Artículo científico

Genetics and economics: Studies and perspectives

Genética y economía: estudios y perspectivas

Alfredo M. Navarro †
Full member of the National Academy of Economic Sciences of Argentina., Argentina

Resumen: Este trabajo analiza la relación entre la genética y la economía y los efectos de la primera y del medio ambiente sobre el fenotipo de los seres humanos. Se describen, entre otras cuestiones, la forma en que se pueden usar los datos del genoma humano para explicar ciertas características económicas, como la existencia de espíritu empresario, de actitud ante el riesgo, el nivel de los ingresos, la propensión a invertir y las facilidades para recibir educación. Se analizan una serie de trabajos relacionados con esta cuestión y se describe las distintas maneras de extraer conclusiones de los datos existentes. Concluye afirmando que este tipo de estudios está en sus comienzos, pero que, dado que ayuda a comprender mejor, juntamente con la evaluación del medio ambiente, ciertas características de la conducta humana, abre un interesante campo de estudios para la ciencia económica.

Palabras clave: genética, genoma humano, comportamiento humano.

Abstract: This paper analyzes the relationship between genetics and economics and the effects of both genetics and the environment on the phenotype of human beings. Among other issues, it describes how data from the human genome can be used to explain certain economic characteristics, such as the existence of entrepreneurship, attitude towards risk, income level, propensity to invest, and ease of receiving education. A series of works related to this issue are analyzed, and the various ways of drawing conclusions from existing data are described. The paper concludes by stating that this type of study is in its infancy, but that it helps to better understand, along with the evaluation of the environment, certain characteristics of human behavior and opens an interesting field of study for economics.

Keywords: genetics, human genome, human behavior.

INTRODUCTION

Although we know that children share certain physical or behavioral characteristics with their parents, there has always been a question as to whether this similarity is due to the fact that they share the same environment, the same education, and the same social group or to the transfer of a certain genetic code, or if both factors influence a proportion that we do not know but that we can try to determine.

This question dates to the beginnings of modern biology. Charles Darwin argued that evolution is possible because living beings transmit their characteristics to their descendants, incorporating increasingly complex mechanisms, which make different species to change towards more advanced forms.

In 1975, Edward O. Wilson published *Sociobiology,* a work that analyzes the social relationships of all living beings. In the last chapter Wilson studies the differences among humans, and unleashes a controversy, because while Wilson argues that genetic input plays a very important role in them, most social science scholars of that time attributed them to the environment and believed that they were due not to hereditary reasons, but rather to some individuals belonging to social groups that had the necessary economic means to enjoy better food, health and housing and, above all, education.

These scholars believed that culture had replaced genes as the cause of the evolutionary process, since these could not change so quickly as to explain the dizzying transformation of the world which occurred in the last two centuries, while changes in cultural elements had operated drastically. This discussion, which was very heated, raises a very important question: what defines our phenotype, the factors which we inherit through the genes transmitted to us by our ancestors, or those we receive from the environment in which we operate?

It is worth pausing for a moment on this question, for which we must keep in mind the work of Steven Pinker (2002) *The Blank Slate*, in which he explains the nature of the, at times, violent discussion that caused the appearance of Wilson's work, and whose title refers to the concept of *clean slate,* as a characteristic of all human beings, whose differences do not come from nature, but from the society in which they have developed and from the education they have received. He points out that John Locke (1632-1704), who laid the foundations of empiricism, maintains that everything comes from experience and that therefore there is no innate difference between men, since all souls are equal in principle, so there is no support for class differences, monarchy, women discrimination or slavery. This idea, which had already been anticipated by René Descartes (1596-1650), is picked up by Jean-Jacques Rousseau (1712-1778) and coincides with Marxist thought. This explains the reaction against Wilson and the emergence of the debate that was called *nature vs. nurture*, which established the competition between the two to explain our characteristics: we have *inherited them from our ancestors, or we have acquired them from the society in which we live in*. The greater the weight of the latter, the greater the result we can obtain by modifying the social system in which we are immersed.

The appearance of genetics had a decisive impact on this question. Not *everything is nurture* anymore. we now have evidence that a part of our characteristics come from our ancestors, and we are no longer all the same in the event that we have a different genetic endowment. If we think that equality between human beings is something we want to move towards,[1] having a clearer vision of what human nature and society is will surely help us build a more egalitarian world to assist those who have inherited deficiencies. It is also necessary to make it clear that the effects of the environment cannot be ignored, and that, depending on the problem analyzed, this fact has more or less importance in the explanation of the different phenotypes.

In this paper we have proposed to review some of the works that we have considered most relevant and, based on this, to make an evaluation of the prospects of this line of research.

INITIAL EXPERIENCES

During the last quarter of the last century, many studies were carried out which tried to explain the presence of a certain characteristic based on hereditary and environmental factors by examining the differences between own or adopted children and between monozygotic and dizygotic twins. The former share all of their DNA, but the latter share only one half.[2] Therefore, if we analyze certain characteristics of the offspring that correlate with those of their parents in greater proportion in the monozygotic than in the dizygotic twins, we can assume that this is due to hereditary factors, since the environment in which the twins develop is very similar because they are raised in the same family in most cases. This method of analyzing the problem is known as "heritability," and both Sacerdote et al. (2011) and Manski (2011) present simple mathematical models to define the task performed with this methodology. The latter tells us that the model to be estimated is of the form:

$$y = g + e \qquad [1]$$

where y is a characteristic of the phenotype, g represents the inherited factors, and e represents the influence of the environment. Assuming that the covariance between g and e is zero, it considers $\mathrm{Var}(y) + \mathrm{Var}(g) + \mathrm{Var}(e)$, defining heritability as the ratio $\mathrm{Var}(g)/\mathrm{Var}(y)$. Let's look at some examples of these preliminary studies.

1. Following this methodology, one of the first attempts to explain the differences exhibited by individuals in the labor market, considering the cultural conditions of the environment and the transmission of conditions through genes, was carried out by Behrman and Taubman (1976) who conducted a study based on a sample of approximately 2,000 twins, some of whom were identical, and some non-identical. They were able to observe that identical twins had more equivalent results in their educational performance and income level than those who were not identical. They came to the provisional conclusion that both cultural context and genetic transmission influence the years of education, socioeconomic status, and income level of the individuals in the group studied.

2. Ebstein et al. (2010) report the results of a study that also analyzes the different correlations between the phenotypes of the two types of twins and their parents, in which they conclude that issues such as prosocial behavior, stress, infidelity, empathy, political attitudes, leadership, aggression, parental care and risk behavior are influenced by hereditary factors.

3. Another example of these preliminary studies was the debate that occurred in the United States regarding the effect of cigarette smoking on health.[3] Several studies demonstrated the association between smoking and lung cancer, circulatory problems and other diseases. American tobacco companies turned to a renowned geneticist, Clarence Little, president of the University of Michigan, who claimed that lung cancer was acquired by genetic inheritance, and not necessarily through smoking. But while it has been shown that a particular gene is associated with lung cancer, it has also been statistically proven that smoking contributes to making it more likely that those with that genetic characteristic will get the disease.

This error is a consequence of considering only genetic factors and leaving aside those related to the environment in which each person lives, beyond suspicions of the existence of economic interests.

4. Benjamin et al. (2010) examined the use of genetic information to analyze economic problems. This represents a significant advancement over previous methodologies, which lacked the current understanding of the human genome. According to these authors, economics can elucidate how market forces react to genetic factors, while genetics can aid economists in identifying and quantifying key causal relationships, potentially leading to alternative suggestions for economic policy.

5. While Manski (2011) argues that the analysis of heritability should be abandoned due to its lack of policy design utility, Hyytinen et al. (2019) advocate for the continued use of this methodology. Although agreeing with Manski's objections regarding its application in public policy, they assert the value of further research in this area to enhance understanding of inheritance effects. Their study presents a table comparing the outcomes of 21 studies applying this methodology across various countries, highlighting significant differences between monozygotic and dizygotic twins. The research, using data from Finland on twins born before 1958, is distinctive in that it analyzes the income of the sample over twenty years (representing permanent income), considers gender, and involves a relatively large sample: 1292 monozygotic women and 1012 men, and 2438 dizygotic women and 2316 men. The findings indicate that 40% of the variance in women's income is attributable to genetic factors, and this percentage increases to just over 50% for men, depending on variations in the estimated model,[4] with statistically highly significant results. The study also reveals a greater hereditary factor in men and an influence of environmental conditions, both shared and non-shared. This is particularly noteworthy in a country like Finland, known for being more egalitarian than other developed nations.

6. Furthermore, Bingley et al. (2023) analyzed a sample from Denmark using a similar methodology, examining the effects of heredity and environment, whether shared or not, on education years, income, and wealth. They introduced modifications to the ACE model, both in terms of altering its assumptions and incorporating the analysis of these variables on the children and spouses of monozygotic and dizygotic twins. Their findings suggest a greater influence of the environment, whether shared or not, indicating that the shared environment accounts for 50% of the variance in years of education[5]. While the results with these modified assumptions seem reasonable, the need for continued meticulous analysis to obtain reliable outcomes is evident.

7. Finally, to conclude this topic, it is pertinent to refer to the work of Sacerdote (2011), who synthesizes a plethora of studies employing the variance decomposition methodology described above to analyze determinants of IQ, income, years of education, and certain personal traits. The findings he reports place greater emphasis on genetic traits over the environment, according to various studies utilizing metaanalysis[6] techniques. These studies indicate that the IQ correlation among non-adopted siblings ranges from 0.44 to 0.54, and between 0.85 to 0.91 for identical twins, suggesting that genetic factors explain between 50 and 60% of this variable. This leads to the conclusion that environmental interventions might be futile, contradicting the evidence we observe. Hence, he advocates using alternative methods and employs multiple regression models, wherein he regresses the aforementioned variables against parental education, gender, and other variables. He also describes studies using models that account for the non-zero correlation between g and e in equation [1]. He concludes that while these studies continue to acknowledge the significance of genetic traits, they find the environment plays a much more substantial role than previously attributed in the earlier studies he references.

WORKS USING THE HUMAN GENOME

The general framework of analysis of the relationship between genotype and phenotype is exposed in the work of Benjamin et al. (2012), which is complemented by Cesarini and Visscher (2017), where they present the following equation:

$$Yi = \sum_{j=1}^{J} X\, ij\beta j + U \qquad [3]$$

Where the explained variable Y is a certain characteristic of the phenotype (e.g., diabetes, learning difficulty, or obesity); j are the different locations of the SNPs; X, is the value of the allele, which can take values of 0, 1 or 2; β_j is the impact that this allele has on the explained variable, i are the different observations of the sample and U a variable, or a set of them, that express environmental conditions. Initially, the first term on the right side of the equation [1] was replaced by the relationship between relatives, which is also a manifestation of inheritance, as we have seen in the works analyzed above.

Genetics has been studied for a long time, but the work of Crick and Watson (1953), in which the famous spiral is described, produced an explosion in the amount of related research.[7]

We now know that the human genome, whose complete sequence has been available since the year 2000, has 23 pairs of chromosomes, inherited one from the father and the other one from the mother, and they form two strands that have the shape of a double helix, composed of basic elements called *nucleotides*, which are the *steps* that form the classic spiral. They contain one of four bases: adenine (A), cytokine (C), thymine (T) or guanine (G), so there are four different classes of nucleotides that are always grouped in pairs, with adenine being associated with thymine (AT) and cytokine with guanine (CG).[8] The human genome has 3,000 million of these base pairs, which make up about 22,000 genes, which are grouped into the 23 chromosomes.[9] Humans share our genetic basis in about 99%, but we differ in what is called *genetic polymorphism* (of which the simplest are the so-called SNPs) which are the parts[10] of DNA in which there are differences between one individual and another. A gene can have hundreds of SNPs, and these can also be in areas that are not part of the genes. It is assumed that there are hundreds of millions of SNPs, and[11] in each of them three possibilities: when we inherit one allele from our father and another from our mother, we can have no minor allele, one minor allele, or two minor alleles. This circumstance, for statistical purposes, is identified with the numbers 0, 1 and 2, values that are used in the regressions.[12]

In the beginning the analysis of the human genome was used in medicine because the mutation of certain genes was related to diseases and the relationship of a single gene with a phenotype was studied. An example is the case of Huntington's disease, a neurological disease that produces, among other symptoms, involuntary movements in the extremities. It was possible to determine which is the gene in which the mutation that produces it is found and then allowing the diagnosis, predict the age at which the disease will appear and its severity, as well as allow better treatment design.

These types of studies began to develop rapidly since 2005 when three researchers,[13] from the *Broad Institute*, Harvard University and MIT, developed a fairly complex statistical analysis method, called GWAS (*Genetic Wide Association Studies*), which is a procedure that seeks to find associations between genetic variants and phenotypes in a population, in order to identify genes or SNPs that may contribute to the appearance of a disease or a certain characteristic. While it is used to study the relationship between genes and different diseases, it can also be used to analyze characteristics of people that have to do with economics,

such as wages, income, risk behavior, entrepreneurship, learning and some other issues.[14] A population that has a certain characteristic is compared to one that does not have it and then by performing regressions (sometimes up to hundreds of thousands of them) where the explained variable is a characteristic (e.g. income level) and explanatory statements of SNP´s or combinations of them. The results are expressed by means of a graphical representation called "Manhattan graph", where on the abscissa axis the 23 chromosomes appear along with the genes and SNPs that compose them, and on the ordered axis the value of the probability assigned to each one, which allows to identify which are the SNPs that are related to a certain characteristic of the phenotype.[15] The GWAS works in interaction with various databases, of which *The Human Gene Mutation Database (HGMD)* stands out,[16] whose function is to compare the relationship between alterations in genes and certain types of alterations of the phenotype related to diseases.

In instances where the phenotype requires the convergence of various genetic variations to exhibit certain characteristics, it is feasible to construct a variable comprising a weighted aggregation of these genetic traits (using Genome-Wide Association Studies, or GWAS). Subsequently, this variable can be utilized as regressors in an equation that elucidates a specific phenotypic characteristic. This method is referred to as regression with *polygenic indexes*.

Soon after it was realized that this new knowledge could also be useful to the social sciences, so they began to study certain characteristics of individuals related to the subjects they study.[17] In instances where certain phenotypic characteristics require the concurrent presence of various genetic variations, it is feasible to construct a variable comprising a weighted combination of these genetic characteristics (utilizing Genome-Wide Association Studies, or GWAS). This variable can then be employed as regressors in an equation aimed at elucidating a specific phenotypic trait. This approach is referred to as regression with polygenic indices.

Thus, was born the discipline known as *genoeconomics* (expression that is due to Benjamin et al. (2010), which studies the use of genetic information to analyze economic problems. It represents an important advance over the methodology that was previously used, when the knowledge of the human genome that we have today was not yet available. Economics, according to these authors, can explain how market forces respond to genetic factors, while genetics can help economists identify and measure important causal relationships, and alternative suggestions to economic policy could emerge.

During the last ten years several works have appeared in which the relationship between the genetic code of individuals and their economic behavior is studied, through two alternative forms of work: the first is to establish *a priori* which molecular markers can produce certain characteristics and then, if people with that marker really have them.

For example, as we know that oxytocin contributes to a behavior tending to solidarity, we can analyze the relationship between this behavior and the presence of a certain molecular marker. The other way is to take a set of people of whom we know certain characteristics and their genetic code and look for correlations between them, which requires the management of an enormous amount of data, since we work with samples from several thousand individuals of which several hundred thousand molecular markers are analyzed.[18] Let us analyze some of the works that we have considered most representative.

1. An example of the first way to study the relationship between genes and phenotype is the experiment described by Benjamin et al. (2007), carried out on data from a work dedicated to analyzing cardiological diseases in Iceland and which began in 1967, in which 30,795 men and women born between 1900 and 1935, living in Reykjavík, were analyzed. In 2002 was studied the genome of 2,300 survivors, who were questioned about some economic issues, such as their years of education and income, their attitude to risk and their intertemporal preference. Then they made a list of SNPs that were suspected candidates to explain these characteristics, where they analyze markers related to the production of dopamine, serotonin, cognitive ability, intelligence level (IQ), memory, and other issues and seek their correlation with the economic

phenotypes. There are certain characteristics in economic behavior that are easy to associate with genes, such as impulsivity, risk aversion, and solidarity, and others that while they may be influenced by genes, which are more distant, such as education level and wealth or income, for which sets of genetic characteristics must be analyzed.

2. Ding et al. (2006) use a database generated to study the effect of cigarette smoking on adolescent behavior, composed of 893 students who were surveyed to determine whether they suffered from obesity, hyperactivity, inattention, drug and tobacco use. In addition, saliva was extracted to analyze their genetic code, with the idea that certain genes had to do with addictions, and these in turn with the state of health that influenced their academic behavior, which in turn did so on their occupation, their income and their family life. Instead of directly analyzing the relationship between the selected genetic markers and the academic result obtained, they build a model where five conveniently chosen genes constitute the explanatory variables. The brain's reward system is linked to addictions, and certain neurotransmitters make us susceptible to contracting them. They describe how an activation of the area of the brain called the *ventral area* releases dopamine, and neurotransmitters carry the signal to the *nucleus accumbens* (limbic part of the brain) to be transmitted from there to the cerebral cortex, which is where decisions are made. These neurotransmitters explain smoking addiction, depression, hyperactivity and other issues, and their abundance is determined by certain genetic markers,[19] which are included in the model used to explain the academic behavior of the adolescents studied. This model is composed of three equations,[20] in two of which the genetic endowment appears as an explanatory variable and is significant in the estimates made. In this way, the genetic factor is incorporated not directly but as one more variable of those that make up the model,[21] which makes the work an interesting methodological contribution. If both the genetic code and the environmental conditions are significant variables in the explanation of certain economic circumstances, we can consider that the debate referred to above has been overcome, since the empirical evidence seems to confirm that both conditions, heredity and environment, influence our characteristics to be determined in each case.

3. Benjamin et al. (2012) study the relationship between the genetic code, permanent income and wealth,[22] and find results that allow us to assume that permanent income is influenced by the genetic code, given the high correlation found between the income of parents and that of children when they are monozygotic twins, higher than that found in dizygotic twins. At the same time, they analyze the methodological problems facing genoeconomics and consider that there is evidence indicating that certain behavioral characteristics are heritable, that the context in which an individual develops is less important to explain their behavior than genetic endowment, and that there is a very important part that is not explained by either of the two factors. Finally, they warn that there is a long way to go, when it comes to risk aversion, the intertemporal discount rate and altruism, among many other issues where genetics can help economists.

4. The second way of working, referred to earlier, is the one used by Beauchamps et al. (2011) which present by way of example, a study[23] on the influence of the environment and genetic transmission. In a first stage they estimate the following equation:

$$Edu = \beta 0 + \beta 1k\, SNPk + \beta 2\, PC + \beta 3\, X + \varepsilon \qquad [4]$$

where Edu are the years of education; SNP_k is the number of copies of the smallest allele (0, 1 or 2); PC is a vector of principal components; [24]X a vector with control variables; ε is a random error term with zero mean and constant variance and β the parameters to be estimated. They worked with a sample of 8,496 individuals and 363,776 SNPs, for which they ran this last number of regressions, which gives an idea of the volume and complexity of the statistical work involved in this type of studies. They selected the twenty SNPs that were most significant considering that, given the enormous number of regressions, false positives would appear. The results obtained were satisfactory but then, in a second stage in which they applied the way of working referred to above as the first of the methodologies in use, they tried to verify the results obtained and analyzed the relationship of those twenty SNPs with the years of education in another sample and the results were quite discouraging, since of the twenty SNPs analyzed, in only nine cases the signs of coefficient coincided. This tells us that we must be pay attention to false positives, and that all statistical studies must be very well confronted and verified.

5. Regarding education and genetics, the work of Lee et al. (2018) stands out, which was conducted by a large number of authors (77) and laboratories and consortia (3), who analyze a sample composed of 1.1 million people, and is one of the most complex studies we have had access. As education is related to health, in addition to problems studied by the social sciences in general and economics in particular, the databases created to analyze problems related to diseases have data on education, which made it possible in that case, to study a large number of observations. They applied the GWAS method and found numerous SNPs related to the years of education received and discovered that most of the genes involved are those related to communication between neurons and genome characteristics explain 11% of the variance of results in the education received.

6. In this same approach to the question, van der Loos et al. (2011) analyze whether the conditions for someone to be a successful entrepreneur have something to do with their genetic code. They consider as an entrepreneur one who is self-employed (without considering the size of his enterprise) in several different alternatives: never employed, only sometimes or always, and explain the way in which, joining several studies come to gather a sample composed of about 70,000 individuals, whose genotype is known, and 500,000 SNP can be analyzed. They use the GWAS procedure and apply the meta-analysis technique, obtaining satisfactory results but still unreliable given the small sample size.

7. More recently, Krammer and Gören (2021) analyzed the relationship between entrepreneurship, which they express with an indicator called *Total-Early Stage Entrepreneurial Activity (TEA)*, which measures entrepreneurial activity in a society and is elaborated for several countries, by quantifying the adult population that is involved in the creation of a new business or one that is in developing, which is explained by several independent variables. On the one hand, a variable showing the effect of the genetic code, which is expressed by DRD4 gene,[25] related to the production of dopamine, which has effects on personality and competitiveness, and is linked to entrepreneurial temperament.[26] On the other hand it uses other independent variables, among which are the GDP, the capital stock, the degree of openness of the economy, the urbanization rate, corruption, legal certainty, unemployment, the population age and its growth rate, immigration, some *dummy* variables to capture the geographical location and some others. They make a cross-sectional estimate using data from 97 countries. They estimate the equation by Ordinary Least Squares[27] and obtain satisfactory results because almost all the variables are significant, which indicates that the entrepreneurial spirit is linked to the genetic code (*nature*) and also to the environment (*nurture*), which indicates that both elements influence the explanation of the dependent variable, thus contributing to the clarification of the debate to which we referred above.

8. A novel way to study these issues is the one used by Molins et al. (2022), who apply an alternative way of investigating the relationship between certain genetic characteristics and risk aversion. For this, they identify 23 works through *Pubmed and Science Direct*,[28] following the method PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analysis*), a procedure used to analyze the quality of papers. They begin by pointing out that both risk and loss aversion involve behavior that might not be rational. Most of the studies they have analyzed support the hypothesis that risk aversion is associated with the pathway followed by serotonin and dopamine, although not its totality. However, they do not find agreement as to the genetic cause of loss aversion, although there are indications that the two are related[29] and claim that there is a long way to go. The progress made so far is not enough to have a final idea on this issue.

9. The relationship between genetics and income is studied by Hill et al. (2019)[30] who analyze a sample composed of 286,301 people and identify 30 *loci*[31] (of which 29 were previously unknown). They conduct a complex study, as they study genes that are present *in pleiotropy*[32] and find relationships between environment, genes, nerve transmitters, brain structure, intelligence, education and income. They start from a large number of genes, which they discard until they are left with only 24, of which 18 are linked to intelligence, which in turn they consider as one of the determining factors of income and consider that their work shows that genetics has to do with the existing inequalities in Great Britain.

10. Barth et al. (2020) analyze for a sample with data from the United States the relationship between the wealth of individuals at the time of their retirement and some of their genetic characteristics. They use the technique of "polygenic indexes" which we have described above, creating an index based on the results obtained through GWAS for a group of genetic characteristics related to the years of education received, and they obtain results that allow them to argue that there is a significant relationship between both. Their results also suggest that the observed difference between income and wealth is due to individuals with different indexes having different appreciations of risk and being more or less inclined to make investments in the stock market, to interpret the macroeconomic situation, and to evaluate prospects based on subjective probability.[33]

SELFISHNESS, COOPERATION AND ALTRUISM

Wilson (1975) opened an interesting debate about the presence of altruism in human societies, which questioned the basic assumption of social sciences and biology. How can it be compatible for human beings to pursue other goals than their own well-being? What room is left for altruism? Wilson's answer is that, for human groups to survive, there must be individuals willing to sacrifice for others so groups with more altruistic animals would have a better chance of survival.[34]

Becker (1976) argues that a society composed of individuals who maximize their individual benefits should generate a better situation for the whole because, although at first we might think that if the altruists transmit part of their income to the selfish the total income would be the same, the selfish will be encouraged to perform or omit actions that improve the situation of altruists since that will benefit them, and so the total situation of society should improve. In this way he argues that economic theory has found an explanation for the benefits of altruism.[35]

It would be a major step forward to be able to determine whether certain genes are related to cooperation and altruism, meaning reciprocal help in the first case and unilateral aid in the second. Let us look at some work on this topic.

1. Garretón and Salinas (2007) argue that there are numerous cases in biology where it has been established that behavior of this type, in which some individuals sacrifice themselves for the survival of the entire species, is linked to the existence of a certain characteristic of the genome,[36] but that there are so far no studies that allow such a statement to be made in human beings.

2. Ebstein et al. (2010) start from the premise that the size of the brain of human beings is more related to the interaction between them than to the challenges of the environment and consider that the growth of brain size is due to this situation. They analyze the genetic mechanism that makes us prone to interact with our peers, and for this purpose they study the relationship between certain genetic markers[37] and social behavior. They refer to various research related to oxytocin and arginine, which are known to be related to social behavior, including some pathological behaviors, such as autism. Likewise, it has been possible to link its presence with certain SNPs. Among them we can highlight the tendency to social behavior, aggression, popularity, altruism in decision-making, affection for children, sexual behavior, life planning, infidelity, romanticism, leadership capacity, political attitudes and appreciation for music and dance.

3. Finally, it would be convenient to continue with the line of research proposed by Dawkins (2016) in his work *The Selfish Gene*, whose first edition was published in 1976 and which he has been improving up to the present. He argues that living beings are only a support for genes to survive and endure through their transmission, which leads him to question whether humans are selfish, altruistic or both simultaneously. Genetics, with its recent advances, could identify genes that explain some of these behaviors.

EPIGENETICS

On the other hand, a discipline has recently appeared, derived from epigenetics, which is known as *behavioral epigenetics*[38] which analyzes how the environment, social context and behavior can influence the *silencing* of certain genes.[39] So far it has been used in medical matters, but it is interesting to note that studies have been carried out on alcohol and drug addiction. Their consumption produces the silencing of the genes that protect us from them, so it generates more consumption of these substances. Perhaps in the future these techniques can be used to analyze harmful economic behaviors. On the other hand, the silencing of some genes *(methylization)* has the characteristic of being transmitted by inheritance. This would give new life to the theory of evolution of Lamarck (1744-1829), who proposed it half a century before Darwin did. But while the latter believed that natural selection was the engine of evolution, the former argued that it was adaptation to the environment that caused different species to change.[40] But if the silencing of genes were inheritable, in such a way that it would influence the phenotype of the offspring, it would produce changes in the theory of evolution (there is already talk about *evolution of the theory of evolution)*, and this would have implications for the social sciences, among other things because it would give new arguments to those who, in the debate *nature vs. nurture* to which we have referred above, take sides with the second, since it would recognize the effect of the environment on our genotype. But it would also constitute a novel and relevant topic of study for economists. If poor nutrition, in addition to causing problems for those who suffer from it, generates conditions that are transmitted to the offspring, it would produce an economic cost, in addition to the human cost, that would lead us to a revision of the need to eliminate extreme poverty in our society, as well as to strengthen the fight against drugs and alcoholism.

DISCUSSION AND FINAL REFLECTIONS

Everything we have seen leads us to the conclusion that the *nature vs. nurture* debate has been overcome, since we now know that both hereditary factors and the environment contribute to define the characteristics of the phenotype in a proportion that varies from one circumstance to another, but that fortunately we can measure empirically with the methodology that is available.[41]

On the other hand, progress has been made in the analysis of the relationship between the genetic code and economic behavior, and as we have seen, genetics influences to a variable extent characteristics of the genotype such as entrepreneurship, risk aversion, addictions, and education, among others.

However, these procedures are still in their initial stages of development and deserve important questions, the most significant of which, in our opinion, is related to the choice of probability for accepting or rejecting the hypotheses considered. If we choose an extremely low value, we risk making a type I error, and otherwise falling into a type II error. This choice cannot be made on a single criterion but depends on the nature of the model we want to estimate, so it has a high degree of subjectivity, and sometimes of arbitrariness. On the positive side, these techniques are evolving rapidly and maybe we can be more confidence on their results in the near future. We do not know what will happen with this line of studies, but we believe it is worth continuing to work on it to better understand the reality in which we are immersed, and although there is a risk that it may be a dead end, we understand that it is worth navigating the path.

If we could have more conclusive results, perhaps we could design economic and social policies on clearer and more consistent bases, ¿can the knowledge we have of genetics be used to design policies, considering the different starting conditions of each human being, specifically to help those less endowed? It is worth considering the opinion of Manski (2011), who argues that with the models used before the possibility of studying genes, which primarily worked with monozygotic and dizygotic twins, they could not be used for policy design because the g value of equation [1] cannot be modified by interventions.[42] However, with the emergence of genes or combinations thereof as explanatory variables for phenotype characteristics, it is possible that the covariance between g and e is not zero, and this allows the assumption that this knowledge could be used to guide certain types of policies. Suppose a genetic characteristic predisposes to diabetes, but it is necessary for the person analyzed to be obese or have a sedentary lifestyle. This would make it possible to act on e and achieve a better situation. But as this author states, this assertion is made with great caution, because we still do not have sufficient knowledge and because we have obtained contradictory results, so we consider that it is necessary to follow a perhaps long path until genetic studies like those described can guide interventions. considering the different starting conditions of each human being, precisely to help those who are less gifted.[43]

Of course, all of this has an ethical side that must be considered. Is it legitimate for us to intrude into the deepest part of a human being, such as their genetic code, and then treat them based on what we have interpreted? But on the other hand, if we can help those who have a different conformation and we do not do so, are we not missing an opportunity for them to be in a better situation? These are questions for which we will have answers when our knowledge of these issues is much more precise, and we have advanced much further along this line of knowledge. Auffray (2004) wonders, "Doesn't exploring the genome run the risk of promoting a reducing view of humanity and justifying new forms of inequality? Will the ethical procedure allow for balanced cooperation among all social actors, respectful of human dignity, a factor of well-being and peace for future generations?" But there is something we are sure of, and that is that if we want to reduce inequality, we should first fully understand the mechanisms that produce it, and genetics can be useful in this task.

Finally, we wish to point out that economics should be more aware that human beings are fundamentally living organisms and have a closer relationship with biology, from which it has distanced itself over time. This shows that the sciences are interdependent and, while it is necessary for specialization to be increasing, a certain degree of generality is also needed, which would lead economics to adopt a different methodological perspective.

REFERENCES

Albert, B., Johnson, A., Lewis, J., Raff, M., Robert, K. y Walter, P. (2002). Molecular biology of the cell. *Annals of Botany, 91*(3), 401-405 http://doi.org/10.1093/aob/mcg023

Altshuler, D., McCarthy, M. y Rich, S. (2005). Genome-wide association studies for complex traits. *Nature Reviews Genetics, 9*, 356–369. http://doi.org/10.1038/nrg2344

Akerlof, G. y Shiller, R. (2015). *Phishing for phools: The economics of manipulation and deception*. Princeton University Press. https://doi.org/10.2307/j.ctvc777w8

Auffray, C. (2004) *El genoma humano.* Siglo XXI.

Barth, D., Papageorge, N. y Thom, K. (2020). Genetic endowments and wealth inequality. *National Bureau of Economic Research Working Paper, 24642.* https://doi.org/10.3386/w24642

Bingley, P., Cappellari, L. y Tatsiramos, K (2023). On the origins of socio-economic inequalities: Evidence from twin families. *Institute of Labor Economics. Discussion Papers, 16520.*

Beauchamps, J., Cesarini, D., Johannesson, M., van der Loos, J., Koellinger, P., Groenen, J., Fowler, J., Rosenquist, P., Thurik, R., y Christakis, N. (2011). Molecular genetics and economics. *Journal of Economic Perspectives, 25*(4), 57-82. https://doi.org/10.1257/jep.25.4.57

Becker, S. G. (1976). Altruism, egoism and genetic fitness: Economics and sociobiology. *Journal of Economic Literature (14)*, 817-826.

Behrman, J. y Taubman, P. (1976). Intergenerational transmission of income and wealth. *American Economic Review, 66*(2), 436-440.

Benjamin, D., Chabris, C., Glaeser, E., Gudnason, V., Harris, T., Laibson, D., Launer, L. y Purcell, S. (2007). Genoeconomics. En M. Weinstein, J. W. Vapuel y K. W. Wachter (Eds.), *Biosocial Surveys* (pp. 304-335). https://nap.nationalacademies.org/read/11939/chapter/20

Benjamin, D., Cesarini, D., Chabris, C., Glesser, E., Leibson, D., Guonason, V., Harris, T., Launer L., Purcell, S., Smith, A., Johanesson, M., Magnusson, P., Beauchamp, J., Christakis, M., Atwood, C., Hebert, B., Freese, J., Hauser, R., Hauser, T., Grankvist, A., Huffman, C. y Lichtenstein, P. (2012). The promises and pitfalls of genoeconomics. *Annual Economic Review,* (4), 627-662. https://doi.org/10.1146/annurev-economics-080511-110939

Carey, G. (2003). *Human genetics for the social sciences*. SAGE Publications, Inc.https://doi.org/10.4135/9781452229591

Carey, N. (2011). *The epigenetics revolution: How modern biology is rewriting our understanding of generic, disease and inheritance*. Columbia University Press.

Cesarini, D. y Visscher, P. (2017). Genetics and educational attainments. *Nature Science of Learning. 2*(4). https://doi.org/10.1038/s41539-017-0005-6

Collard, D (1978). Altruism and economics: A study in non-selfish economics. *Southern Economic Journal, 46*(1), 321-322. https://doi.org/10.2307/1057034

Dawkins, R. (2016). *The selfish gene*. Oxford University Press. (Trabajo original publicado en 1976).

Ding, W., Lehrer, S., Rosenquist., y J., McGovern J., (2006). The impact of poor health on education: New evidence using genetic markers. *National Bureau of Economic Research Working Paper, 12304*. https://doi.org/10.3386/w12304

Ebstein, R., Salomon I., Chew, S., Zhong, S. y Knafo, A. (2010). Genetics of human social behavior. *Neuron, 65*(6), 831-844. https://doi.org/10.1016/j.neuron.2010.02.020

Francis, R. (2011). *Epigenetics: How environment shapes our genes*. W.W. Norton and Company Inc.

Garreton, V. Blanco, F. Holuigue, L y Salinas, P. (2007). Salicylic acid and reactive oxygen species in the activation of stress defense genes. En S. Hayat y A. Ahmad (Eds.), *Salicylic acid: a plant hormone* (pp. 197-246). Springer. https://doi.org/10.1007/1-4020-5184-0_8

Gunes, F. (2018). *Penalized regression method for linear models in SAS/STAT*. SAS Institute Inc Working Paper.

Hardy, J. y Singleton, A. (2009). Genome wide association studies and human disease. *New England Journal of Medicine. 360*(17), 1759-1768. https://doi.org/10.1056/NEJMra0808700

Hill, W., Davies, D., Ritchie, S., Skene, N., Brois, J., Bell, S., Di Angelantonio, E., Roberts, D., Xuenyi, S., Davies, G., Liewald, D., Porteous, D., Hayward, C., Butterworth, A., Mclntosh, A., Gale, C., y Deary. I. (2019) Genome-wide analysis identifies molecular system and 149 genetic loci associated with income. *Nature communications,* (10). https://doi.org/10.1038/s41467-019-13585-5

Horvath, S. y Raj, K. (2018). DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews of Genetics(19)*, 371-384. https://doi.org/10.1038/s41576-018-0004-3

Hyytinen, A., Ilmakunas, P., Johansson, E. and Tolvanen, O. (2019). Heritability of lifetime earnings. *The Journal of Economic Inequality, (17)*, 319-335. https://doi.org/10.1007/s10888-019-09413-x

Kauffman, M. (2018). *Genetic variations, common diseases and human behavior. Economía y Biología.* Academia Nacional de Ciencias Económicas.

Keynes, J. (1921). A treatise on probability. *The Philosophical Review 31*(2). 180-186. https://doi.org/10.2307/2178916

Krammer, S. y Gören, E. (2021). Wired in? Genetic traits and entrepreneurship around the world. *Technological Forecasting and Social Change,* (168) https://doi.org/10.1016/J.TECHFORE.2021.120788

Lamarck, J. B. (2018). *Zoological philosophy*. Macmilan and Co. (Trabajo original publicado en 1809).

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Linnér, R. K., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., ... Cesarini, D. (2018). Gene discovery and polygenetic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics* (50), 1112-1121. https://www.nature.com/articles/s41588-018-0147-3

Manski, C. (2011). Genes, eyeglases and social policy. *Journal of Economic Perspectives, 25*(4), 83-94. https://doi.org/10.1257/jep.25.4.83

Mirsky, S. (2009). What´s good for the group. *Scientific American, 300* (1).

Molins, F., Fatmanur, S. y Serrano, M. (2022) The genetics of risk aversion: A systematic review. *International Journal of Environmental Research and Public Health. (19).* http://doi.org/10.3390/ijerph192114307

Navarro, A. (2018). *Sociobiology and social science. Biology and Economics*. National Academy of Economic Sciences.

Pinker, S. (2002). *The Blank Slate: The modern denial of human nature.* Penguin Press

Sacerdote, B. (2011). Nature and nurture effects on children's outcomes: What have we learned from studies of twins and adoptees? En J. Benhabib, A. Bisin, M. O. Jackson (Eds), *Handbook of social economics*. North Holland. https://doi.org/10.1016/B978-0-444-53187-2.00001-2

Van der Loos, M. J. H. M., Koellinger, P. D., Groenen, P. J. F., Rietveld, C. A., Rivadeneira, F., van Rooij, F. J. A., Uitterlinden, A. G., Hofman, A., y Thurik, A. R. (2011). Candidate gene studies and the quest for the entrepreneurial gene. *Small Business Economics* (37), 269-275. https://doi.org/10.1007/s11187-011-9339-2

Tinbergen, J. (1968). *Política económica*. Fondo de Cultura Económica. (Trabajo original publicado en 1952).

Watson, J. and Crick, F. (1953). Molecular Structure of Nucleic Acids. *Nature, (171)*. 737-738. https://www.nature.com/articles/171737a0

Wilson, E. (2000). *Sociobiology: The new synthesis*. Harvard University Press. (Trabajo original publicado en 1975). https://psycnet.apa.org/record/1975-28678-000

Notes

1. Jan Tinbergen [1952] (1968) proposes, with general acceptance, five objectives for economic policy. One of them is to reduce inequality in society. The other four are: maintaining full employment, stable prices and balanced external accounts; eliminating poverty; maintaining peace and preserving individual freedom as long as it is compatible with the other ends.

2. When the egg and sperm select half of the genome to join with their counterpart, they do so randomly, which is why this is the average value of a probability distribution.

3. See Akerlof and Shiller (2015).

4. The equation they estimate using Ordinary Least Squares (OLS), which is also applied in other studies, is as follows:

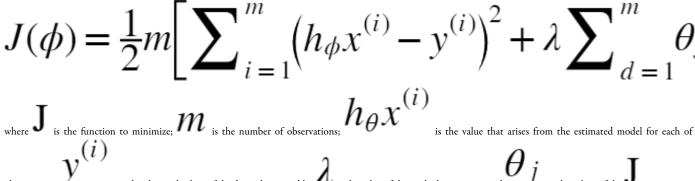$$INC_{ij} = \beta_0 + \beta_1 INC_{ji'} + \beta_2 R_j + \beta_3 (INC_{ji'}R_j) + \varepsilon_{ji} \qquad [2]$$

Where INC_{ji} represents the measurement of twin i in pair j, $INC_{ji'}$ is the measurement of the other twin from the same pair j, R stands for the coefficient of relation (1 for monozygotic and 0.5 for dizygotic twins), and ε_{ij} is an error term. This model is widely used and is known as the Additive Common Environment (ACE). In this case, some alternative formulations are made, leading to the results presented in the work

5. They find statistical evidence that monozygotic twins receive more similar treatment from their parents compared to dizygotic twins.

6. The procedure known as meta-analysis consists of creating a pool composed of papers analyzing the same question, thus increasing the sample size.

7. Kauffman (2018) states that the first indications about genetic studies can be traced back to Hippocrates and Aristotle, and that was Mendel who in 1868 discovers that genetic variations are discrete and lays the foundations of modern genetics being his work recognized recently at the beginning of the last century.

8. The readers interested in deepening on this question should consult the work of Carey (2003) if they want to see an accessible text, oriented to scholars of the social sciences, or to Albert et al. (2002) if they prefer a more in-depth work. The complete human genome can be found on the John Hopkins website University, while it is medically oriented, it gives us an idea of the operation of this type of database.

9. The number of genes that make up each of the chromosomes is variable. While chromosome 1 has 3000 genes, the 21 has only some 300. Each of the genes has a name, usually related to the protein it orders to produce, which is assigned by a committee called *Gene Nomenclature Committee* (GNC). So do each of the SNPs that have been studied.

10. SNP stands for the expression *Single Nucleotide Polymorphism*, which is a variation in a single nucleotide. Some people have an A where others have a G, which is a special type of allele, which is somewhat broader, since it refers to a different version of a gene. Sites in the genome where individuals differ in a simple base, those called SNP. There are 140 million referenced to date, a figure that is growing progressively. The 1% of our genetic code that we don't share with other individuals in our species is what makes us not be identical.

11. It can be found in the database *dbSNP*, which is freely accessible.

12   The estimation of the parameters of models with a large number of regressors that are often correlated and present overfitting which makes traditional econometric method appropriate, so some alternatives have appeared to overcome these problems. The most used is the method called LASSO (*least absolute shrinkage and selection operator*) known as *regression Penalized* which consists of introducing a constraint in the Ordinary Least Squares model, so the equation to be estimated is as follows:

$$J(\phi) = \frac{1}{2}m\left[\sum_{i=1}^{m}\left(h_\phi x^{(i)} - y^{(i)}\right)^2 + \lambda\sum_{d=1}^{m}\theta\right.$$

where J is the function to minimize; m is the number of observations; $h_\theta x^{(i)}$ is the value that arises from the estimated model for each of observations; $y^{(i)}$, are the observed values of the dependent variable; λ is the value of the applied constraint; and θ_j are the values of the J parameters to this constraint applies. In this way, biased parameters are obtained, the variance of the model is reduced, and overestimation is avoided (*overfitting*), which optimizes the between the two. An amount to be determined from the parameters that take zero value and arrive at a model that forecasts better outside the sample used for the estimation this, the sample is divided into three parts. In the first, which is usually 60% of the available data, the estimate is made using the Artificial Intelligence procedure known as *learning*; in the second, composed of 20% of the available data, the values of the estimated parameters are validated, and the remaining part is used to evaluate the mod Gunes, 2018).

13   See Altshuler et al. (2005). A description of how GWAS operates can be found at Hardy and Singleton (2009).

14   The genetic basis of about three thousand diseases, but only about a dozen of the characteristics of the phenotype that have to do with economic issues.

15   It has been used to analyze the genetic components of thousands of diseases, but its use in social sciences it much less. It is not exempt from criticism, of which the most common is related to type I errors (obtaining false positives) and type II errors (by applying a too strict probability criterion, such as the so-called *Criterion of Bonferroni*). On the other hand, it does not consider environmental factors since it analyzes the genetic causes of the characteristics of the phenotype analyzed.

16   We thank *The QIAGEN Digital Insights Team* that have provided we with access to their database.

17   See Benjamin et al. (2012).

18   That is to say that the study is conducted without any previous theoretical basis, so this procedure resembles what the econometricians call *data mining*, which consists of looking for correlations and then explaining them, like someone digging wells for oil without any plan, expecting to find it accidentally. This is what GWAS does.

19   The genes DRD2 and SLC6A3 are related to dopamine and CYP2B6 is related to another neurotransmitter called tryptophan hydroxylase.

20   The three equations are as follows:

(1)
$$A_{ijT} = \beta_0 + \beta_1 X_{iT} + \beta_2 H_{iT} + \beta_3 Q_{jT} + \varepsilon_{iT}$$

(2)
$$k_{iT} = \delta_0 + \delta_1 X_{iT} + \delta_2 H_{iT} + \delta_3 G^{k_i} + \varepsilon_{iT}$$

(3)

where A is academic achievement; X features of the individual and the family; H the state of health of that person; Q the categorization of the school; k cigarette smoking; G the genetic component; i the number of the individual making up the sample; and T represents the course attended by the adolescent considered. As the error terms were correlated with the endogenous variables and the covariance between the three error terms was assumed not zero, they estimated the model applying, in addition to Least Squares Ordinary, Minimal Squares in two and three stages. The results of the estimates can be seen in the appendix of the cited work.

21   The inclusion of variable G avoids the existence of bias due to the omission of explanatory variables in the model, while highlighting that the genetic markers chosen to have explanatory power over the variable we are trying to analyze, which is academic behavior.

22   This work, as can be seen in the references, has 22 authors who belong to different universities or research institutes located in different countries and research centers, dedicated to economic, biological or medical studies, which gives an idea that it is an interdisciplinary task, which brings together scholars located in places very distant from each other.

23   Framingham Heart Study, started in 1948 and extended over three generations.

24   In the statistical appendix of the paper we are discussing where both variables are defined precisely.

25   The DRD4 gene has four exons, of which the alleles of the third are of interest (which are segments of the ADN that contain information to produce a certain protein and make up the mRNA).

26   The way they construct the values of this variable is explained in detail in Seeing (2016).

27   He also does some tests with the Instrumental Variables method. It is necessary to point out that they do not use GWAS, because they start from a certain gene and analyze if it is statistically significant, while this method starts from the characteristics of the phenotype and looks for which genes they are related.

28   These are databases that allow you to find works in scientific journals, which are among the best known and used by scholars of all disciplines.

29   They analyze in detail diverse types of genes that are presumed to be related to risk aversion and loss, and their respective alleles. Among them are the SLC6A4, DRD4, and ANKK1 among others. The first is the gene that produces the serotonin transporter protein.

30   17 co-authors appear in this work who affirm that have collaborated equally. It is noteworthy that all are geneticists, and none is an economist.

31   Is a specific location on a chromosome where a particular gene or genetic marker is located.

32   It is the ability of some genes or genetic variants to affect the phenotype in different ways, i.e., they influence more than one characteristic observed in the phenotype.

33   See Keynes [1921] (1973).

34   See Navarro (2018).

35   See also Collard (1978).

36   There are many examples of this behavior, one of which are described in Mirsky (2009), where it is explained how if one group of bacteria sacrifice for the others, these can survive adverse conditions that would otherwise cause their death.

37   Although they also use GWAS, they refer to the results obtained by analyzing twin monkeys and dizygotic, which are consistent with those.

38   Epigenetics can be defined as the set of modifications that our genetic material undergoes that changes the way genes are activated or deactivated, but without altering themselves. "Epi" in the word Epigenetics derives from Greek "above", that is, above genetics (Carey, 2011). A good description of this discipline and its medical applications can be seen in Horvath and Raj (2018).

39   For example, see Francis (2011).

40   Lamarck [1809], states "It is not the organs, that is, the nature and shape of the parts of an animal's body, which have given rise to its habits and its faculties. It is on the contrary, their habits, their way of living and the circumstances in which the individuals from whom they come have found themselves, are those that over time have constituted the form of their body, the number and state of an organ, and the faculties, in short, that they enjoy. [...] It is known that this animal [the giraffe], the tallest of mammals, lives in the interior of Africa, where the arid and grassless region forces it to browse the trees. From this habit, sustained after a long time, it turned out that his front legs have become longer than those of the back, and that his neck has been elongated in such a way, that the animal raises its head and reaches it six meters in height without rising on the hind legs".

41   Just as economists have developed a set of statistical techniques to process economic data since the 1930s, what we know as *econometrics*, genetics also requires procedures that allow for the processing of available statistical information. Perhaps, in the near future, this may also lead to the formation of a new discipline, *genometrics*.

42   The difference in the results obtained in the studies referenced in sections 5 and 6 of the second part of this work exemplifies what we discuss regarding the use of these studies in designing public policies.
43The consequence, in principle, of the above is that the results we have today should not be used for policy design.

43   Barth et al. (2020) concur with this viewpoint, asserting, "Importantly, demonstrating a genetic basis for behavioral outcomes in no way precludes the possibility of effective public policies. A better understanding of why individuals with higher polygenic scores achieve better results may allow for a better design of policies and educational environments that help to improve outcomes. For example, it may be that children with lower polygenic scores begin to face challenges at particular ages or struggle to meet specific educational milestones. In that case, we could better target educational policies to help alleviate these roadblocks. In this manner, the future of genetic research is likely to be just as concerned with nurture as it is with nature. In short, studying how genes are connected to choices and behavior is important because it provides guidance for creating the kinds of environments where everyone, regardless of genetic endowments, has the opportunity to thrive."

ADDITIONAL INFORMATION

Clasificación JEL: A10, A11, A12, B41, C13, D31, D33, I14, I24, Z13.
En la memoria de Alfredo Martin Navarro †: Full member of the National Academy of Economic Sciences of Argentina.

Alfredo M. Navarro †
Genetics and economics: Studies and perspectives
Genética y economía: estudios y perspectivas