An Information Theoretic Job Satisfaction Analysis

Enrico Ciavolino
Dipartimento di Filosofia e Scienze Sociali, Università del Salento, Italy

Abstract: The aim of this study is to analyze a Job Satisfaction (JS) model by using an information theoretic approach based on the semi-parametric Generalized Maximum Entropy (GME) estimator. The GME is used in order to estimate the parameters of the Structural Equation Model (SEM), which represents the theoretical representation of the relationships between the human being and his job. Moreover, thanks to the entropy index measure, the theoretical model is analyzed in its some particular aspects, which can be seen as sub-structures in the relationships.

Key words: Generalized maximum entropy, job satisfaction, structural equation model, normalized entropy measure

INTRODUCTION

The focus of this study is on the SEMs, applied to the JS study by using the GME estimator. More specifically, the information theoretic estimator is used to analyze a main model of JS, while in the second step, the normalized entropy measure (Golan et al., 1996) is used to evaluate the contribution of each sub-model, as part of the main model, which is interesting from a JS theoretical point of view and supported by the entropy index measure.

ESTIMATION METHODS FOR THE STRUCTURAL EQUATION MODELS

It’s well known in literature that the estimation methods for analyzing data for SEM in a soft and in a hard way are, respectively, Partial Least Squares (PLS) and Maximum Likelihood Estimation (MLE).

PLS and MLE are in the specification and in the restrictions conceptually different, as numerically different in the parameters estimation values (Wold, 1975), so, the definition of conditions to integrate the two estimation methods in a unique strategy of analysis, is not an easy practicable way. As a consequence of these distinct features, the two approaches has been reasonably considered as alternative instruments to be used according with the data availability, the conceptual assumptions, the theory background, the objective of the research and so on.

A middle way between the soft and hard modelling is represented by the GME estimation method. This method allows to explore the data by using a soft approach that can give a general view of the relationships among the variables and some indications about aspects that can be deeply investigated. This exploration should be made without taking into account restrictions about sample size, number of variables or distributional assumptions on the errors terms.

In the next paragraphs the GME estimation method and the use of normalized entropy index are presented to evaluate the sub-structure in the data.

The generalized maximum entropy estimation method: The GME method (Golan et al., 1996) represents a semi-parametric estimation method for the SEM (Al-Nasser, 2003; Ciavolino and Al-Nasser, 2009; Ciavolino, 2009) which works well in case of ill-behaved or nonlinear data. The SEM (Jöreskog, 1973), can be formulated in a system of equations as follows:

\[ \eta_{(n,i)} = B_{(n,m)} \cdot \eta_{(n,i)} + \Gamma_{(m,n)} \cdot \xi_{(n,i)} + \tau_{(n,i)} \]  \hspace{1cm} (1)

\[ y_{(p,i)} = A_{(m,n)}^{y} \cdot \eta_{(n,i)} + e_{(p,i)} \]  \hspace{1cm} (2)

\[ x_{(n,i)} = A_{(n,q)}^{x} \cdot \xi_{(n,i)} + S_{(n,i)} \]  \hspace{1cm} (3)

These equations consist of two main parts: the first part is the Structural Model (1) and represents the linear relationships among the latent variables; where \( \eta \) is the vector of the \( n \) endogenous latent variables, \( \xi \) the vector of the \( n \) exogenous latent variables, \( x \) and \( y \) respectively the vectors of the \( q \) and \( p \) manifest exogenous and endogenous variables.

The coefficient matrices \( B \) and \( \Gamma \) contain the path coefficients of the effects of the endogenous on the endogenous variables and the exogenous on the endogenous variables.
The second part is the Measurement Model, representing the relationships between the manifest and the latent variables: endogenous in Eq. 2 and exogenous in Eq. 3.

The coefficient matrices $\Lambda'$ and $\Lambda''$ measure the relationships between the manifest and latent variables, endogenous and exogenous.

The vectors $\tau$, $\varepsilon$ and $\delta$, are the structural and the measurement errors vectors. For the definition of the whole model, four other matrices have to be defined. The co-variance matrices $\Phi$, between the latent variables $\zeta$, $\Psi$, between the error term $\tau$; $\Theta^e$, between the error term $\varepsilon$; $\Theta^\delta$, between the error term $\delta$. The three SEM equations can be re-formulated as a unique function model:

$$y_{p,0} = \Lambda_{\varepsilon} \tau + \{I_{\mu,0} - P_{\mu,0}\} \tau + \{I_{\mu,0} - \delta_{\mu,0}\} + b_{p,0}$$

With $m$ endogenous latent variables, $n$ exogenous latent variables and $p$ and $q$ manifest endogenous and exogenous variables. The GME approach considers the re-parameterization of the unknown parameters and the disturbance terms as a convex combination of expected values of a discrete random variable. The coefficient matrices, $B$, $\Gamma$, $\Lambda'$, $\Lambda''$ and the co-variance matrices, $\Phi$, $\Psi$, $\Theta^e$, $\Theta^\delta$, are all re-parameterized as expected values of discrete random variable with M fixed points' for the coefficients and $N$ for the errors: $B_{(m,m)} = Z_{(m,m-M)} \cdot P_{(m,m)}$;

$$\Gamma_{(m,n)} = Z_{(m-M,n-M)} \cdot P_{(m-M,n-M)}$$

$\Lambda'_{(p,m)} = Z_{(p-M,m-M)} \cdot P_{(p-M,m-M)}$;

$\Lambda''_{(p,m)} = Z_{(p-M,m-M)} \cdot P_{(p-M,m-M)}$;

$$\tau_{(m,n)} = V_{(m,n-M)} \cdot w_{(m,n)}$$

$$\varepsilon_{(m,n)} = V_{(m,n-M)} \cdot w_{(m,n)}$$

The matrices $Z$ and $V$ are diagonal and the generic matrix element is represented respectively by the vectors $Z_{k} = [-C -C^2 0 0 2 C]$ with $|k|=1,\ldots,m$; $p,q$ and $V_{h} = [-b -b 2 0 b 2] with |h|=1,\ldots,n, p,q$. These vectors define the support space and are called fixed points. Usually they are composed by five elements ($M=N=5$), but the fixed points $v_i$ related to the sample observed, can be chosen by Chebychev's inequality (Pukelsheim, 1994).

Given the re-parameterization and the re-formulation, the GME system can be expressed as a constrained non-linear programming problem. The coefficients and the error terms are estimated by recovering the probability distribution of the discrete random variables set. The vectors $p^i = \text{vec}(P^i \cdot P)$, $p^{\delta} = \text{vec}(\delta \cdot P)$, are obtained by using the vec operator of the matrices $P$, $\delta$. The probability vectors: $P^i$, $P^\delta$, $P^{\delta}$, $W^i$, $W^\delta$, are calculated by maximizing the following Shannon (1948) entropy function:

$$H(p^i, P^\delta, p^{\delta}, W^i, W^\delta) =$$

$$-p^i_{(j)} \cdot \ln p^i_{(j)} - p^{\delta}_{(i,j)} \cdot \ln p^{\delta}_{(i,j)} - p^{\delta}_{(j)} \cdot \ln p^{\delta}_{(j)} +$$

$$-w^i_{(j)} \cdot \ln w^i_{(j)} - w^{\delta}_{(i,j)} \cdot \ln w^{\delta}_{(i,j)} - w^{\delta}_{(j)} \cdot \ln w^{\delta}_{(j)}$$

subjected to the consistency constraint, which represents the information about the model reported in the Eq. 4 and normalization constraints, that means, the sum of both coefficients and the error terms probability vectors have to be equal to 1:

$$I_{m,m} \cdot \Gamma_{(m,m)} \cdot P_{(m,m)} = I_{m,m}$$

$$I_{m,m} \cdot \Omega_{(m,m)} \cdot P_{(m,m)} = I_{m,m}$$

$$I_{n,m} \cdot \Phi_{(n,m)} \cdot P_{(n,m)} = I_{n,m}$$

$$I_{n,m} \cdot \Psi_{(n,m)} \cdot P_{(n,m)} = I_{n,m}$$

$$I_{n,m} \cdot \Theta^e_{(n,m)} \cdot P_{(n,m)} = I_{n,m}$$

$$I_{n,m} \cdot \Theta^\delta_{(n,m)} \cdot P_{(n,m)} = I_{n,m}$$

The GME provides, also, the normalized entropy measure (Golan et al., 1996; Golan, 2008) that quantifies the level of information in the data, giving a global measure of the goodness of relationships, hypothesized in a model.

The normalized entropy measure: The normalized entropy measure can be expressed by the following formulation:

$$S(p) = -p^i \cdot \ln p^i$$

$$K \cdot \ln M$$

The normalized index is a measure of the uncertainty information, where $-p^i \cdot \ln p^i$ is the Shannon's entropy function, defined only for the coefficients probabilities; $K$ is the number of the coefficients to be estimated and $M$ is the number of the fixed points.

The quantity $K \cdot \ln M$ represents maximum uncertainty, since in the case of no contribution of one variable, the entropy coefficient is that of a uniform distribution of the $K$ variables with $M$ outcomes:

$$\ln M = M - 1 + \frac{1}{M}$$

If $S(p) = 0$, there is no uncertainty, while $S(p) = 1$ means total uncertainty. The normalized entropy measure is used to derive the log-likelihood ratio statistic:

$$W(\text{ME}) = 2 K \ln(M) [1 - S(p)]$$
Which, under the null hypothesis that all the parameters are zero, converges in distribution to a $\chi^2$ with K degrees of freedom, where the degrees of freedom are the number of constraints imposed. Moreover a Pseudo-$R^2$ can be defined as:

$$R^2 = 1 - S(\hat{p})$$  \hspace{1cm} (9)

The $R^2$ is a measure to derive a goodness of fit of the model, where the value 0 implies no informational value of the data and 1 implies perfect certainty or perfect in-sample prediction. This measure is the same as the information index in Soofi (1992).

The normalized entropy measure can be used also as a method for the selection of the explicative variables. In fact, it is possible to calculate the Shannon’s entropy function for one predictor (with K=1) as $-p_k \cdot \ln p_k$ so, the uncertainty information of the kth explicative variable is the following:

$$S(\hat{p}_k) = -\frac{p_k \cdot \ln p_k}{\ln M}$$ \hspace{1cm} (10)

The normalized entropy measure can determine the uncertainty information of a specified sub-structure in the data. Let consider a SEM where according with some theoretical assumptions, can be divided into S elements, whose generic $S^k$ element defines a sub-model with a specific number of coefficients $c_{ik}$ such that $c_{ik} \cdot K$. The level of the uncertainty information of the specified Sth sub-model can be defined as follow:

$$S(\hat{p}_S) = \sum_{i=1}^{S} -\frac{p_{ik} \cdot \ln p_{ik}}{c_{ik} \cdot \ln M}$$ \hspace{1cm} (11)

The quantity above defined expresses the contribution of the $S^k$ part of the model in making the uncertainty information of the whole SEM system. The normalized entropy measure of whole SEM, in case the sub-models are mutually exclusives, can be expressed as the weighted average of the normalized entropy measure of each sub-model, as follow:

$$S(\hat{p}) = \sum_{i=1}^{S} \frac{S(\hat{p}_S) \cdot c_i}{\sum_{i=1}^{S} c_i} = -\frac{p \cdot \ln p}{K \cdot \ln M}$$ \hspace{1cm} (12)

This measure can be a suitable tool to understand the percentage of information explained by the different sub-models that can be specified as a part of the principal model.

**THE JOB SATISFACTION EVIDENCE**

In order to show an application of the GME as method to detect SEM sub-structures in the data and to understand the human labor aspects, a job satisfaction model is presented. Data were gathered from a sample composed of 87 employees working in the private sector in the province of Leece, a small town situated in the South of Italy.

The tool for studying the JS was a questionnaire, face to face administered and subdivided into specific evaluation areas for investigating on JS related to workers’ attitudes.

The questionnaires were made up of 20 items, used for measuring 6 latent constructs (Leadership, Future Perspective, Decision Autonomy, Salary, Professional Improvement and Job Satisfaction) which represent both intrinsic and extrinsic aspects of one’s job (Carpita, 2009). Each item is an indicator (manifest variable) having the form of a positive statement concerning several elements of the working reality of the subjects interviewed.

The workers were asked to evaluate the statements by using an ordinal scale consisting in five response choices ranging from 1 (total disagree) to 5 (maximum agree). Table 1 shows the Manifest Variables (MVs) related to each Latent Variables (LVs) for the job satisfaction measurement model Ciavolino (2009).

The main model is reported in Fig. 1, where the SEM graphical representation considers: latent variables, defined by using the Greek letters and drawn by circles; manifest variables, defined by Latin letters and drawn by rectangles.

**The main model of job satisfaction:** The Fig. 2 shows the GME estimated structural coefficients ($\beta$ and $\gamma$) of the theoretical JS model above defined. All the causal relationships are significant. The GME provides first step of explorative analysis in order to evaluate the relationships among the selected latent and manifest variables, without requiring any assumptions, with respect to the errors distribution, the sample size and the multi-collinearity.

---

2Professional Improvement concerns workers’ chances to grow up in professional sense, by boosting their skills; the Future Perspective are the worker’s expectations, depending on the management and on the organization of the workplace; Salary is the variable more recognizable and appraisable by the workers because they immediately act on the workers’ lifestyle; the Leadership represent relational aspects in the working life. Decision Autonomy measure the degree of perceived autonomy that workers enjoy during their job and JS includes indicators of the overall satisfaction.
Table 1: Latent and manifest variables

<table>
<thead>
<tr>
<th>LV</th>
<th>Manifest Variables (MV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership</td>
<td>• $X_1$: Management informs the employees about the functions and the aims</td>
</tr>
<tr>
<td></td>
<td>• $X_2$: The communication between management and employees is real possible</td>
</tr>
<tr>
<td></td>
<td>• $X_3$: The communications between management and employees are clear and punctual</td>
</tr>
<tr>
<td>Decision autonomy</td>
<td>• $X_4$: The employees take decision autonomy</td>
</tr>
<tr>
<td></td>
<td>• $X_5$: The employees can choose their own method of working</td>
</tr>
<tr>
<td></td>
<td>• $X_6$: In the enterprise, innovations and changing organization are encouraged</td>
</tr>
<tr>
<td>Future perspectives</td>
<td>• $X_7$: People’s merits are recognised and rewarded</td>
</tr>
<tr>
<td>Salary</td>
<td>• $X_8$: Salary is fair in relation to my responsibility</td>
</tr>
<tr>
<td></td>
<td>• $X_9$: Salary is fair in relation to the working hours</td>
</tr>
<tr>
<td></td>
<td>• $X_{10}$: Wage is fair in relation to my educational qualification</td>
</tr>
<tr>
<td></td>
<td>• $X_{11}$: In the enterprise there are real opportunities to obtain wage rise</td>
</tr>
<tr>
<td>Professional</td>
<td>• $X_{12}$: The enterprise in which I am working offers some opportunities to improve my</td>
</tr>
<tr>
<td>improvements</td>
<td>skills</td>
</tr>
<tr>
<td></td>
<td>• $X_{13}$: The enterprise organizes training courses</td>
</tr>
<tr>
<td></td>
<td>• $X_{14}$: The tasks are professionalized</td>
</tr>
<tr>
<td>Job satisfaction</td>
<td>• $X_{15}$: I am satisfied about my work</td>
</tr>
<tr>
<td></td>
<td>• $X_{16}$: My task and my qualification are correlated</td>
</tr>
<tr>
<td></td>
<td>• $X_{17}$: My task is suit to my capabilities</td>
</tr>
<tr>
<td></td>
<td>• $X_{18}$: My task reflects my expectation</td>
</tr>
<tr>
<td></td>
<td>• $X_{19}$: My task is rewarding</td>
</tr>
<tr>
<td></td>
<td>• $X_{20}$: My functions are planned in a good way</td>
</tr>
</tbody>
</table>

Fig. 1: The path diagram of the job satisfaction main model

Fig. 2: The estimate coefficients of the job satisfaction main model

2The LV Salary is referred to the equity aspects in the work life, as a matter of fact, many authors define it as distributive fairness (Carpita, 2009).

*The items of the job satisfaction measure as overall satisfaction respect to the qualification, capabilities, expectations, rewarding and planning of the work functions.
The first Job Satisfaction model hypothesized involves a remark about the Leadership style (Ciavolino and Dahlgaard, 2009; Dahlgaard and Ciavolino, 2007). The main challenges for leaders are to build a long-term vision, to increase commitment, to build teams andcoalitions in order to reach required organizational objectives (Kuean et al., 2010). This consideration lie at the bottom of the decision to link the latent variable Leadership with the Decision Autonomy (0.479) and the future perspectives (0.556) constructs.

The future perspectives in their turn influence the perception about Salary (0.668) and Professional Improvements (0.679). It is reasonable to hypothesize that, if an enterprise offers to the employees the chance to obtain promotions and embraces a meritocratic principle by recognizing and rewarding workers’ merits (considering a long term perspective), employees are incited to improve their skills and to better their performance. Salary (0.249) and Professional Improvements (0.328) act on Job Satisfaction. The Salary represents one of the most important extrinsic feature, taking into account monetary aspect, but the LV above defined is more related to the distributive fairness and consequently the non-monetary (e.g., working hours or professional development) aspects of the work. Moreover opportunity to achieve professional improvements seems to have an higher impact than the salary. In the model theorized Decision Autonomy, influenced by Leadership, acts on Job Satisfaction (0.377). An enterprise democratically regulated, in which workers can propose suggestions and choose their own working method taking part in the decision-making process.

The sub-models analysis: Table 2 reports the values of the normalized entropy for each manifest variable that measures the reduction in uncertainty information of each manifest variable. The quantity $S(\hat{p}_i)$ gives information about the distance between the probability distribution of each coefficient and a uniform distribution. The value $S(\hat{p}_i) = 0$ implies a strong relationship, while the value $S(\hat{p}_i) = 1$ implies the variable is extraneous from the model.

Table 3: The normalized entropy measures of each sub-model

<table>
<thead>
<tr>
<th>Models</th>
<th>$c$</th>
<th>$c$ ln(M)</th>
<th>$\sum p_i \cdot \ln p_i$</th>
<th>$S(\hat{p}_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1 = Motivational$</td>
<td>12</td>
<td>19.335</td>
<td>9.162</td>
<td>0.474</td>
</tr>
<tr>
<td>$s_2 = Pseudo-monetary$</td>
<td>13</td>
<td>20.923</td>
<td>11.034</td>
<td>0.527</td>
</tr>
<tr>
<td>$s_3 = Relational$</td>
<td>14</td>
<td>22.532</td>
<td>14.317</td>
<td>0.635</td>
</tr>
<tr>
<td>The main model</td>
<td>27</td>
<td>43.455</td>
<td>21.391</td>
<td>0.492</td>
</tr>
</tbody>
</table>

This table could be also interpreted as a further item-scale analysis which takes into account the casual relationships among the LVs. In this case it is possible to read that same items, as for instance the $X_{13}$ for the Professional Improvements or $X_{14}$ for the Job Satisfaction, could be considered not belonging the this JS model. These results suggest us that additional considerations can be done on the specification of the measurement model, but for the aim of this paper we will take all the manifest variables.

The first explorative study allows to test the validity of the theoretical hypotheses based on the main model, by analyzing the statistical features of the structural and measurement relationships, but, at the same time, it allows to evaluate the contribution of each sub-model. Leaving out the interpretation of the manifest variables, we focalize on how the normalized entropy measure can reflect the uncertainty information of each above specified structure in the model (supra §2.2).

By reminding this, three sub-models have been individuated with the following structure: The sub-models contain the LV Future Perspective as exogenous variable which, in the first sub-model, acts on the Professional Improvements; in the second sub-model acts on the Salary. These endogenous variables, in both cases, have a direct and casual action on the JS. The former sub-model has been named Motivational Model ($s_1$) the latter is the Pseudo-Monetary ($s_2$) one. The third sub-model is the so called Relational model ($s_3$) linking the exogenous variable Leadership to the Decision Autonomy, related to Job Satisfaction.

Table 3 contains the results of the normalized entropy measures related to each sub-model.

\footnote{It is called pseudo-monetary model because it is reported to the equity aspects in the work life.}
The number of fixed points $M$ is equal to 5. The first two columns \( c_i \) and $c_i \ln(M)$ report, respectively: the number of constraints on the variables, that means the number of coefficients estimated and the entropy value in case of maximum uncertainty condition.

The following two columns \( \sum_{k=1}^{c_i} p_k \cdot \ln(p_k) \) and $S(\hat{p}_i)$, represent respectively: the entropy estimated and normalized entropy measure to evaluate the informational content of the estimates with 1 reflecting uniformity (complete ignorance) of the estimates and 0 reflecting perfect knowledge (supra §2.2).

The first three rows report the results of the three sub-models ($s_1$=Motivational, $s_2$=Pseudo-Monetary, $s_3$=Relational), while in the last row there is the result of the Main Model.

In this case the sub-models are not disjoint, that means, some parts of one specific model are also a part of another model (the common parts are the future perspective and the job satisfaction), so the whole SEM system normalized entropy measure cannot be obtained as an entropy weighted average of each sub-models (Eq. 12).

The JS main model has a goodness of fit equal to 0.508, which represents a good level of reduction of uncertainty generated by this theoretical model. The contributions of each sub-model in reduction of uncertainty are respectively $S(\hat{p}_1) = 0.474$, $S(\hat{p}_2) = 0.572$, $S(\hat{p}_3) = 0.635$.

Although the model are not disjoint, it is interesting to see that same sub-models have a value of $S(\hat{p}_1)$ smaller than the main model, meaning that the introduction of the other variables and coefficients, improve the interpretability of the JS model and the reduction of the uncertainty.

In particular the Relational Model, with an $S(\hat{p}_3) = 0.635$ seems to give no so strong contribution in the reduction of the uncertainty information of the main model.

On the other hand, Motivational Model and Pseudo-Monetary are the most important parts in the contribution in making the information of the whole SEM system. As the matter of the fact, the contribution of each sub-model is quite similar and very close the entropy expressed by the main. The interpretation is that these two part of the main model represent two of the main important aspects to be considered in the explanation of the JS, because the introduction of other parameters don't give a so high impact in the reduction of uncertainty.

**CONCLUSIONS AND DISCUSSION**

This study has focused on a study of the JS based on a survey composed by 87 Italian employees working in the private sector, in order to show the GME as method to detect SEM sub-structures in the data and to understand the human labor aspects.

It has been shown how normalized entropy measure can be a suitable tool to improve the interpretative overall results by dividing and measuring the contribution of each part of the model in making the reduction of the uncertainty information.

This method gave us the chance to strengthen the theoretical substratum of the model and to pick out some sub-models for better understanding the working dynamics for JS (Hosseinian et al., 2008).

Through the Motivational Model it is emerged that the Future Perspective emphasizes the importance of the Professional Improvements during the working life to improve the JS. Instead, through the Pseudo-Monetary Model it is emerged the Future Perspective influence on the perception about own Salary. In this theoretical position, Salary is a non monetary variable because it doesn't contain monetary measures but social measures. Besides, given the Future Perspective causal action on the Salary, this last has been defined such as dynamic subjective variable which changes in according to future expectations.

The last local structure analysis of Job Satisfaction has been developed on the model specifying causal relationships between Leadership, Decision Autonomy and Job Satisfaction. It is emerged that the so called Relational Model in the reduction of uncertainty information of the main model is not so strong. It is undoubtedly that the quality of interpersonal exchange between superiors and subordinates can make up an important element of the employee's job satisfaction and for the performance of the enterprise, for this reason this sub-model of the main JS model should be better studied or measured in the survey used for the analysis, to improve the contribution in the specification of the JS.

In conclusion, the results obtained by conducting the SEM GME analysis allowed fixing some relevant points to be taken into account in implementing policies aiming at realizing an efficient and satisfying organization of work. In this sense individual components of work (and their satisfaction) are to be considered a focal resource for enterprises which decide to invest on human capital giving rise to a series of activities which enable working
people and their employing organizations to agree about the objectives and nature of their working relationships and secondly, ensures that the agreement and goals are fulfilled.

REFERENCES


