging and the Direction of Technical Change, December 2009


Claudia Buch, Cathérine Tahmee Koch and Michael Kötter, Margins of International Banking: Is there a Productivity Pecking Order in Banking, too?, December 2009


Johannes Binswanger and Daniel Schunk, What is an Adequate Standard of Living during Retirement?, December 2009

Armin Falk and James J. Heckman, Lab Experiments are a Major Source of Knowledge in the Social Sciences, December 2009


Ben J. Heijdra and Laurie S. M. Reijnders, Economic Growth and Longevity Risk with Adverse Selection, December 2009

Johannes Becker, Taxation of Foreign Profits with Heterogeneous Multinational Firms, December 2009
2900 Douglas Gale and Piero Gottardi, Illiquidity and Under-Valuation of Firms, December 2009


2902 Arno Riedl, Behavioral and Experimental Economics Can Inform Public Policy: Some Thoughts, December 2009


2904 Marcel Gérard and Fernando M. M. Ruiz, Corporate Taxation and the Impact of Governance, Political and Economic Factors, December 2009


2906 Xavier Vives, Asset Auctions, Information, and Liquidity, January 2010

2907 Edwin van der Werf, Unilateral Climate Policy, Asymmetric Backstop Adoption, and Carbon Leakage in a Two-Region Hotelling Model, January 2010

2908 Margarita Katsimi and Vassilis Sarantides, Do Elections Affect the Composition of Fiscal Policy?, January 2010

2909 Rolf Golombek, Mads Greaker and Michael Hoel, Climate Policy without Commitment, January 2010

2910 Sascha O. Becker and Ludger Woessmann, The Effect of Protestantism on Education before the Industrialization: Evidence from 1816 Prussia, January 2010

2911 Michael Berlemann, Marco Oestmann and Marcel Thum, Demographic Change and Bank Profitability. Empirical Evidence from German Savings Banks, January 2010

2912 Øystein Foros, Hans Jarle Kind and Greg Shaffer, Mergers and Partial Ownership, January 2010

2913 Sean Holly, M. Hashem Pesaran and Takashi Yamagata, Spatial and Temporal Diffusion of House Prices in the UK, January 2010

2914 Christian Keuschnigg and Evelyn Ribi, Profit Taxation and Finance Constraints, January 2010

2915 Hendrik Vrijburg and Ruud A. de Mooij, Enhanced Cooperation in an Asymmetric Model of Tax Competition, January 2010

2916 Volker Meier and Martin Werding, Ageing and the Welfare State: Securing Sustainability, January 2010
2917 Thushyanthan Baskaran and Zohal Hessami, Globalization, Redistribution, and the Composition of Public Education Expenditures, January 2010

2918 Angel de la Fuente, Testing, not Modelling, the Impact of Cohesion Support: A Theoretical Framework and some Preliminary Results for the Spanish Regions, January 2010

2919 Bruno S. Frey and Paolo Pamini, World Heritage: Where Are We? An Empirical Analysis, January 2010

2920 Susanne Ek and Bertil Holmlund, Family Job Search, Wage Bargaining, and Optimal Unemployment Insurance, January 2010

2921 Mariagiovanna Baccara, Allan Collard-Wexler, Leonardo Felli and Leeat Yariv, Gender and Racial Biases: Evidence from Child Adoption, January 2010

2922 Kurt R. Brekke, Roberto Cellini, Luigi Siciliani and Odd Rune Straume, Competition and Quality in Regulated Markets with Sluggish Demand, January 2010

2923 Stefan Bauernschuster, Oliver Falck and Niels Große, Can Competition Spoil Reciprocity? – A Laboratory Experiment, January 2010

2924 Jerome L. Stein, A Critique of the Literature on the US Financial Debt Crisis, January 2010

2925 Erkki Koskela and Jan König, Profit Sharing, Wage Formation and Flexible Outsourcing under Labor Market Imperfection, January 2010

2926 Gabriella Legrenzi and Costas Milas, Spend-and-Tax Adjustments and the Sustainability of the Government’s Intertemporal Budget Constraint, January 2010

2927 Piero Gottardi, Jean Marc Tallon and Paolo Ghirardato, Flexible Contracts, January 2010

2928 Gebhard Kirchgässner and Jürgen Wolters, The Role of Monetary Aggregates in the Policy Analysis of the Swiss National Bank, January 2010

2929 J. Trent Alexander, Michael Davern and Betsey Stevenson, Inaccurate Age and Sex Data in the Census PUMS Files: Evidence and Implications, January 2010

2930 Stefan Krasa and Mattias K. Polborn, Competition between Specialized Candidates, January 2010

2931 Yin-Wong Cheung and Xingwang Qian, Capital Flight: China’s Experience, January 2010

2932 Thomas Hemmelgarn and Gaetan Nicodeme, The 2008 Financial Crisis and Taxation Policy, January 2010
2933 Marco Faravelli, Oliver Kirchkamp and Helmut Rainer, Social Welfare versus Inequality Concerns in an Incomplete Contract Experiment, January 2010

2934 Mohamed El Hedi Arouri and Christophe Rault, Oil Prices and Stock Markets: What Drives what in the Gulf Corporation Council Countries?, January 2010

2935 Wolfgang Lechthaler, Christian Merkl and Dennis J. Snower, Monetary Persistence and the Labor Market: A New Perspective, January 2010

2936 Klaus Abberger and Wolfgang Nierhaus, Markov-Switching and the Ifo Business Climate: The Ifo Business Cycle Traffic Lights, January 2010

2937 Mark Armstrong and Steffen Huck, Behavioral Economics as Applied to Firms: A Primer, February 2010

2938 Guglielmo Maria Caporale and Alessandro Girardi, Price Formation on the EuroMTS Platform, February 2010

2939 Hans Gersbach, Democratic Provision of Divisible Public Goods, February 2010

2940 Adam Isen and Betsey Stevenson, Women’s Education and Family Behavior: Trends in Marriage, Divorce and Fertility, February 2010

2941 Peter Debaere, Holger Görg and Horst Raff, Greasing the Wheels of International Commerce: How Services Facilitate Firms’ International Sourcing, February 2010

2942 Emanuele Forlani, Competition in the Service Sector and the Performances of Manufacturing Firms: Does Liberalization Matter?, February 2010

2943 James M. Malcomson, Do Managers with Limited Liability Take More Risky Decisions? An Information Acquisition Model, February 2010

2944 Florian Englmaier and Steve Leider, Gift Exchange in the Lab – It is not (only) how much you give …, February 2010

2945 Andrea Bassanini and Giorgio Brunello, Barriers to Entry, Deregulation and Workplace Training: A Theoretical Model with Evidence from Europe, February 2010

2946 Jan-Emmanuel De Neve, James H. Fowler and Bruno S. Frey, Genes, Economics, and Happiness, February 2010