

Financial capital measure with item response theory: A didactic approximation

Henry Sebastián Rangel Quiñonez, Alvaro Mauricio Montenegro Díaz,
Luisa Fernanda Arenas Estevez

ABSTRACT

Objective: The objective of the article is to present in a didactic and concise way the fundamental concepts of item response theory (IRT) and its possible application in the economic sciences and show the bias problem that occurred when estimating a latent variable such as financial capital in microentrepreneurs through IRT, assuming normal distribution in an unfounded a priori way.

Research Design & Methods: We introduce a Bayesian hierarchical IRT model for graded responses where the latent traits have a skew normal distribution. Financial capital was measured by a survey applied to 384 microentrepreneurs from the metropolitan area of Bucaramanga (Colombia). The preliminary statistical analysis of data hints that the latent trait is not symmetric. Models that include a normal and a skew normal distribution were tested.

Findings: We detected that assuming the distribution of the normal trait may overestimate the calculation of financial capital in microentrepreneurs, which would cause loans to be assigned without support.

Implications & Recommendations: When applying IRT to economic matters such as in the measurement of financial capital, it is recommended to review the assumptions that this technique handles, especially the normality of the latent trait, since if assumed without verification or theoretical support can cause bias in parameters.

Contribution & Value Added: An improvement is presented to the graded response model with normal distribution in IRT for the measurement of financial capital, paying special interest in providing a pedagogical explanation for a public related to economics.

Article type: research article

Keywords: IRT in economics; Skew model; Latent trait; Graded Model; Bayesian method; Financial Capital

JEL codes: G21, C11, C52

Received: 10 January 2021

Revised: 24 June 2021

Accepted: 7 September 2021

Suggested citation:

Rangel Quiñonez, H.S., Montenegro Díaz, A.M., & Arenas Estevez, L.F. (2022). Financial capital measure with item response theory: A didactic approximation. *Entrepreneurial Business and Economics Review*, 10(1), 65-80. <https://doi.org/10.15678/EBER.2022.100105>

INTRODUCTION

The item response theory (IRT) has been widely used in education and psychology but little has spread to economics. There are two common types of latent variables in economics. First encompasses those that do not have a unit of measure such as quality of life, food security, financial capital, human capital, among others. Second – those that despite having a unit of measure, do not have sufficient information for its calculation, such as hectares of illicit crops and daily gross domestic product (GDP); both types are considered latent variables. In these cases, economics resorts to alternatives such as the measurement of proxy variables, e.g. years of education, professional experience, skills proxies of human capital, or the construction of coincident indexes. The novelty of this research is the didactic presentation of the IRT focused on the economic sciences and the proposal to adapt the classic IRT model with

normal distribution to an IRT model with asymmetric distribution, appropriate for economic variables with theoretical indications of asymmetry.

Financial capital among family entrepreneurs is an example of a variable considered conceptually asymmetric since it is intimately related to the distribution of wealth, which usually presents a positive bias, as is the case of Colombia. Financial capital comes mainly from external sources in the form of debt (Parker, 2009) and constitutes one of the variables that directly affect the performance of small businesses (Oppedal Berge, Bjorvatn, & Tungodden, 2015; Atmadja, Su, & Sharma, 2016). This is constituted by the monetary and physical assets owned by the business such as money, credits, or assets, which can be easily converted into cash or resources for the company (Mzid, Khacklouf, & Soparnot, 2019; Kaztman, 2000). This variable is constantly measured by banks, which in their financial intermediation activity capture monetary surpluses and redistribute them among the economic agents that require them – but not before ensuring that their investment is returned. They usually base their credit studies on a weighted sum of a set of questions ranging from accounting to the socioeconomic characteristics of the borrower. However, they do not have a single or comparable scale between banking entities; although many of them have similar constructs, it is common to find cases in which an applicant is subject to credit in one financial entity and not in another, possibly due to the lack of unification of the construct and the different weightings given to each measured item. These facts show the importance of allocating a single and comparable measure of financial capital for micro entrepreneurs. The success of the banking system and the guarantees for a good functioning of the circular flow of the economy depend on the correct granting of loans.

Statistical software for IRT commonly program their algorithms of the estimation of parameters from the frequentist approaches, assuming normality of the latent trait, but what would happen if the normal distribution of the latent trait is erroneously assumed, as could happen with financial capital, which is conceptually asymmetric? Incorrect adoption of trait distribution may cause biases in the estimated parameters of the items (Xu & Jia, 2011). Moreover, we may see that the hypothesis of normality is usually taken in the calculation for convenience and not for conceptual support, contrary to what is proposed in this article. Hence the relevance of developing and validating IRT models that assume the distribution of the traits to be different from normal, such as the model proposed here applied in the measurement of financial capital in a sample of microentrepreneurs in the metropolitan area of Bucaramanga (MAB). Therefore, our research presents in a didactic and concise way the fundamental concepts of IRT and its possible application in the economic sciences. Therefore, it is relevant to relax the normality assumption of latent trace; commonly accepted in psychology and education applications. Contradictory to the usual asymmetry found in economic variables.

For this reason, this article must be accessible to a public that is not necessarily expert in IRT – especially in economics – as it was designed to be applied and understood by academics in other fields. In this sense, the paper is divided into five parts, together with discussion and attachment. The first part will briefly present state of the art of measuring latent variables. The second part will succinctly compile parameter estimation strategies under the frequentist and Bayesian approaches from IRT models. The third part will show the specification of a hierarchical Bayesian model with a distribution of the skew latent trait with the explanation of data processing. The last part will offer the estimation results of the model and their comparison with the normal model so as to finish with discussion. The attachments contain an explanations for the skew normal distribution function of Azzalini (1985), a verification of the unidimensionality of the financial capital construct, and the adjustment to the empirical distribution of the scores of the financial capital construct.

LITERATURE REVIEW

The measurement of latent or unobservable variables has occupied much attention since the beginning of modern science. Within the elevation of the scientific method and the Copernican turn, mathematics has become the epistemological posture with the highest esteem to attain knowledge. Its prestige began to take root with the end of medieval philosophy and the birth of Descartes's rationalism in the middle of the seventeenth century. Subsequently, Immanuel Kant and his work *The Critique*

of Pure Reason consolidated the bases of mathematics as a method. These works, still far from mathematical formalization, are seminal for the positivism of Auguste Comte, who proposed the scientific method, and with it the entry of mathematics to spheres other than the natural sciences, such as sociology or economics, *i.e.* the social sciences

Some of the first attempts to measure social variables were found in the economic current of nineteenth-century marginalism. The mathematical models of Cournot, Bertrand, Hotelling, and the foundations of microeconomic theory, such as marginal utility, are a clear sign of interest in modelling the behaviour of economic agents in the same way that natural sciences do. Variables such as human capital, utility, job satisfaction, entrepreneurship, or empathy do not have a unit or instrument of direct measurement, such as the height or weight of an object does.

That is the reason why economists use instrumental or proxy variables. It is assumed that they are directly associated with the latent variable we intent to measure. For instance, years of education or years of experience are commonly used as proxy variables to measure the human capital variable (Schultz, 1961); networks and voluntary organizations have been used to measure the social capital construct (Paldam, 2000). The human development construct is usually measured through a set of observable variables, such as life expectancy, adult literacy rate, or GDP per capita in purchasing power parity (Alkire, 2002). Thus, with the construction of coincident indexes, proxy variables are used as indirect measures of latent variables. Some examples can be consulted in Forni, Hallin, Lippi and Reichllin (2000) and in Issler and Notini (2016). Nevertheless, some of these attempts, somewhat inipient, have not achieved the desired objectives and it is common to find diverse studies about measuring latent variables from different perspectives with divergent opinions and results.

In contrast to the limited methods used in economics, scientists in social sciences have developed statistical methods for measuring latent variables. Particularly in psychometry, the classical test theory (CTT) attributed to Spearman (1904) was the first serious approach to measure mental constructs such as intelligence and skills in mathematics. The CTT was a hegemonic paradigm of psychometry for more than 60 years.

Currently, IRT models constitute the preferred statistical tools of statisticians for measuring latent variables. They are a set of techniques that lose simplicity in calculations by making assumptions more flexible and broadening the scope of CTT. There exist two problems in CTT that do not occur in IRT. First, CTT measurements are not invariant, which means that you cannot compare the scores of different tests of an individual, as they are not measured on the same scale, and second, the absence of invariance of the test means that the difficulty and reliability of the test depends on the sample.

There exist several types of IRT models. From a very general perspective a model can be dichotomous or polytomous, depending on the scale used in a test. In a dichotomous test, the coded responses are only 1 and 0, corresponding to correct or incorrect responses. However, polytomous models are useful when there is an ordinal (or graded) order with three or more categories. Polytomous models are useful to measure latent traits from surveys designed with this objective. This is the case in this research. Furthermore, IRT models can be unidimensional, if the measurement instrument is designed to measure a unique latent trait. Alternatively, the model must be multidimensional. In this research, financial capital is considered a unidimensional latent trait, so unidimensional polytomous models may be useful in our approach. Even though the unidimensionality of the latent trait is theoretically assumed, the expert must always verify that the measurement instrument measures one construct. According to Reckase (2009), the basic assumptions of IRT models are:

1. the location of an examinee, on the scale of measurement of the latent trait, does not change during the test;
2. the characteristics of the items remain constant in any situation in which the test is presented: the items are independent form the examinees that parameters such as the difficulty of the item do not change, even under the study of different representative samples;
3. the response of a person to an item is independent of his/her response to another item, which implies that the solution to one item does not provide information to respond to other items;
4. the response of a person to an item is not related to the responses of other people to that same item;

5. the relationship between the probability of answering in a particular category of an item and the location of individuals on the trait scale can be represented with a continuous mathematical function;
6. the probability of responding in high categories of an item increases according to latent trait increases; thus, for any item, the probability of a response in high categories is an increasing function of the latent trait.

In this work, we use unidimensional polytomous IRT models. In these models, the items have more than one category ordered or not, as the Likert or nominal scales. Some of them are the graded model, the partial credit model, and the nominal model. Such models provide more trait information than dichotomous ones. Furthermore, tests built with polytomous items are less costly and time-consuming to apply, which can have a positive effect on examinees (Ostini & Nering, 2006, pp. 7-8). This section will focus on ordinal-type response models, as they are the most relevant to the type of data treated here.

As mentioned above, the applications of the IRT models are few in economics. However, some research has shown their applicability in measuring human development (Morán & Álvarez, 2001), inequality, poverty, and welfare (Deutsch, J., Silber, J., Xu, Y., & Wan, G. 2020), multidimensional poverty (Fusco & Dickes, 2008), financial risk (Caviezel, V., Bertoli Basrsott, L., & Lozza, S. O., 2011), organizational innovation (Wongtada, & Rice, 2008), and managerial ability (Schellhorn, C., 2013).

Classical approach parameters estimation

The first methods for estimating parameters in IRT assume a known ability or latent trait. This process is called quantal response bioassay (QRB; Lawley, 1943; Finney, 1944). The fundamental idea is to assume the existence of groups of examinees with known latent skill scores located across the entire scale of the trait and allowing correct and incorrect response proportions to be formed within each group. Another approach similar to QRB is to assume that the parameters of the items are known and the latent traits unknown. This estimation process is frequently used in adaptive computerized tests as some language proficiency tests, which have calibrated items parameters and focus on trait calculation (Baker & Kim, 2004; Reckase, 2009). Commonly, estimates of item parameters and predictions of latent traits are required at the same time. For this reason, the joint maximum likelihood estimation method was introduced. The technique consists of estimating at the same time the parameters of items and predicting the latent traits of individuals. In 1968, Birnbaum proposed the basic concepts of this paradigm (Baker & Kim, 2004, p. 18). The model has two fundamental assumptions: 1) the response of the items is independent of each other, and 2) the response vectors of individuals are independent from each other. Baker and Kim (2004, p. 108) identify some problems in the application of this paradigm: the estimation of the parameters of the items can be complex in some IRT models because likelihood equations may not converge if initial values are set too far from optimal values. Moreover, others verified that the estimates of this method are not consistent with the increase in sample size (Baker & Kim, 2004, p. 157; Neyman & Scott, 1948).

To overcome these issues, Bock and Lieberman (1970) proposed the algorithm called item parameter estimation via marginal maximum likelihood, which provides consistent estimators when faced with an increase the sample size. However, the method only worked with very small sets of items. Afterwards, Bock and Aitkin (1981) implemented the Expectation-Maximization algorithm (EM) as reformulation to Bock and Lieberman (1970) making it more efficient computationally. In addition, the EM algorithm could be applied to real test sizes. The latter method assumes that the distribution of the latent trait is a standard Gaussian distribution, which requires a technique of numerical methods to calculate integrals, such as the Hermite-Gauss quadrature, with the limitation that it only works when the distribution is normal. For this reason, Bayesian hierarchical methods may be a good alternative for estimating parameters when the latent trait cannot be assumed to be normal.

The Bayesian approach relies on the Bayes theorem, which combines the probability obtained from the likelihood function based on observed data, with prior information on the distribution of parameters (Baker & Kim, 2004, p. 179). This approach considers all parameters to be random variables, contrary to the classical approach which assumes that parameters are fixed but unknown. From this perspective, there are two types of parameters: interest and nuisance. The former are the ones with which

the researcher wishes to make an inference, while the latter are – as their name indicates – noise or nuisance parameters that will not be subject to analysis.

- H:** The latent variable “financial capital” has a skew normal distribution that can cause its estimation to be biased if it is done by a traditional method that a priori assumes it is normality.

RESEARCH METHODOLOGY

The hierarchical Bayesian gradual response model with asymmetric normal trait distribution (HBGRM ASN) has a categorical response function (CRF) defined as:

$$P[u_{ij} = g | \theta_j, \xi] = \text{logit}^{-1}(\eta_{g-1}) - \text{logit}^{-1}(\eta_g) \quad (1)$$

in which $P[u_{ij} = g | \theta_j, \xi]$ is the probability that respondent j answers category g at the item i , given the ability θ_j and the parameter vector of item ξ . $\text{logit}^{-1}(\eta_{g-1}) = (1 + e^{-\eta_{g-1}})^{-1}$, $\eta_{g-1} = (\alpha_i * \theta_j - \alpha_i * \beta_{i,g-1})$ and $\text{logit}^{-1}(\eta_g) = (1 + e^{-\eta_g})^{-1}$, $\eta_g = (\alpha_i * \theta_j - \alpha_i * \beta_{i,g})$. These expressions form cumulative category response functions (CCRF) of logistic distribution. Figure 1 shows the change of the CCRF for an item with five categories in which $\alpha = 1$ and values of $\beta = (-2, -1, 0, 7, 2)$.

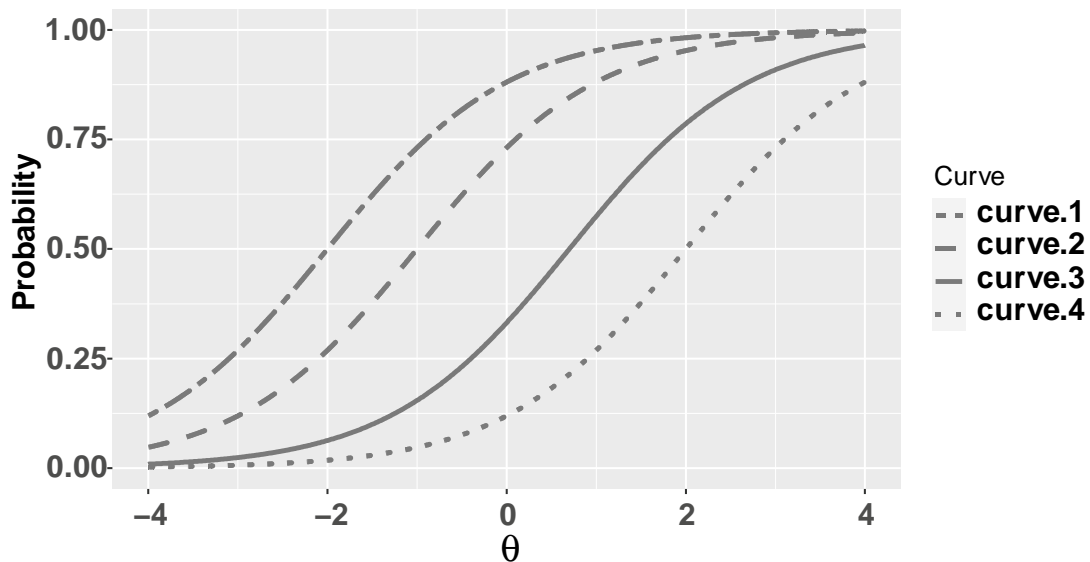


Figure 1. Cumulated categorical response function

Source: own elaboration of Reckase (2009).

We can interpret α_i as the discrimination parameter of item i and $\beta_{i,g}$ which represents the upper cut point for the category g of item i , also known as difficulty parameters; $\beta_{i,g} = \beta_{i,1}, \dots, \beta_{i,k-1} \in \mathbb{R}^{k-1}$, ordered $\beta_{i,g} < \beta_{i,g+1}$. One must assume the extreme cases of $\beta_{i,0} = -\infty$ and $\beta_{i,k+1} = \infty$, which implies that $\text{logit}^{-1}(\eta_0) = 1$ and $\text{logit}^{-1}(\eta_{k+1}) = 0$. $k \in \mathbb{N}$, with $k > 2$ representing the number of categories of the item i , in the following way $g \in \{1, \dots, k\}$. So that $P[u_{ij} = g | \theta_j, \xi]$ can assume different expressions, such as

$$[u_{ij} = g | \theta_j, \xi] \begin{cases} 1 - (1 + e^{-\eta_1})^{-1} & si \quad g = 1, \\ (1 + e^{-\eta_{g-1}})^{-1} - (1 + e^{-\eta_g})^{-1} & si \quad 1 < g < k, \\ (1 + e^{-\eta_{g-1}})^{-1} & si \quad g = k, \end{cases} \quad (2)$$

The above equation represents the CRF of the logistic model graded for the homogeneous case. The CRFs corresponding to the CCRFs in Figure 1 are shown in Figure 2.

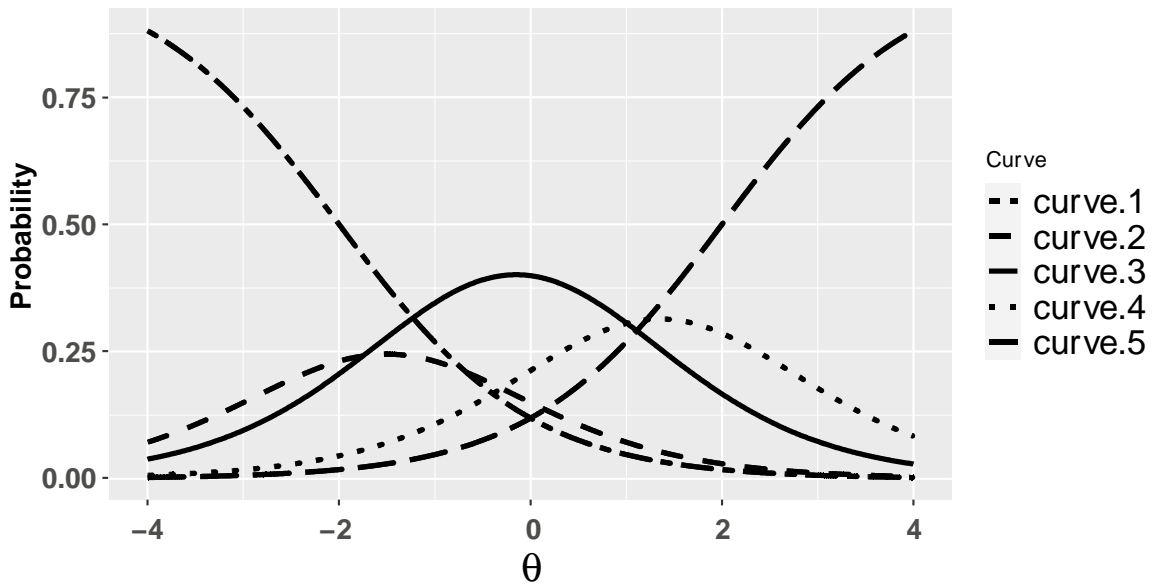


Figure 2. Categorical response function for items with five options

Source: own elaboration of Reckase (2009).

Figure 2 shows how the probability of responding to option 1 (represented by curve 1) is inversely proportional to the trait, opposite to the probability of responding to category 5 (represented by curve 5), which is directly related to the value of the trait. Note that the highest point of the probability of response for each item occurs at different points on the trait scale. For example, category 3 has a higher probability of response for traits close to zero, different from category 2 which has the highest probability for traits close to -2, this characteristic is linked to the information provided by each category. The Bayesian hierarchical model assumes prior distributions for parameters and hyperparameters. Consequently, the prior densities assigned here are:

$$\begin{aligned}
 \theta_j &\sim SN(0,1,\rho), \\
 \rho &\sim N(0,1), \\
 \alpha_i &\sim N(0,5), \\
 \beta_{ig} &\sim N(\mu_{beta_{ig}}, \sigma_{beta_{ig}}), \\
 \mu_{beta} &\sim N(0,1), \\
 \sigma_{beta} &\sim Cauchy(0,2).
 \end{aligned} \tag{3}$$

For the prior distribution of the latent trait θ_j a skew normal distribution of Azzalini (1985) is assumed, the scale and location parameters are fixed in order to guarantee the model's identifiability. The parameter ρ is the hyperparameter of asymmetry normal distribution and standard normal distribution is assumed. The skew normal distribution is assumed since it is the one that best fits in a first inspection of the empirical latent trait (see Attachment A). This parameter is the centre of attention in this model because its significance depends on the success of this proposal on the classical model with normal distribution. Restrictions are imposed by truncation of the discrimination parameter α_i at 0 for the monotonicity of the CRF and the hyperparameter σ_{beta} because it is a deviation.

Data

The data in this article were taken from the joint project of the Industrial University of Santander and a Santander financial institution, whose main interest in that research was to measure the levels of psychological, financial, physical, social, and human capital among microentrepreneurs of the MAB. In that research, the calculation of the level of capital was based on the principal component analysis (PCA) technique, which has advantages due to broad existing documentation and the ease of its computational application. However, PCA is based on a decomposition of the covariance matrix of the items assuming that these are of a continuous nature, characteristics that are not met by most surveys measured using Likert-type categorical scales. These limitations were overcome with the IRT models, such as the one proposed here.

The analysed microenterprises have little productive capacity (neighbourhood stores, small clothing, and footwear factories, etc.), they lack organizational structure, accounting books, or bank credit history. Most workers belong to the same family, which is why they are known in Colombia as *famiempresas* (family businesses). This level of informality makes it difficult to access bank loans, forcing them to take out informal loans with individuals. Therefore, formal credit agencies in Colombia have chosen to evaluate the risk of non-payment together with the owner's family, since it is assumed that their strengths and weaknesses will be transmitted to their companies. Therefore, several survey type instruments have been developed to measure the financial capital that commonly – in larger companies – would be measured using the financial statements. The application of the proposed model will be done with data from the financial capital module of 384 microentrepreneurs. The underlying latent trait was assumed to be asymmetric, due to its strong relationship with the income distribution, which in developing countries usually presents large asymmetries.

The construct of financial capital was made up of eight polytomous ordinal items that measured the financial response capacity of the family to obtain liquidity and face business and family activities. The original survey consisted of 77 dichotomous questions recoded into polytomous items. To achieve this objective, different methods were applied, including principal component analysis (PCA), cluster analysis, and expert judgment. Details of the process are available in the article by Rangel Quiñonez and Yáñez Canal (2018). The final instrument version can be consulted at the end of this document in attachment B.

Parameters estimation of the proposed model was implemented in the STAN language version 2.2.0, together with the Rstan package of the programming language R (coding details available in Rangel, 2019). Stan is a new Bayesian statistics programming language written in C++, which provides statistical inference of parameters through Markov Monte Carlo strings such as the *No-U-Turn samplers*. The Stan language can be used in other languages such as Python, R, Matlab, Mathematics, Julia, and Stata, and it runs under Windows, Mac OS, and Linux operating systems by using interfaces built into each of them (Luo & Jiao, 2018).

RESULTS AND DISCUSSION

The results obtained from the implementation of the model show that the highest discrimination parameters α_i were those corresponding to items 1, 5 and 7, as opposed to the lowest alphas of items 3, 4 and 8. The $Rhat$ in all cases was 1, which evidences good convergence of the parameters, and similarly the ratio $\frac{n_{eff}}{N} > 0.001$ – in which N is the number of interactions – for this case equal to 4000, it evidences a good Markov chain. None of the 95% credibility intervals contains 0. This is an indicator of good model specification.

The calculation of the difficulty parameter requires the division of the $\beta_{i,g}$ calculation by Stan by the discrimination parameter α_i . Table 2 presents the true values of the $\beta_{i,g}$. The lowest average of the estimated parameters for betas was $\beta_{3,1} = -6.49$ and the highest was $\beta_{3,5} = 12.89$. The closeness of the cut points parameters in a single item, suggests that category unification is possible. These parameters had good $Rhat \approx 1$ values and $\frac{n_{eff}}{N} > 0.001$, also no credibility intervals contain 0.

Table 1. Descriptive statistics of discrimination parameters

α_i	μ	5%	25%	50%	75%	95%
α_1	4.01	2.46	3.17	3.81	4.66	6.17
α_2	0.71	0.40	0.58	0.70	0.83	1.04
α_3	0.45	0.17	0.33	0.45	0.56	0.74
α_4	0.70	0.38	0.57	0.69	0.82	1.03
α_5	1.91	1.34	1.65	1.88	2.14	2.59
α_6	0.72	0.39	0.59	0.72	0.86	1.08
α_7	3.43	1.20	2.29	3.23	4.35	6.33
α_8	0.52	0.15	0.35	0.50	0.66	0.93

Source: own study.

Table 2. Difficulty parameters $\beta_{i,g}$

ITEM	$\beta_{i,1}$	$\beta_{i,2}$	$\beta_{i,3}$	$\beta_{i,4}$	$\beta_{i,5}$
1	2.29	2.95	X	X	x
2	1.39	5.36	x	X	X
3	-3.94	0.47	0.51	X	X
4	-1.07	3.46	4.14	X	X
5	-0.27	-0.18	0.34	X	X
6	-1.71	0.83	2.56	x	x
7	-1.54	1.34	4.03	6.96	x
8	-6.49	-0.09	5.38	8.69	12.89

Source: own study.

Figure 3 presents a scatter plot between the classical test scores (sum scores of financial capital) and the averages of the individual parameters, showing the existing linear relationship, corroborated by a Pearson’s correlation coefficient equal to 0.85. Estimated density and classical score functions are also shown, both cases show positive asymmetry.

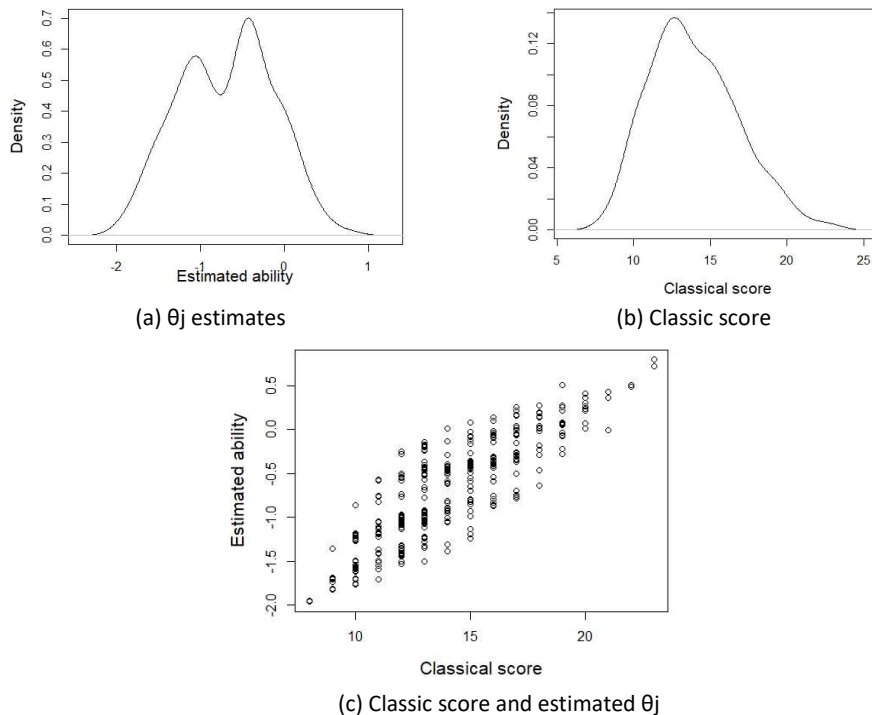


Figure 3. Comparison between the latent trait distribution of the IRT model and the distribution of the classical score

Source: own elaboration.

Descriptive statistics of the latent trait shows that the average of θ_j for financial capital assume the skew normal trait is 0.599, the minimum is -1.961, and the maximum is 0.796. On the other hand, the minimum 0.999 and the maximum 1.009 in the Rhat for the θ_j parameters reaffirms the good convergence since the average and the median are close to 1. The estimated values of the hyperparameters are shown below.

Table 3. Descriptive statistics of the estimation of the hyperparameters

Hyperparameter	μ	5%	25%	50%	75%	95%	n_{eff}	Rhat
ρ	-1.90	-2.88	-2.27	-1.87	-1.50	-1.04	338	1.01
μ_{β}	1.20	0.38	0.87	1.20	1.52	1.99	3073	1.00
σ_{β}	2.55	1.94	2.25	2.51	2.80	3.32	4130	1.00

Source: own study.

Special attention deserves parameter $\rho = -1.9$ since it is the estimated asymmetry of the density function of the latent trait. This indicates that the asymmetry parameter is significant as the credibility interval does not contain zero. To corroborate this result, we require a Bayesian hypothesis test of the simulations of hyperparameter ρ .

Parameters comparison estimated by the asymmetric model with the normal model

With the idea of verifying changes in the magnitudes of the parameters estimated under the assumptions of asymmetry and normality, we prepared point plots to verify overestimates or underestimates of these parameters. Figure 4 shows the difference between the parameters estimated with skew normal distribution and the parameters estimated with standard normal distribution. The horizontal line at zero of the ordered axes is taken as a reference to observe the changes in the estimate. The discrimination parameters (α_i) do not present large changes in six items, but items 3 and 5 are underestimated when estimated with the assumption of normality. The difficulty parameters (β_i) present a moderate overestimation at the cut points of the first six items and a relevant overestimation at the cut-off points of items 7 and 8. Moreover, the individuals' parameters show a marked overestimation at the moment of estimation by means of the normal models instead of the asymmetric model.

Now, to contrast the suitability of the model, the information criterion *efficient approximate leave-one out cross validation-loo* of the software package loo R was used (Vehtari, Gelman, Gabry, & Yao, 2019).

Table 4. Information criterion Loo

	Skew Model	Normal Model
Loo	5408.400	5420.400

Source: own study.

Table 4 shows that the best model is skew normal because it has less information criteria. To support the above result the difference of the likelihood and the standard error was calculated with the *loo_compare* function of the R-stan package. The comparison results are $elpd_{diff} = -6.0$ with $SE = 1.7$, which indicate that the first model fits better the data. This case corresponds to the skew model (Vehtari *et al.*, 2019, p. 6).

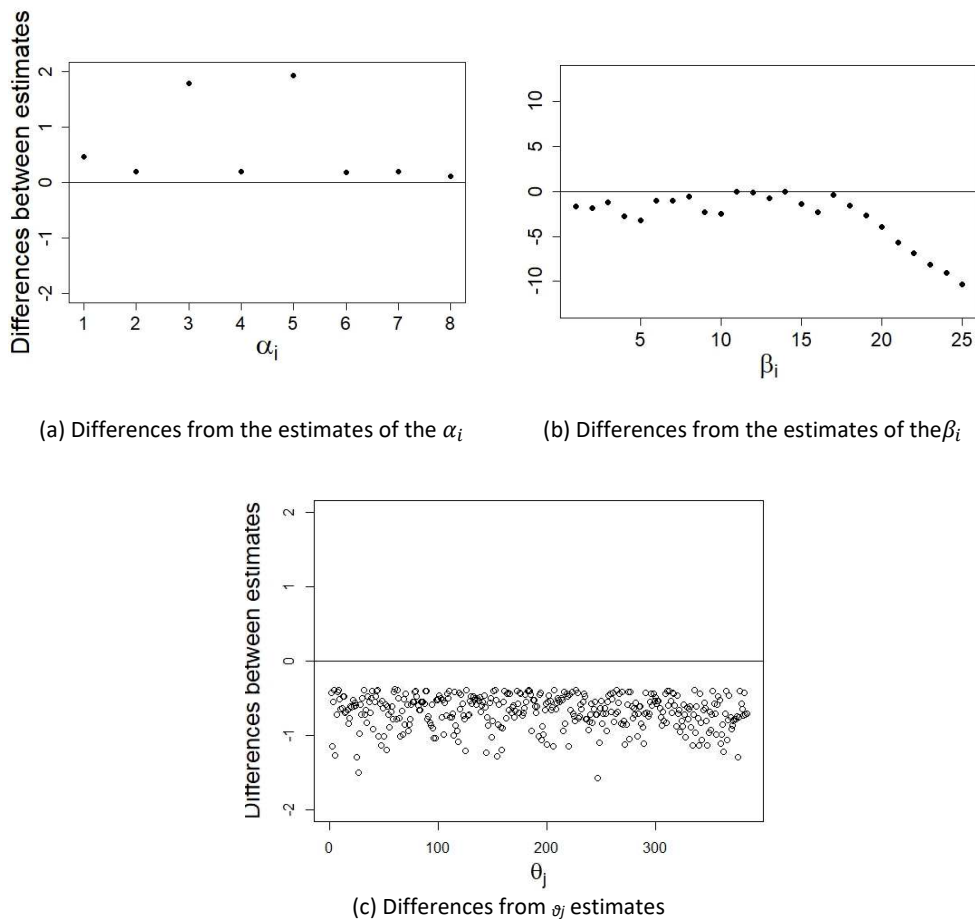


Figure 4. Difference between parameters α_i , β_i and θ_j estimated under skew normal and normal distribution
Source: own elaboration.

CONCLUSIONS

This research emerged from the frequent need of economists to measure latent or unobservable variables, so it is introduced to the item response theory as an alternative to the limited tools usually used in economics. In this article, we outline the main models of IRT by highlighting their different uses and theoretical assumptions. Perhaps the most restrictive assumption is that of the normality of the latent trait since a large part of economics variables are asymmetric. For this reason, we proposed an alternative for estimating a graded IRT model that relaxes this assumption. The application here exposed is made to obtain formal measures of Financial Capital. In common practice, the results of surveys to measure Financial Capital are treated trivially and without any mathematical rigor by financial institutions when they conduct credit studies of their clients. For example, ratings are taken as real numbers – which in fact is a categorical measure – without a comparable metric since the score depends on the construct between entities. The approximation used in Attachment B suggested that the measured trait was not symmetric, so the hypothesis of normality did not seem adequate and should therefore be treated with an asymmetric distribution by selecting the skew normal distribution (Azzalini, 1985) as suitable for this problem. The technical part reported that at the time, there was no software available that implemented – from the classical approach – the polytomous TRI models with asymmetric normal distribution as an alternative to the problem, which is the solution the Bayesian approach. With the help of STAN, a probabilistic programming language, considered the gold standard in the industry today, a hierarchical Bayesian model of gradual response with distribution of the skew normal trait

was built. We concluded that the skew model presents an improvement in the Bayesian Loo information criterion compared to the normal model. This reaffirms what Xu and Jia (2011, p. 15) found when assuring that the asymmetric normal distribution presents advantages with respect to the normal model of a 2pl model, when the distribution is extremely asymmetric or presents multimodality, which is something that cannot be rejected by the data in this article. Thus, the statement of Xu and Jia (2011) can be extended to the polytomous case with the findings of this paper: the mean value of the skew parameter ρ estimated by MRJBRG ASN is -1.9, which is significant since its credibility interval does not contain zero, corroborated by the Bayesian hypothesis test.

Let us note that if zero were contained, the asymmetric normal distribution would be equal to the normal distribution, leaving the preference of the proposed model over the normal model unsupported. All the parameters of discrimination are significant, the largest average value is α_3 corresponding to question 1 (financial savings or investment service or product), and the one with the least discrimination is α_8 (ability to pay through the financial sector). This indicates that small changes in the latent trait produce large changes in the probability of answering a category in question 3 – as opposed to question 8 – while items with α_i close to 1 are preferable and indicate a smooth discrimination not as abrupt as items 3 and 4. In the same manner, all the difficulty parameters are significant as it can be verified that items 1, 3, 4, and 8 have some β_{ig} close together: item 1 could group categories b and c; items 3 and 4 could group categories c and d; and for item 8 it is suggested to join categories d and e. The estimation of these parameters can be used for the validation of the construction, thus replacing the more used confirmatory and exploratory factorial analysis. When comparing the Loo statistic of the asymmetric and normal model, it is concluded that the asymmetric model is better than the normal model. Which begs the question: What is the change of the parameters under normal estimation? We observed that in the α_i , there appeared an underestimation by the normal model, especially in items 3 and 5, opposed to the overestimation of the β_{ig} , particularly in the items with more categories (7 and 8), finally in the person parameters θ_j that were overestimated in the normal model.

This leads us to the conclusion that the normal model – when applied to the data of financial capital – overestimates this variable that is linked to the payment capacity of family businesses, causing banks to take an incorrect signal from the entrepreneurs. Thus, the normal model would classify entrepreneurs with high risk of non-payment as companies subject to credit. This can be seen in the median of the latent trait in the asymmetric model (-0.59), which is negative and lower compared to the median of the trait in the standard model (0.02), indicating that the standard model gives half of the family businesses a positive trait compared to the asymmetric standard model, which gives half of the family businesses a score greater than or equal to -0.59. Let us note that we observed the sensitivity of the significance of the asymmetry parameter ρ to changes in the priori distribution of the parameter α_i within the tests conducted for the construction of this model, since when an informative priori such as $\alpha_i \sim N(0, 1)$ was included, it meant that ρ was not significant. Contrary to what happened with the non-informative priori of $\alpha_i \sim N(0, 5)$. Thus, future research should focus on: 1) verifying the impact of the a priori distributions of the parameters, especially the discrimination parameter in the Bayesian approach and 2) performing simulation to verify consistency in the results of the Bayesian model.

REFERENCES

- Alkire, S. (2002). Dimensions of human development. *World Development*, 30(2), 181-205. [https://doi.org/10.1016/S0305-750X\(01\)00109-7](https://doi.org/10.1016/S0305-750X(01)00109-7)
- Atmadja, A. S., Su, J.-J., & Sharma, P. (2016). Examining the impact of microfinance on microenterprises performance (implications for women-owned micro enterprises in Indonesia). *International Journal of Social Economics*, 43(10), 962-981. <https://doi.org/10.1108/IJSE-08-2014-0158>
- Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 171-178. Retrieved from <https://www.jstor.org/stable/4615982?seq=1> on July 21, 2021.
- Baker, F.B., & Kim, S.H. (2004). *Item Response Theory: Parameter Estimation Techniques*. Second Edition (2nd ed). CRC Press. <https://doi.org/10.1201/9781482276725>

- Bock, R.D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, *46*(4), 443-459. <https://doi.org/10.1007/BF02293801>
- Bock, R.D., & Lieberman, M. (1970). Fitting a response model for n dichotomously scored items. *Psychometrika*, *35*(2), 179-197. <https://doi.org/10.1007/BF02291262>
- Caviezel, V., Bertoli Basrsott, L., & Lozza, S.O. (2011). Measuring risk profile with a multidimensional rasch analysis. *Journal of Applied Quantitative Methods*, *6*(4), 14-29 Retrieved from <https://www.semanticscholar.org/paper/Measuring-risk-profile-with-a-multidimensional-Caviezel-Bar-sotti/60bfa7d043e478a3f8c2d02d14fb6b93425b56e3> on July 21, 2021.
- Deutsch, J., Silber, J., Xu, Y., & Wan, G. (2020). Measuring inequality, poverty, growth and welfare via the use of asset indexes: The case of Armenia, Azerbaijan and Georgia. *The Singapore Economic Review*, *65*(supp01), 7-33. <https://doi.org/10.1142/S0217590819440016>
- Finney, D. (1944). The application of probit analysis to the results of mental tests. *Psychometrika*, *9*(1), 31-39. <https://doi.org/10.1007/BF02288711>
- Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2000). The Generalized Dynamic-Factor Model: Identification and Estimation. *The Review of Economics and Statistics*, *84*(4), 540-554. Retrieved from <https://www.jstor.org/stable/2646650> on July 21, 2021.
- Fusco, A., & Dickes, P. (2008). The Rasch model and multidimensional poverty measurement. In N. Kakwani & J. Silber (Eds.), *Quantitative Approaches to Multidimensional Poverty Measurement* (1st ed., pp. 49-62). United Nations Development Programme. https://doi.org/10.1057/9780230582354_3
- Issler, J.V., & Notini, H.H. (2016). Estimating Brazilian monthly GDP: A state-space approach. *Revista Brasileira de Economia*, *70*(1), 41-59. <https://doi.org/10.5935/0034-7140.20160003>
- Kaztman, R. (2000). Notas sobre la medición de la vulnerabilidad social. In: CEPAL, BID and BIRF, ed., *Taller Regional sobre la Medición de la Pobreza: Métodos y Aplicaciones*, (5th ed. [online] pp.275-301) Retrieved from: <<https://repositorio.cepal.org/handle/11362/31545>> on 20 July 2021.
- Lawley, D. N. (1943). On problems connected with item selection and test construction. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, *61*(3), 273-287. <https://doi.org/10.1017/S0080454100006282>
- Luo, Y., & Jiao, H. (2018). Using the stan program for bayesian item response theory. *Educational and Psychological Measurement*, *78*(3), 384-408. <https://doi.org/10.1177/0013164417693666>
- Neyman, J. & Scott, E. (1948) Consistent Estimates Based on Partially Consistent Observations. *Econometrica*, *16*(1), 1-32. <https://doi.org/10.2307/1914288>
- Morán Álvarez, J.C., & Álvarez Martínez, P. (2001). Medida del desarrollo humano para los países de la América Latina. Trimestre Económico, *68*(2), 195-208. Retrieved from <https://www.jstor.org/stable/20857057> on July 21, 2021
- Mzid, I., Khacklouf, N., & Soparnot, R. (2019). How does family capital influence the resilience of family firms?. *Journal of International Entrepreneurship*, *17*(2), 249-277. <https://doi.org/10.1007/s10843-018-0226-7>
- Oppedal Berge, L.I., Bjorvatn, K., & Tungodden, B. (2015). Human and Financial Capital for Microenterprise Development: Evidence from a Field and Lab Experiment. *Management Science*, *61*(4), 707-722. <https://doi.org/10.1287/mnsc.2014.1933>
- Ostini, R., & Nering, M.L. (2006). *Polytomous item response theory models*. Quantitative Application in the Social Sciences Series. SAGE Publications, Inc. <https://www.doi.org/10.4135/9781412985413>
- Paldam, M. (2000). Social capital: One or many? Definition and measurement. *Journal of Economic Surveys*, *14*(5), 629-653. <https://doi.org/10.1111/1467-6419.00127>
- Parker, S. C. (2009). *The Economics of Entrepreneurship*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511817441>
- Rangel Quiñonez, H. S., & Yáñez Canal, G. (2018). Clasificación de una muestra de microempresarios del área metropolitana de Bucaramanga, a partir de los capitales de salud, educativo, social, físico y financiero, utilizando el análisis de correspondencia múltiple. *Ensayos de economía*, *28*(53), 165. <http://dx.doi.org/10.15446/ede.v28n53.75018>
- Rangel Quiñonez, H. (2019). Modelo Político Unidimensional de Teoría de Respuesta al ítem con distribución asimétrica del trazo latente. Master's thesis in statistical sciences, National University of Colombia. Bogotá, D.C
- Reckase, M. (2009). *Multidimensional item response theory*. Berlin: Springer.
- Schellhorn, C., & Sharma, R. (2013). Using the Rasch model to rank firms by managerial ability. *Managerial Finance*, *39*(3), 306-319. <https://doi.org/10.1108/03074351311302818>

Schultz, T. (1961). Investment in Human Capital. *American Economic Review*, 51(1), 1-17. Retrieved from <http://www.jstor.org/stable/1818907> on July 21, 2021.

Spearman, C. (1904). General intelligence objectively determined and measured. *The American Journal of Psychology*, 15(2), 201-292. <https://doi.org/10.2307/1412107>

Vehtari, A., Gelman, A., Gabry, A., & Yao, Y. (2019). Package ‘loo’. Retrieved from <https://cran.r-project.org/web/packages/loo/loo.pdf> on July 21, 2021

Wongtada, N., & Rice, G. (2008). Multidimensional latent traits of perceived organizational innovation: Differences between Thai and Egyptian employees. *Asia Pacific Journal of Management*, 25(3), 537-562. <https://doi.org/10.1007/s10490-008-9085-4>

Xu, X., & Jia, Y. (2011). The sensitivity of parameter estimates to the latent ability distribution. *ETS Research Report Series*, 2, i-17. <https://doi.org/10.1002/j.2333-8504.2011.tb02276.x>

Appendix A:

The skew normal distribution function Azzalini (1985) can be described as proportional to the product of the density function times the cumulative function of the normal distribution:

$$SN(x) = 2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)^2} \frac{1}{2} Erfc\left(-\rho \frac{x-\mu}{\sigma\sqrt{2}}\right),$$

then,

$$SN(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)^2} Erfc\left(-\rho \frac{x-\mu}{\sigma\sqrt{2}}\right),$$

in which ρ is the asymmetry parameter, μ is the location parameter, σ the scale parameter and x is the value of the random variable. It is understandable that if $\rho = 0$ then $Erfc = 1$ colluding to normal distribution.

Adjusting the distribution to the latent trait

To verify the unidimensional nature of the construction is applied a Principal Component Analysis (PCA) on the eight items.

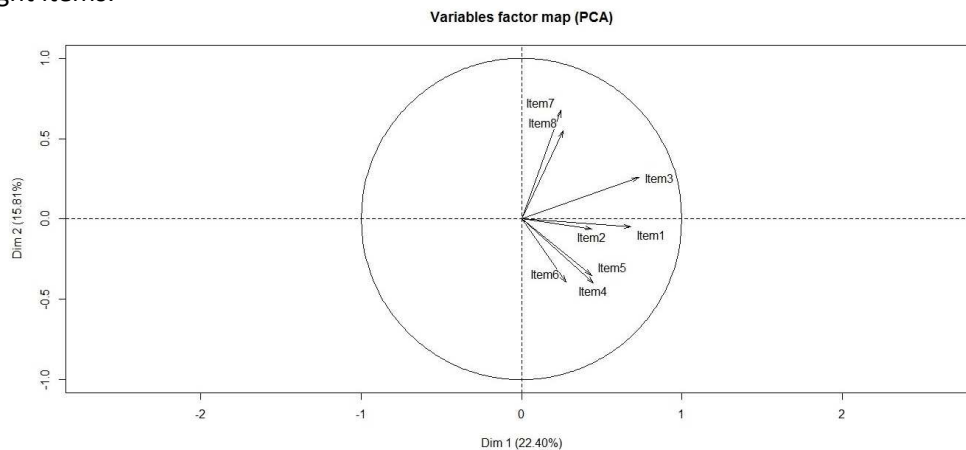


Figure A-1. PCA for Financial Capital

Source: own elaboration.

Table A-1. Own values of the 8 items of the construct

Variable	Eigenvalues	Percentage of variance	Percentage of accumulated variance
comp 1	1.75	21.91	21.91
comp 2	1.18	14.78	36.70
comp 3	1.09	13.59	50.29

Source: own study.

In the previous Figure A-1 it is evident that the eight items have only one direction that would support the unidimensionality of the construct, Table A-1 debates this conclusion by presenting the values of the eight components, which affirms that at least three dimensions exist. This is not surprising since the original questionnaire had 77 questions, so one of the previous works to this investigation was the selection and conformation of the financial capital construct. However, these data serve as an example of the use of the algorithm proposed here for economic issues. With the above exceptions, the financial capital score is calculated by means of the sum of the items' responses and, on the other hand, by means of the respondents' projections in the first component of the PCA.

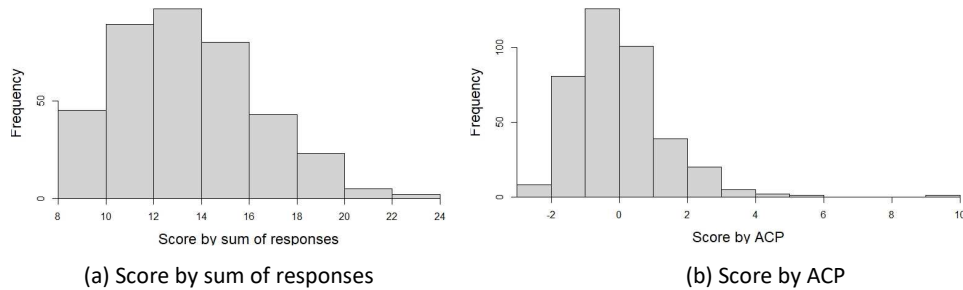


Figure A-2. Histograms of construct scores

Source: own elaboration.

As shown in Figure A-2, the scores by the two methods show positive asymmetry. Therefore, we proceed to look for the best distribution that fits them.

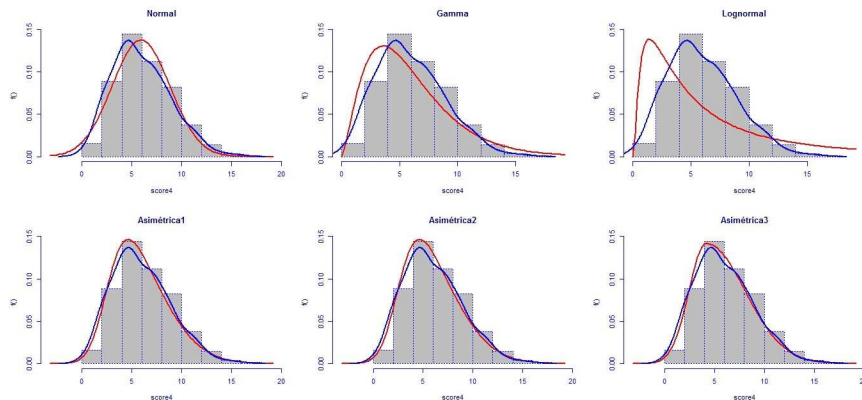


Figure A-3. Adjusting various density functions to the histogram of construct scores

Source: own elaboration.

Figure A-3 shows the score by the classic method and different settings by means of the Gamlss package of the R software. The red curve on the histogram represents the empirical distribution and the blue curve the distribution adjusted according to the distributions: normal, gamma, lognormal, and asymmetric 1, 2, and 3. As you can see the asymmetric distribution 1 fits very well to the classical score.

Appendix B:

To facilitate the programming, the questions of the construct have been organized in an ascending way.

- 1) Capability of earning income from the sale of company assets:
 - a) had not obtained income in the last month from sales of machinery or equipment;
 - b) had obtained income one time;
 - c) had obtained income two times.
- 2) Family's ability to receive income from real estate as urban, agricultural, or recreational property:
 - a) had not obtained income in the last month from real estate;
 - b) had obtained income two times;
 - c) had obtained income three times.
- 3) The degree of financial savings or investment service or product:
 - a) lack of financial savings;
 - b) having savings accounts, current accounts, and programmed savings;
 - c) having current savings, programmed savings and communal savings accounts;
 - d) having shares, term deposit certificates, and savings accounts.
- 4) The degree of insurance of the family and the business:
 - a) do not have insurance;
 - b) basic insurance such as funeral, life, or debtor's insurance in the case of having a credit;
 - c) in addition to having basic insurance have some of the intermediate insurance such as insurance against theft, earthquake insurance, home insurance, car insurance and unemployment;
 - d) have basic life and death insurance for the debtor along with housing, fire and earthquake insurance, insuring the most asset of family members, the home.
- 5) The kind of indebtedness of the family:
 - a) do not have loans;
 - b) informal loans such as "drop-by-drop" loans, loans with pawnbrokers, and trust companies;
 - c) having loans with friends, suppliers, educational loans, car loans, and loans with NGOs or foundations;
 - d) loans with banks.
- 6) Measures the level of family benefits:
 - a) less than or equal to 600 000 COP;
 - b) more than 600 000 COP but less than or equal to 1 200 000 COP;
 - c) more than 1 200 000 COP but less than or equal to 2 400 000 COP;
 - d) more than or equal to 2 400 001 COP.
- 7) The ability of family businesses to earn extra income different from the profits of the main business:
 - a) have no extra income;
 - b) one extra;
 - c) two extra;
 - d) three extra;
 - e) four extra.

The sources of income that were investigated were: income from the work activity of family members, income from alimony, income from money contributed by family members who are resident or not in the country, income from old age or disability pensions, income from businesses other than the one surveyed, income from deposits, income from stock gains, income from the sale of business machinery, income from the sale of household goods, income from donations other than to the government or from family members, income from family subsidies, and income from state subsidies.

- 8) The capacity of family businesses to make payments by means of the financial sector:
 - a) do not use means of payment from the financial sector;
 - b) one means;
 - c) two means;
 - d) three means;
 - e) four means;
 - f) five means.

The means of payment considered for this classification were: credit card, debit card, automatic account payment, transfer of funds via the Internet and telephone, prepaid card, factoring, and other.


Authors

The contribution share of author is 70% H. Rangel, 20% A. Montenegro, 10% L. Arenas

Henry Sebastián Rangel Quiñonez

Economist and Philosopher, MSc in Statistics. Universidad Nacional de Colombia (Colombia). Lecturer at Universidad Santo Tomás and Universidad Industrial de Santander. His research interests include item response theory, economic development, behavioral economics, and epistemology of science.


Correspondence to: Mr. Henry Sebastián Rangel Quiñonez, Universidad Industrial de Santander, Carrera 27 #180 - 395, Floridablanca Santander Edificio Santander, postal code 681003, Colombia, e-mail: sebasrangel@gmail.com

ORCID  <http://orcid.org/0000-0002-6745-6753>

Alvaro Mauricio Montenegro Díaz

Mathematician and Statistician. MSc in Statistics. PhD in Statistics. Professor at Universidad Nacional de Colombia (Colombia). His research interests include item response theory and Bayesian statistics.


Correspondence to: Dr. Alvaro Mauricio Montenegro Díaz, Universidad Nacional de Colombia, Carrera 30 # 45-03, edificio 405, oficina 364, postal code 111321, Colombia, e-mail: ammontenegrod@unal.edu.co

ORCID  <http://orcid.org/0000-0001-8512-9038>

Luisa Fernanda Arenas Estevez

MSc in public policy from Universidad Nacional de Colombia (Colombia). Lecturer at Universidad Pontificia Bolivariana (Colombia). Her research interests include public policy, inequality, economic development, and data analysis.

Correspondence to: Ms. Luisa Fernanda Arenas Estevez. Postal address: Autopista Piedecuesta, Km 7 Edificio E, Of 202, postal code 681007, Colombia, e-mail: luisafernandarenas@gmail.com

ORCID  <http://orcid.org/0000-0001-5022-1854>

Acknowledgements and Financial Disclosure

The authors would like to thank Ian P. Roche to Texas A&T University for his style corrections and translation help as well as to Jorge Enrique Garcia, William Gonzales Calderón and Diego Nicolas Bernal for the comments to the document.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright and License



This article is published under the terms of the Creative Commons Attribution – NoDerivs (CC BY-ND 4.0) License
<http://creativecommons.org/licenses/by-nd/4.0/>