

NAKE

Nieuws

Volume 15, number 1

April 2003

Netwerk Algemene en Kwantitatieve Economie

Netherlands Network of Economics

Board

Prof. dr. J. Muysken (UM), Chairman

Prof. dr. ir. G. van der Laan (VU), Vice-Chairman & Secretary

Prof.dr. B.J. Heijdra (RUG), Treasurer

Prof. dr. H.P. Boswijk (UvA)

Prof. dr. E.E.C. van Damme (UvT)

Prof. dr. H. Folmer (LUW)

Prof. dr. M. Janssen (EUR)

Director

Prof. dr. J.E. Ligthart (UvT)

NAKE Secretariat

Ms. Astrid van Hemert

Faculty of Economics

Tilburg University

PO Box 90153

5000 LE TILBURG

Phone: 013-466 2516 / Fax: 013-466 3042

E-mail: nake@uvt.nl

Homepage

<http://few.uvt.nl/nake>

Preface

Spring has arrived. Besides dealing with the usual day-to-day activities, the NAKE team has been working hard on compiling a list of potential Utrecht courses for the academic year 2003-2004 and preparing for the June NAKE workshop. In addition, the NAKE educational committee got together in March 2003 to discuss the course program for the academic year 2004-2005. Work is under way to set up nine field courses—of 10 weeks each—from which students could choose in their second year of the research masters program.

We are very happy to announce that we have found 57 fellows/lecturers willing to offer 49 courses. Because some NAKE students aim to complete their NAKE course work in one year, we have included a number of the popular courses from last year's program as well. Moreover, a number of new courses will be offered to which we would like to draw your attention: **Erwin Bulte** and **Kees van Kooten** on "Natural Resource Economics," **Jacob Engwerda** on "Linear Quadratic Differential Games" and "Optimal Control of Dynamic Systems" and **Klaus Wälde** on "Economic Fluctuations and Cycles." NAKE students, we value your opinion! Please fill out the questionnaire so that we can better tailor the NAKE course program to your preferences. You can go to www.nake.nl to fill out an electronic form.

The next NAKE workshop will be held from June 10-13 at Erasmus University Rotterdam. We have invited three leading experts to introduce you to cutting-edge developments in their respective fields. **Manuel Arellano** (CEMFI) will lecture on "Panel Data Econometrics," **Gilles Saint-Paul** (Université des Sciences Sociales de Toulouse) on "The Future of Labor" and **Mark P. Taylor** (University of Warwick) on "The Economics of Exchange Rates". The workshop program can be found in this booklet and to register, you can visit our website www.nake.nl. Please register soon.

In December 2002, we had a successful workshop at Tilburg University. Two reports on the courses that were taught are included in this issue of NAKENieuws. **Robert Inklaar** (RUG) and **Ward Romp** (RUG) jointly wrote an excellent report on the lectures of Lucrezia Reichlin, which dealt with “Dynamic Factor Models.” Lucrezia’s lectures inspired them to write a paper on this topic, which they intend to submit to a journal. **Martin Kahanec** (UvT) wrote a very interesting overview of Patrick Bolton’s lectures on “Incomplete Contracts and the Theory of the Firm.”

Best regards,

Jenny Ligthart

NAKE Workshop
10 - 13 June 2003
Erasmus University Rotterdam

PROGRAM

Tuesday, June 10

10.30 - 11.30 Registration/coffee
11.30 - 12.45 Mark P. Taylor
 12.45 - 13.45 Lunch
13.45 - 15.00 Manuel Arellano
 15.00 - 15.15 Break
15.15 - 16.30 Gilles Saint-Paul
 16.30 - 17.30 Welcome reception
 18.00 Excursion

Wednesday, June 11

09.00 - 10.45 Gilles Saint-Paul
 10.45 - 11.00 Break
11.00 - 12.45 Mark P. Taylor
 12.45 - 13.45 Lunch
13.45 - 15.30 Manuel Arellano
 15.30 - 15.45 Break
15.45 - 17.15 Gilles Saint-Paul
**17.30 - 19.00 Individual
 consultation**

Thursday, June 12

09.00 - 10.45 Manuel Arellano
 10.45 - 11.00 Break
11.00 - 12.45 Gilles Saint-Paul
 12.45 - 13.45 Lunch
13.45 - 15.30 Mark P. Taylor
 15.30 - 15.45 Break
15.45 - 17.15 Manuel Arellano
 19.30 Workshop dinner

Friday, June 13

09.00 - 10.45 Mark P. Taylor
 10.45 - 11.00 Break
11.00 - 12.45 Manuel Arellano
 12.45 - 13.45 Lunch
13.45 - 15.30 Gilles Saint-Paul
 15.30 - 15.45 Break
15.45 - 17.15 Mark P. Taylor
 17.30 - 18.30 Closing drinks

Factor Models in Large Panels of Time Series

Report by Robert Inklaar and Ward Romp*

1 Introduction

In recent years increasingly large datasets on economic time series have become available. However, many of the existing statistical tools are not well suited to analyzing such datasets. One of the tools that is frequently used to extract relevant information from large datasets is factor analysis. The basic idea behind factor analysis is that movements in a large number of series are driven by only a limited number of common shocks. This type of information extraction has been successfully applied to many fields of research, ranging from asset pricing theory (Ross, 1976), forecasting (Stock and Watson, 2002) to business cycle research (Altissimo and others, 2001).

In factor models, each series of a dataset is decomposed into a common component that is driven by only a few common shocks and into an idiosyncratic component. In this report we focus specifically on a recent identification and estimation scheme as proposed by Forni and others (2000). The common components in their so-called *generalized dynamic factor model* are derived by assuming a dataset with an infinite number of series and an infinite number of time series observations. The method they propose combines the model of Chamberlain (1983) and Chamberlain and Rothschild (1983), which allows for correlated idiosyncratic components, but is static and the models of Sargent and Sims (1977) and Geweke (1977) that are dynamic but assume orthogonal idiosyncratic components. For most time series analyses the dynamic character of the model is especially important as common shocks may not have the (main) impact on a series contemporaneously but with a lead or lag.

This report is structured as follows. Section 2 of this report discusses the traditional static factor model and its estimation. In this model, the common shocks only have a contemporaneous effect on each of the series. Next, in Section 3, we show how Forni and

*Department of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, e-mail: r.c.inklaar@eco.rug.nl

others (2000) extend this model to allow for both leads and lags of the common shocks to influence the series as well as to allow for a more general relationship between the idiosyncratic components of different series. Section 4 deals with some of the problems in selecting the number of common shocks. Finally, Section 5 discusses the application of this model to business cycle analysis, an avenue of research we pursue in a related paper (see Inklaar, Jacobs and Romp, 2003).

2 The Static Factor Model

The simple static factor model has been used for a long time but it provides a useful introduction into this class of models. The basic idea is that a vector of N stationary time series can be decomposed into a common component and an idiosyncratic component. The common component is driven by only a few common shocks, denoted by Q ($< N$). In the static factor model, the common shocks affect the series only contemporaneously. The basic model is given by

$$\mathbf{X} = \mathbf{U}\mathbf{B} + \mathbf{\Xi} = \mathbf{\Phi} + \mathbf{\Xi}, \quad (1)$$

where \mathbf{X} is the $(T \times N)$ matrix of observations with N time series of length T . The series are normalized to have zero mean and unit variance. \mathbf{U} is the $(T \times Q)$ matrix of Q common shocks and \mathbf{B} is the $(Q \times N)$ matrix of factor loadings, which determines the impact of common shock q on series n . The common shocks and the factor loadings together make up the common component $\mathbf{\Phi}$. After the influence of common shocks has been removed, only the idiosyncratic component $\mathbf{\Xi}$ remains.

To estimate the common component, we have to find a linear combination of the series in \mathbf{X} explaining as much as possible of total variance, in effect minimizing the idiosyncratic component.¹ The parallel with least squares estimation is clear from this formulation, but the fact that the common shocks are unobserved complicates the problem. A further problem is that it is not known beforehand how many common shocks Q affect the series in \mathbf{X} . We will not deal with this issue here, but discuss it in Section 4.

The standard way to extract the common component in the static case is to use principal component analysis.² In principal component analysis, the first Q eigenvalues and eigenvectors are calculated from the variance-covariance matrix of the dataset \mathbf{X} .

¹For some purposes it is useful to separately identify the common shocks and factor loadings, but as only the product of the two is known, a rotation has to be chosen. As we are only interested in estimating the common component we will not treat this problem.

²See Brillinger (1975) for more details.

The common component is then defined as:

$$\hat{\Phi} = \mathbf{X}\mathbf{V}\mathbf{V}', \quad (2)$$

$$\mathbf{V} = [p_1, \dots, p_Q], \quad (3)$$

where p_i is the eigenvector corresponding with the i^{th} largest ($i = 1 \dots Q$) eigenvalue of the covariance matrix of \mathbf{X} .

3 The Generalized Dynamic Factor Model

In the static factor model the implicit assumption is made that all series are only contemporaneously affected by the common shocks. In the real world, however, this assumption is unlikely to hold. For example, in the analysis of business cycles, leading, coincident and lagging variables can be distinguished. The coincident variables show the current state of the economy, while leading and lagging variables give information about future and past developments, respectively. This rich dynamic structure of the dataset, however, will not be adequately represented by a model where shocks only have a contemporaneous impact on the series.

Models that relax this assumption are not new. Dynamic factor models were introduced by Geweke (1977) and Sargent and Sims (1977), but they assume orthogonal idiosyncratic components. This latter assumption was relaxed in the models of Chamberlain (1983) and Chamberlain and Rothschild (1983), but these are again static models. This report focuses on the model and estimation scheme of Forni and others (2000), which is both dynamic and allows for limited cross-correlation between the idiosyncratic component. For identification they assume an infinite cross-sectional dimension as in Chamberlain (1983) and Chamberlain and Rothschild (1983).

The generalized dynamic factor model as proposed by Forni and others (2000) has a similar representation as the static factor model, but with dynamic factor loadings:

$$\begin{bmatrix} x_{11} & \dots & x_{2N} \\ \vdots & \ddots & \vdots \\ x_{T1} & \dots & x_{TN} \end{bmatrix} = \begin{bmatrix} u_{11} & \dots & u_{2Q} \\ \vdots & \ddots & \vdots \\ u_{T1} & \dots & u_{TQ} \end{bmatrix} \begin{bmatrix} b_{11}(L) & \dots & b_{2N}(L) \\ \vdots & \ddots & \vdots \\ b_{Q1}(L) & \dots & b_{QN}(L) \end{bmatrix} + \begin{bmatrix} \xi_{11} & \dots & \xi_{2N} \\ \vdots & \ddots & \vdots \\ \xi_{T1} & \dots & \xi_{TN} \end{bmatrix}, \quad (4)$$

or in matrix notation

$$\mathbf{X} = \mathbf{U}\mathbf{B}(L) + \mathbf{\Xi}, \quad (5)$$

with notations as above and L is the lag operator. Again, the series in \mathbf{X} are normalized to have a zero mean and unit variance. Forni and others (2000) analyze this model for $T, N \rightarrow \infty$.

The latter mentioned authors, make four assumptions about the characteristics of the common and idiosyncratic shocks. First of all, they assume that the common shocks ($u_{1t} \dots u_{Qt}, t = 0, \dots, T$) are white noise and that the idiosyncratic components are stationary processes, uncorrelated with past, present and future values of the common shocks. Second, the spectral density matrix of the observation matrix \mathbf{X} exists. Crucially, they make two assumptions about the eigenvalues of the spectral density matrix. First, they assume that the first Q eigenvalues go to infinity as the number of series N goes to infinity and second, they assume that all remaining eigenvalues remain bounded. Under these four assumptions it can be shown that the common component of each series is uniquely identified and consistently estimated as both N and T go to infinity.

To use information of leading and lagging variables, the time series are converted to the frequency domain. Using the Fourier representation of the observed series, calculations are both simpler and more intuitive. The proposed estimation scheme consists of four steps. In the first step, the information in the time domain is transformed to the frequency domain. Second, a filter is constructed that maximizes the variance explained by the common component just as in the static model by using principal component analysis. Then the filter is transformed back to the time domain and applied to the time series to obtain the common component of each series.

More precisely, their estimation scheme is as follows. First, calculate for some selected integer M the $2M + 1$ autocovariance matrices $\mathbf{\Gamma}_m$, $m = -M, \dots, M$. The integer M represents the number of leading and lagging observations that contain information on the current common component. To obtain a consistent estimator, M must go to infinity and the ratio M/T must go to zero as T tends to infinity. Forni and others (2000) propose to use $M = \text{round}(\frac{2}{3})T^{1/3}$. Next the spectral density $\mathbf{\Sigma}(\theta_t)$ is estimated using the Bartlett kernel estimator.

$$\mathbf{\Sigma}(\theta_t) = \sum_{m=-M}^M \mathbf{\Gamma}_m \omega_m e^{-im\theta_t}, \quad t = 0, 1, \dots, 2M, \quad (6)$$

where $\omega_m = 1 - [|m|/(2M + 1)]$ is the Bartlett kernel.³

In the second step the Q largest (real) eigenvalues are calculated as well as the corresponding (complex) eigenvectors $p_q(\theta_t)$ of the spectral density matrix at frequency θ_t .

³This means the spectrum is evaluated at $2M + 1$ frequencies. In other words $\theta_t = 2\pi t/(2M + 1)$.

The weights of the filter in the frequency domain are now given by

$$\mathbf{K}(\theta_m) = V(\theta_m)V(\theta_m)', \quad m = -M, \dots, M, \quad (7)$$

with V' the transposed complex conjugate of V .

The weights of the filter in the time domain are obtained by the inverse Fourier transform

$$\mathbf{K}_k = \sum_{m=0}^{2M} \mathbf{K}(\theta_m) e^{ik\theta_m}, \quad k = 0, 1, \dots, 2M. \quad (8)$$

Applying this filter to the dataset gives an estimator of the common component

$$\hat{\Phi} = \frac{1}{2M+1} \sum_{k=-M}^M L^k(\mathbf{X}) \mathbf{K}_k, \quad (9)$$

where observations are replaced by zeros if observations are not available at the start and end-points. That is, each common component is a moving average (M lags and leads) of the series in question. The weights of this moving average are determined by the eigenvectors of the spectral densities. For $M = 0$, this model reduces to the static model of the previous section.

As this is a two-sided filter, it suffers from the usual start and end-point problems. To calculate the estimator for the last observation, one needs M future observations which are clearly not yet available. This makes the estimation scheme as presented here unsuitable for prediction. But as Forni and others (2003) show, the scheme can be adapted to use the information in the leading variables so the generalized dynamic factor model is also suitable for prediction.

4 The Choice of Q

Both in the static and in the generalized dynamic factor model we have assumed the number of common shocks Q is known. In practice, of course, Q is not predetermined and must be chosen or ideally estimated based on the dataset at hand. As Forni and others (2000) point out, there is no formal statistical test to help identify the number of factors although for the static model Bai and Ng (2002) propose information criteria that can determine the number of factors from observed data. For the generalized dynamic model, Forni and others (2000) propose two additional heuristic criteria, which are also applicable to a dynamic setting.

The first criterion looks at the value of the eigenvalues of the spectrum averaged over all frequencies. Given a dataset of N series, the eigenvalues can be calculated for a subset

of $R(< N)$ series. If the first Q eigenvalues diverge and the $(Q + 1)^{th}$ does not diverge when R increases, the number of factors is correctly identified. The second criterion relates the choice of Q to the variance explained by the i^{th} eigenvalue (averaged over all frequencies). If the model assumptions are fulfilled, there should be a substantial gap between the variance explained by the Q^{th} eigenvalue and the $(Q + 1)^{th}$. Forni and others (2000) propose to include factors as long as they explain a certain percentage of total variance, for example, 5 percent. Altissimo and others (2001) use a cutoff percentage of 10 percent.

To illustrate the difficulties in selecting the correct number of factors, we calculated the eigenvalues of the spectrum for a simulated dataset as well as a dataset of economic time series for the Euro area.⁴ To ensure comparability, both datasets contain 38 series of 189 observations. Figure 1 shows the eigenvalues for the simulated dataset which includes four common shocks by construction.

Along the horizontal axis, we show the number of series $R = 2, \dots, N$. As is immediately clear, the first four eigenvectors diverge quickly as more series are included while the remaining eigenvalues remain relatively small. So with simulated data there is no problem in identifying the number of factors. However, Figure 2 shows the eigenvalues of the spectrum but this time for a dataset of actual data. Here the identification of the number of factors is much less straightforward as there is no clear divergence of the eigenvalues. This illustrates the practical problems in implementing the first criterion.

Using the second criterion also has drawbacks, both conceptual and practical. From a conceptual point of view, the theory of the generalized dynamic factor model only predicts that the explained variance of eigenvalue Q relative to $Q + 1$ should grow, but it does not yield any predictions about the absolute value of the eigenvalues. Secondly, it is obvious that using this rule with a 10 percent cutoff percentage, at most 10 common shocks can be identified before running out of variance. In practice the maximum number of shocks this rule can identify is probably much smaller. Since the shocks are ordered by the amount of variance they explain, the first common shock explains at least as much of total variance as the second. Forni and others (2000) show that including a larger number of factors than the ‘true’ Q gives no problems as the surplus factors become increasingly less important as the dataset grows. However, not including enough factors is likely to cause problems as not enough relevant information is incorporated in the common components.

⁴See Inklaar, Jacobs and Romp (2003) for details.

5 Applications and Avenues for Further Research

The generalized dynamic factor model looks like an important new tool for analyzing large datasets as it can summarize the information from such large datasets in a convenient way. This factor model also reduces many of the problems of earlier factor models. One area where this model has already been applied is in the field of business cycle analysis. Altissimo and others (2001) use this model to construct a business cycle indicator for the Euro area based on nearly 1000 economic time series. This indicator tracks Euro area GDP, but gives a more timely and more frequent picture of the Euro economy than can be obtained from the National Accounts. This business cycle indicator can therefore serve as a useful early-warning signal for recessions.⁵ Further applications such as in the area of forecasting have also started to appear, see Forni and others (2003).

However, Section 4 and the example of the business cycle indicator suggest additional avenues of research. First of all, it would be useful to have more formal statistical criteria to select the correct number of common factors in a dynamic factor model. Extending the information criteria suggested by Bai and Ng (2002) to dynamic factor models would prove useful as would the design of a formal statistical test.

As mentioned above, Altissimo and others (2001) use a dataset of nearly 1000 time series to construct their business cycle indicator. Such a large dataset naturally comes closer to the necessary infinite number of series than most other datasets do. The questions this raises though is whether all these series are really needed or if maybe a good indicator can be constructed with a much a smaller dataset. Forni and others (2000) show that the common component can be estimated reasonably reliably for datasets of much smaller dimensions. In a more practical setting, we show in Inklaar, Jacobs and Romp (2003) that from a dataset of only 38 series a good business cycle indicator can be constructed as long as the series are selected based on sound economic reasoning.

⁵For more details on this business cycle indicator see <http://www.cepr.org>.

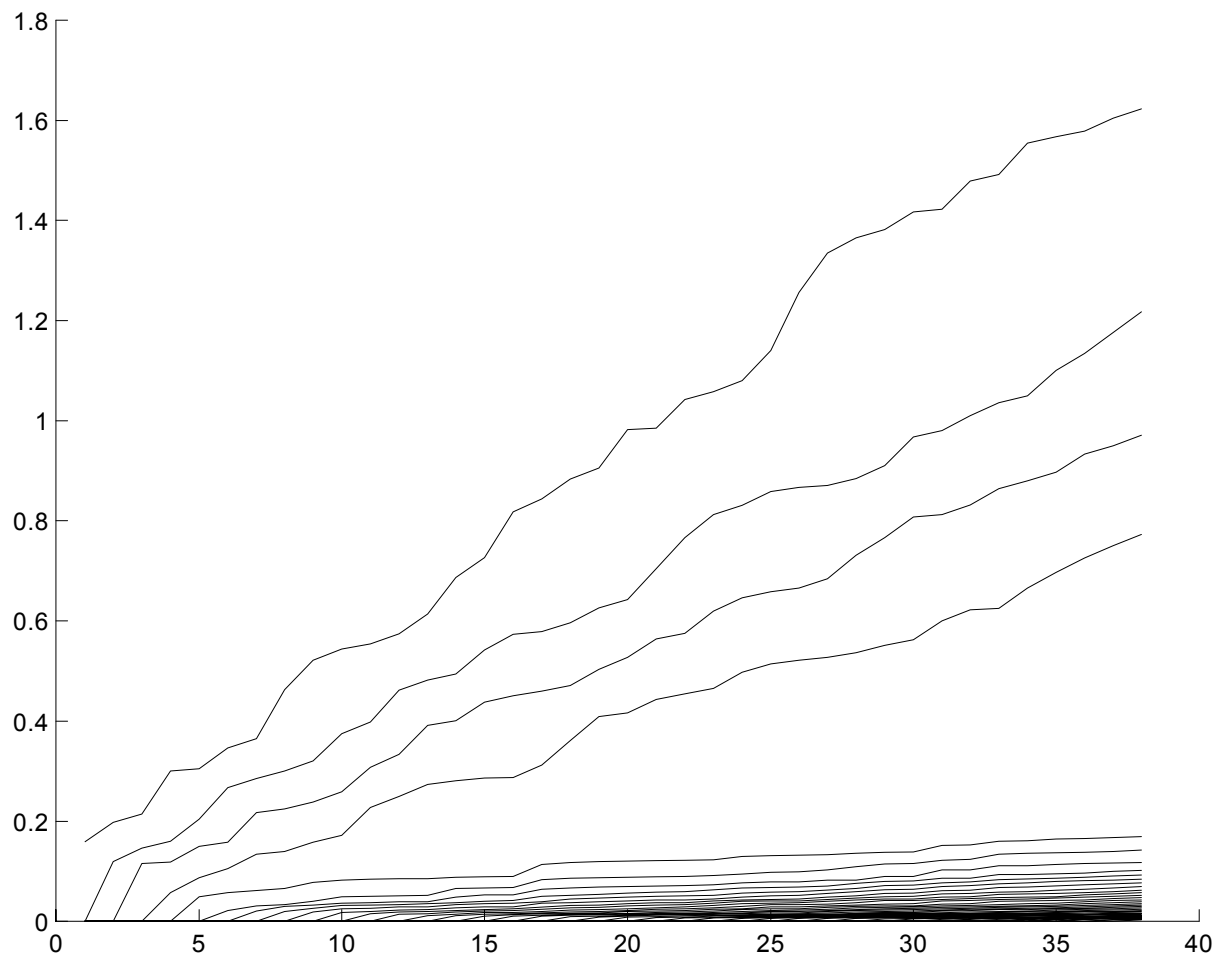


Figure 1: Eigenvalues of the spectrum averaged over the frequencies for a simulated four-factor model (38 series, 189 observations)

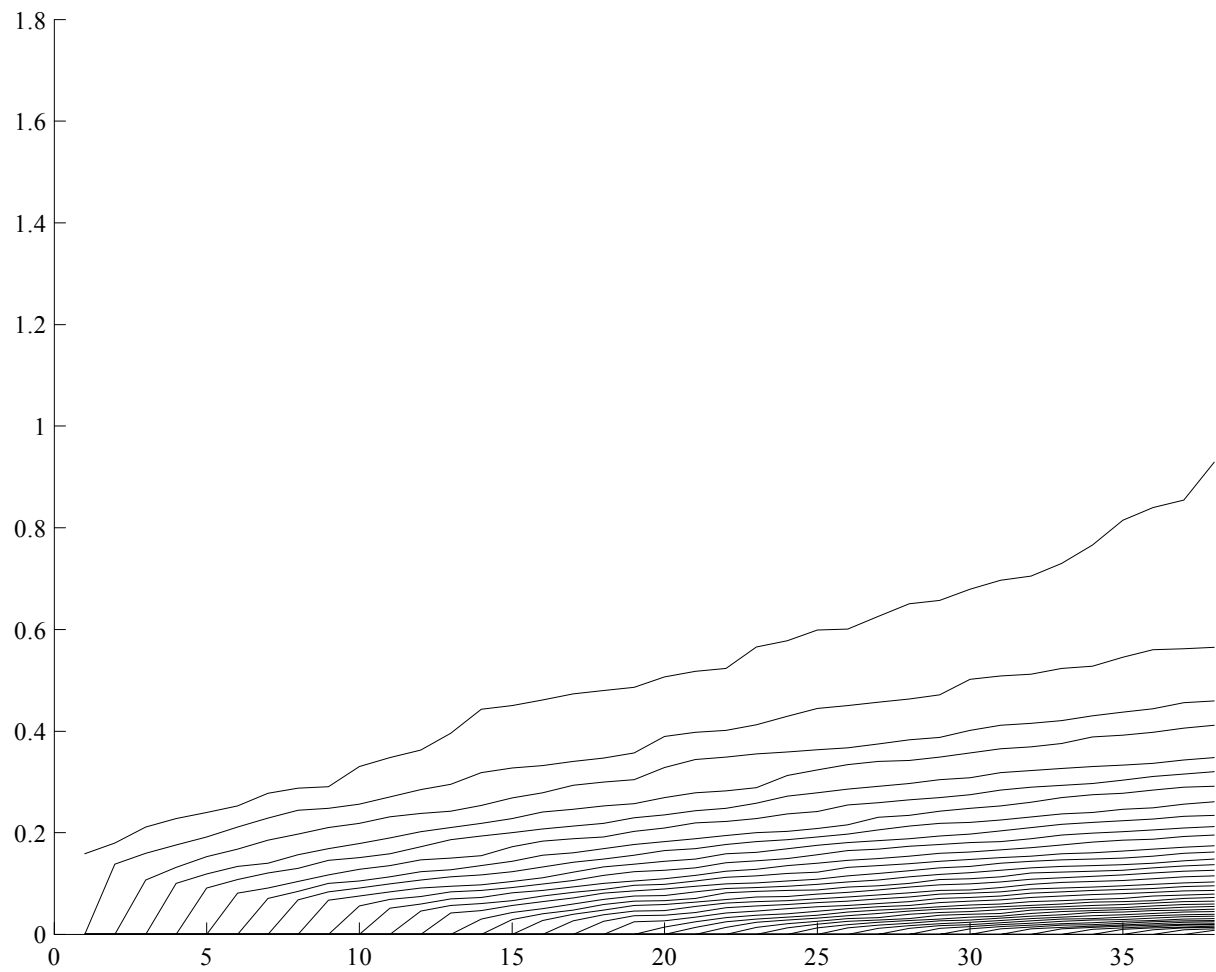


Figure 2: Eigenvalues of the spectrum averaged over the frequencies for Euro area data (38 series, February 1987 – October 2002)

References

- [1] Altissimo, F. and Bassanetti, A., Cristadoro, R., Forni, M., Lippi, M., Reichlin, L., and G. Veronese (2001), "EuroCOIN: A Real Time Coincident Indicator of the Euro Area Business Cycle", CEPR Discussion Paper, No. 3108.
- [2] Bai, J. and S. Ng (2002), "Determining the number of factors in approximate factor models", *Econometrica*, Vol. 70, no. 1, pp. 191-221.
- [3] Brillinger, D. R. (1975), *Time Series: Data Analysis and Theory*, Holt, Rinehart and Winston: New York.
- [4] Chamberlain, G. (1983), "Funds, Factors, and Diversification in Arbitrage Pricing Models", *Econometrica*, vol. 51, no. 5, pp. 1281-1304.
- [5] Chamberlain, G. and M. Rothschild (1983), "Arbitrage, Factor Structure and Mean-Variance Analysis in Large Asset Markets", *Econometrica*, Vol. 51, no. 5, pp. 1305-1324.
- [6] Forni, M., Hallin, M., Lippi, M. and L. Reichlin (2000), "The Generalized Dynamic-Factor Model: Identification and Estimation", *Review of Economics and Statistics*, Vol. 82, no. 4, pp. 540-554.
- [7] Forni, M., Hallin, M., Lippi, M. and L. Reichlin (2003), "Generalized Dynamic Factor Model, One-sided Estimation and Forecasting", downloadable from <http://www.dynfac.org>.
- [8] Geweke, J. (1977), "The Dynamic Factor Analysis of Economic Time Series" in Aigner and Goldberger (eds.) *Latent Variables in Socio-Economic Models*, North-Holland: Amsterdam.
- [9] Inklaar, R., Jacobs, J., and W.E. Romp (2003), "Business cycle indicators: Does a heap of data help?" downloadable from <http://www.eco.rug.nl/medewerk/inklaar>.
- [10] Ross, S. (1976), "The Arbitrage Theory of Asset Pricing", *Journal of Economic Theory*, Vol. 13, pp. 341-360.

Incomplete Contracts and the Theory of the Firm

Report by Martin Kahanec

1 Introduction

Contract theory originated in the works of Edgeworth in the early twentieth century. During that time, the complexity of the studied matter was reduced to the characterization of efficient exchange contracts in the world of certainty and perfectly informed agents. It took scientists approximately one quarter of a century to take account of the intricacies of uncertain contractual situations. As a result, the focus of the analyses shifted to topics like insurance, risk management, and portfolio choice. Developments in the microeconomic theory another twenty years later facilitated analyzing contractual situations where agents' access to information is imperfect and where contractual parties have asymmetric information about contractually relevant variables. These developments brought about the introduction of incentive and participation constraints and the theory of incentive contracting was born. In the 1980s, the novel notion of repeated contracting came into being and concepts such as commitment, renegotiation, and renegotiation-proofness gained scientific prominence.

It is the latest developments in contract theory and the theory of the firm that Patrick Bolton discussed during the 33rd NAKE Workshop in Tilburg. As he explained, the most recent theories take the incompleteness of contracts as given while control, ownership, authority, and tasks are all control variables to be allocated among contracting parties. The corollary of this view on contract theory is that procedural and institutional design questions are at least as fundamental and worth studying as the issues of asymmetric information, uncertainty, and repeated contracting. That institutional and procedural settings become control variables is a substantive methodological shift the latest theories hinge on. In this paper, attention is paid to two main topics of Bolton's lecture series: how the theory of incomplete contracts explains what institutional and contractual arrangements between, first, service supplier and consumer and, second, between business partners arise in the equilibrium.

2 The Employment Relation: A Model of Incomplete Contracting

How the notion of incomplete contracts enriches our understanding of the employment relation was discussed during the first part of the seminar series. The first formal model of this relation built within the framework of incomplete contracts appeared in 1951 in the pioneering work of Herbert Simon. In this model, Simon compared two possible long-term contracts in the situation where the buyer would like the seller to provide some service at some point in the future. In particular, two contracts are compared, these being the ‘sales contract’ that precisely specifies the service to be provided and the ‘employment contract’ that specifies the limits within which the buyer can demand the future service, the acceptance set. Simon shows that it is the degree of uncertainty at the time when the contract is signed and the degree of seller’s indifference among elements of the acceptance set that determine buyer’s and seller’s preferences over the set of possible contracts, respectively. It is the limitations of the set of possible contracts in Simon’s analysis that urges us to ask some challenging questions. Why do the parties need to agree to some trade before the contractual uncertainty is resolved? Why don’t they rather wait and write a spot contract after the resolution of the uncertainty? Why can’t they write a possibly superior state-contingent contract? And why is the contract specifying an acceptance set less costly to write than a contract specifying a state contingent delivery plan?

An important branch of literature evolved around these questions and, in particular, focused on the ‘hold up’ problem that may arise whenever the buyer has some (unobservable) investment possibilities to raise her payoff from the future service. If she undertakes these investments, she may be ‘held up’ by the seller when they negotiate over a spot contract specifying the terms of service provision. In particular, consider the situation in which the buyer makes an observable contract-specific investment that determines her future payoff from interaction with the seller. Afterwards, the two parties bargain. If the bargaining power of the seller is such that he can extract the entire surplus and there is no post-contractual commitment, the buyer, knowing this, has no incentives to invest ex ante and in equilibrium will not invest at all. The underlying reason is that the seller does not compensate the buyer for ex-ante investment that is sunk costs at the bargaining stage. In this extreme case, the hold up problem prevents the buyer from taking any contract-specific investment. The hold up problem impedes the buyer’s investment even when it is

unobservable under the condition that the bargaining between the two agents is non-recurring and hence the buyer is unable to gain a stronger bargaining position.

In such a situation, the buyer protects her investment against ex-post opportunism on the seller's side by writing a long-term contract before the contract-specific investment is undertaken. For it may be uncertain what kind of service the buyer will need after the uncertainty is resolved, a contract specifying an acceptance set may be the optimal contract in this situation. The next subsection formalizes these ideas in a model that shows how ex post opportunism may result in inefficient ex ante investment and answers some of the aforementioned questions. The model abstracts from ex post inefficiencies that may arise from contractual disputes.

2.1 The Model

Consider a situation in which a risk neutral buyer and seller wish to contract a service out of n types of service that the seller can provide. Suppose the buyer needs only one service ex post and he does not ex ante precisely know which one. Service provision is verifiable and observable ex post. In addition, abstract from any other contractual possibilities the parties may have outside of the considered situation. More formally, assume that if the parties trade a service a_i at price p_i then the buyer gets $v(a_i, \theta) - p_i$ and the seller obtains $p_i - c(a_i, \theta)$ where $i \in n$ denotes the type of the service the seller can provide, θ denotes the state of nature, there are $l \in \Omega$ states of the nature such that $\Omega > n$, and $c(\cdot)$ denotes the cost function. By investing I ex ante, the buyer can increase $v(\cdot)$. Investment is costly and the returns to investment are decreasing. For the sake of tractability, let $v(a_i, \theta, I) \in \{0, v(I), V(I)\}$ and $c(a_i, \theta) \in \{0, c, C\}$ where $0 < V(I) - C < v(I) - c < V(I) - c$ for all $I > 0$. In addition, assume that for any given θ any valuable service has positive costs, the costs of service increase in its value, and there ex post always exists a service such that $v(a_i, \theta, I) = v(I)$. By solving this situation for the optimal contract, we arrive at the first order condition that the marginal value of investment to the buyer must equal 1.

$$\text{Max}[I \geq 0] \{v(I) - I\} \Rightarrow v'(I^*) = 1, \quad (1)$$

where I^* denotes the optimal value of I . If we assume that in any realized state there is at least one service with the highest and one with the lowest value to the buyer and that the parties cannot write a perfect state contingent contract, as simple as this condition may

appear to be, it may be practically difficult to satisfy it. Under these circumstances, the parties have four options:

Option 1: Wait until the contractual uncertainty is resolved and write a spot contract (denoted by s).

Option 2: Write a sales contract (S) that uniquely specifies the service to be provided ex post at a pre-specified price.

Option 3: Write a buyer-employment contract (BE) that specifies an acceptance set from which the buyer can choose ex post at pre-specified terms.

Option 4: Write a seller-employment contract (SE) specifying an acceptance set from which the seller can choose ex post at pre-specified terms.

2.2 Main Results

Based on the modeling strategy we adopted, we may now effectively compare the four options. It is straightforward to see that under Option 1 the first-order condition would be $v'(I^s) = 2$, so that, upon inspecting (1), it follows that $I^s < I^*$. A more mathematically involved analysis reveals that under the Option 2 the optimal investment level must satisfy $v'(I^S)[(n+1)/n] + V'(I^S)[(n-1)/n] = 2$. The buyer-employment contract under Option 3 boils down to: $v'(I^{BE}) + V'(I^{BE}) = 2$. Finally, after some manipulation we realize that the seller-employment contract effectively turns out to be equivalent to the ex post spot contract in terms of ex ante optimal investment I^{SE} .

From the above sketch of options, we see that an important question arises, namely, what is the optimal contract? To answer this question, we need a criterion of optimality. The natural criterion in our case is the magnitude of the expected surplus from trade $v(I^k) - c$, where $k \in \{s, S, BE, SE\}$. Adopting this criterion, we observe that as long as $v'(I) \geq V'(I) \geq 0$, both the sales and buyer-employment contract (weakly) dominate the ex-ante spot contract (and the seller-employment contract). To compare the sales and the buyer-employment contract is more intricate. As showed by Simon, the sales contracts tends to generate provision of services that do not maximize the ex-post net surplus while the buyer employment contract often results in provision of excessively costly services. Simon, considering a number of exogenous variables that may affect the choice, gives a contingent answer to the question which of the two contracts is preferable. In contrast to Simon, we adopt here the ex-ante efficiency concept and look at the efficiency of the buyer's ex-ante incentives to invest. In this framework we can develop the condition that whenever the

marginal return on investment in the inefficient undertaking is higher than that for the efficient action $V'(I^S) > v'(I^S)$, the buyer-employment contract typically dominates the sales contract, and vice-versa. It is our allowing for renegotiation in the present context that makes the comparison between the two contracts more intricate. An important critique of the whole framework comes from Alchian and Demsetz (1972), who do not find any difference between an employment contract and a relationship between a customer and a shop assistant.

The analysis above argues that the difference lies in the distinction between ex-ante and ex-post (spot) contracting. The above section shows how we can assess and compare efficiency of various kinds of contracts between the buyer and the seller. Assuming that it is the most efficient contracts that the potential contracting parties write, the analysis above also predicts what contracts typically emerge in equilibrium. The next section, focusing on the second pivotal topic of Bolton's seminar, discusses the theory of efficient ownership and the firm.

3 Ownership and Property Rights: A Model of the Firm

While the early incomplete contract theory focused on the employment relation, it was the ownership title that attracted economists' attention in the late 1980s and the 1990s. Spurred by the work of Grossman and Hart (1986), an important strand of the incomplete contracts literature evolved around the notions of property rights, ownership, and control. Perhaps the most important contribution of Grossman and Hart was that they changed our understanding of what property rights are. While the earlier works for example by Alchian and Demsetz (1972) and Meckling and Jensen (1976) defined property rights in terms of cash-flow rights, Grossman and Hart define them in terms of 'residual rights of control.' In other words, an ownership title gives the owner the right to divest himself of the property as he considers right, within some contractually or legally specified limits. Grossman and Hart (1986) and Hart and Moore (1990) formalize this definition of property rights within the incomplete contract framework. To do so, they postulate that the owner of the firm has the right to exclude others from using the firm's assets, thereby protecting himself from ex post opportunism. The following example, due to Grossman and Hart, illustrates why these theories predict that the ownership is allocated to the party requiring the most protection against ex post opportunism.

Consider a situation where we only have two firms; take the example of a publisher and a printer. If these firms are owned separately, they are free to write a long-term supply contract, which we assume to be incomplete. In this situation, it is assumed that whenever unexpected contingencies arise, the contract is efficiently renegotiated ex post. Assuming that the parties undertake contract specific investments that are costly to reverse, the ex post negotiating situation differs from the ex ante situation because the parties may now be locked in the contract relation to a varying degree. If the parties foresee the possibility of the lock-in, the one that expects a worsened negotiating position may abstain from taking the contract specific investment even in the case it is socially efficient to take it.

On the contrary, if the two parties were integrated so that the party that potentially ends up with a weakened negotiating position as described above owns them both and effectively becomes the ‘employer’, this party’s ex-ante incentives may be quite different. Correspondingly, in this case the other party becomes the ‘employee’. In particular, the stronger bargaining position of the employer in the employment contract, as compared to the long-term supply contract, may secure or at least make it more likely that any efficient investment on the employer’s side is undertaken. On the other hand, it is important to realize the other side of the coin and see that under the employment contract the employee may be discouraged to take ex ante efficient investments. Based on the logic of this example, Grossman and Hart predict that the equilibrium ownership allocation is determined by the relative value of each party’s ex-ante investments. In particular, they predict that it is the party that requires the most protection against ex-post opportunism, the one that has the most valuable ex-ante investment opportunities, obtains the property rights on the integrated firm and becomes the ‘employer’. This simple example has cast some light on the pros and cons of integration. The next section illustrates the argument more formally as explored by Hart and Moore (1990).

3.1 The Model

Assume a situation in which N agents in the first stage make ex-ante investments x_i at a cost of $c(x_i)$. These investments are assumed to be specific to a subset of assets $A \subseteq A^*$, where A^* denotes the set of all possible assets in the economy. In the second stage, $S \subseteq N$ agents form a coalition and combine their investments with the set of assets A to produce an ex-post surplus of $V(S, A | x)$. To simplify the analysis, we assume away long-term trade contracts and ex-post inefficiency. In addition, we postulate that the surplus from trade is

divided according to the Shapley Value concept. In this simplified case the only way the ownership structure affects the final contractual arrangement is through affecting the ex-post division of $V(\cdot)$ that, in turn, affects the ex-ante investment incentives. To make the following illustration tractable, assume $N = 2$ and $A^* = \{a_1, a_2\}$.

In the described setup, consider again the ‘employer’ and ‘employee’ as above. Then, three ownership arrangements are possible: non-integration, employer integration, and employee integration. Investigating the non-integration case, we realize that the surplus, call it $v(x)$, can only be created if the parties agree to use the two assets together. The Shapley Value assigns benefits, B_i , of $v(x)/2$ to each of the parties. Under employer integration, assuming that the employer still benefits from cooperating with the publisher, hence assuming the following: $V(2, \emptyset | x) = 0$, $V(1, \{a_1, a_2\} | x) = \pi_1(x_1) < v(x_1, x_2)$, the Shapley Values are $[v(x) - \pi_1(x_1)]/2 + \pi_1(x_1)$ for the employer and $[v(x) - \pi_1(x_1)]/2$ for the employee. Correspondingly, in the reverse situation of employee integration the Shapley Values are $[v(x) - \pi_2(x_2)]/2 + \pi_2(x_2)$ for the employee and $[v(x) - \pi_2(x_2)]/2$ for the employer.

3.2 Main Results and Generalizations

Knowing the surplus shares across situations, we can look for the corresponding ex-ante investments. Under the usual regularity assumptions about $v(x)$, $\pi(x)$, and $c(x)$, the ex-ante non-observable¹ (or non-verifiable) investments are chosen competitively to maximize each agent’s surplus: $Max[x_i] \{ B_i(S | x) - c_i(x_i) \}$. The Nash equilibrium investment levels are determined by the resulting first-order conditions. Utilizing this simple analysis under the usual assumptions, particularly that $\pi_i'(x_i) > 0$, we obtain the result that the employer has greater incentives to invest for any given level of employee’s investment in the case of employer integration, than under non-integration and employee integration. Similar results apply to the incentives of the employee. This theory of property rights sheds some light on the costs and benefits of integration. We see that integration induces investment on the side of the party that potentially loses more bargaining power in case of some contingencies (employer in our example) at the costs of investment of the party that loses

¹ If investments were observable or verifiable we would not escape the Coase prediction

relatively less (here the employee). Thus, the property rights theory predicts an equilibrium ex-ante efficient allocation of property rights if we believe that enforcing ownership titles is easier and cheaper than writing and enforcing trade contracts².

Generalizing the framework to the multi-agent setting, Hart and Moore show that under some practically mild conditions for any ex-ante allocation of ownership rights there is equilibrium under investment. Moreover, they derive three simple rules for optimality of ownership allocation. First, if only one agent invests ex ante, he or she should be the owner of all property. Second, joint ownership of assets is inefficient. And, third, strictly complementary assets should be owned by one agent. Bolton and Whinston extend the analysis to the case of substitutable investments. They discuss the efficiency of ownership allocation as determined by the degree of competitiveness among agents. They show that in the case of ex-post competition the equilibrium ownership structures may be, in case they are concentrated, inefficient due to the strengthened market power of the owner. The benefits in terms of the protection against ex-post opportunism may become overrun by the costs of the excessive market power.

The simple illustration above shows that the property rights theory can explain some aspects of the efficiency of equilibrium ownership structures, the costs and benefits of integration, and the effects of ownership structure on economic incentives and decisions. In addition, if we believe that the property rights are allocated most efficiently in equilibrium, the property rights theory predicts what ownership structures typically arise in various situations in the corporate world. However, it is important to bear in mind the simplifications we imposed, particularly the symmetric information setup, invariant centralized bargaining under any ownership allocation, the Shapley-type bargaining outcomes, etcetera.

4 Conclusion

This paper summarizes two central topics of Bolton's seminar series. While Bolton extended these topics in many directions and elaborated on numerous details, this paper conveys some important messages of the seminar series. First, the incomplete contracts theory advances our understanding of economic structures and permits their deeper studying. Second, we should consider contractual and ownership structures as control

variables while the incompleteness of contracts is given. Third, institutional arrangements affect economic efficiency. The efficiency of institutional arrangements can be measured by the potential societal surplus. Fourth, we have to make distinction between ex post and ex ante contracting. This distinction is fundamental when assessing the efficiency of various contracts in diverse situations. Fifth, some ownership structures are more efficient than others in certain situations. Sixth, if we believe that efficient institutional arrangements are more likely to arise than the inefficient ones in the equilibrium, we can predict typical institutional arrangements in various situations. Finally, we have learned that the incomplete contracts theory is by no means complete and much is left for future research.

References

- Alchian, A. and H. Demsetz (1972), "Production, Information Costs, and Economic Organization." *American Economic Review*, 62(5): 777-95.
- Bolton, P., 2002, Lecture Notes NAKE Workshop Tilburg.
- Bolton, P. and M. Whinston (1993), "Incomplete Contracts, Vertical Integration, and Supply Assurance", *Review of Economic Studies*, 60:121-48.
- Grossman, S. and O. Hart (1986), "The Cost and Benefits of Ownership: A Theory of Vertical and Lateral Integration", *Journal of Political Economy*, 94: 691-719.
- Hart, O. and J. Moore (1990), "Property Rights and the Nature of the Firm", *Journal of Political Economy*, 98: 1119-58.
- Jensen, M. and W. Meckling (1976), "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", *Journal of Financial Economics*, 3 (4): 305-60.
- Simon, H. (1951), "A Formal Theory of the Employment Relationship." *Econometrica*, 19:293-305.

² Two assumptions are necessary: each party has enough wealth to trade any amount of ownership titles and there are no post-trade gains to re-trading the ownership rights.

Upcoming....

<i>Month</i>	<i>Day</i>	<i>NAKE activity</i>
June 2003	10 - 13	Workshop at Tinbergen Institute – Rotterdam
September 2003	5	Start of Utrecht courses, block I 2003/2004
October 2003	24	NAKE day 2003 - Dutch Central Bank Amsterdam
October 2003	31	Start of Utrecht courses, block II 2002/2003
December 2003	08 - 12	Workshop at the University of Groningen

Information about these activities can be obtained from Astrid van Hemert, nake@uvt.nl