REVSTAT – Statistical Journal
Volume 7, Number 3, November 2009, 265–290

TOWARD THE DEFINITION OF A STRUCTURAL EQUATION MODEL OF PATENT VALUE: PLS PATH MODELLING WITH FORMATIVE CONSTRUCTS

Authors: Alba Martinez-Ruiz
– Department of Statistics and Operations Research,
Technical University of Catalonia, Barcelona, Spain
Universidad Católica de la Ssma. Concepción, Chile
alba.martinez-ruiz@upc.edu

Tomas Aluja-Banet
– Department of Statistics and Operations Research,
Technical University of Catalonia, Barcelona, Spain
tomas.aluja@upc.edu

Received: October 2009 Revised: October 2009 Accepted: November 2009

Abstract:

• This paper aims to propose a structural equation model which relates the variables that determine the patent value. Even though some patent indicators have been directly used to infer the private or social value of innovations, the results suggest that patent value is a more complex variable that may be modeled as an endogenous unobservable variable in a first- and in a second-order model, and which depends respectively on three and four constructs. Such variables include the knowledge used by companies to create their inventions, the technological scope of the inventions, the international scope of protection, and the technological usefulness of the inventions. The model allows the conceptualization of patent value into a potential and a recognized value of intangible assets, aiming toward an index construction approach. Partial least square (PLS) path modelling is performed as an exploratory model-building procedure. We use a sample of 2,901 patents granted in the United States in the field of renewable energy.

Key-Words:

• *patent value; patent indicators; PLS path modelling; structural equations models.*

AMS Subject Classification:

• 62H25, 62P20.

1. INTRODUCTION

Patents are one of the main sources of technological information. A patent is an exclusive right granted to inventors by a state only when the invention fulfils three basic requirements: the invention is new, it involves an inventive activity and it is useful for industry. Until now research involving patent data has been associated with the analysis of information contained in the patent document, such as backward and forward citations or number of claims, and the relationship between patents and research and development (R&D), innovation or economic growth. In recent years, patent indicators have been used to study the economical value of patents. In most cases, analytical approaches have been based on standard econometric analysis techniques such as probit or logit models, and survey analysis. However, patent value may be seen as a complex construct depending on a variety of elements. General and specific market conditions, countries' legal frameworks, geographic proximity or accumulated scientific and technological knowledge are different dimensions that have shown to affect patent value.

This paper proposes that a holistic and multidimensional model may offer a robust understanding of the different variables that determine patent value. For the moment, and considering patent document information, two path models are built considering five dimensions represented by five constructs. They are: patent value, technological usefulness of the invention, knowledge stock used by the company to create the technology, technological scope of the invention, and international scope of protection. The models are strongly based on the theory developed by the technological change scientific community and a thorough review of the literature on patent valuation. Each construct is associated with a set of observable variables. So, they can be estimated by these indicators. Manifest variables are mainly built from information contained in patent documents. A set of patents granted in the United States (U.S.) in the area of renewable energies was retrieved from Delphion database. The proposed path models are replicable because they could be repeated for different technological fields or countries. Moreover, the models may allow one to distinguish between: (a) those variables related to patent value at the time of application, i.e. those variables that could deliver a measure of potential value of patents, and (b) those that determine the value after the patent's application.

In the literature, research that addresses patent value using a structural equation model (SEM) approach is quite scarce. Moreover, rather traditional methods based on multivariate normal distribution assumption have been implemented. The advantage of SEM is flexibility in working with theory and data, approaching the whole phenomenon, and a more complete representation of the complex theory. Additionally, and contrary to a covariance-based approach such as the linear structural relation model (LISREL), PLS path modelling is theorybuilding-oriented and causal-predictive-oriented. Therefore, the exploratory nature of this procedure allows for the first formulation of a structural model of patent value. Finally, the PLS path modelling algorithm is a powerful technique for the analysis of skewed or long-tail data, such as patent data. Therefore, we also attempt to show the benefits of PLS path modelling as a tool for exploration and prediction of skewed data.

In this research, the models specification is made from a PLS perspective. So, we are posing PLS models. Section 2 provides background on patent indicators and constructs, and section 3 reviews the PLS path modelling procedure for hierarchical component models with repeated manifest variables and formative constructs. Section 4 addresses the first- and second-order model formulation, while also postulating on the indicators, latent variables (LVs) and causal relationships among variables. In particular, formative and reflective relationships among manifest and latent variables are justified. A description of patent data is given in Section 5. Section 6 reports the results, and shows the performance and effectiveness of PLS path modelling when working with patent data characterized by long tails. Finally, section 7 gives final remarks and some directions for future research.

2. PATENT INDICATORS AND CONSTRUCTS

Patent indicators have been used by scientific communities to study phenomena such as technological change or the growth of science and technology. Forward citations, i.e. the number of times that each patent has been cited by another patent, are the most widely used indicator to measure the value or importance of patents. Nevertheless, other indicators have also been introduced as a measure of value, such as family size, number of claims, number of international patent classification (IPC) codes where the patent is classified, and backward citations. Here, family size refers to the number of countries where a patent is sought for the same invention [27]. As a general patenting strategy, companies protect their inventions in their local countries first and then in other jurisdictions. Patents with a large family size tend to be more valuable or important [21], although Guellec *et al.* reported that this relationship might sometimes be inaccurate and "may reflect a lack of maturity of the applicant" [18, p. 114]. Even so, family size may be proposed as a proxy variable for the international scope of patent rights, and as a measure of patent value. The number of backward citations or references in a patent represents "all of the important prior art upon which the issued patent improves" [35, p. 318], and allows one to demonstrate that the invention is genuinely new. Claims are made in a special section in the patent document, where the thing that is being protected is specified. The claims section consists of a numbered list. Therefore, the number of claims is in fact the number of inventions protected [42, p. 134]. Patents with a large number of claims have a higher likelihood of being litigated, so they can be considered more valuable [22, 28, 38]. International patent classification classes were introduced as a proxy variable for the scope of protection by Lerner [31]. An invention with a larger technological scope should be more valuable due to its broader potential applications. The number of inventors and the number of applicants have also been used as indicators of the patent value [38].

Most patent indicators have been used to explain a conceptual variable or a construct. The relationship between patent citations and patent value has been deeply studied [1, 4, 18, 20, 21, 37, 38, 43]. Carpenter *et al.* [4], Albert *et al.* [1] and Harhoff *et al.* [20] have successfully shown that those patents that are related to important technological developments are most highly cited. Harhoff *et al.* [21] was the first to use backward and forward citations together as proxy variables for patent value, and Trajtenberg [43] established the role of citations as an indicator of the value of innovations. Patent citations and patent value have also been associated with market value and/or the R&D expenditures of companies [10, 15, 19, 31]. The relationship among patent value and patenting strategy, technological diversity (through the IPC), domestic and international R&D collaborations and/or co-applications (analyzing the country of residence of the authors) and the mix of designated states for protection (through the family size), have been studied by Guellec and van Pottelsberghe [18]. Reitzig [37, 38] studied the factors that determine an individual patent value. Analyzing the results of a questionnaire, he found that novelty and inventive activity are the most important factors in patents that are used as "bargaining chips". Connolly *et al.* [10] showed that patent statistics are significantly related to companies' market value. In addition, Griliches [15] found a significant relation among companies' market value, the book value of R&D expenditures and the number of patents. He based his research on a time-series cross-section analysis of United States firm data. Lerner [31] reported that patent scope has a significant impact on the valuation of firms, while Hall *et al.* [19] investigated the trend in US patenting activities over the last 30 years, finding that the ratios of R&D to asset stock, patents to R&D, and citations to patents significantly affect companies' market value.

On the other hand, some of these indicators have been related to other constructs. The number of inventors and applicants, backward citations and the number of claims have been related to patent novelty, i.e. the technological distance between a protected invention and prior art. A patent's protection level or its technological scope or breadth can be measured by the number of claims or number of IPC classes into which the patent is classified [31]. Furthermore, patent stocks or knowledge stocks have been associated with the economic growth of a country as well as the economical activity [16], research and development results [29] and the value of innovation [40] and technological performance [42]. In this last case, the researchers found that the number of claims is a better indicator than the number of patents in the national technological capacity.

Finally, little research has reported on the structural relationship among latent variables which influence patent value using a multidimensional approach. The recent investigations of Harhoff [21, 22] and Reitzig [37, 38] used a large number of indicators of patent value aimed mainly at estimating the probability of opposition to a patent. In most cases, analytical approaches have been based on standard econometric analysis techniques (probit or logit models) or survey analysis. One reason that could explain why a multidimensional and structural approach has not been applied to technology/patent valuation is that more general structural models are based on maximum likelihood estimation and the multivariate normal distribution of data. Patent indicators are very heterogeneous and asymmetric, and, in general, they exhibit a large variance and skew. Consequently, assuming that this type of data has a multivariate normal distribution may lead to biased results. As seen below, PLS path modelling overcomes this drawback because it is an iterative algorithm that makes no assumptions about data distribution. Moreover, unlike other methods such as probit or logit models, it allows researchers to depict the relationship among a set of latent variables. Thus, we have the possibility of modelling the patent value as an unobservable variable.

2.1. Patent value

Patents are intellectual assets that do not necessarily have an immediate return. A patent may protect a product that can be manufactured and sold. But a patent may also protect technologies which, together with other technologies, enable the manufacture of a final product. In both cases, to obtain an economic value from patents may be extremely difficult. In studying patent value, different approaches have been taken throughout the literature. Some of the approaches focus on the private value of a patent while others concentrate on a patent's social value. Lanjouw *et al.* [27, p. 407] defined the private value of a patent in terms of "the difference in the returns that would accrue to the innovation with and without patent protection". The magnitude of this difference would be crucial in applying or renewing the protection. Reitzig [38] also focused on the private value of patents, and specifies the need to consider the patent value as a construct. Technical experts were surveyed and, according to them, the research showed that the factors that determine patent value are: state of the art (existing technologies), novelty, inventiveness, breadth, difficulty of inventing, disclosure and dependence on complementary assets[1]. Additionally, Trajtenberg [43] showed that patent data was highly correlated with some indicators of the social benefits of innovations. Guellec *et al.* [18] presented a value scale proposing

[1]We attempt to consider these variables as constructs in the proposed structural model. However, recall that in this research, the manifest variables are mainly obtained from the patent document. So, latent and manifest variables are subject to this constraint.

that technology increases its own value as it passes through different stages: from invention to application, examination, publication and decision to grant, and finally to the high value stage if the patent is granted. The distinction is made between the intrinsic value of the patent simply for being granted (and thereby having proven novelty, inventive activity and applicability) and the potential value of technology (dependent on its potential for generating future returns).

Some patent indicators have been used to directly infer the patent's value, such as forward citations or family size (see Table 1). Even though this may be useful and may give an approximation of the patent value, many elements may affect the invention and protection process. We consider some of these factors based on the presented background, and represent their interactions proposing a multidimensional analysis of the problem. It is worth noting that this research does not seek to determine the value of an individual patent or to obtain a monetary value of the assets. Rather, the patent value is proposed in terms of the technological usefulness of the inventions. This model, however, allows us to compare and rank the value of a company's patent portfolios. We address the question of what variables determine the patent value and how they relate to each other. These variables are modeled as unobserved variables. So, they and their relationships set up a structural equation model.

Table 1: Brief summary of different approaches used to study the patent value.

Author	**Construct**	**Indicators**	**Dependent variable**	**Method**
Trajtenberg (1990)	Social value of innovations	Patent count weighted by citations	Consumer surplus	Multinomial logit model
Guellec *et al.* (2000)	Patent value, patenting strategy, technological diversity, R&D collaboration	Number of IPC, family size, dummy variables, etc.	Probability that a EPO patent application is granted	Probit model
Reitzig (2003)	Patent value, novelty, inventive activity, invent around, disclosure	—	'present patent value'	Survey, probit model
Harhoff *et al.* (2003)	Private value of patents, value of renewed patent protection and asset value of patent right	Survey of patent-holders, backward and forward citations, family size, IPC, outcome of opposition proceedings	Patent right as a price to sell the patent right	Survey, probit model
Hall *et al.* (2005)	Market value	Patent citations, R&D expenditures, total assets	Tobin's q	Tobin's Q equation

3. THE PLS PATH MODELLING APPROACH FOR MODEL FORMULATION

PLS path modelling is a component-based procedure for estimating a sequence of latent variables developed by the statistician and econometrician Herman Wold [45, 46, 47]. During the last few years, it has proved to be useful for estimating structural models, in marketing and information system research in particular, and in the social sciences in general [6, 12, 23, 24, 33, 41]. Some of its features have encouraged its use, such as: (1) it is an iterative algorithm that offers an explicit estimation of the latent variables, and their relationships, (2) it works with few cases and makes no assumptions about data distribution — in contrast with LISREL that makes strong assumptions about data distribution and where hundreds of cases are necessary for its application, and (3) it overcomes the identification problems when formative measurement models are included. Wold [47] emphasizes that "using prior knowledge and intuition the investigator is free to specify the LVs, to design the inner relations, and to compile a selection of indicators for each LV" [p. 582]. The path model "is usually tentative since the model construction is an evolutionary process. The empirical content of the model is extracted from the data, and the model is improved by interactions through the estimation between the model and the data and the reactions of the researcher" [45, p. 70].

In a PLS path modelling approach, the structural model or inner model — also called the inner relations and substantive theory — depicts the relationship among latent variables as multiple regressions:

$$\xi_j = \beta_{j0} + \sum_i \beta_{ji}\, \xi_i + \nu_j \tag{3.1}$$

where ξ_j and ξ_i are the endogenous and exogenous latent variables, respectively, and β_{ji} are called path coefficients and measure the relationship among constructs. The arrangement of the structural model is strongly supported by theory at the model specification stage. So, PLS path modelling is used to explore if these relationships hold up or whether other theory-based specifications, that may be proposed, help in providing a better explanation for a particular phenomenon. The condition imposed is $E(\xi_j/\xi_i) = \sum_i \beta_{ji}\, \xi_i$. There is no linear relationship between predictor and residual, $E(\nu_j/\forall \xi_i) = 0$ and $\mathrm{cov}(\nu_j, \xi_i) = 0$.

The measurement model or outer model — also called the outer relations — describes the relationship between latent (ξ_i) and manifest (x_{ih}) variables in two different ways: mode A and mode B. "Mode A is often used for an endogenous LV and mode B for an exogenous one. Mode A is appropriate for a block with a reflective measurement model and mode B for a formative one" [41, p. 268]. Reflective relationships seek to represent variance and covariances between the manifest variables that are generated or caused by a latent variable. So, observed variables are treated as an effect of unobserved variables [2, 9]. In a reflective measurement model, the manifest variables are measured with error. Alternatively, formative relationships are used to minimize residuals in the structural relationships [14], and here, manifest variables are treated as forming the unobserved variables. MacCallum and Browne [32] said that observed variables in a formative model are exogenous measured variables. In a formative outer model the manifest variables are presumed to be error-free and the unobserved variable is estimated as a linear combination of the manifest variables plus a disturbance term, so they are not true latent variables (as in the traditional factorial approach). As in this case all variables forming the construct should be considered, the disturbance term represents all those non-modeled causes.

In mode A or in reflective relationships, manifest and latent variables relationships are described by ordinary least square regressions:

$$x_{ih} = \pi_{ih0} + \pi_{ih}\,\xi_i + \epsilon_{ih} \;. \tag{3.2}$$

The parameters π_h are called loadings. The condition imposed is $E(x_h/\xi) = \pi_{h0} + \pi_h\,\xi$, ϵ_h with zero mean and uncorrelated with ξ. Loadings indicate the extent to which each indicator reflects the construct, and represent the correlation between indicators and component scores.

In mode B or in formative relationships, unobserved variables are generated by their own manifest variables as a linear function of them and a residual:

$$\xi_i = \sum_h w_{ih}\,x_{ih} + \delta_i \;. \tag{3.3}$$

The parameters w_h are called weights, and allow us to determine the extent to which each indicator contributes to the formation of the constructs. Each block of manifest variables may be multidimensional. The condition imposed is $E(\xi/x_h) = \sum_h w_h\,x_h$. This implies that the residuals δ_i have zero mean and they are uncorrelated with the manifest variables x_i.

Wold's basic-design of PLS path modelling [45, 46, 47] does not consider higher-order latent variables. Therefore, in Wold's algorithm each construct must be related to a set of observed variables in order to be estimated. However, Lohmöller [30] proposed a procedure for the case of hierarchical constructs; that is to say, for cases where there is a construct that does not have a block of measurement variables, or more simply: it is only related to other constructs. In hierarchical component modelling, manifest variables of first-order latent variables are repeated for the second-order latent variable. So, a set of "auxiliary" variables is introduced for estimation purposes. After that, the model is estimated using PLS path modelling in the usual way. Hence, the specification of PLS has an additional equation that Lohmöller [30] called the cross-level relation:

$$y_{jl} = \pi_{jl0} + \pi_{jl}\,\xi_j + \epsilon_{jl} \;. \tag{3.4}$$

The condition imposed is $E(\xi_j \epsilon_{jl}) = 0$. We are interested in this type of model because, as seen below, the patent value construct may be modeled as a second-order latent variable, i.e. the value can only be estimated through linear relations with other latent variables.

Reliability of reflective measurement models is evaluated by examining loadings. A rule of thumb generally accepted is 0.7 or more. This implies that "there is more shared variance between construct and variable than error variance" [24, p. 198]. A low value in a loading factor suggests that the indicator has little relation to the associated construct. All indicators of a block of variables must reflect the same construct. Therefore, there should be high collinearity within each block of variables. Thus, the internal consistency of a reflective measurement model is related to the coherence between constructs and their measurement variables. The unidimensionality of the block of variables may be assessed by using Cronbach's alpha coefficient (should be > 0.7), and composite reliability (should be > 0.7). According to Chin [6, p. 320] "alpha tends to be a lower bound estimate of reliability whereas composite reliability is a closer approximation under the assumption that the parameter estimates are accurate".

To represent the extent to which measures of a given construct differ from measures of other constructs (discriminant validity), the average variance extracted (AVE) may be calculated. Therefore, as suggested by Fornell and Larcker [13], the percentage of variance captured by the construct in relation to the variance due to random measurement error is computed (should be > 0.5). Likewise when models have more than two reflective constructs, cross loadings may be obtained by calculating the correlations between component scores and indicators associated with other reflective constructs. If an indicator has higher correlation with another latent variable instead of the associated latent variable, its position should be reconsidered in the model. Therefore, each indicator has to be more related to its construct than another one in the same model. To assess the significance of loadings, weights and path coefficients, standard errors and t-values may be computed by bootstrapping (200 samples; t-value > 1.65 significant at the 0.05 level; t-value > 2 significant at the 0.01 level).

The inner model is assessed by examining the path coefficients among latent variables. The value of path coefficients provides evidence regarding the strength of the association among latent variables. Moreover, the coefficient of determination (R-square) of each endogenous variable gives the overall fit of the model or the percentage of variance explained by the model. In this research, PLS path modelling and bootstrapping were carried out in SmartPLS [39] with a centroid weighting scheme.

3.1. A brief overview of formative and reflective outer models

The distinction between reflective and formative measurement models for structural equation models is an issue that has been addressed by several scientific communities. Major contributions have been made by researchers from statistics [9], psychology and sociology [2, 3], information science [36], and business and marketing research [11, 14]. There are some decision rules criteria to determine if a relationship should be modeled as formative or reflective (mode B or mode A in the Wold's PLS approach). The guidelines can be summarized in five points as follows [9, 14, 34]. (1) The strong theory and the previous knowledge of a phenomenon under study should help to clarify the generative nature of the construct. When a formative relationship is considered, manifest variables must cover the entire scope of construct. (2) Correlations among manifest variables. In a reflective outer model, manifest variables have to be highly correlated; in contrast this condition must not be applied in a formative outer model. (3) Within-construct correlations versus between-construct correlations. This is a common practice in the model specification stage by means of cross-validation; the applied rule is that the former should be greater than the latter. However, Bollen and Lennox [2] show that this may lead to an incorrect indicator selection for reflective and formative outer models, because this rule may have exceptions. So, the condition must be applied with caution. (4) Sample size and multicollinearity affect the stability of indicator coefficients, and they are a frequent problem in multiple regressions. So, multicollinearity will influence the quality of the estimates in formative relationships. (5) Interchangeability. This concept refers to whether or not the manifest variables share the same concept [11, 25]. All manifest variables in a reflective model explain the same construct. So, removing an indicator from the block of variables should not have a significant effect on the construct. The situation is completely different when considering formative outer models. The indicators do not have to be interchangeable or share the same concept. That is what [2] called "sampling facets of a construct"; in other words manifest variables of a formative block of variables should represent all the aspects that form the concept. Finally, Gudergan *et al.* [17] recently proposed a procedure based on tetrad analysis to distinguish between a reflective and formative measurement model in a component-based approach. However, when an outer model has less than four observed variables, this procedure requires adding manifest variables from other outer models. Therefore, the discussion on the reflective and formative nature of the constructs studied here is based mainly on the five rules presented previously.

4. PATENT VALUE MODELS

Two models were tested. First of all, we are interested in knowing the relationships among patent indicators, patent value, and different constructs which up to now have been studied and identified as patent value determinants[2]. In previous research, these constructs have not been modeled as unobservable variables, such as in a structural equation model approach. So, the model formulation began by defining the patent value as an endogenous latent variable, since it is the primary variable to be estimated in the model. Summarizing the results of previous researchers, three unobserved variables related to the dependent variable were identified as exogenous: the knowledge stock of the patent, the technological scope of the invention, and the international scope of the protection (see Figure 1). We took into account all of the measurement variables found in the state of the art, and which can be computed from information contained in the patent document. Nevertheless, indicators constructed from the patent text, such as from the abstract or technical description, are excluded from this study.

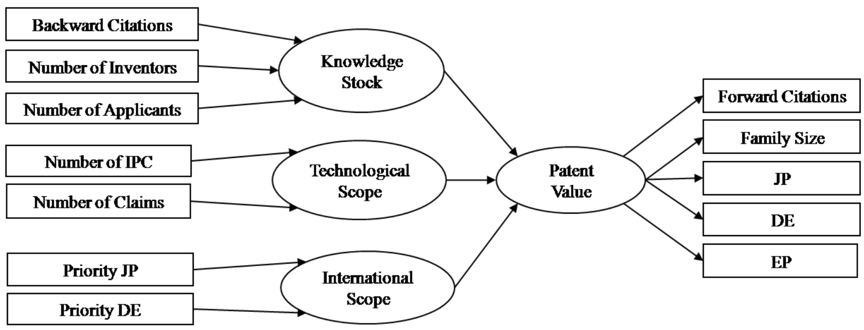

Figure 1: First-order model of patent value; patent value is an endogenous latent variable; knowledge stock, technological scope, and international scope are formative exogenous constructs.

The knowledge stock represents the base of knowledge that was used by the applicant to create an invention. This would be the content domain. This existing knowledge encourages the inventive activity and may come from within or outside the company. We would like to find those indicators that are value determinants, and that companies may use to make decisions. Since we are considering the patent document as the main data source, the applicants and inventors — that have contributed their knowledge to the creation of the invention — may be considered as forming this construct. The same applies to the backward citations.

[2]It is worth noting that we are not interested in explaining the variance and covariance among manifest variables as in a covariance-based approach, at least not at this stage.

The previous works, cited in the patent document, are the scientific and technical knowledge units that must exist before the creation of an invention, and they may be used as knowledge inputs within the invention process. Moreover, backward citations represent the prior art, and demonstrate that the invention had not been protected before. These three indicators have been related to the patent value for other authors (see for instance [38]). However, they still have not been used to estimate an unobserved variable as they are in a structural equation model.

From a theoretical standpoint, the knowledge stock is an exogenous latent variable, and affects the value of a patent. Keeping in mind the backward citations, it seems reasonable to think that an invention that is protected in an area where a lot of inventions are applied — hence with a large knowledge stock — will have less value than a potential radical innovation or a breakthrough invention, and therefore having a smaller knowledge stock. The number of inventors and applicants are revealed first in time, and cause a change on the knowledge stock, and not vice-versa. Additionally, it is not difficult to see that there is no covariance among backward citations, and the number of inventors and applicants. For instance, a patent may contain a large number of references, but the invention may be created only by one inventor or by one applicant. So, a reflective approach would fail to meet the unidimensionality condition. For this construct, however, multicollinearity would not be a problem. Hence, a formative mode is suitable for modelling the relationship between the indicators and the knowledge stock.

The technological scope of the invention is related to the potential utility of an invention in some technological fields. So, the manifest variables for this construct are the number of four-digit IPC classes where the patent is classified, and the number of claims of the patent. The IPC classes allow us to know the technical fields related to the invention, and therefore the number of potential application fields. This does not mean that an invention ultimate use is restricted to a determined area. A company may protect an invention for strategic purposes, for example to prevent its being used by a competitor. Here, the underlying issue is that the larger the number of classification codes, the larger the number of potential application fields, and hence, the greater the technological scope of the patent. On the other hand, and according to Tong and Frame [42, p. 134], "each claim represents a distinct inventive contribution, so patents are, in effect, bundles of inventions". Claims are a description of what the inventors actually claim to have invented and describe the potential application of the invention. As seen in the literature review, the number of claims should reflect the inventive activity of the invention. So, under the assumption that a highly sophisticated invention will require much inventiveness, the patent will also have a considerable amount of claims. Thus, this variable will also give information about the technological scope of the patents. It is arguable that this is not always so. Probably there are sophisticated inventions that have not required a large number of claims to be protected. But this may be unusual in the renewable energy field. As seen in Table 1 below, the number of claims is a skewed variable (skewness = 4.29, kurtosis = 43.65), with median 14. Following the rules presented before to distinguish between formative and reflective outer models, in this case, the manifest variables are revealed first, and cause a change in the technological scope of the inventions. When defining the manifest variables determining the technological scope. Probably, inventors have an idea of the applicability of the invention long before the time of protecting it. But, it is the patent value, therefore the protected invention, that is being analyzed here. So, a formative relationship is modeled between the indicators and the constructs. Additionally, as with the knowledge stock, there is no collinearity among manifest variables, and the block of variables is not one-dimensional.

The international scope refers to the geographic zones where the invention is protected. Inventions are usually protected in the local country first and then in others, as part of the companies' patenting strategy. All the patents considered in the sample are granted in the U.S. So, we defined two dummy variables that consider whether the invention had been protected in Japan (priority JP) or in Germany (priority DE) during the priority period. Japan and Germany are large producers of renewable energy technologies. Hence, it is interesting to examine whether these variables affect the patent value. Variables indicating whether inventions have been protected through the European Patent Office (EPO) or by the World Intellectual Property Organization (WIPO) have been excluded from the analysis because they provide little information. This means that for the international scope, not all the variables that could form the construct are being considered. So, higher disturbance terms are expected in this case. The international scope is clearly caused by the manifest variables. Here, again there is no collinearity among manifest variables, the block of variables is not one-dimensional. Therefore, formative relationships are considered in this block of variables.

On the other hand, the importance of a patent for future technological developments will be reflected in the number of times that the patent is cited, since the patent is useful for the development of other technologies [18], and in the patenting strategy pursued by the company over time. The latter is measured by taking into account the size of the patent family or the number of countries where the protection is sought. For the block of variables of patent value, a reflective relationship is considered between manifest and latent variables. As in this case all the indicators should explain the same construct (aside from the variables that have traditionally been used to infer the patent value), dummy variables are defined by considering whether the patent has been protected in Japan (JP), Germany (DE) or through the European Patent Office (EP). So, in this research, the first analyzed case is a first-order model composed by four constructs: knowledge stock, technological scope, international scope, and patent value (Figure 1).

It is worth noting that the first three constructs — knowledge stock, technological scope, and international scope — give an *a priori* value of patents. Thus, the intrinsic characteristics of the patent at the time of its application, along with the patenting strategy of the company in the priority period, may give a preliminary idea of patent value. In contrast, patent value estimated through forward citations and family size gives an *a posteriori* value for patents. This value (recognized value) is obtained over time and is given by others through the number of times that the patent is cited and the number of countries where the protection is sought. Estimating the patent value only through these manifest variables seems too ambitious. Rather, it is reasonable to think that the patent value is jointly given by those variables that determine the *a priori* and the *a posteriori* patent value. Using this approach, the influence of the *a posteriori* relative to the *a priori* patent value may also be assessed. Hence, the indicators that were initially related to the patent value are also associated with a fifth underlying latent variable related to the potential usefulness of the patent. The more useful a patent is, the more it is cited by others and the more important it is to the company's patenting strategy. We call this latent variable "technological usefulness". From a methodological standpoint, this means that the patent value is not directly related to a block of observed variables. So, this construct is regarded as a second-order latent variable that is influenced by all of the other constructs in a second-order model. The proposed model is shown in Figure 2. We explore the veracity of the assumptions with PLS path modelling.

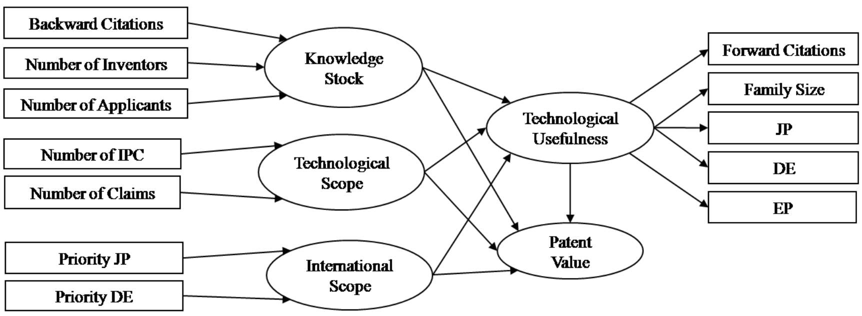

Figure 2: Hierarchical component model of patent value; patent value is an endogenous second-order latent variable; technological usefulness is a reflective endogenous latent variable; knowledge stock, technological scope, and international scope are formative exogenous constructs.

5. PATENT DATA

Renewable energy patents include wind, solar, geothermal, wave / tide, biomass, and waste energy. To select suitable patent data, we use the IPC classes for renewable energies listed by Johnstone *et al.* [26]. The sample comprises a total of 2,901 patents (sample 1), published in 1990–1991, 1995–1996, 1999–2000 and 2005–2006, and granted in the U.S. (source: Delphion database). We retrieved these data, and the indicators described above were computed. The number of claims was collected manually for each patent.

Table 2 provides descriptive statistics for patent indicators. The results indicate that some variables are very heterogeneous and asymmetric, and they also exhibit large variance. So, normality is not a good assumption. Positive values of skewness indicate positive/right skew (notice how the medians are always smaller than the means). Likewise, positive kurtosis indexes show distributions that are sharper than the normal peak.

Table 2: Descriptive statistics of patent data.

Manifest Variable	Mean	Standard Deviation	Minimum	Mediam	Maximum	Skewness	Kurtosis
Number of applicants	1.04	0.29	1	1	9	12.85	260.81
Number of inventors	2.21	1.58	1	2	14	1.76	4.23
Backward citations	15.36	18.97	0	11	327	5.54	50.79
Number of IPC	6.28	4.52	1	5	48	2.09	7.71
Number of claims	17.02	15.08	1	14	279	4.29	43.65
Priority JP	0.19	0.39	0	0	1	1.54	0.37
Priority DE	0.08	0.27	0	0	1	3.09	7.55
Forward citations	5.63	10.16	0	2	158	5.3	46.83
Family size	8.53	11.62	1	6	202	5.58	51.27
Dummy JP	0.44	0.49	0	0	1	0.23	−1.95
Dummy DE	0.32	0.46	0	0	1	0.75	−1.44
Dummy EP	0.43	0.49	0	0	1	0.25	−1.94

Additionally, the priority countries of these patents are U.S. (59%), Japan (19%), Germany (9%), Great Britain (2%), France (1%) and so on. Patents belong to 1,581 applicants. Patents have been granted to companies (69%), individuals (25%) and universities, research centers or governmental institutions (6%). Due to the manner in which the sample was selected, the sample is homogenous in terms of technological area and the country where the patents were granted. However, the sample is heterogeneous in terms of the type of applicant or the industry in which the companies are classified, and this heterogeneity could affect the results. This also means that there are companies belonging to different industries that are interested in developing renewable energy innovations. At any rate, it is worth noting that at this stage, the patent value model is being tested in general at the level of renewable energy technologies. We estimate the model using the total sample (2,901 patents, sample 1). However, providing that time is an important factor that may affect the findings, three additional samples were taken. Patent indicator matrices were selected in the following application years: 1990–1991 ($N = 129$, sample 2), 1995–1996 ($N = 128$, sample 3) and 1999–2000 ($N = 536$, sample 4). So, in order to analyze whether it is possible to find a pattern in the parameter estimates, the proposed models were estimated with all data, and with time-period data (notice that cases are different in each time-period).

6. RESULTS

The internal consistency of reflective outer models, technological usefulness and patent value was assessed by using Cronbach's alpha and composite reliability. For the first-order, the Cronbach's alpha coefficients for patent value are 0.68, 0.79, 0.76 and 0.68 for samples 1, 2, 3 and 4, respectively. Moreover, composite reliability coefficients are 0.77, 0.85, 0.84 and 0.79 for each sample, respectively. So, the patent value is unidimensional. AVE scores are 0.48, 0.56, 0.54 and 0.48 for patent value and for samples 1, 2, 3 and 4, respectively. So, the constructs capture on average more than 50% of the variance in relation to the amount of variance due to measurement error. In the second-order model, technological usefulness has the same Cronbach's alpha and composite reliability coefficients that patent value has in the first-order model. Cronbach's alpha coefficients for the patent value are 0.59, 0.68, 0.7 and 0.58 for samples 1, 2, 3 and 4, respectively. Composite reliability coefficients are 0.72, 0.76, 0.79 and 0.71 for each sample, respectively. Therefore, both technological usefulness and patent value are unidimensional. The technological usefulness captures on average a 54% of the variance in relation to the amount of variance due to measurement error (see the AVE scores for patent value in the first-order model). However, AVE scores for patent value (second-order latent variable) are quite different, 0.24, 0.29, 0.3 and 0.22 for samples 1, 2, 3 and 4, respectively. So, this block of variables is unidimensional, and the latent variable captures on average a 26% of the variance in relation to the amount of variance due to measurement error. This low percentage may be because reflective and formative indicators have been repeated for the second-order latent variable.

Table 3 reports the cross loadings for the reflective block of variables in the second-order model of patent value in the three analyzed time-periods. Forward citations, family size and dummy variables JP, DE and EP are slightly more correlated in the three time-periods, with the technological usefulness of the patents rather than the patent value itself. In regards to other indicators, quite the opposite happens: the correlation between indicators and patent value are always higher than the correlation between indicators and technological usefulness. This is adequate even though patent value indicators are used as auxiliary variables in order to estimate the model. It is worth noting that cross loadings of some variables are very similar over time, suggesting a pattern. This phenomenon is interesting because it indicates that the number of inventors; the number of IPC classes; dummy variables JP, DE and EP; forward citations and family size are strongly and constantly correlated with the patent value and its technological usefulness throughout time. This empirical evidence supports the relationships between latent and manifest variables as proposed in the models.

Table 3: Cross loadings between indicators for reflective block of variables.

Manifest Variable	1990–1991		1995–1996		1999–2000	
	Patent Value	Technological Usefulness	Patent Value	Technological Usefulness	Patent Value	Technological Usefulness
Number of inventors	0.572	0.279	0.611	0.424	0.492	0.135
Backward citations	0.064	0.129	0.092	0.067	0.141	0.091
Number of IPC	0.587	0.387	0.465	0.357	0.495	0.228
Number of claims	−0.074	−0.027	0.403	0.257	0.131	0.048
Dummy priority JP	0.527	0.258	0.391	0.253	0.414	0.162
Dummy priority DE	0.205	0.127	0.103	0.127	0.154	0.136
Forward citations	0.229	0.292	0.295	0.29	0.085	0.085
Family size	0.775	0.894	0.741	0.825	0.714	0.859
Dummy JP	0.816	0.836	0.818	0.833	0.727	0.774
Dummy DE	0.692	0.775	0.754	0.808	0.559	0.681
Dummy EP	0.666	0.818	0.739	0.799	0.658	0.809

Tables 4 and 5 present the standardized loadings and weights by PLS estimation and t-values by bootstrapping for the first- and second-order models, respectively. Loadings and weights reveal the strength of the relationship between manifest and latent variables. The number of inventors, the number of IPC classes and the dummy priority variables JP and DE are strongly and significantly related to their constructs in all cases in the first- and in the second-order models. Some authors [5, 7, 44] have studied the performance of the PLS path modelling algorithm using Monte Carlo simulations. Among others, the factors analyzed have been the sample size and the number of manifest variables per latent variable. In general, researchers agree and recommend having at least three indicators per construct. However, only Chin *et al.* [8] considered in their study the case of two observed variables per latent variables in their study of interaction effects with reflective outer models. However, as a result of their simulation study, Vilares *et al.* [44, p. 13] reported that "PLS always produces good estimates for perceived value loadings [a latent variable with two indicators, the author]. This is an interesting result, since PLS is presented as being 'consistent at large' ...". In the formative outer models analyzed here, there are few indicators available per construct. However, the magnitudes of the weights are large enough to infer that there may be a formative relationship between indicators and constructs. Additionally, these results suggest that the patent value and the technological usefulness are evident since the patent is applied. Therefore, the value can be assessed at an early stage. The number of claims shows a weaker association with the technological scope than the number of IPC classes. Perhaps this indicator is more related to the "quality" of the invention, not in the sense of how inventions have an impact on different technological fields (scope) but rather on how important this impact is in a given technological field. Regarding the international scope, this variable seems to be formed by its indicators. The manifest variables are statistically significant in all cases in the two analyzed models. So, this could mean that in the renewable energy field, besides protecting the invention in the U.S., it is important as a value determinant for early protection of the inventions which originate in the other two largest producers of these technologies: Japan and Germany.

Table 4: Standardized loadings and weights for outer models for the first-order model of the patent value, t-values in parenthesis, * at the 0.01 significance level, ** at the 0.05 significance level.

Construct	Indicator	Sample 1	1990–1991	1995–1996	1999–2000
Knowledge stock	Backward citations	0.541* (1.860)	0.420* (1.688)	0.128 (0.791)	0.499* (1.670)
	Number of inventors	0.807** (3.054)	0.920** (4.937)	0.988** (9.086)	0.872** (2.794)
Technological scope	Number of IPC	0.966** (5.935)	0.997** (13.746)	0.803** (5.455)	0.985** (4.502)
	Number of claims	0.176 (0.756)	−0.058 (0.364)	0.529 (1.432)	0.103** (0.354)
International scope	Priority JP	0.802** (3.662)	0.909** (5.492)	0.904** (7.844)	0.847** (3.630)
	Priority DE	0.725** (2.814)	0.512** (2.043)	0.502** (2.479)	0.660** (2.422)
Patent value	Forward citations	−0.108 (0.940)	0.274** (2.041)	0.299* (1.693)	0.096 (0.524)
	Family size	0.840** (9.464)	0.893** (36.017)	0.813** (15.126)	0.845** (5.297)
	Dummy JP	0.777** (6.593)	0.843** (19.572)	0.841** (21.277)	0.802** (4.549)
	Dummy DE	0.690** (5.530)	0.777** (11.126)	0.811** (18.389)	0.671** (4.087)
	Dummy EP	0.780** (7.921)	0.808** (11.975)	0.794** (12.513)	0.786** (5.272)

On the other hand, patent value and technological usefulness are always strongly and significantly reflected in their explanatory variables. Forward citations, patent family and dummy variables constantly reflect patent value in the first-order model and technological usefulness in the second-order model. The forward citations are not significant in the models evaluated in 1999–2000. But, this may be due to the fact that in recent years patents have been cited less, and the variable is less informative than in previous years. Moreover, loadings for the relationship between forward citations and technological usefulness are smaller

Table 5: Standardized loadings and weights for outer models for the second-order model of the patent value, *t*-values in parenthesis, * at the 0.01 significance level, ** at the 0.05 significance level.

Construct	Indicator	Sample 1	1990–1991	1995–1996	1999–2000
Knowledge stock	Backward citations	0.439 (1.619)	0.248 (1.103)	0.122 (0.991)	0.357 (1.114)
	Number of inventors	0.871** (3.828)	0.976** (8.060)	0.989** (24.728)	0.938** (3.214)
Technological scope	Number of IPC	0.952** (6.544)	0.995** (18.078)	0.761** (4.633)	0.974** (4.140)
	Number of claims	0.220 (1.028)	−0.078 (0.546)	0.584** (3.139)	0.150 (0.516)
International scope	Priority JP	0.867** (4.090)	0.931** (10.601)	0.947** (7.863)	0.915** (4.096)
	Priority DE	0.639** (2.422)	0.465** (2.709)	0.401* (1.701)	0.548* (1.943)
Technological usefulness	Forward citations	0.762** (6.833)	0.836** (22.739)	0.834** (24.167)	0.774** (5.177)
	Family size	0.795** (10.667)	0.818** (11.800)	0.799** (18.126)	0.809** (11.499)
	Dummy JP	0.705** (7.983)	0.775** (11.891)	0.809** (18.318)	0.681** (6.256)
	Dummy DE	−0.052 (0.488)	0.292** (2.280)	0.290** (2.190)	0.085 (0.616)
	Dummy EP	0.853** (13.577)	0.894** (36.226)	0.825** (21.104)	0.859** (11.526)
Patent value	Backward citations	0.232 (1.511)	0.064 (0.564)	0.092 (1.005)	0.141 (0.735)
	Number of inventors	0.476** (3.477)	0.572** (5.964)	0.611** (8.825)	0.492** (3.016)
	Number of IPC	0.549** (5.909)	0.587** (7.837)	0.465** (4.820)	0.495** (3.420)
	Number of claims	0.185 (1.296)	−0.074 (0.748)	0.403** (3.193)	0.131 (0.810)
	Priority JP	0.387** (2.723)	0.527** (5.466)	0.391** (3.604)	0.414** (2.461)
	Priority DE	0.202** (5.318)	0.205** (2.262)	0.103 (1.269)	0.154 (1.191)
	Forward citations	−0.085 (0.861)	0.229* (1.944)	0.295** (2.453)	0.085 (0.659)
	Family size	0.730** (8.250)	0.775** (15.351)	0.741** (11.612)	0.714** (5.952)
	Dummy JP	0.711** (6.083)	0.816** (20.264)	0.818** (18.295)	0.727** (4.349)
	Dummy DE	0.586** (5.318)	0.692** (8.318)	0.754** (11.977)	0.559** (4.349)
	Dummy EP	0.672** (7.196)	0.666** (6.650)	0.739** (13.752)	0.658** (6.341)

than, for instance, loadings for the relationship between family size and technological usefulness. These results may mean that the longitudinal nature of this variable — citations that are received throughout the time — is an important factor that should be taken into account when considering this indicator in the models. The quality of each outer model is measured through the communality index, i.e. the proportion of variance in the measurement variables accounted for by the latent variable. For the second-order model, communality indexes for patent value are 0.29, 0.30 and 0.22 for the 1990–1991, 1995–1996 and 1999–2000 models, respectively. Therefore, indicators have approximately 30% of the variance in common with its latent variable. As seen above, this low percentage may be because reflective and formative indicators have been repeated for the second-order latent variable. The communality indexes for technological usefulness are 0.57, 0.55 and 0.49 for each time-period, also giving evidence of an important percentage of shared variance.

Tables 6 and 7 show the findings for the inner relationships (standardized beta coefficients, significance levels and coefficients of determination) for the first- and second-order models respectively. Path coefficient of knowledge stock, technological scope and international scope as related to patent value are significant at 0.01 levels in almost all cases. Therefore, the patent value may be formed by constructs estimated from reliable patent indicators. The first-order model allows us to obtain an estimate of the patent value "in time equal to zero". As showed in the second-order model, the knowledge stock, the technological scope and the international scope are also related to technological usefulness. Moreover, technological usefulness and patent value are significantly related, indicating how the former is an important variable in the prediction of the latter. The second-order model allows us to obtain the patent value as the sum of the value in time equal to zero, and the value given by others, that is the technological usefulness.

Table 6: Standardized path coefficients for the first-order model of patent value, t-values in parenthesis, * at the 0.01 significance level, ** at the 0.05 significance level.

Latent Variable	**Sample 1**	**1990–1991**	**1995–1996**	**1999–2000**
Knowledge stock to Patent value	0.115 (1.248)	0.202* (1.987)	0.306** (2.263)	0.091 (1.040)
Technological scope to Patent value	0.238** (2.892)	0.314** (4.221)	0.335** (3.084)	0.200** (2.278)
International scope to Patent value	0.243** (3.199)	0.154* (1.998)	0.251** (3.044)	0.220** (2.420)
R^2 of patent value	0.161	0.234	0.35	0.114

Table 7: Standardized path coefficients for the second-order model of patent value, t-values in parenthesis, * at the 0.01 significance level, ** at the 0.05 significance level.

Latent Variable	Sample 1	1990–1991	1995–1996	1999–2000
Knowledge stock to Patent value	0.280** (9.979)	0.226** (9.510)	0.229** (12.349)	0.293** (8.281)
Technological scope to Patent value	0.278** (8.811)	0.227** (10.737)	0.226** (8.870)	0.271** (7.620)
International scope to Patent value	0.212** (5.505)	0.232** (11.314)	0.166** (7.659)	0.236** (5.160)
Knowledge stock to Technological usefulness	0.104 (1.162)	0.180* (1.752)	0.299** (3.771)	0.072 (0.783)
Technological scope to Technological usefulness	0.237** (2.686)	0.315** (3.290)	0.334** (3.387)	0.207** (2.133)
International scope to Technological usefulness	0.225** (2.486)	0.142 (1.376)	0.236** (3.042)	0.200** (2.252)
Technological usefulness to Patent value	0.683** (14.511)	0.668** (16.951)	0.697** (20.558)	0.698** (11.207)
R^2 of patent value	0.998	0.998	0.999	0.997
R^2 of usefulness	0.148	0.219	0.338	0.103

The determination coefficient for patent value is 0.9 in the second-order models, i.e. the model fit the data in an acceptable way. This result is not surprising; it confirms the aforementioned findings and indicates how the data is better explained by second-order models as compared with first-order models. However, we must consider this result carefully, because the patent value is estimated considering all the measurement variables of the models. Another explanation for this is that in the second-order models, the contribution of the recognized value of patents (technological usefulness) is considered, and this would help fit the data better. Unlike patent value, technological usefulness has a moderate coefficient of determination. Perhaps other indicators should help to better explain the model, or again the longitudinal nature of the forward citations is an important factor to be considered. However, we think that the results are acceptable, taking into account the literature review and the goodness of fit obtained using other models in the analysis of patent data. It is worth noting that the structural relationships are significant.

7. FINAL REMARKS

This research relates manifest variables that come from information contained in the patent document with latent variables into a single replicable model. The magnitude of this relationship and the importance of each construct are known, including the influence of knowledge stock, the technological and international scope in the value of the technology. In the first-order model, the variables that most affect the patent value are the technological and the international scope. In the second-order model, the technological usefulness is also important.

A distinction between two patent values can be made: an *a priori* and intrinsic value, which the patent has at the moment of its application (the potential value of the patent); and an *a posteriori* value that the patent acquires over time through the actions of a company or others (the value that is recognized). The potential value depends on the characteristics of the patent at the time of application -such as the patenting strategy of a company, the technological applicability of the patents in different technological fields and the base of knowledge that is necessary for the creation of a new invention. As time passes, the patent potentiality is recognized and reflected in the number of times that it is cited and in the number of countries where it is protected. This recognition is a reflection of its technological usefulness. Even though companies can assess the importance or impact of their inventions, these results and the procedure for obtaining them are becoming a tool for improving the strategy of developing new products and inventions, improving intellectual property policy and for comparing technologies with other competitors. The stability of results over time augur that this may be possible.

In order to assess companies' patent portfolios using a model that can be replicated, a follow-up to this research will study patent value evolution as well as the market-patent relationship and its implications. Furthermore, there are other indicators related to patent value that have been previously studied, but they cannot be computed from the information contained in the patent documents, such as the number of renewals and the number of opposition cases. Nevertheless, these variables could be related to another latent variable in the model, or be a reflection of the technological usefulness of an invention. Finally, PLS path modelling has proven to be a suitable approach for analyzing patent data.

ACKNOWLEDGMENTS

This research was supported by a grant from the Universidad Católica de la Ssma. Concepción and the Comisión Nacional de Investigación Científica y Tecnológica (CONICYT), Chile. We are grateful to the participants in the ISSI conference, Madrid, June 2007, the DRUID conferences, Aalborg, January 2008, and Copenhagen, June 2008, and the COMPSTAT conference, Porto, August 2008, for comments.

REFERENCES

[1] Albert, M.; Avery, D.; Narin, F. and McAllister, P. (1991). Direct validation of citation counts as indicators of industrially important patents, *Research Policy*, **20**(3), 251–259.

[2] Bollen, K. and Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective, *Psychological Bulletin*, **110**(2), 305–314.

[3] Bollen, K. and Ting, K.F. (2000). A tetrad test for causal indicators, *Psychological Methods*, **5**(1), 3–22.

[4] Carpenter, M.; Narin, F. and Woolf, P. (2005). Citation rates to technologically important patents, *World Patent Information*, **3**(4), 160–163.

[5] Cassel, C.; Hackl, P. and Westlund, A. (1999). Robustness of partial least-squares method for estimating latent variable quality structures, *Journal of Applied Statistics*, **26**(4), 435–446.

[6] Chin, W. (1998). *The partial least squares approach for structural equation modelling*. In "Modern Methods for Business Research" (G. Marcoulides, Ed.), Lawrence Erlbaum Associates, 295–336.

[7] Chin, W.W. and Newsted, P.R. (1999). *Structural equation modelling analysis with small samples using partial least squares*. In "Statistical strategies for small sample research" (R.H. Hoyle, Ed.), Sage Publications, California, 307–341.

[8] Chin, W.W.; Marcolin, B.L. and Newsted, P.R. (2003). A partial least squares latent variable modelling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study, *Information Systems Research*, **14**(2), 189–217.

[9] Cohen, P.; Cohen, J.; Teresi, J.; Marchi, M. and Noemi, C. (1990). Problems in the measurement of latent variables in structural equations causal models, *Applied Psychological Measurement*, **14**(2), 183–196.

[10] Connolly, R. and Hirschey, M. (1988). Market value and patents: A Bayesian approach, *Economics Letters*, **27**(1), 83–87.

[11] Diamantopoulos, A. and Winklhofer, H. (2001). Index construction with formative indicators: An alternative to scale development, *Journal of Marketing Research*, **38**, 269–277.

[12] Esposito, V. (2007). *The PLS approach to data exploration and modelling: An everlasting matter of dispute or a playground for integrating different cultures?*. In "PLS and Related Methods: Proceedings of the PLS'07 International Symposium" (H. Martens, T. Naes and M. Martens, Eds.), Matforsk, As, Norway, 8–11.

[13] Fornell, C. and Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error, *Journal of Marketing Research*, **18**(1), 39–50.

[14] Fornell, C. and Bookstein, F.L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory, *Journal of Marketing Research*, **19**(4), 440.

[15] Griliches, Z. (1981). Market value, R-and-D, and patents, *Economics Letters*, **7**(2), 183–187.

[16] Griliches, Z. (1981). Patent statistics as economic indicators – A survey, *Journal of Economic Literature*, **28**(4), 1661–1707.

[17] Gudergan, S.P.; Ringle, C.M.; Wende, S. and Will, A. (2008). Confirmatory tetrad analysis in PLS path modelling, *Journal of Business Research*, **61**(12), 1238–1249.

[18] Guellec, D. and van Pottelsbergue, B. (2000). Applications, grants and the value of patent, *Economics Letters*, **69**(1), 109–114.

[19] Hall, B.; Jaffe, A. and Trajtenberg, M. (2005). Market value and patent citations, *The RAND Journal of Economics*, **36**(1), 16–38.

[20] Harhoff, D.; Narin, F.; Scherer, F. and Vopel, K. (1999). Citation frequency and the value of patented inventions, *Review of Economics and Statistics*, **81**(3), 511–515.

[21] Harhoff, D.; Scherer, F. and Vopel, K. (2003). Citations, family size, opposition and the value of patent rights, *Research Policy*, **32**(8), 1343–1363.

[22] Harhoff, D. and Reitzig, M. (2004). Determinants of opposition against EPO patent grants – The case of biotechnology and pharmaceuticals, *International Journal of Industrial Organization*, **22**(4), 443–480.

[23] Henseler, J.; Ringle, C.M. and Sinkovics, R. (2009). The use of partial least square path modelling in international marketing, *Advances in International Marketing*, **20**, 277–319.

[24] Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research, a review of four recent studies, *Strategic Management Journal*, **20**, 195–204.

[25] Jarvis, C.B.; Mackenzie, S.B.; Podsakoff, P.M.; Mick, D.G. and Bearden, W.O. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research, *Journal of Consumer Research*, **30**(2), 199.

[26] Johnstone, N.; Hascic, I.; Clavel, L. and Marical, F. (2007). Renewable energy policies and technological innovation: Empirical evidence based on patent counts, *Empirical Policy Analysis Unit OECD Environmental Directorate*.

[27] Lanjouw, J.; Pakes, A. and Putnam, J. (1998). How to count patents and value intellectual property: The uses of patent renewal and application data, *The Journal of Industrial Economics*, **46**(4), 405–432.

[28] Lanjouw, J. and Schankerman, M. (2001). Characteristics of patent litigation: a window on competition, *The RAND Journal of Economics*, **32**(1), 129–151.

[29] Lanjouw, J. and Schankerman, M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators, *Economic Journal*, **114**(495), 441–465.

[30] Lohmöller, J. (1989). *Latent Variables Path Modelling with Partial Least Squares*, Physical-Verlag, Heidelberg.

[31] Lerner, J. (1994). The Importance of patent scope – An empirical analysis, *The RAND Journal of Economics*, **25**(2), 319–333.

[32] MacCallun, R. and Browne, M. (1993). The use of causal indicators in covariance structure models: Some practical issues, *Psychological Bulletin*, **114**(3), 533–541.

[33] Martensen, A. and Grøhnholdt, L. (2003). Improving library users' perceived quality, satisfaction and loyalty: An integrated measurement and management system, *The Journal of Academic Librarianship*, **29**(3), 140–147.

[34] McDonald, R.P. (1996). Path analysis with composite variables, *Multivariate Behavioral Research*, **31**(2), 239–270.

[35] Narin, F.; Hamilton, K. and Olivastro, D. (1997). The increasing linkage between US technology and public science, *Research Policy*, **26**(3), 317–330.

[36] Petter, S.; Straub, D. and Rai, A. (2007). Specifying formative constructs in information systems research, *MIS Quarterly*, **31**(4), 623–656.

[37] Reitzig, M. (2003). What determines patent value? Insights from the semiconductor industry, *Research Policy*, **32**(1), 13–26.

[38] Reitzig, M. (2004). Improving patent valuations for management purposes – Validating new indicators by analyzing application rationales, *Research Policy*, **33**(6-7), 939–957.

[39] Ringle, C.; Wende, S. and Will, A. (2005). SmartPLS 2.0 (beta), `www.smartpls.de`.

[40] Sherry, E. and Teece, D. (2004). Royalties, evolving patent rights, and the value of innovation, *Research Policy*, **33**(2), 179–191.

[41] Tenenhaus, M.; Esposito, V.; Chatelin, Y. and Lauro, C. (2005). PLS path modelling, *Computational Statistics & Data Analysis*, **48**, 159–205.

[42] Tong, X. and Frame, J. (1994). Measuring national technological performance with patent claims data, *Research Policy*, **23**(2), 133–141.

[43] Trajtenberg, M. (1990). A penny for your quotes – Patent citations and the value of innovations, *The RAND Journal of Economics*, **21**(1), 172–187.

[44] Vilares, M.; Almeida, M. and Coelho, P. (2009). *Comparison of likelihood and PLS estimators for structural equation modelling: A simulation with customer satisfaction data*. In "Handbook of partial least squares: Concepts, methods and applications" (V. Esposito Vinzi, W.W. Chin, J. Henseler and H. Wang, Eds.), Springer, Berlin, in print.

[45] Wold, H. (1980). *Model construction and evaluation when theoretical knowledge is scarce – Theory and application of partial least square*. In "Evaluation of Econometric Models" (J. Kmenta and J.G. Ramsey, Eds.), Academic Press, New York, 47–74.

[46] Wold, H. (1982). *Soft modelling: The basic design and some extensions*. In "Systems under indirect observation, part II" (K.G. Jöreskog and H. Wold, Eds.), North-Holland, Amsterdam, 1–54.

[47] Wold, H. (1985). Partial least square, *Encyclopedia of Statistical Sciences*, **6**, 581–591.