Appendix A

Historical Overview of PLS-PM

Statistical methods are not created instantly or magically. Sometimes we may get that false impression of "spontaneous generation" but obviously it is just a distorted view of the reality. Like any other technology and methodology, the family of PLS methods did not appear out of the blue. Herman Wold, the father of PLS, didn't sit down one day and said: "Oh, I'm so bored today. I need to do something, but what? Oh, I know!, let's use least squares regressions to develop an algorithm for calculating principal components".

My humble opinion is that you cannot fully understand the PLS methods without taking into consideration their history and development. It is true that you can apply the PLS methods and live happily without knowing their origins. But if you are like me, that is not enough when you want to completely understand not only the nuts and bolts behind PLS but also the motivations and context under which the PLS toolkit has been developed.

The truth is that statistical textbooks don't usually talk about the historical framework and developments that give rise to new methods. Admittedly, textbook authors are mathematicians, statisticians, analysts, but not historians. I am not a historian either, but I've spent a decent amount of time researching the history of PLS. The content of this appendix is the result of that quest. I'm not pretending to cover all the events and every single historical aspect behind PLS-PM. Instead, my goal is to provide enough material so we can have the "big PLS picture".

1 Preliminaries

The family of PLS methods has its roots in the amazing work undertaken by the Swedish statistician Herman Wold and his research group. Over a long period of time, various aspects of the PLS methods were extracted from other techniques and then refined and adapted to tackle different types of data analysis tasks. As a matter of fact, Wold and his team spent about two decades, from middle 1960s to early 1980s, developing and improving a series of algorithms based on Ordinary Least Squares regressions to handle several modeling problems. Likewise, different names for these methods were proposed over the years such as *Nonlinear Iterative Least Squares* (NILES), *Nonlinear Iterative Partial Least Squares* (NIPALS), *Partial Least Squares basic design*, and *PLS soft modeling*, among others.

This first period of evolution comprised a variety of analysis and modeling problems with data from social sciences (e.g. economics, psychology, sociology) that were tackled with different versions of least-squares-based iterative algorithms. Then in the early 1980s, Herman's son, Svante Wold, took over the development of the PLS principles and shifted the applications to chemistry and the food industry. This change of the PLS methods was focused on regression problems and soon after the PLS regression movement became firmly established.

Overall, the road traveled by the PLS methods along its development process is an intertwined path that has resulted nowadays in two main different schools of PLS analysis: 1) PLS regression models (PLS-R) and 2) PLS path modeling (PLS-PM). Within the field of chemometrics and related disciplines, the PLS acronym is synonym of PLS regression techniques. In contrast, within the fields of marketing research, information technology and other social science disciplines, the term PLS is synonym of PLS path modeling. Moreover, the contributions of several disciplines to the PLS landscape have created a comprehensive literature in which is not unusual to find different terms referring to the same concept while at the same time, and probably worst, find identical names used to designate different notions.

As you can imagine, the ramification of the PLS principles into different branches and fields of application, the lack of uniformity in terminology, combined with the fact of having two main authors (father and son) with the same last name, are the perfect mix for a puzzling collage. These factors are the primary source of confusion and misunderstanding that many users and readers, both inside and outside the PLS communities, stumble upon when dealing with the family of PLS methods.

What is usually told

If you don't have access to a couple of references that talk about the history of PLS, the idea that you get from most articles and publications may lead you to get an incomplete picture of what happened. The condensed fairy tale version of the story that usually appears in a variety of texts follows this pattern:

> Once upon a time, a Swedish professor called Herman Wold was working on econometric models. Then he developed the preliminary idea of partial least squares in his paper of 1966 where the NILES procedure was introduced. This procedure was later named NIPALS in the early 1970s and it kept that name for several years. Then in the late 1970s, the PLS approach to path models with latent variables was formally introduced, although it was presented under the label

> of soft modeling. A modification of the PLS ideas were applied in chemometrics by Svante Wold, and the first published paper on PLS regression appeared in 1983. This PLS regression version became such a hit that its popularity and success eclipsed the other PLS methods ... And the PLS methods lived happily ever after.

Although this synopsis is not false, it is very far from being an accurate story. We can find a better story by checking what Herman Wold says in his memoirs:

What you probably didn't know

(extract from Wold, 1982):

> "Sociology in the early 1960s opened a new era in quantitative system analysis by O. T. Duncan's merger of the latent variables in psychology and the path modeling with manifest variables in econometrics ... Karl G. Joreskog in 1967 was the first to bring the computer to operative use for Maximum Likelihood (ML) estimation of factor models, and in 1970 Joreskog launched his LISREL algorithm for ML estimation of path models with latent variables. When seeing the LISREL algorithm I realized that principal components and canonical correlations can be interpreted as path models with one and two latent variables, respectively, and with each latent variable explicitly estimated as a weighted aggregate of its indicators; this gave me the incentive to extend the Fix-Point algorithm to general path models with latent variables indirectly observed by multiple indicators. There were two stumbling blocks in this endeavor: the step from two to three latent variables, and the step from one to two inner relations between the latent variables. Once these stumbling blocks were overcome it was easy to design general algorithms for iterative estimation of path models with latent variables. The ensuing algorithms I first called NIPALS, later PLS. An array of somewhat different PLS versions was designed; in late 1977 I arrived at what I regard as the end station: the basic design for PLS estimation of path models with latent variables, or briefly PLS soft modeling."

I am not expecting you to understand everything in the above paragraph. But I think it gives us enough hints of the origins of PLS. If you already know about the history of structural equation models, some of the events described by Wold should be familiar to you. To connect the dots and get a broader vision of the origins of PLS, we'll begin with a short biography of Herman Wold. Then, we will cover the antecedents of PLS, describe its evolution, and mention the current state in the PLS-PM arena.

2 Herman Wold

Let's begin with Herman Wold, the creator of PLS. Who was Herman Wold? The father of PLS was the great Swedish statistician and econometrician Herman Wold. In fact, his complete name was Herman Ole Andreas Wold, and he was not born in Sweden but in Norway, in the small town of Skien in december 25th, 1908. Herman was the youngest of a family with six brothers and sisters that moved to Sweden in 1912 near Stockholm where he spent the rest of his life. Herman began his college studies in 1927 in the University of Stockholm, and he took courses in physics, mathematics, and economics. After graduating in 1930, he started to work for an insurance company gathering data about fluvial precipitation in Sweden in order to design a fee for insurances against damages caused by rains. However, Herman didn't have the intention to stay in the actuarial field and he went back to the academia for studying a PhD under the guidance of Harald Cramér.

Herman received his PhD in 1938 and he was a professor of mathematical statistics and actuarial mathematics until 1942 when he was offered the chair of the Department of Statistics at the University of Uppsala. Indeed, he remained in that position until 1970 when he moved to the University of Gothenburg where he worked until his retirement in 1975.

In 1940 he married Anna-Lisa Arrhenius and they had three children, two girls and one boy (Svante Wold, born in 1941). After his retirement in 1975 he returned to Uppsala to live, but it was an active retirement; he received support for his continuing research from the Volkswagen Foundation and spent much time at the University of Geneva and the University of Uppsala as Professor Emeritus.

Wold was elected to Fellowship of the Institute of Mathematical Statistics in 1946, the American Statistical Association in 1951, Honorary Fellowship of the Royal Statistical Society London in 1961, and an Honorary Doctorate from the Technical Academy of Lisbon in 1965. He was also elected to membership of the Swedish Academy of Sciences in 1960, and served as a member of the Nobel Economic Science Prize Committee from 1968 to 1980. Wold was also active in international statistical econometric affairs. He served as Vice-President of the International Statistical Institute from 1957 to 1961, was elected to Presidency of the Econometric Society in 1966, and Honorary Membership of the American Economic Association and the American Academy of Arts and Sciences in 1978. Herman Wold died on february 16th, 1992.

Wold's Work Periods

The work of Herman Wold can be roughly divided in four main periods:

- (1932-1938) undergraduate studies and doctoral thesis on stationary time series
- (1938-1952) econometric analysis of consumer demand

- (1945-1970) econometrics and path models with directly observed variables
- (1966-1992) systems analysis and path models with indirectly observed variables

The most remarkable feature is that throughout these 60 years his major contributions were based on the least squares (LS) principle. One of the best testimonies of his involvement with LS can be found in the following excerpt from Wold's autobiographical notes (1982):

> "Since the 1930s the mainstream of contemporary statistics has been the Maximum Likelihood (ML) principle. ML is a highly developed body of joint probability distribution, usually a multivariate normal distribution. In the triumphant ML evolution, Least Squares (LS) was left in a backwater. LS being an autonomous optimization principle, it suddenly dawned upon me that the rationale of LS could be established on a distribution-free basis, by subjecting the predictive relations of the model to predictor specification (1959-1963). In the ensuing LS framework I initiated first the fix-point method for estimation of interdependent systems (1956, 1966) and then the partial least squares (PLS) method of path models with latent variables indirectly observed by multiple indicators (1971-72). The PLS approach, or briefly soft modeling, has kept me busy ever since".

2.1 The Econometrician

Herman Wold was an econometrician and a statistician. A lot of his work had to do with econometrics, and it is not surprise that the development of the PLS methods is related with some of the econometric projects that Herman Wold was working on.

If we were to consult any book about the history of econometrics, we would see that the period of emergence of this discipline began in the early twentieth century with two main events. One took the form of the development of **time-series analysis**. The other took the form of **demand analysis**. Actually, the first two work periods of Herman Wold coincided precisely with these two main research areas in econometrics. While his graduate studies were related to time-series analysis, his postdoctoral activities (from 1938-1952) were focused on demand analysis and econometric model building. In both topics, and for the rest of his professional career, his research was based on the Least Squares principle for estimation and modeling purposes.

You have to understand that Wold got involved in econometrics when that field wasn't even named like that. Today we find entire bookshelfs dedicated to econometrics in any economics section of a college library. But back in the days when Herman was working in that area, there were no such defined notions of econometrics as today. The term was coined by Ragnar Frisch in 1926, but it didn't catch the attention until years later.

Time Series Studies (1932-1938)

What Herman Wold studied in his graduate years were topics about probability and time series. As a matter of fact, his doctoral dissertation *A Study in the Analysis of Stationary Time Series*, is about stationary stochastic processes and it contains his famous theorem on time series decomposition.

From the historical point of view, it is worth mentioning the use that Wold made of the least squares principle for his doctoral work. Herman studied the one-step prediction of a time series using linear least squares, an approach later known as the Wold decomposition: the decomposition of a time series into the sum a purely non-deterministic component and a deterministic component (for which there is an exact linear predictor). To naked eyes this might not seem to be very relevant but the truth is that Wold had already started to embrace least squares as one of his favorite analytical tools, a tool that would be the foundation for the rest of his work until the development of Partial Least Squares.

Consumer Demand Analysis (1938-1952)

In the summer before presenting his doctoral dissertation, Wold was appointed by the Sewdish government to perform an econometric analysis of consumer demand on the basis of available Swedish statistics. After his dissertation, Wold carried out the study of consumer demand analysis from 1938 to 1940. The main line of approach was to combine the analysis of family budget data and market statistics (time-series data) so as to obtain a unified picture of the demand structure in Sweden in 1920-38. Although the government commission only lasted two years, Wold spent about 14 years doing research and time series analysis of the collected data, as well as publishing several articles between 1938 and 1947. All this work was published in the dual form of a research report, first, and then a specialized textbook on econometrics: *Demand Analysis* (published in 1952).

2.2 The Least Squares Affair

Directly related with his work on Time Series and Demand Analysis, Wold got involved in a peculiar confrontation that happened within econometrics during the 1940s: Ordinary Least Squares against Maximum Likelihood. This period is crucial for the development of PLS because this is the time when Wold, somewhat stubbornly, embraced the Least Squares principle against all other methods, especially Maximum Likelihood. To understand why, we need to talk a bit more about econometrics and demand analysis.

During the first half of the twentieth century, one of the greatest challenges in econometrics was the estimation of demand analysis equations and simultaneous equation systems. Before the 1940s, the main analytical tool used for this task was ordinary least squares (OLS).

Although it was not the perfect tool, OLS was able to get the job done in most occasions. However, as models started to become more sophisticated, there were more and more cases where OLS simply didn't seem to work. Consequently, the method of least-squares started to being criticized and some econometricians began to stress out that the device was not foolproof if applied uncritically.

Partly due to lack of understanding, partly due to conceptual confussions, econometric theorists were burdened with a myriad of problems that took many years to be solved or to be aware of. We're talking about things like *identification*, *measurement errors*, *multicollinearity*, *model choice*, and so forth, that are known to affect not only OLS but many other estimation methods when applied to economic time-series and simultaneous equations. These issues are now taught in any basic econometrics course, but back in Wold's days all these problems were the cause of so many headaches for econometricians trying to figure out how the economic systems work.

Haavelmo's wholesale dismissal of OLS

The main protagonist of this period was Trygve Haavelmo (Nobel prize in Economics, 1989) who highlighted the problems of Least Squares when applied for estimating simultaneous equation systems. Haavelmo first noticed that there were problems in the OLS method for estimation in a simultaneous equation model. For instance, let us consider a very simple model with two simultaneous equations like this one:

$$Y = aX + e_1$$

$$X = bY + e_2$$

where e_1 and e_2 are error terms that follow a multinormal distribution $N(\mathbf{0}, \mathbf{V})$ with

$$\mathbf{V} = \begin{bmatrix} \sigma_{11}^2 & 0 \\ 0 & \sigma_{22}^2 \end{bmatrix}$$

If we use X to explain Y, the OLS solution of the first equation would imply that:

$$E(Y|X) = aX$$

However, solving Y anatically we get that:

$$Y = \frac{e_1 + ae_2}{1 - ab}$$

and the correct conditional expectation allowing for simultaneity should be:

$$E(Y|X) = \frac{b\sigma_{11}^2 + a\sigma_{22}^2}{b^2\sigma_{11}^2 + \sigma_{22}^2} X$$

To solve the problem, Haavelmo turned to the principle of Maximum Likelihood (ML) whose appealing properties soon caught the attention of most econometricians. Consequently, Haavelmo's revolution in econometric models and his advocacy for ML was perceived as almost equivalent to his discovery of the OLS problem.

Wold's effort to rescue OLS

Having used OLS regression extensively in his earlier works, Herman Wold felt deeply disturbed by Haavelmo's wholesale dismissal of OLS. Being in shock by the criticisms, Wold decided to challenge Haavelmo's ML approach by checking whether OLS was always inconsistent when applied to simultaneous equations. If that were the case, OLS would also be inconsistent when applied to his consumer demand models. In 1946, Bentzel and Wold first distinguished whether a system of simultaneous equations was **recursive** (later called *causal chain systems*) or **non-recursive** (later called *interdependent systems*). They then proved the equivalence of LS estimation with ML estimation in a recursive system when the disturbances of different equations were independently and normally distributed.

Behind Wold's stance against ML, there were deeper conceptual reasons for his opposition. To understand this let us consider a basic demand and supply model formed by the following simultaneous equations system:

$$Quantity_t^D = a_1 Price_t + a_2 Cost_t \qquad \text{demand}$$

$$Quantity_t^S = b_1 Price_t + b_2 Income_t \qquad \text{supply}$$

Notice how both quantity demanded and quantity supplied are determined by price, but there is no way in which price is determined. Wold disagreed with this model of demand and supply in which both relationships were determined within the same time period. He found it difficult to believe that the economic system was determined simultaneously. For him, causal forces only worked in one direction at any one time and these forces should be reflected in the models. Instead, Wold preferred to rearrange the variables in a time-based sequence like the following one:

$$Quantity_t^S = b_0 + b_1 Price_{t-1} \qquad \text{supply}$$

$$Price_t = a_0 + a_1 Quantity_t^D \qquad \text{demand}$$

Herman Wold had a particular perception of econometric models: for him the models should follow a **causal chain** format (i.e. a recursive system). Wold put emphasis on knowing the interrelations of all the variables in terms of the time lapse between causes and effects. In this way, one could reduce the system of dynamic causal relations to one final form equation. The condition to achieve such reduced form involved no simultaneously determined equations like Haavelmo suggested.

In summary, the main stream in econometrics for estimating simultaneous equation systems was the one based on the method of Maximum Likelihood. Wold was not part of this group. On the contrary, he always tried in one way or another to find a solution by least squares. Making distinction between simultaneous equations of recursive and non-recursive type, Bentzel and Wold showed that Haavelmo's wholesale dismissal of OLS estimation of the structural equations was applicable to non-recursive systems only, but not for recursive systems. For approximately the following 15 years, Wold went on a tortuouse mission to find an OLS solution for non-recursive systems.

2.3 The Philosopher of Science

Herman Wold was also deeply involved in topics related to *Philosophy of Science*. During the 1950s and 1960s Wold's research interest broadened from econometrics to other non-experimental analysis in general, and to the philosophy of science. In particular, a lot of his work in this field spinned around the prickly notion of *causality*. Although it occupies a special place within the philosophical arena, it is in the Philosophy of Science scenario where causality is one of the rockstars, a very controversial one by the way.

Through the 1950s Herman Wold dedicated a considerable part of his work to discuss notions of *model building* and *causality*. Keep in mind that these topics that were fundamentally controversial at that time. By the early 1950s, Wold began to place emphasis on the operative use of OLS regressions for causal purposes. Specially when applied to econometric models, he focused on distinguishing causal chain (or recursive) systems versus interdependent (or non-recursive) systems. He put emphasis on two points:

- the causal interpretation and operative use of the structural form of recursive systems (causal chain systems).

- the obscurities around the structural form of interdependent systems

As we saw, the choice of regression in econometrics was a very unstable ground to step on and OLS was having a lot of bad press. On top of that, there was the delicate issue of causal interpretation of OLS regressions. Throughout the 1940s causal arguments were still practically taboo in the socioeconomic and behavioral sciences, and these arguments were mostly discussed under the table. At the same time, Sewall Wright's path models in genetics were also started to catch the attention of sociologists for purposes of causal analysis. In addition, a gradual comeback of causal notions gained attraction from several independent but almost simultaneous publications by Paul Lazarsfeld, Herbert Simon, and Wold.

The Causality Crisis

How does all this relate to PLS? Well, the period of time at the end of the 1950s is crucial for the posterior development of PLS. Wold had already proved that Least Squares could be applied for recursive models. His big obstacle, however, was the estimation of non-recursive systems, or what Wold used to called: interdependent models. This was one of the biggest challenges he had to deal with in his research career, at least during the 1950s.

In his quest to find an OLS solution for non-recursive systems, but also in his journey across the causal labyrinth, all the roads he explored were dead-end roads. One attempt after another, Wold was not able to solve the non-recurvie problem and he finally ended up in a blind alley. By the end of 1950s, realizing that his discourse about the notions of causality and modeling had little or no impact in his analyses, he finally gave up. During a visiting period to the University of Columbia between 1958 and 1959, he suffered what we could call an "intellectual crisis".

After this melt down, he agreed to push the causal arguments on causal chains and interdependent systems aside. He made a radical decision: give up all what he had done and start all over again from a new conceptual and philosophical perspective. Wold made a new start on the basis of what he denominated **predictor specification**. This new approach in turn led very soon to the Fix-Point method and eventually to Partial Least Squares.

What's behind this stuff of *predictor specification*? Basically, Wold decided to leave causality aside and asked himself what could be achieved with a predictive relation. The joint multivariate distribution assumptions typically used in Maximum Likelihood were also to be left aside. The only thing that Wold wanted to take into account was the systematic part of a predictive relation expressed as the corresponding conditional expectation —which he called *predictor specification*— In this way, by requiring just mild supplementary assumptions, Wold achieved to use OLS regression for estimating coefficients that are consistent in the large-sample sense. Moreover, he extended the notion of predictor specification to models with multiple relations of blocks of variables, which in turn would lead to PLS.

3 Pre-PLS Era

After giving up the causal arguments for the estimation of structural equations, Herman Wold started to work on a new methodology that he dubbed the **Fix-Point** (FP) method. The FP method was developed as his new proposal to estimate interdependent (ID) systems by least squares.

Around the beginning of the 1960s Herman had achieved some progress with his work on ID systems using his concept of predictor specification. At that time, he was also starting to show his work to the public in some conferences. In December of 1964 he travels to the

United States on a tour of seminars to present his fix-point method for analyzing *Generalized Interdependent* (GEID) systems. Don't worry about these systems. The important thing is that during one of his seminars at the University of North Carolina, professor G. S. Tolley asked Wold whether it was possible to apply his procedure to compute principal components. After a more detailed discussion with Tolley and R. A. Porter, Wold got the idea of applying his fix-point method to calculate principal components. Years later, Wold would write that this suggestion was for him one of those moments of good luck that allowed him to open new research lines and develop the framework for the Partial Least Squares methods.

Iterative Least Squares

The first paper that presents the basis of what a decade later would give rise to PLS Path Modeling is an article published in the book *Research Papers in Statistics. Festschrift for J. Neyman* printed in 1966 and edited by F. N. David. The paper written by Herman Wold is *Nonlinear Estimation by Iterative Least Squares Procedures*. Under the acronym NILES, Wold introduces the framework and problems analyzed by what he calls **iterative least squares procedures**. It is worth noting that in this first publication, Wold does not mention the term *partial* nor PLS. He only talks about NILES procedures. This article is a catalogue of applications that can be treated by different iterative algorithms based on least squares regressions. These repertoire of applications illustrates the problems that his team in the University Institute of Statistics at University Uppsala had been working on such as:

- Principal components
- Canonical correlations
- Hybrid models of principal components, canonical correlations, and multiple regression
- Principal components in the case of partial information (i.e. with missing data)
- Interdependent systems
- Quotient regression
- Factor analysis
- A servomechanism that involves errors both in the equations and the variables

Wold recognizes that until that moment, there is nothing that accounts for the analyzed problems that his iterative least squares methodology is capable of. He mentions that:

> "The validity and possible advantages of the NILES approach have to be judged from case to case. At this early stage of development it is too early to venture a broad appraisal. It is safe to say however that the NILES procedures, thanks

to their simplicity and their adaptability to the mathematical structure of the various models, open up new vistas in the domain of nonlinear estimation"

PCA by Least Squares

Directly realted to the first publication of NILES we find another publication that same year (1966) on the *Estimation of Principal Components and Related Models by Iterative Least Squares*. In the early 1950s Wold had the chance to meet the famous psychometrician Louis Thurstone and his wife Thelma. Impressed and inspired by the work done by Thurstone and the developments in Psychometrics of that time, Wold decided to organize in Uppsala an international symposium about Factor Analysis in march 1953: the *Uppsala Symposium on Psychological Factor Analysis*. As a matter of fact, Wold gave a talk on Monte Carlo Simulation with principal components: *A Monte Carlo illustration of principal components analysis*. This talk in turn, was inspired by the paper of Paul Whittle (one of Wold's students): *On principal components and least squares methods of factor analysis*. This would be something irrelevant if it wasn't because this simulation study will be used more than a decade later for his paper on PCA with least squares published in the book *Multivariate Analysis III, Proceedings of the Third International Symposium on Multivariate Analysis (383 - 407)*.

The acronym NILES is used by Wold as an umbrella term to group different linearizing devices to estimate non-linear models. Even though he still does not use the word Partial, this concept is already reflected in his procedures: the idea of split the parameters of a model into subsets so they can be estimated in parts.

NIPALS Period

A couple of years later, in 1969, another paper **Nonlinear iterative partial least squares (NIPALS) estimation procedures** appears in *Bull Inst Internat Statis*. This is the publication where the words *Partial Least Squares* are used for the first time but still it does not imply an approach for latent variable models. Wold starts using the acronym NIPALS instead of the NILES acronym previously used in 1966. Then in 1973 another publication about the iterative algorithms that Wold's team was working on is published under the name: **Nonlinear Iterative Partial Least Squares (NIPALS) Modelling: Some Current Developments**.

The term *NIPALS modeling* is used to cover a number of different but closely related approaches to solve the analytic problems of special sciences:

> "The present report, drawing from the experiences that have emerged in the theoretical and applied work with NIPALS modeling, focuses on the potential use of the NIPALS approach in the recent many-faceted development of path models in the social sciences."

4 Soft Modeling

In 1971, Karl Joreskog introduces his LISREL approach for structural equation models with latent variables. Motivated by this type of models, Herman Wold adapts his NIPALS modeling to approach the SEM models:

> "I see the birth of PLS soft modeling as another case of serendipity. In 1971 when I saw Joreskog's LISREL algorithm for Maximum Likelihood estimation of path models with latent variables, it struck me that the same models could be estimated by suitable elaboration of my two algorithms for principal components and canonical correlations"

A series of publications in 1970s reflect the transformation and evolution of ideas behind the PLS methods:

- Wold H (1975) *From hard to soft modelling.* in H. Wold (ed.) Modelling in complex situations with soft information. (Paper. Third World Congress on Econometrics; Toronto, Canada)
- Wold H (1975) *Path models with latent variables: The Non-Linear Iterative Partial Least Squares (NIPALS) approach.* In J. Gani (ed.), Perspectives in probability and statistics: Papers in honour of M.S. Bartlett on the occasion of his sixty-fifth birthday (117-142).
- Wold H (1976) *Open path models with latent variables: The NIPALS (Nonlinear Iterative Partial Least Squares) approach.* Uppsala, Sweden: University, Department of Statistics.

At the end of the 1970s the NIPALS term is replaced by the PLS acronym. Then, in October of 1980 Wold and Joreskog organized a workshop at Cartigny, Switzerland, about the two structural equation modeling approaches: the distribution-based maximum likelihood (LISREL), and the distribution-free least squares approach (PLS). The results from that workshop were edited by Joreskog and Wold, and published in 1982 as two volumes: *Systems Under Indirect Observation: causality, structure, prediction*; with Part I dedicated to LISREL, and Part II dedicated to PLS.

The 1990s Gap

In the mid-1980s the PLS methods seemed to be living an era of expansion and consolidation. On one hand, PLS regression methods are being successfully developed in chemometrics. On the other hand, Jan-Bernd Lohmoller develops LVPLS, the first computer program integrating the basic PLS-PM algorithm and the extensions made by Lohmoller. Moreover, the publication of Lohmoller's book at the end of the 1980s gives the impression that it is a matter of time for PLS to be fully applicable in the social sciences. However, the decade of 1990s was a period that witnessed a metamorphosis of the PLS framework in which the PLS approach to latent variable path models was forgot, and only the PLS regression method was conserved for further development. The PLS path modeling "leftovers" were left aside as a mere curiosity, only appreciated by a handful of researchers in USA working primarily in marketing research and information technologies. Perhaps the oly remarkable milestone of the 1990s is the deveopment by Wynne Chin of PLS-Graph, the first PLS-PM program with a graphical user interface for drawing path diagrams.

2000s: The Consolidation

Although PLS-PM suffered from lack of popularity in the 1990s, the first decade of 2000s witnessed the consolidation of PLS Path Modeling. Part of the reason PLS-PM acquired an increasing popularity during the the first decade of 2000s is because of the impulse given by the French school of data analysis. The other major reason is because of the *European Satisfaction Index System* (ESIS) project that boosted the development of PLS-PM software and its application in marketing studies. Also, the series of international symposia exclusively decidated to PLS methods contributed to its widespread around the world.

5 PLS-PM Related Timeline

1889 Sewall Wright's birth, December 21 (father of path analysis).

1908 Herman Wold's birth, December 25 (father of PLS).

1918 First paper on path analysis *On the Nature of Size Factors by Sewall Wright*.

1920 First path diagram published in *The Relative Importance of Heredity and Environment in Determining the Piebald Pattern of Guinea Pigs* by Sewall Wright.

1925 First application of path analysis in econometrics *Corn and hog correlations* by Sewall Wright.

1938 Wold's Doctoral dissertation: *A Study in the Analysis of Stationary Time Series*.

1946 Wold and Bentzel publish the paper *On statistical demand analysis from the viewpoint of simultaneous equations* where they prove that recursive systems estimated by OLS have an equivalent solution to ML estimates.

1950s

1953 Book about econometrics by Herman Wold in association with Lars Jureen: *Demand Analysis. A study in econometrics.*
Celebration of the *Uppsala Symposium on Psychological Factor Analysis* organized by Herman Wold.

1959 Herman Wold suggests Karl Joreskog to do a dissertation on factor analysis.

1960s

1963 Herman Wold conceives the Fix-Point method at the *Econometric Week of the Vatican Academy of Sciences.*

1964 During the first seminars on the Fix-Point method at the University of North Carolina in December 1964, the estimation of principal components by the NILES procedure was brought to Wold's attention by G. S. Tolley and R. A. Porter.

1965 Talk *On the estimation of principal components and related models by iterative least squares* presented by Wold in the International Symposium on Multivariate Analysis held at the University of Dayton, Ohio, USA.

1966 First publications of NILES procedures: (1) Nonlinear Estimation by Iterative Least Squares Procedures and (2) Estimation of Principal Components and Related Models by Iterative Least Squares.
Introduction of path analysis into sociology with the publication of Path Analysis in the American Journal of Sociology by Otis D. Duncan.

1970s

1970 Inspired by the work of Arthur S. Goldenberger, Karl Joreskog conceives the idea of combining econometric and psychometric models into a single mathematical model.

1971 Inspired by the work of Karl Joreskog's LISREL for path models with latent variables, Wold conceives the idea of estimate the same models by his iterative algorithms of least squares regressions.

1973 Wold changes the name of the NILES procedures to NIPALS modeling with the publication of *Nonlinear Iterative Partial Least Squares (NIPALS) modeling: Some current developments.*
Publication of the books that set the foundations of the French data analysis school *L'Analyse des Données*. Tome 1: La Taxinomie et Tome 2: L'analyses des Correspondances by Jean-Paul Benzécri.

1980s

1981 Publication of the book *The Fix-Point Approach to Interdependent Systems* edited by Herman Wold. Presentation of the program LVPLS 1.6 by J. B. Lohmoller.

1982 *Systems under Indirect Observation: Causality, Structure, Prediction*, vols I and II edited by Karl Joreskog and Herman Wold.
Publication of *A Second Generation of Multivariate Analysis*, Vols I and II, edited by Claes Fornell.
First application of PLS-PM in marketing with a paper published in Journal of Marketing Research: *Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory* by Claes Fornell and Fred Bookstein.

1985 Partial Least Squares in Encyclopedia of Statistical Sciences by Herman Wold.

1989 Lohmoller's book: *Latent Variable Path Modeling with Partial Least Squares*.

1990s

1992 Death of Herman Wold, 16 February.
Publication of *A Primer for Soft Modeling* by R. F. Falk and N. B. Miller.
Designing of the Swedish Customer Satisfaction Barometer and foundation of what would become Claes Fornell International (CFI) Group.

1993 Wynne W. Chin presents his PLS-Graph software, the first PLS-PM software with a Graphical User Interface for drawing path diagrams.

1994 Article *Partial Least Squares* by Claes Fornell and Jaesung Cha.

1995 Publication of *The Partial Least Squares (PLS) Approach to Causal Modeling: Personal Computer Adoption and Use as an Illustration* by Barclay, Thompson, and Higgins.

1998 Publication of the book *La Régression PLS: Théorie et Pratique* by Michel Tenenhaus.
Publication of *The Partial Least Squares Approach to Structural Equation Modeling* by Wynne W. Chin.

1999 First International Symposium on PLS and Related Methods, Paris, France.
Structural Equation Modeling Analysis With Small Samples Using Partial Least Squares by Wynne W. Chin and Peter R. Newsted.

2000s

2001 Second International Symposium on PLS and Related Methods, Anacapri, Italy.

2003 Third International Symposium on PLS and Related Methods, Lisbon, Portugal.

2005 Fourth International Symposium on PLS and Related Methods, Barcelona, Spain.

2007 First Workshop, HEC and ESSEC, Paris, France.
Fifth International Symposium on PLS and Related Methods, Aas, Norway.

2009 Sixth International Symposium on PLS and Related Methods, Beijing, China.

2010 Publication of the *Handbook of Partial Least Squares. Concepts, Methods and Applications* edited by Esposito Vinzi, Wynne Chin, Jorg Henseler, and Huiwen Wang.

2012 Seventh International Symposium on PLS and Related Methods, Houston, USA.

6 Reading List

- **Hermand Wold** by Herman Wold (1982). As part of a series of autobiographies in the book *The Making of Statisticians*, Wold tells his memories and tales. The most complete biographical text of Wold.

- **Notes on the History and Nature of Partial Least Squares (PLS) Modelling** by Paul Geladi (1988). Besides the autobiographical text of Herman Wold, this paper by Paul Geladi in the *Journal fo Chemometrics* narrates the main events and happenings that gave rise to Partial Least Squares.

- **Obituary: Professor Herman Wold** by Peter Whittle (1992). Written by one of Wold's former PhD students, Peter Whittle provides a short but rich biography of Herman Wold. With touching memories from his friendship with Herman Wold, Whittle describes a concise picture of all the stuff Wold did and achieved in his life.

- **Obituary: Herman Wold, the father of PLS** by Paul Geladi (1992). Short obituray of Herman Wold in the journal .

- **The ET Interview: Professor H.O.A. Wold 1908 - 1992** by David F. Hendry and Mary S. Morgan (1994). In this paper, economic historians David Hendry and Mary Morgan publish the content of a couple of interviews with Herman Wold. This is a must read document to understand the econometrician facet of Wold and his work and contributions to econometrics.

- **The History of Econometric Ideas** by Mary S. Morgan (1990). Directly related with *The ET Interview*, this is one of the best books that talks about the history of econometrics, and one of the few texts that helps understand the role Wold played in the development of econometrics. Although the book doesn't contain notes about PLS, it does help to understand the tasks that Wold was working on and the problems he was trying to solve, which eventually led to the development of PLS.

- **Chemometrics and Intelligent Laboratory Systems 58** Special edition of the famous chemometrics journal, this issue is dedicated to PLS methods. The forewords by Tormod Naes and Michel Tenenhaus, as well as the personal memories of Svante Wold and Harald Martens, provide an interesting perspective on the historical development of PLS methods.