Bernd-O. Heine, Matthias Meyer and Oliver Strangfeld (2005)

Stylised Facts and the Contribution of Simulation to the Economic Analysis of Budgeting

Journal of Artificial Societies and Social Simulation vol. 8, no. 4
<http://jasss.soc.surrey.ac.uk/8/4/4.html>

For information about citing this article, click here

Received: 02-Oct-2005 Accepted: 02-Oct-2005 Published: 31-Oct-2005

Abstract

The application of computer simulation as a research method raises two important questions: (1) Does simulation really offer added value over established methods? (2) How can the danger of arbitrariness caused by the extended modelling possibilities be minimised? We present the concept of stylised facts as a methodological basis for approaching these questions systematically. In particular, stylised facts provide a point of reference for a comparative analysis of models intended to explain an observable phenomenon. This is shown with reference to a recent discussion in the "economic analysis of accounting" literature where established methods, i.e. game theory, as well as computer simulations are used: the susceptibility of the "Groves mechanism" to collusion. Initially, we identify six stylised facts on the stability of collusion in empirical studies. These facts serve as a basis for the subsequent comparison of four theoretical models with reference to the above questions: (1) We find that the simulation models of Krapp and Deliano offer added value in comparison to the game theoretical models. They can be related to more stylised facts, achieve a better reproduction and exhibit far greater potential for incorporating yet unaddressed stylised facts. (2) Considered in the light of the stylised facts to which the models can be related, Deliano's simulation model exhibits considerable arbitrariness in model design and lacks information on its robustness. In contrast, Krapp demonstrates that this problem is not inherent to the method. His simulation model methodically extends its game theoretical predecessors, leaving little room for arbitrary model design or questionable parameter calibration. All in all, the stylised-facts-concept proved to be very useful in dealing with the questions simulation researchers are confronted with. Moreover, a "research landscape" emerges from the derived stylised facts pinpointing issues yet to be addressed.

Keywords:
Computer Simulation, Stylised Facts, Methodology, Groves Mechanism, Collusion, Game Theory

Introduction

1.1

Modelling social phenomena using computer simulation is a promising research method. Simulation researchers claim that its major benefit is its extended modelling possibilities.[1] However, they still have to convince their orthodox peers in the various related research fields. In particular, they must provide answers to the following two questions. Firstly, does simulation really offer added value over the established methods? And if so, is it possible to identify precisely its specific merit? Secondly, extended modelling potential also implies higher degrees of freedom in model construction. Given these increased degrees of freedom in modelling, the question arises as to how the danger of arbitrariness can be prevented.

1.2

A current discussion in the field of management accounting provides a good illustration of this general problem. In the "economic analysis of accounting" literature, game theory is currently a well-established method for analysing the behavioural implications of different accounting rules and instruments.[2] Recently, there has been a lively debate as to whether the Groves mechanism, a method for capital allocation, is susceptible to collusion. Apart from game theory approaches, simulation models are used to analyse the dynamics of the collusion problem.

1.3

In this discussion, the same two questions arise again, but in the context of a specific research topic: Do the simulation models applied offer added value in comparison to traditional game theory analyses, and if so, where exactly is this contribution located? Moreover, the standard rationality assumption in game theory appears quite restrictive for many researchers, but these restrictions have some advantages relating to model construction. In particular, the rationality assumptions are a clear guide for the individual researcher and therefore individual contributions are well aligned with the overall research approach.[3] If one decides to model agents differently, the individual researcher must decide what characteristics should be ascribed to an agent. How can the numerous modelling decisions involved be evaluated? Is there any guide available to the individual researcher and the broader scientific community to judge which assumptions are appropriate and which are arbitrary?

1.4

In this paper, we introduce the methodological concept of "stylised facts" to tackle these two questions. First we consider this concept in detail, so that it can be used as a basis for assessing simulation models (section 2). Then we show, with reference to the discussion on the Groves mechanism, how this concept can be applied to answer the two questions posed above and discuss the contribution of simulation to this current research problem (section 3). Since the use of the concept is not limited to this particular case, we include a general methodological perspective in our final conclusion (section 4).

The Concept of Stylised Facts as Basis of Assessment

2.1

In order to introduce the concept of stylised facts and to explain its usefulness for this paper, we begin with the basic idea behind this concept. Subsequently, we elaborate on the two ways of using it, the ex ante (model construction) and ex post (model evaluation) use. Finally, we explain how the latter form of use helps to answer the questions stated in the introduction.

2.2

The basic idea behind the concept of stylised facts is to guide the researcher in his endeavour to build an adequate model of the phenomenon under investigation. One important difficulty in this respect is to find an appropriate level of abstraction. The model should be parsimonious enough to avoid distraction by minor details and at the same time rich enough to capture the relevant aspects of the phenomenon. The economist Kaldor was facing this obstacle when he intended to propose a research agenda in macroeconomic growth theory that was geared towards the explanation of observable phenomena. In this context, Kaldor (1961/1968) introduced the concept of stylised facts to offer a way to identify and communicate key

observations that demanded scientific explanation. He coined the term "stylised facts" for stable patterns that emerge from many different sources of empirical data.

2.3

This approach has important methodological implications beyond macroeconomics. It provides a means of clarifying the objective of positive research[4] based on the extraction of the characteristic features of the phenomenon under investigation. This is of particular importance, because the more complex the phenomenon, the more difficult it is to capture its relevant aspects with scientific means of observation. According to Kaldor, in particular statistical representations of empirical observations "are always subject to numerous snags and qualifications" (Kaldor 1961/1968:178). Consequently, it is often difficult to directly identify the relevant aspects of a phenomenon, which should be captured in a model.

2.4

To handle this problem, Kaldor suggested a "stepping stone" in the process of modelling: "[T]he theorist should be free to start off with a stylized view of the facts" (Kaldor 1961/1968:178) in order to clarify the objective of his research first. Such a stylised view concentrates on broad tendencies and ignores individual details so as to identify robust patterns across different observations. On this basis the researcher can then proceed to construct a model, which is able to explain these stylised facts (without distractions caused by minor, contradictory details in empirical studies and in other material).

2.5

To provide an illustration of the concept, some of Kaldor's stylised facts on macroeconomic growth are listed in the following quotation:

> As regards the process of economic change and development in capitalist societies, I suggest the following 'stylized facts' as starting point for the construction of theoretical models: (1) The continued growth in the aggregate volume of production and in the productivity of labour at a steady trend rate; no recorded tendency for a falling rate of growth of productivity. (2) A continued increase in the amount of capital per worker, whatever statistical measure of 'capital' is chosen in this connection. (...) (Kaldor 1961/1968:178).

2.6

Two of the methodological implications of the concept should be highlighted. Firstly, it introduces a different notion of the "as if"-approach (cf. Boland 1987/1994:535). The neoclassical "as if"-approach uses simplistic assumptions *as if* they were true (cf. Friedman 1953/1989), implying that the assumptions necessary for a chosen level of abstraction should be accepted and considered from the perspective of the purpose of the whole model, in particular with reference to the predictions it can generate. The concept of stylised facts differs in that it states facts *as if* they truly represented the phenomenon under investigation, thereby guiding model construction. Hence, they refer to an earlier step in the scientific process and their purpose is to facilitate the choice of an appropriate level of abstraction. Secondly, it complements both the realist and instrumentalist point of view: For realism, stylised facts are, on the one hand, compatible with the conviction that "facts about the world" can be known (cf. Mäki 1998:407), on the other hand, they express that models are also instruments of thought and not a mere description of reality, thereby favouring an analytical approach towards model construction in contrast to naïve realism. For instrumentalism, stylised facts on the one hand, are compatible with the neoclassical "as if"-approach in that they do not stipulate a certain choice of abstraction, but rather clarify the objective of theorizing for the subsequent choice of appropriate abstractions. On the other hand, the concept of stylised facts suggests that mere prediction is not enough. The researcher should also be concerned about the generative mechanisms[5] behind his results ("looking under the hood", cf. Hausman 1994/1995). Hence, the concept of stylised facts offers productive implications from both methodological perspectives.

2.7

Unfortunately, these benefits carry two major caveats. The researcher is now in a position to influence two important steps in the scientific process. He states the stylised facts he wishes to explain and then constructs a model based on specific abstractions, yielding certain hypotheses which potentially fit the stylised facts. From a methodological perspective, this allows the production of research results in at least two unproductive ways (cf. Boland 1987/1994:536). Firstly, the researcher can claim that empirical falsifications of his results are meaningless, because they are "noise" vis-à-vis the broad tendencies captured by his stylised facts ("immunisation strategy"). Secondly, he can make his stylised facts fit the results of his model. One could call this "ad hocery" relating to the derivation of stylised facts. This potential danger of abuse was noted by Solow in direct response to Kaldor, when he quoted his stylised facts of macroeconomic growth with the comment "[t]here is no doubt that they are stylized, though it is possible to question whether they are facts." (Solow 1969/1988:2).

2.8

Although quite critical, the quotation also points to a central aspect of the productive use of the concept. The explicit statement of a specific set of stylised facts by an individual researcher should not be the end, but the *beginning* of a critical discussion among the experts in a particular field. This discussion makes explicit and open to criticism, what is otherwise only used implicitly by individual researchers ("their view of the phenomenon"). Ideally, at the end of such a process, a consensus emerges, at least with regard to some stylised facts. Such a discussion has to begin with existing empirical data and may even stimulate further empirical research. "Triangulation" is one important concept in this respect. By combining different empirical research methods, the respective shortcomings of one individual method can be overcome.[6] Consequently, the derivation of stylised facts should be based transparently on empirical investigations *and* should be subject to discussions among experts in the field. Ultimately, good stylised facts are the result of a critical discussion within the expert community, with a high level of consensus that they represent a robust pattern in empirical observations.[7] Before such a consensus has been reached, the transparency of derivation, the amount and consistency of empirical results and the independence from specific models or streams of literature may serve as supporting indicators of stylised fact quality.

2.9

Hence, the use of stylised facts does not substitute empirical testing of the resulting theories and other falsification attempts, but helps *ex ante* to decide on the direction of research and to make this choice transparent to peers.[8] Stylised facts thus render the objective of positive research explicit to a research community. This allows to focus research (*ex ante*) on making a contribution to the understanding of observable phenomena and to provide guidance for model construction, in particular for formulating productive abstractions based on the observed characteristics of these phenomena. It should be noted that this does not imply that stylised facts already latently "contain" a model, which the researcher simply has to make explicit. Model construction still remains an art. But in the process of model construction, the consideration of stylised facts prevents a researcher from building models that do not yield relevant and testable hypotheses about reality.[9]

2.10

These virtues apply, if a set of stylised facts is used for model construction. But stylised facts can bring to bear their virtues also from an *ex post* perspective. Applied from a distance from the process of model construction,[10] stylised facts can serve as basis for evaluating a collection of already-existing models intended to explain the same phenomenon. Using a given set of stylised facts of a phenomenon, models can be analysed comparatively, with a focus on their productive implications without distraction by minor issues that may also be covered by the models: "[A]s long as we can come to an agreement regarding the 'stylized' facts, the comparative appropriateness of competing explanatory abstractions can be brought into clear and decisive focus" (Boland 1987/1994:535-536).[11]

2.11

A comparative analysis of this kind is well suited to answer the issues raised in the introduction about (1) the value added and (2) the degree of arbitrariness of simulation models. The *value added* by simulation models can be assessed by comparing the explanatory power of these models with respect to the relevant stylised facts to those using established methods.[12] From the stylised-facts-perspective, the contribution of different models is their ability to explain the stylised facts of the phenomenon. A model that makes a contribution to explaining a stylised fact would, therefore, be regarded as more valuable than another model geared towards side issues (cf. Boland 1987/1994:536). If there are several stylised facts, the contribution of a model increases with its ability to cover more stylised facts and the absence of contradictions relating to the other stylised facts. Marcet and Nicolini (2003:1477-1478) provide a good illustration of this aspect by demonstrating the value added of their model with respect to a list of well-established stylised facts concerning hyperinflations. They show that existing models are not consistent with all of these stylised facts and their model therefore adds value by being the first that explains them all consistently.[13]

2.12

In order to assess the *degree of arbitrariness* of a model, the concept of stylised facts can be used to *focus* evaluation. It offers a means of reducing the complexity in the analysis of a model to a level at which unambiguous statements are possible. A particular stylised fact provides a means of isolating the corresponding generative mechanism (cf. Lawson 1989:62,66) and thereby enables a subsequent validation that is sufficiently focussed to yield significant results. Such an analysis allows the identification of unproductive ways of generating stylised facts, such as the use of unfalsifiable assumptions,[14] and thereby complements the comparative analysis of value added.

2.13

The issue of arbitrariness is of particular importance, because of the extended modelling possibilities offered by computer simulation. In this respect, a mere reproduction of stylised facts (cf. Lawson 1989:61-62) would not distinguish models that introduce productive abstractions from ones that produce almost any result after some "tweaking". In order to separate the productive from unproductive methods of reproducing stylised facts and to suggest a general approach towards the assessment of arbitrariness in simulation model construction, it must be analysed *how exactly* stylised facts are reproduced by the model under investigation. For this purpose, the stylised facts can be interpreted as a spotlight, i.e. with reference to a set of stylised facts, it is possible to scrutinize the model mechanics systematically. The concept offers a clear guide as to how one should "look under the hood" of model results: First, for each stylised fact the mechanism producing it has to be identified. Second, for each mechanism its implementation with regard to (a) model design and (b) calibration has to be checked. The former refers to qualitative ("what are the constitutive elements of a model?"), the latter to quantitative (e.g. parameter values) decisions of the modeller. In the context of such quantitative degrees of freedom, usually assessed by the application of sensitivity analyses, we use the term "robustness".

2.14

Hence, the concept of stylised facts is well suited to provide a basis for the assessment of value added and degree of arbitrariness of simulation models. Moreover, it acknowledges the potential of the simulation method to arrive at richer models of a phenomenon, but at the same time reveals, where this is unnecessary or comes at the price of arbitrariness. Stylised facts therefore provide a solid point of reference to assess a model with regard to the phenomenon it serves to explain, independent from the capabilities of the modelling method used.

Application of the Stylised-Facts-Concept

3.1

To illustrate the usefulness of the stylised-facts-concept, we provide an example of its

application. Our example, the current discussion about the susceptibility of the Groves mechanism to collusion, is taken from our research area of management accounting. Most investigations of the Groves mechanism, a sophisticated incentive scheme for capital allocation, are part of the "economic analysis of accounting" literature. In this research area, game theory based on the rational choice model is currently *the* established method for analysing the behavioural implications of different accounting rules/instruments. Recently, some studies used computer simulation to consider the collusion phenomenon from a different angle. Using this approach, the rigid assumption of the rational choice framework can be relaxed and the dynamics of the collusion phenomenon analysed. Among the four different models that contribute to the debate, two use game theory, i.e. the rational choice model, and the remaining two are based on computer simulations.

3.2

This setting is well suited to illustrate the use of the stylised-facts-concept to investigate the specific contribution of simulation models.[15] We start with an introduction to the debate about the Groves mechanism and collusion (3.3-3.7). Then, we derive the stylised facts of collusion (3.8-3.17) and finally use them to conduct a comparative analysis of the models in order to assess the contribution of simulation (3.18-3.44).

The Current Discussion about the Groves Mechanism

3.3

In order to decide on an effective capital allocation, central management needs information about the investment opportunities in its different divisions. Due to its distance from divisional business, it suffers from an informational disadvantage and must rely on information given by divisional management. Unfortunately, the success of investments depends on many unpredictable factors, so that an estimate of an investment project's rate of return cannot be attributed to any specific factor. This leaves division managers with degrees of freedom, allowing them to pursue their personal interests: When competing for the firm's limited capital, divisional managers may try to use their informational advantage by overstating the returns from their division's investment opportunities as to receive a higher share of the firm's capital. This would increase their division's profit and in turn their compensation[16], but at the expense of the firm as a whole, since other, more profitable, projects whose return rates were not overstated, would receive less capital.

3.4

To overcome this dilemma between individual managerial motivation and truthful reporting, the Groves mechanism, among others[17], has been developed. It was introduced by Groves (1973) and Groves and Loeb (1979) as a method of capital allocation in decentralised firms with limited capital. The aim of the Groves mechanism is to compel managers to report truthfully about the expected returns of their investments and at the same time motivate them to strive hard to fulfil or even exceed their estimates. This is achieved by dividing their compensation into two parts. Accordingly, a division manager's compensation G_i consists of part a, that only depends on the result actually achieved by the manager M_i himself, whereas part b depends on the sum of the reported returns of the other managers (see figure 1).

$$G_i = \underbrace{s_i^a(c_i)}_{a} + \underbrace{\sum_{j \neq i}^{n} s_j^r(c_j)}_{b}$$

Figure 1. A divisional manager's compensation according to the Groves scheme

3.5

To illustrate the Groves mechanism, we use a simple example with two managers competing for capital c. Manager M_1 expects the return s^r_1 and M_2 expects s^r_2 where $s^r_1 < s^r_2$ (e.g. $s^r_1 = 7\%$ and $s^r_2 = 10\%$). Each of the projects needs the total amount of capital c. Given true reports by both managers, their compensation will be: $G_1 = 0 + s^r_2 * c$ and $G_2 = s^a_2 * c + 0$. M_1 will receive no capital and his compensation will be based solely on the reported return rate s^r_2 of M_2. M_2 receives all the capital and a compensation based on his realised return rate s^a_2. Because his report was truthful, and abstracting from risk, this will be the return reported. Now, we examine what happens if M_1 reports an exaggerated expectation $s^{r'}_1 > s^r_2$. Therefore, he will now receive all the capital. The managers' compensation will now be $G'_1 = s^a_1 * c + 0$ and $G'_2 = 0 + s^{r'}_1 * c$. Because $s^a_1 < s^r_2$, M_1 will now receive a *lower* compensation than he would have received if he had reported truthfully. In contrast, M_2 will receive a *higher* compensation than in the case of both managers reporting truthfully (as $s^{r'}_1 > s^a_2$). This result also holds for more managers and more projects, so we can use it to describe the main characteristic of the Groves mechanism: independent of the reports of other managers, a manager will always receive the highest compensation when reporting truthfully. Consequently, the Groves mechanism achieves its goal of giving managers an incentive to report truthfully.

3.6

This result has to be reconsidered, since the analysis has, so far, not taken into account the effect of coordinated behaviour among divisional managers. We use our above example and assume that the amount of total capital is now 2*c and therefore supports two projects. The return rates of four available projects (each requiring c units of capital) are $s_{1.1} < s_{2.1} < s_{1.2} < s_{2.2}$ and therefore with M_1 (responsible for $s_{1.1}$ and $s_{1.2}$) and M_2 (responsible for $s_{2.1}$ and $s_{2.2}$) reporting the truth, the compensations will be $G_1 = s^a_{1.2} * c + s^r_{2.2} * c$ and $G_2 = s^a_{2.2} * c + s^r_{1.2} * c$. If both managers now increase the expected returns stated in their reports by the same amount $\epsilon > 0$ (e.g. 4%), the compensations will rise by $c * \epsilon$. Because no change in the allocation of the capital will take place, the firm's total profit will not decrease, but the managers will receive a higher proportion of it. Therefore, if the managers succeed in finding a way to agree on jointly reported increased return rates, what we will call collusion, they can raise their compensation without producing higher outputs. However, it should be noted that the above characteristic of the Groves mechanism still holds: if one manager withdraws from collective exaggeration, he can increase his compensation, so that collective exaggeration is structurally unstable.[18]

3.7

This question of the stability of collusive agreements using the Groves mechanism has been addressed in a debate with consecutive contributions by Budde, Göx and Luhmer (1998), Kunz and Pfeiffer (1999), Krapp (2000) and Deliano (2000). The first two contributions use models based on classical game theory, whereas the latter two authors use the simulation method. Before these models can be compared, the stylised facts that serve as basis for the comparative analysis have to be derived.

Derivation of Stylised Facts of Collusion

3.8

In order to identify stylised facts about the stability of collusion, we must gather empirical observations of this phenomenon and subsequently extract the conspicuous broad tendencies. As a necessary first step, we need to establish a precise definition of the term "collusion" and then identify sources of relevant empirical observations. Finally, we derive stylised facts on this

basis.

3.9

The term collusion was introduced in antitrust economics.[19] It refers broadly to cooperative behaviour between firms, which yields higher profits for the colluding firms at the expense of the rest of the economy (cf. Roberts 1987/1994:482). Since most forms of collusion are unlawful, explicit contracts are almost non-enforceable and might even be used as evidence in trial. However, under specific conditions, no contracts are needed to facilitate collusive behaviour. Loose agreements or even mutual signalling are sufficient. Because a firm applying the Groves mechanism would not tolerate contractual agreements, we focus on this implicit form of collusion. Hence, we can adapt the more general definition from antitrust economics to our purposes: collusion is the cooperation of individuals to maximise their own profits against and at the expense of an institution without an enforceable contract. This definition is consistent with most of the literature on collusion (cf. Che and Kim 2004:2) and at the same time covers the relevant features of collusion in the context of the Groves mechanism, i.e. that a group of managers colludes against their firm to increase their compensation without the possibility of enforcing the collusive behaviour of peers by contract, breaking the (at least implicit) rules of the firm and impairing its financial resources.

3.10

Unfortunately, there has not yet been a broad discussion in the scientific community concerning the stylised facts of the collusion phenomenon. Consequently, we have to start with the identification of the patterns emerging from the existing empirical data of the collusion phenomenon to derive a list of stylised facts. We are aware that this limits the quality of the stylised facts, since a broad discussion of the stylised facts in the community of experts is desirable. In addition, the empirical studies have not been conducted with the explicit aim of identifying stylised facts, so that a lack of empirical data in some areas has to be expected. However, we believe that a careful and transparent derivation of stylised facts on the basis of the existing studies is an unavoidable first step in the cumulative progress of research in the field. Moreover, the analysis remains meaningful for the sake of illustration and as a contribution to the discussion in the field, especially because it can serve as blueprint for subsequent analyses.

3.11

For the derivation of stylised facts it seems obvious to look at empirical studies of the Groves mechanism. To our knowledge, only the experiment of Waller and Bishop (1990) uses empirical methods to examine the Groves mechanism. However, such a narrow approach would not take into account that from the perspective of collusion as an empirical phenomenon, the Groves mechanism merely provides a specific setting for collusive behaviour. Collusion, according to our definition, has also been studied in other settings ranging from auctions to oligopolistic competition, regulation, and internal organisations (cf. Che and Kim 2004:2). Consequently, broad tendencies about the stability of such collusive behaviour can be observed there as well. This allows us to draw on a broader set of empirical data, including triangulation between statistical analyses, experiments and case studies. In addition, the integration of different perspectives and research backgrounds is particularly helpful as specific facts cannot be observed directly in all settings.[20]

3.12

The identification of broad tendencies in a set of individual empirical findings and the subsequent condensation to a list of stylised facts includes inductive elements that necessarily involve subjective judgement.[21] As discussed in the methodological section, we address this issue by making our reasoning as transparent as possible. We draw on 12 articles reporting empirical results about the stability of collusion in various economic settings.[22] From these sources, we extracted all findings about collusion as a basis for the derivation of stylised facts. In order to cope with the diversity of these findings and to provide a basis for understanding our judgement, we grouped the findings along the constitutive elements of the collusion

phenomenon: (1) the characteristics of the colluding *individuals/firms*, (2) the *situation* in which collusion potentially occurs and (3) the *course of events* leading to or destabilising collusion. For each finding, we state the implications leading to our stylised facts. Findings with particular complex implications not fitting into any stylised fact were treated separately in the group (4) *miscellaneous aspects* to discuss them in greater detail. See table 1 for the grouped list of findings, which is followed by a discussion showing the derivation of stylised facts. For the sake of readability the sources of the findings were moved to a separate table. To link the findings in table 1 to their sources each finding is followed by a number. This number can be used to identify the source of the finding in table 3 in the appendix.

Table 1.1: Empirical findings about the stability of collusion by cluster

Cluster	Findings about Collusion	Implications	Stylised Fact
Individuals/ Firms	The setting that stabilises collusion most strongly is a relatively small number of competing firms. (F13, *statistical analysis*)	Small group size stabilises collusion.	SF "Group Size"
	"Previous [experimental, the authors] studies indicate that collusion sometimes occurs in duopolies, but is very rare in markets with more than two firms." The same tendency is observed in the current experiments. (F20, *experiment*)	Small group size stabilises collusion.	
	There is high concentration in the market: the four colluding firms have a market share of 99%. (F22, *case study*)	High concentration implies that a small number of firms dominate the market, stabilising collusion.	
	Even if there are about 50 buyers in total, if there are few large buyers, there is collusion among the large buyers. (F1, *case study*)	High concentration implies that a small number of firms dominate the market, stabilising collusion.	
	Collusion is more likely when the affected firms have a large market share. (F8, *statistical analysis*)	High concentration implies that a small number of firms dominate the market, stabilising collusion.	
	Even with many participants (nine), collusion can occur if the product characteristics limit cross-regional shipping. (F21, *case study*)	Limited transportation implies the formation of sub-markets with a small number of participants ("niches") and therefore stabilises collusion.	
	Tendering procedures for individual projects support collusion even in markets with many participants because only a few firms are participating in the tendering. (F3, *case study*)	A single tendering procedure can be considered as a sub-market, in which a small group size stabilises collusion.	

Table 1.2: Empirical findings about the stability of collusion by cluster

Cluster	Findings about Collusion	Implications	Stylised Fact
Situation	"Collusion often breaks down in the last round." (F17, *experiment*)	The imminent and certain termination of the interaction destabilises collusion.	SF "Time Horizon"
	As the discount rate increases, the percentage of colluding subjects declines. (F6, *experiment*)	An increase in the discount rate shortens the relevant time horizon, thereby destabilising collusion.	
	Market environments relatively free of complications such as product diversity or demand variability stabilise collusion. (F14, *statistical analysis*)	Uncomplicated market settings stabilise collusion.	SF "Setting"
	More information about the market (demand and cost functions) favours collusion, as players tend to use strategies that are the best responses to the behaviour of the other firms in the previous round. (F15, *experiment*)	The more information is available, the more effectively actors can avoid or handle complications in the market setting. Therefore, information about the market stabilises collusion.	
	The empirical results show that collusion is more likely when the firms in question are regionally concentrated. (F11, *statistical analysis*)	Short distances between the participating firms reduce complexity in terms of regional diversity and allow for easy monitoring, thus stabilising collusion.	
	The empirical results show that collusion is more likely when there are large cooperatives in the market. (F12, *statistical analysis*)	The presence of large umbrella organisations reduces market complexity for the relevant players.	
	More information about quantity supplied and profits of other firms in the market leads to increased competition as firms tend to increase quantity to imitate the rival firm with the highest profit (in comparison to aggregate information only). (F16, *experiment*)	In the experimental setting reported, specific information inhibits the formation of collusion, because the firms try to imitate the most successful player.	

Table 1.3: Empirical findings about the stability of collusion by cluster

Cluster	Findings about Collusion	Implications	Stylised Fact
Situation	Potential risk reduction in terms of pricing and fluctuations in the workload favours collusion. (F5, *case study*)	Benefits of collusion include risk reduction with respect to pricing and capacity utilisation.	SF "Benefits"
	Collusion is more likely when market demand is price inelastic. (F7, *statistical analysis*)	Price inelasticity implies that higher prices do not affect demand proportionally, leading to higher benefits	

		from collusion that sets higher prices.
	At the time the collusion commenced, prices rose significantly although raw material prices remained unchanged. (F23, *case study*)	A high impact on prices indicates that collusion yielded substantial benefits.
	When collusion broke down, market prices declined substantially. (F25, *statistical analysis*)	A high impact on prices indicates that collusion yielded substantial benefits.
	Collusion is more likely when the importance of the particular market for the sellers (in which collusive arrangements are proposed) is neither excessively low (the market is not really important to the seller) nor excessively high (there are no other potential markets). (F*10, statistical analysis*)	Very low importance implies minimal benefits to be gained. Excessively high importance implies high risk, potentially offsetting the benefits of collusion.

Table 1.4: Empirical findings about the stability of collusion by cluster

Cluster	Findings about Collusion	Implications	Stylised Fact
Course of Events	In experiments with fixed pairs, partial collusion can be observed, but this does not apply to randomly matched pairs. (F18, *experiment*)	Constant group composition stabilises collusion.	SF "Group Composition"
	New and different participants lower the risk of collusion in auctions. (F2, *statistical analysis*)	Inhomogeneous and changing group composition destabilise collusion.	
	Participants in collusive agreements try to avoid destabilisation because of market entry by collective acting. (F4, *case study*)	New entrants destabilise collusion.	
	Collusion is more likely when crops have long cultivation times. (F9, *statistical analysis*)	Long cultivation times represent a barrier to entry, thereby stabilising group composition.	
	After a price war, collusion was successfully re-established and maintained for a relatively long period. (F26, *statistical analysis*)	The reported price war can be interpreted as means of demonstrating the ability to harm defectors. The increased credibility of the threat to punish defectors stabilises collusion.	SF "Enforcement"
	When the cartel suspected that cheating had occurred, the cartel cut prices for a time so as to enforce behaviour consistent with the collusive agreement, and then returned to collusive price. (F24, *statistical analysis*)	The price cut demonstrates the ability to harm defectors, establishing a credible threat which stabilises collusion.	

Miscellaneous Aspects	Sequential (in comparison to simultaneous) price setting destabilises collusion in duopolies. (F19, *experiment*)	The reported experimental results are based on the models of Stackelberg and Cournot, implying a very specific setting that restricts the validity of the findings. Therefore, the findings do not constitute stylised facts.	See text

3.13

In the *first* cluster "individuals/firms", F13 and F20 refer to the influence of group size on the stability of collusive agreements. The three findings F22, F1 and F8 all constitute evidence that high concentration in a market favours collusion. This can be linked to the first two findings: Because the influence of small players in concentrated markets is low, the market is dominated by a few large players. Concentration therefore reduces effective group size. Another way of reducing the effective group size in crowded markets is the formation of niches referred to by F21 and F3. We can integrate these findings into the following statement: small effective group size stabilises collusion (SF "Group Size"). This stylised fact is supported by seven findings based on all three empirical methods.[23]

3.14

In the *second* cluster "situation", different aspects to which the findings refer can be distinguished: (1) relevant time horizon, (2) complexity of the setting and (3) potential benefits.

1. Concerning time horizon, F17 refers to the destabilisation of collusion by an imminent and certain termination of the interaction. F6 addresses a similar aspect, since a higher discount factor implies a shorter relevant time horizon. The following statement integrates these findings: long time horizons stabilise collusion (SF "Time Horizon"). This stylised fact is supported by comparatively few findings, all of them based on experiments.[24]
2. With respect to the complexity of the setting, F14 (which is already based on data of different industries) states, that uncomplicated market settings favour collusion. This tendency is supported by more specific findings such as the availability of market information (F15), short geographic distances between the firms (F11) and the existence of large umbrella organisations (F12), all referring to aspects that reduce or allow a more efficient coping with market complexity. These findings can be condensed into the following statement: uncomplicated settings (in terms of low dynamics, low complexity, high transparency etc.) stabilise collusion (SF "Setting"). F16 points in the opposite direction by reporting an experimental finding where more information destabilises collusion. Since the experiment considers a very specific type of information corresponding to particular economic models, a contradiction to the tendency above seems acceptable. SF "Setting" is supported by a substantial number of findings from experiments and statistical analysis.[25]
3. Regarding the potential gains from collusion, F5 directly addresses the benefits of collusion, in the form of reduced risk in pricing and capacity utilisation, that stabilise collusion. The price inelasticity of demand (F7) also implies high profit potential when firms manage to raise prices by collusion. F23 and F25 both report that switches between collusive and non-collusive behaviour have a high impact on prices, indicating substantial benefits from successful collusion. F10 can be interpreted in the same way, indicating that potential benefits from collusion decline with a lower level of importance of the respective market and higher risk.[26] The four findings based on statistical analyses and case studies allow us to derive SF "Benefits": high attainable benefits stabilise collusion.[27]

3.15

The *third* cluster "course of events" contains all findings on the development of collusive behaviour over time, addressing the aspects of (1) group composition and (2) enforcement.

1. The stabilising effect of constant group composition on collusion over time is addressed by F18, F2 and F4. This tendency is supported further by F9, because high barriers of entry keep out new entrants. Supported by an experiment, a case study and statistical analyses, we state SF "Group Composition" as follows: stable group composition stabilises collusion.
2. The observation of price wars for a limited period of time in F26 and F24 can be interpreted as a demonstration of the ability of colluding firms to harm defectors. This stabilises collusion by establishing a credible threat that defectors from collusion will be punished. Thus SF "Enforcement" states: effective enforcement strategies stabilise collusion. Both findings are based on statistical analyses of a specific industry.[28]

3.16

Finally, the finding F19 must be discussed under "miscellaneous aspects", as the underlying experiment is based on models of Stackelberg and Cournot duopolies. The result relating to the destabilising effect of sequential price setting is therefore restricted to a very specific setting.[29] Moreover, there is no further evidence of a similar effect, so that this finding does not qualify as indicator of a broad tendency.

3.17

To summarise, table 2 provides an overview of the stylised facts that have been derived. As discussed, this list represents the judgement of the authors on the basis of empirical research on collusion. For the purpose of this paper, this list should be considered as the result of a transparent procedure applying the methodological statements of Section 2 in practice.[30]

Table 2: Overview of the stylised facts derived

Name of Stylised Fact	Description
SF "Group Size"	Small effective group size stabilises collusion.
SF "Time Horizon"	Long time horizons stabilise collusion.
SF "Setting"	Uncomplicated settings (low dynamics, low complexity, high transparency etc.) stabilise collusion.
SF "Benefits"	High attainable benefits stabilise collusion.
SF "Group Composition"	Stable group composition stabilises collusion.
SF "Enforcement"	Effective enforcement strategies stabilise collusion.

Model Assessment

3.18

In this section, we assess the four models introduced in 3.7 in order to evaluate the contributions of the simulation models. The structure of assessment corresponds to the issues raised in the introduction: Firstly, the explanatory power of the different models with reference to the stylised facts will be assessed to allow for a specification of the value added offered by simulation models (3.19–3.32). Secondly, we address the arbitrariness issue by using stylised facts to analyse the way the models produce their results in terms of model design and parameter calibration (3.33–3.44).

Specification of the Value Added Offered by the Simulation Models

3.19

To analyse the value added, we will look sequentially at the different models that analyse the stability of collusive agreements under the Groves mechanism. In each case, we start with a condensed description of the model and then discuss which stylised facts the model relates to and finally investigate to what degree these model results fit the stylised facts. Finally, we consider the value added offered by the simulation models. Before the detailed analysis, figure 2 provides an overview of the comparative analysis, relating the results of each model to the different stylised facts. The first two models use traditional game theory, the other two the simulation method. This structured overview already demonstrates that only four stylised facts can be linked to the entire collection of models, which indicates that there is still considerable potential for future work. Additionally, this "research landscape" seems, at first glance, to be quite symmetrical in terms of contributions by method. However, closer inspection shows that this is not correct.

	Analyses			
	Traditional Game Theory		Simulation	
Stylised Facts	Budde/Göx/ Luhmer (1998)	Kunz/Pfeiffer (1999)	Krapp (2000)	Deliano (2000)
SF "Group Size"	☐	---	---	◘
SF "Time Horizon"	---	◘	◘	---
SF "Setting"	---	---	◘	---
SF "Benefits"	---	---	---	---
SF "Group Composition"	---	---	---	---
SF "Enforcement"	---	◘	■	☐

--- SF not addressed
☐ SF addressed, with contradicting results
◘ Results reproduce SF to some extent
■ Results according to SF

Figure 2. Overview of comparative analysis

3.20

Budde, Göx and Luhmer (1998) use standard game theory to show that the problem of collusion within the Groves scheme has a prisoners' dilemma structure for the two managers in question. Looking at a one-period game of this form, they come to the conclusion that the Groves mechanism is resistant to collusion (cf. Budde, Göx and Luhmer 1998:13-15).

3.21

The model can be related only to SF "Group Size". The authors explicitly discuss the influence of group size and demonstrate that there will be no stable collusive agreements for n players in a one-shot game (cf. Budde, Göx and Luhmer 1998:17-18). This result contradicts SF "Group Size" and moreover suggests that collusion is not a problem at all for the Groves mechanism.[31]

3.22

Kunz and Pfeiffer (1999) refer directly to this result and construct a multi-period model to show that the Groves mechanism can lose its collusion resistance, assuming repeated

interactions among managers. They modify Budde, Göx and Luhmer's prisoners' dilemma model by allowing one manager to threaten the other to terminate cooperation (threat model) or to build up a reputation for cooperative behaviour (reputation model). In the *threat* model, managers additionally know with a certain probability whether their interaction will last for another round.[32] Because of the iterated structure of the game, a manager can now "punish" the other manager by defecting for the rest of the interaction. Threatening with this "trigger" strategy as an extreme form of punishment can be used as a means of enforcing collusive agreements. In this setting, the authors prove that collusive agreements can be implemented, if the probability does not exceed an endogenously determined threshold (cf. Kunz and Pfeiffer 1999:213-215). The *reputation* model is based on the assumption that both managers know that, with a probability greater zero, the other participant is behaving cooperatively, i.e. is following the "tit for tat" strategy.[33] Using the concept of sequential equilibrium[34], they show that the number of rounds in which both parties defect, is limited by a constant k. If t, as the number of rounds to be played, exceeds k, managers can cooperate for at least t-k rounds (cf. Kunz and Pfeiffer 1999:215-217). Therefore, collusive agreements can emerge at least for some periods.

3.23

The results of the two models can be related to SF "Time Horizon" and SF "Enforcement". The *threat* model shows that if the probability of ending the game is below a certain threshold, collusive behaviour offers a higher payoff to the manager than reporting truthfully. This result links directly to SF "Time Horizon", because a decreasing probability can be interpreted as an increasing time horizon. Moreover, using the "trigger" strategy for punishment can be related very generally to SF "Enforcement". The *reputation* model shows that in earlier rounds, there is a strict incentive for a manager to imitate the behaviour of a "tit for tat" player, which induces cooperative behaviour on the part of the other manager (cf. Kunz and Pfeiffer 1999:217). This leads to mutual co-operation for at least t-k rounds. This can be interpreted as establishing a reputation for acting cooperatively and also reproduces a very general aspect of SF "Enforcement".

3.24

Krapp (2000) extends this strand of literature in at least two important directions:[35] Firstly, he analyses the influence of different strategies relating to the stability of collusion. For this purpose, he uses the approach of Axelrod (1984) and simulates a tournament of 16 plausible strategies, encompassing, besides "tit for tat" and "trigger" already used by Kunz and Pfeiffer, also others such as "constant defection" or "constant collusion".[36] These different strategies play a round robin[37] and are ranked according to their performance (cf. Krapp 2000:274-275). Secondly, Krapp introduces "noise" to the model, which makes it more difficult to sustain cooperation: a manager cannot observe directly whether or not the other has cooperated. This is modelled by introducing uncertainty, i.e. output is not only a result of a manager's efforts, but also subject to stochastic influences. Therefore, a manager can only *estimate* whether the other manager did behave cooperatively or not (cf. Krapp 2000:267).

3.25

The results of this simulation model can be related to SF "Enforcement", SF "Setting" and SF "Time Horizon". A first general result with respect to SF "Enforcement" is that agents reacting to the actions of the other manager are more successful than agents with independent behaviour (e.g. constant collusion). Complementary to this aspect, it allows for a detailed analysis of the impact of different strategies on the ability of managers to extract surplus from collusive arrangements (cf. Krapp 2000:269-271,274-275). Moreover, it is possible to examine specific interaction processes with "high resolution" and to analyse their development in depth, which yields additional insights. For instance "tit for tat" does not perform particularly well in this tournament, due to the fact that it may punish even though the other player did cooperate. Such misjudgements are due to the stochastic influence on the outcome and the strategy is not well adapted to this problem (cf. Krapp 2000:269-271).[38] This result can be related to the SF "Setting", because the introduction of uncertainty as a situational property increases the

complexity of signal exchange about the willingness to collude. These features can be considered as significant value added by the simulation model, compared to the traditional models. Furthermore, SF "Time Horizon" is addressed by a second tournament including a discount factor for the payoffs. This destabilises collusion, reproducing the general aspect of SF "Time Horizon". However, the model does not investigate the influence of different specifications of the discount factor. Additionally, the simulation model facilitates the analysis of the interaction of SF "Time Horizon" and SF "Enforcement": Strategies which attempt to establish reputation are more successful in a setting with no discounts on future payoffs, i.e. reputation mechanisms do particularly well for a long time horizon.[39] This ability to address such interactions can be considered as another substantial added value provided by the simulation model.

3.26

Deliano (2000) uses a different simulation-based approach to overcome the restriction of a limited and discrete set of strategies in Krapp's model. In his model, the managers' behaviour emerges from six characteristics: their impatience, their tendency to punish, forgive, repent, conciliate and trust.[40] At the beginning of a simulation run, the characteristics of the agents are set randomly after which agents adjust them at the end of each round in order to increase their payoff. The simulation has three settings: In the first setting, the agents' characteristics remain constant. These results serve as a point of reference to compare the effects of two learning procedures (cf. Deliano 2000:51-53): In the second setting, agents adjust their characteristics (leading to a change in behaviour) only on the basis of their own history ("individual learning"). In the third setting, the agents can further improve their strategy by also taking into account the past behaviour of the other agents and their respective payoffs, which increases their "knowledge basis" using the experience of the others ("group learning").

3.27

The model results can be related to SF "Group Size" and SF "Enforcement". Deliano reports for various different settings, that an increased group size decreases the tendency towards collusive behaviour.[41] These results conform to SF "Group Size". Comparing the different forms of learning, the percentage of individual managers reporting in a collusive manner decreases, proceeding from "no learning" to "individual learning" and even further to managers using "group learning" (cf. Deliano 2000:133). But looking at the duration of periods, the length of continuous collective collusion increases with the change from no learning to group learning (cf. Deliano 2000:135). The same tendency can be observed even more clearly for the duration of collective truthful reports (cf. Deliano 2000:132). Such increasingly long phases of one type of behaviour might point to a stabilising effect of learning that applies to both collusive and truthful behaviour. This relates to SF "Enforcement" with the interpretation that learning ability allows managers to perceive the disadvantages of not colluding. Unfortunately, Deliano does not investigate, whether the payoffs accumulated over longer collusive phases could compensate for the losses incurred through a decrease in the share of collusive periods.[42] Therefore, the model does not allow to infer a fit with SF "Enforcement". This lacking payoff inclusion also suggests a missed opportunity to address SF "Benefits". Likewise, Deliano's model allows "exchanging" one of the managers after a certain number of rounds by resetting the learning characteristics, but does not study any effects of a change in group composition which could be related to SF "Group Composition" (cf. Deliano 2000:86-87).

3.28

We are now in the position to summarize the results: Both traditional game theory and simulation models can be related to SF "Group Size", SF "Time Horizon" and SF "Enforcement". In addition, Krapp's model can be linked to SF "Setting". Moreover, the simulation models could potentially address the blind spots SF "Benefits" and SF "Group Composition". However, it should be noted that not all models *fit* the stylised facts to which they can potentially be related: The results of Budde, Gök and Luhmer contradict SF "Group Size" and Deliano's model covers enforcement issues, but does not allow any inferences with regard to SF "Enforcement".

3.29

The methods also differ in the depth and generality with which the stylised facts can be addressed. Though Kunz and Pfeiffer demonstrate the existence of an impact of enforcement strategies on the stability of collusion, they are unable to analyse the influence of different enforcement strategies. This is done in considerable detail in Krapp's simulation model. Hence, one central value added of simulation models is their ability to take a closer look at the dynamics and their detailed account of different interaction processes. However, this possibility has only been utilised for SF "Enforcement" so far.

3.30

More generally, one might conjecture that the game theoretical models have already been stretched to the limit,[43] while simulation models have the potential to address stylised facts more extensively. In this respect, it should be noted that Krapp's simulation model can be related to three stylised facts, while Kunz and Pfeiffer had to construct separate models which contributed to the analysis of only two (SF "Time Horizon" and SF "Enforcement"). Additionally, Krapp showed that there is an interaction between these two. Hence, one can expect that computer simulation models facilitate an analysis of more stylised facts within one model, when game theoretical models fail to remain tractable. This potential value of simulation models is of particular importance considering that there are areas of macroeconomics where single models are able to address and reproduce all currently discussed stylised facts (cf. Marcet and Nicolini 2003:1478).

3.31

To exemplify and support the above claims about the potential value of simulation models, we show (1) how the blind spots could be addressed and (2) how all SF could be integrated in one model.

1. Concerning the *blind spots* SF "Benefits" and SF "Group Composition", we first name the problem to be solved in order to model these two SF, then outline the ability of simulation models to do so and point out examples of existing simulation models based on similar mechanisms. The SF "Benefit" is linked in all four models to the payoff structure of the prisoners' dilemma (PD). Different payoffs for identical behaviour of the agents can be seen as different parameterisations of the PD payoff structure. To analyse the impact of the parameter setting on the outcome of the model, the sensitivity of the results to the parameters has to be identified. Using simulation a variation of the parameters is easy to implement by having different simulation runs for the different values of the parameters.[44] The SF "Group Composition" addresses the effect of exchanging agents. To assess the implications of such a disturbance in the situation of the other agents, one has to analyse the impact of an unpredictable change of behaviour of one of the agents after all agents have adapted to the behaviour of one another. In simulation models, no assumptions concerning the behaviour of the agents (like "rational choice") have to be made. Therefore, in order to analyse the impact of exchanging agents in a simulation model, only the relevant attributes of the agents have to be changed. As the behaviour of agents can be implemented in the model as a variable, modelling the SF "Group Composition" is a straightforward operation.[45]
2. To illustrate how all the different SF could be *integrated into one model*, we look at the general problems arising from the integration of different effects and how these can be overcome by using the simulation approach. Subsequently, we give an outline of how such a model could look like for our problem under examination. When integrating different mechanisms into one model, the researcher faces the challenge of an increasing amount of interactions to be taken into account. Due to the higher number of interactions the dynamics tend to increase and analytical approaches are often not able to solve the problem or have to fall back on additional assumptions like the existence of one player using the tit-for-tat strategy to find solutions for the model (cf. Kunz and Pfeiffer 1999:215). Simulation instead offers nearly unlimited capacity to model interactions. Therefore additional assumptions are not necessary to find solutions for the model.

However, the limited ability of researchers to understand very complex models has to be taken into account. As the model of Krapp already integrates three of the six SF, we take it as a starting point for an outline of an integrated model for the problem under examination. In the following paragraphs we describe the mechanisms that could be used to integrate the other SF. The integration of the SF "Benefits" can be realized straightforward by a variation of the payoff structure included in the model of Krapp (cf. Krapp 2000:265), as pointed out before. To include the SF "Group Size", the mechanism must allow for more than two agents to be included in the model. One approach is to select one agent and let him play against the average of the remaining n-1 agents. This reduces the n-agent problem to a two-agent problem. This approach seems suitable as for many situations the agents are not really able to see all individual decisions of their peers, but are confronted with the average outcome. To address the SF "Group Composition", first the implementation of the SF "Enforcement" has to be modelled differently, allowing the agents to adapt their strategies to their environment. This adaptation is part of the generating mechanism of the SF "Group Composition" as shown above. As Krapp (2000:268) uses 16 *fixed* strategies to test for the chance of collusive behaviour to be successful, no adaptation can take place. An example for an adaptation mechanism is Axelrod (1997). He uses evolutionary learning to search for the "best" strategy to be used for the PD. Using such an approach, changes in group composition can be modelled by replacing some of the agents that already have survived several iterations of evolutionary selection by agents following randomly generated strategies.

3.32

The model outlined would be able to address all six SF drawing on the advantages of the simulation method. As the model is path dependent (the behaviour of an agent depends on the previous decision of the other agents) and there are additional disruptions due to the exchange of agents, no mathematical/analytical approach with similar properties can be conceived of so far.

Assessment of Arbitrariness in Model Design and Parameter-Setting Robustness

3.33

In this section, the suggested contribution of the models is assessed at a second and deeper level. Now, the "model mechanics" behind each stylised fact become relevant. For this purpose, each stylised fact can be interpreted as a *spotlight* isolating the respective generative mechanism in each model, thereby facilitating an in-depth discussion of the degree of arbitrariness in model construction and parameter-setting robustness.

3.34

The following assessment examines the models at three different levels. We start by specifying the stylised facts to which the model can be related. From this perspective, we identify the respective generative mechanism in the model and check the model design by considering whether the assumptions made with respect to the implementation of this mechanism are well founded or rather arbitrary. Finally, if parameters have to be set, their calibration is scrutinised. Based on the information given by the authors we assess the sensitivity of the relevant results.

3.35

The model of Budde, Göx and Luhmer (1998) can be related to SF "Group Size". However, it does not reproduce the well-established observation that an increase in group size decreases the stability of collusion. The generating mechanism for this result is the incentive structure of the situation, combined with the assumption of rational behaviour. In particular, a one-shot prisoners' dilemma is modelled and it is shown that an extension in group size to more than two players does not change the equilibrium outcome (cf. Budde, Göx and Luhmer 1998:17-18). In principle, the design of the model is well established, assumptions are very general and the model has been applied to a wide range of problems.[46] However, the model does not have the appropriate level of abstraction to address the stylised facts. Reducing the phenomenon to a

one-shot prisoners' dilemma can, at best, be understood as highlighting the basic incentive structure of the problem and therefore providing a stepping stone for the subsequent discussion. As there are no parameters involved, robustness is not an issue.

3.36

The *threat* model of Kunz and Pfeiffer (1999) can be related to SF "Time Horizon" and SF "Enforcement". The mechanism behind the former is implemented simply by introducing a parameter for the probability that the interaction will be continued in the next period. Lowering the probability can be interpreted as shortening the time horizon of the managers. This is applied widely and can be considered as well founded in the literature. With respect to the calibration of the parameter, the model is robust, because, as for the full range of possible values, the model produces outputs that can be interpreted in accordance with SF "Time Horizon". However, the primary reason for introducing this parameter is different. According to Kunz and Pfeiffer (1999:215), it is implemented to overcome the technical problem that infinite games can, according to the folk theorem, have many equilibria.[47] The mechanism related to SF "Enforcement" is the "trigger" strategy. It is an extreme form of punishment which can be used by a manager to threaten the other and thereby enforce collusive behaviour. The identification of the "trigger" strategy is straightforward, as there are only two ways for a manager to behave: offering cooperation or not doing so. The constant choice of the non-cooperative behaviour for the rest of the game is, in this respect, the clearest form of punishment. This assumption can also be regarded as valid, at least as a first step in modelling enforcement issues. However, this model can only cover the extreme point in a wide spectrum of punishment strategies.

3.37

The *reputation* model of Kunz and Pfeiffer (1999) can only be linked to SF "Enforcement".[48] The constitutive elements in this respect are a probability parameter and the "tit for tat" strategy, which the authors consider to be a robust means to build reputation (cf. Kunz and Pfeiffer 1999:215). The probability parameter is a probabilistic description of the belief of one manager as to whether the other manager is a "tit for tat" player. Using these elements in combination with the concept of sequential equilibrium, it is shown that collusive behaviour is a dominant strategy, at least in t-k rounds (cf. Kunz and Pfeiffer 1999:216). The "tit for tat" strategy has been extensively examined by Axelrod (1984:17-39) and can be considered as generally accepted. This is also true for the concept of sequential equilibrium, although it is quite sophisticated in terms of information requirements (cf. Kreps and Wilson 1982). With respect to calibration, the model is robust, because changes only occur if the parameter changes its value from zero to a value greater than zero. Again, this parameter seems to be introduced primarily for technical reasons, thus enabling the "reputation" mechanism (cf. Kunz and Pfeiffer 1999:217).

3.38

The model of Krapp (2000) can be related to SF "Enforcement", SF "Setting" and SF "Time Horizon". The model design behind the enforcement mechanism is the simulation of different strategies for sustaining collusion. Simulating a tournament of different behavioural strategies, in contrast to the computation of equilibria, allows an observation of the dynamics of interaction between two managers. Because the tournament approach investigates all possible combinations, only the selection of strategies is not exhaustive and could, for this reason, be considered as arbitrary. However, these strategies consider the basic strategies used by Axelrod (1984:192-205) and were the result of two rounds of competitions developed by the scientific community. Krapp adopts them in conformity with his objectives, where the mechanism for SF "Setting" comes into play (cf. Krapp 2000:268). In particular, his adaptation considers difficulties in communication by introducing "noisy" signals. This method of modelling uncertainty is common in agency theory (cf. Milgrom and Roberts 1992:214-218). The calibration of the model is achieved by setting parameters for the participation constraint, the risk aversion of the managers and the variance of random influence on their compensation (cf. Krapp 2000:269). No sensitivity analysis is conducted for changes in these parameters and no information is given

about the variance of tournament results.[49] Therefore, it cannot be determined without rebuilding the model, if the model is robust with respect to these parameters. Finally, in a sub-model, a mechanism which can be related to SF "Time Horizon" can be identified. Its implementation is again quite straightforward and achieved by discounting the future profits (cf. Krapp 2000:269). For the same reasons given for the game theoretical models, this is methodologically sound. However, its value is set to 3% without any reported sensitivity analysis. In summary, Krapp makes solid assumptions to produce additional insights on enforcement dynamics, but unfortunately, the setting and sensitivity of some key parameters is not sufficiently transparent.

3.39

The simulation model of Deliano (2000) can be related to SF "Group Size" and SF "Enforcement". With respect to the latter, the model design is based on a selection of six characteristics and two different forms of learning. The characteristics serve as a starting point to endogenously generate the behaviour of managers in the simulation. The selection of the characteristics is made with reference to Macy (1996) and in part to Axelrod (1984) (cf. Deliano 2000:49). However, no arguments are given as to why these characteristics from different sources *in combination* should be considered as a solid basis. The implementation of the two forms of learning is even more arbitrary: Firstly, a complex model with many parameters is introduced (cf. Deliano 2000:62-67). Subsequently, these parameters (e.g. the number of search runs, length of the memory of agents, degree of adaptation in each run etc.) acquire a "plausible" value (cf. Deliano 2000:83-86). Hence, in many respects, the model design is highly arbitrary.[50] Furthermore, the parameter setting is not tested for sensitivity with respect to the relevant results (cf. Deliano 2000:86-87). For some parameters, the author only informs the reader about the selected value and for others, he argues vaguely with reference to pretests without giving any clear results.[51] Due to this parameter selection, the model's robustness for SF "Enforcement" might be also doubted. Already on a questionable basis, the mechanism behind SF "Group Size" is implemented by including different numbers of agents in the simulation. In terms of design, modelling additional managers is quite sound methodologically, as no additional assumptions are made. Only the number of managers is increased and the additional managers are implemented identically. In comparison to the other models, this can be considered as a generally useful feature of the simulation model (see above). In terms of calibration, only the settings for two and four managers are investigated. No reason for the selection of these values is given (cf. Deliano 2000:84). However, it could be worthwhile to analyse the effects of increasing the number of managers in more detail in order to acquire a better understanding of the relationship and to demonstrate the model robustness with respect to parameter setting. Moreover, it should be kept in mind that these results, although they can in principle be related to SF "Group Size", are based on a model that is open to criticism because of arbitrariness in model design and a lack of robustness in parameter choice. Comprising our assessment of Deliano's model, we believe that he does not handle the increased degrees of freedom provided by simulation very well.[52]

3.40

We are now in the position to summarize our discussion of arbitrariness and robustness. The concept of stylised facts also contributed to the assessment of the models at a deeper level by *focussing* the discussion. Based on the respective stylised facts a model can be referred to, the relevant model mechanics were identified. Subsequently, these were tested for arbitrariness of assumptions and robustness of parameter setting.

3.41

The results are heterogeneous, but some tendencies stand out. The models based on game theory are able to apply an established and familiar method. Because they can apply such standards, they do not appear to involve much arbitrariness at first sight. However, the model of Kunz and Pfeiffer demonstrates well that the capabilities of traditional game theory are sometimes stretched to the limit (see above on SF "Enforcement"). In this respect, they have to restrict themselves to the "trigger" strategy to model a concept of credible threats or assume the

existence of a "cooperative" manager who makes the reputation mechanism work. This mechanism is based on the sophisticated, but restrictive concept of sequential equilibrium. In comparison, the tournament approach investigating the dynamics of interaction of different strategies fits more naturally to the stylised fact and appears less laboured. However, such a fit is not guaranteed, given in particular the lack of established standards. Deliano's model is a case in point, where considerable problems with respect to arbitrariness in model design can be identified. However, this problem is not inherent to his different approach per se, but stems from his particular way of constructing the simulation model.

3.42

A second tendency refers to the use of parameters. A successively increasing number of parameters has been introduced from model to model. However, model results have not been checked systematically for sensitivity. While in the model of Kunz and Pfeiffer, these parameters were introduced primarily for technical reasons and can be considered robust in terms of the stylised facts addressable within the limits of their model, such a test would be worthwhile in the two simulation models, because greater effects can be expected. The respective stylised facts represent a useful point of reference in this regard.

3.43

For an interpretation of the results, it is useful to consider the different models as a successive development. With reference to the stylised facts, a decreasing degree of abstraction can be observed in the sequence of models. However, adding additional assumptions in modelling should be considered well, since this creates uncertainty about the necessity of even more assumptions in order to make the model work effectively (cf. Lindenberg 1992:19). A possible strategy for dealing with the increased degrees of freedom is to stay as close as possible to traditional modelling approaches, as done by Krapp. Complementary, simulation researchers can examine whether their models are well constructed with reference to the stylised facts and make the appropriate improvements.

3.44

This leads us to the general conclusion concerning model assessment. By using the concept of stylised facts, one is in the position to identify systematically the contributions of simulation models. In particular, simulation models have the potential to address *all* stylised facts of collusion derived, as depicted in the "research landscape" of figure 2. While game theoretical models had to be stretched to the limit in addressing the dynamics of the problem, simulation allows addressing several "stylised facts" simultaneously. However, given the associated rise in degrees of freedom available to the researcher, it is also necessary to scrutinise the construction of simulation models more deeply. The stylised-facts-concept proved useful for this purpose as well. By using the concept, on the one hand weaknesses in model design and robustness have been identified and on the other hand strengths have been highlighted.

Conclusion

4.1

This paper demonstrated how the concept of stylised facts can be used to specify and assess the contribution of simulation models. In particular, we explained how it can be used to answer two questions that simulation researchers are typically confronted with, if they wish to convince their more traditional peers of this method: Firstly, how exactly does simulation offer added value over the established methods? Secondly, given the increased degrees of freedom in modelling, how can the danger of arbitrariness be prevented? For this purpose, we elaborated the concept of stylised facts introduced by Kaldor and specified how stylised facts can be used to structure and focus model assessment.

4.2

These general epistemological considerations were subsequently applied to a discussion in the field of management accounting. Recently, there has been a lively debate about the

susceptibility of the Groves mechanism to collusion. From this discussion a collection of models emerged in which both traditional game theory and simulation are used. In a first step, we derived six stylised facts about the collusion phenomenon based on different empirical sources. In a second step, we specified and assessed the contribution of the different models to the explanation of these stylised facts. A major value added by simulation models is their ability to address in detail the enforcement process, an important aspect of the phenomenon. In our comparative analysis, arbitrariness and missing robustness have been located, but one simulation model successfully demonstrates that this problem is not inherent to the method. Thus, stylised facts can be used to increase transparency of model construction decisions and thereby enhance critical discussion.

4.3

This case illustrates how the stylised facts of a phenomenon can be used as a point of reference to structure and focus discussion about the contribution of simulation models. As such, simulation researchers do not have to refer only to abstract arguments about the advantages of the method to convince their peers. With reference to a set of stylised facts, ideally based on a broad set of empirical observations, the value added by using a simulation model can be specified precisely. At the same time, they can also be used to detect (and prevent) arbitrariness in model construction.

4.4

Reflecting on our analysis from a more general epistemological perspective, at least two further positive effects of using the concept can be identified: Firstly, the concept of stylised facts stimulates discussion on the characteristic features of a phenomenon[53] and thereby (re-)directs researchers' attention to the phenomenon to be explained. This inhibits a regularly observed tendency for modelling to become an end in itself. Secondly, a by-product of our comparative analysis was a "research landscape" of an existing discussion giving a systematic and condensed overview. This allows a visualisation of the current state of a research field and draws attention to blind spots, which again productively stimulates and directs further research.

4.5

In this paper it was merely possible to illustrate, by means of a methodological case study, the application and benefits of the stylised facts concept. However, the stylised facts derived can provide a starting point of a critical discussion in the expert community. Moreover, we hope that others make use of this approach in their research areas, allowing for a better assessment of the general applicability of the concept in the future. One crucial issue will probably be the varying availability of empirical data. In this respect one might observe a specific form of "theory-ladeness" concerning the availability of empirical data: The respective dominant theoretical approaches and models in a field probably have a strong influence on which aspects of a phenomenon are studied empirically. For example, economists might often only narrowly look at incentive aspects as a result. However, at the end this is not a problem of the concept of stylised facts itself. Instead the concept of stylised facts pin-points what still has to be done in a research area. In particular, it would produce a demand for more empirical investigations on a broader theoretical and methodological basis.

Appendix

Table 3.1: Extraction of findings about the stability of collusion in selected empirical studies

Study	Empirical Method and Context	Findings about Collusion	No.
Banerji/Meenakshi (2004):	*Case study* of auction	Even if there are about 50	F1

Buyer Collusion and Efficiency of Government Intervention in Wheat Markets in Northern India: an Asymmetric Structural Auctions Analysis	process of two wholesale wheat markets in northern India (supplemented with information from market committee records and personal interviews)	buyers in total, if there are few large buyers, there is collusion among the large buyers (cf. p. 248).	
Beckmann (2004): Art Auctions and Bidding Rings: Empirical Evidence from German Auction Data	*Statistical analysis* of a survey among German auctioneers and auction houses	New and different participants lower the risk of collusion in auctions (cf. p.135).	F2
Doree (2004): Collusion in the Dutch Construction Industry: an Industrial Organization Perspective	*Case study* of collusion in the Dutch construction industry based on interviews and a study of documents	Tendering procedures for individual projects support collusion even in markets with many participants because only a few firms are participating in the tendering (cf. p. 151).	F3
		Participants in collusive agreements try to avoid destabilisation because of market entry by collective acting (cf. p. 151).	F4
		Potential risk reduction in terms of pricing and fluctuations in the workload favours collusion (cf. p. 152).	F5
Feinberg/Husted (1993): An Experimental Test of Discount-Rate Effects on Collusive Behavior in Duopoly Markets	*Experimental study* of the influence on time preference on collusive duopoly equilibria	As the discount rate increases, the percentage of colluding subjects declines (cf. p. 157).	F6

Table 3.2: Extraction of findings about the stability of collusion in selected empirical studies

Study	Empirical Method and Context	Findings about Collusion	No.
Filson (2001): Market Power and Cartel Formation: Theory and an Empirical Test	*Statistical analysis* of the development of marketing orders (specific form of legal cartel) in	The empirical results for legal cartels* show that collusion is more likely when ...	

	the fresh market of the USA		
		... market demand is price inelastic (cf. p. 478).	F7
		... the affected firms have a large market share (cf. p. 478).	F8
		... crops have long cultivation times (cf. p. 478).	F9
		... the importance of the particular market for the sellers (in which collusive arrangements are proposed) is neither excessively low (the market is not really important to the seller) nor excessively high (there are no other potential markets). (cf. p. 478).	F10
		... the firms in question are regionally concentrated (cf. p. 478).	F11
		... there are large cooperatives in the market (cf. p. 478).	F12
		*Because the setting under investigation is a special case where collusion is legal (and therefore easily observable) and enforcement is limited in that a two third majority of all market participants is needed, Filson claims that his findings apply to the collusion phenomenon in general (cf. p. 466-467,477).	
Fraas/Greer (1977): Market Structure and Price Collusion: an Empirical Analysis	*Statistical analysis* of illegal explicit price collusion	The setting that stabilises collusion most strongly is a relatively small number of competing firms (cf. p. 42).	F13
		Market environments relatively free of complications such as product diversity or demand variability stabilise collusion (cf. p. 42).	F14
Huck/Normann/Oechssler (1999): Learning in Cournot Oligopoly: an Experiment	*Experiment* designed to test various learning theories in the context of a Cournot oligopoly	More information about the market (demand and cost functions) favours collusion, as players tend to use strategies that are the best responses to the behaviour of the other firms in the previous round (cf. p. 89,92).	F15
		More information about quantity supplied and profits of other firms in the market leads to increased competition as firms tend to increase quantity to imitate the rival firm with the highest profit (in comparison to aggregate information only) (cf. p. 89,92).	F16

Table 3.3: Extraction of findings about the stability of collusion in selected empirical studies

Study	Empirical Method and Context	Findings about Collusion	No.
Huck/Wieland/Normann (2001): Stackelberg Beats Cournot: on Collusion and Efficiency in Experimental Markets	*Experimental study* conducted with university students	"Collusion often breaks down in the last round." (p. 756)	F17
		In experiments with fixed pairs, partial collusion can be observed, but this does not apply to randomly matched pairs (cf. p. 750).	F18
		Sequential (in comparison to simultaneous) price setting destabilises collusion in duopolies (cf. p. 750).	F19
Huck/Normann/Oechssler (2004): Two are Few and Four are Many: Number Effects in Experimental Oligopolies	Meta analysis of 19 n-firm Cournot *experiments* completed by a series of actual *experiments* studying oligopolies with two, three, four, and five firms	"Previous [experimental, the authors] studies indicate that collusion sometimes occurs in duopolies, but is very rare in markets with more than two firms."(p. 440) The same tendency is observed in the current experiments (cf. p. 443).	F20
Kamerschen/Morgan (2004): Collusion Analysis of the Alabama Liquid Asphalt Markets	*Case study* from 1961-1978	Even with many participants (nine), collusion can occur if the product characteristics limit cross-regional shipping (cf. p. 681).	F21
		There is high concentration in the market: the four colluding firms have a market share of 99% (cf. p. 681).	F22
		At the time the collusion commenced, prices rose significantly although raw material prices remained unchanged (p. 691).	F23
Porter (1983): A Study of Cartel Stability: the Joint Executive Committee (JEC), 1880-1886	*Statistical analysis* of volume and price information about the JEC, a transportation cartel	When the cartel suspected that cheating had occurred, the cartel cut prices for a time so as to enforce behaviour consistent with the collusive agreement, and then returned to collusive price (cf. p. 302).	F24
		When collusion broke down, market prices declined substantially (cf. p. 312).	F25
Slade (1987): Interfirm Rivalry in Repeated Games: an Empirical Test of Tacit Collusion	*Statistical analysis* of the Vancouver retail gasoline market	After a price war, collusion was successfully re-established and maintained for a relatively long period (cf. p. 248).	F26

Acknowledgements

The authors thank the participants of the EPOS workshop 2004 in Koblenz/Germany, particularly Ulrich Frank, Scott Moss and Klaus G. Troitzsch, for helpful comments and discussions. Moreover, we thank Alexandra Matthes for critical feedback on a draft version of this paper and two anonymous referees for their insightful comments.

Notes

[1] With the expression "extended modelling possibilities" the authors primarily refer to the fewer constraints that exist for simulation models and the ability to describe medium-run dynamics compared to long-run equilibria in neoclassic economic models. For a detailed comparison of agent-based simulation models and neoclassic economic models see Chang and Harrington (in press).

[2] For an overview, see Lambert (2001).

[3] Holland and Miller (1991:367) state this problem quite straightforwardly: "Usually, there is only one way to be fully rational, but there are many ways to be less rational."

[4] Positive research, in contrast to normative research, aims at neutrally explaining empirical observations (cf. Blaug 1998).

[5] The term "generative mechanism" is used following Lawson (1989). For the purposes of this paper we only use it to distinguish two levels: A first level of results and second level that leads to these results (generative mechanism).

[6] The aim is not only validation, but also a deeper and wider understanding of a phenomenon (cf. Olsen 2004). For an in-depth discussion of this form of research see Creswell (2003). Modell (2005) gives an exemplary overview of accounting research triangulating between surveys and case studies.

[7] The importance of criticism in science is the central to Popper's understanding of science as exposed already in Popper (1934/1959). Other authors such as Peirce or Habermas have emphasised in particular the social and especially discursive processes enabling critical discussion leading to a convergence in beliefs of scientists (cf. Hands 2001:218-221).

[8] Using the standard epistemological distinction between the "context of discovery" and "context of justification", the ex ante use of the stylised-facts-concept can be located in the former. According to traditional philosophy of science, the emphasis of criticism has to be given to the latter (cf. Popper 1934/1959:27-34). Recently, this view has been criticised, because, if the complete set of scientific hypotheses to be tested has a certain bias, this bias cannot be eliminated by strict tests. Consequently, one should also take into account the context of discovery (cf. Okruhlik 1994/1998:200-205). Against this background, the concept of stylised facts provides a more systematic approach of hypothesis generation in the often neglected context of discovery (cf. Lawson 1989:67).

[9] The German philosopher of science Albert coined the term "Modell-Platonismus" (Model Platonism) for this type of modelling, in which model construction becomes an end in itself (cf. Albert 1963/1967).

[10] From the perspective of model evaluation however, the creating researcher's set of stylised facts should not be used again: an additional reference to the successful explanation of the stylised facts does not add any value, because they can no longer serve as a neutral point of

reference. It has to be assumed that the researcher already tailored his model to the stylised facts, so he intentionally created a fit to them. This will be desirable if the stylised facts are generally accepted, but this would render an explicit model evaluation concerning its "fit" unnecessary. In the other case, were the researcher stated his own view of the stylised facts, as Kaldor did, an explicit model evaluation would be biased.

[11] This also circumvents the caveats of "ad hocery" and "immunisation", since model results are given and evaluated relative to each other.

[12] Explanation is not reduced to the deduction of correct predictions from the model, but also means isolating the "mechanism" that leads to the results. Models explain by producing insights, they exhibit how the results are produced. With reference to simulation models see in particular Chang and Harrington (in press), for a general discussion of this point see also Meyer (2004:9-59).

[13] Nevertheless, the caveat from above applies: the combined use of stylised facts for model construction and evaluation reduces the value of this claim substantially. However, this does not change the structure of the argument (and suggests a division of labour), so that it can still provide an insightful illustration.

[14] More precisely, this means that after successful explanation of stylised facts, the necessity for testing shifts to the next level: the mechanisms a model uses to reproduce stylised facts now have to be tested as well, otherwise the model adds little more value than a thought experiment, "demonstrating only that a hypothesized process could be the source of some stylized fact" (Cohen 1999:375).

[15] This should be considered a methodological case study on how the concept of stylised facts can be used to answer the two questions addressed in the introduction. The selection of studies is representative in the sense, that two different modelling techniques have been used, which allows for a comparison, and it covers a whole debate. However, it is also possible to apply this approach to a discussion in which only one modelling technique has been used so far. Moreover, the discussion on the susceptibility of the Groves mechanism is quite new. Therefore, the number of contributions in this field is limited, thus manageable. Finally, there has been sufficient empirical research on the collusion phenomenon, which is important for the derivation of stylised facts.

[16] Compensation, by assumption, is tied to a division's absolute performance, because compensation for the relative profit per unit of capital invested would lead to massive underinvestment, since each manager then has the incentive to disregard less profitable projects to avoid dilution of the most profitable project.

[17] Other incentive mechanisms are for example the Weitzman scheme, the Osband-Reichelstein scheme, profit pooling or a bonus pool. For an overview, see Burgess and Metcalfe (1999).

[18] If M_2 reports truthfully while M_1 still overstates the returns, M_2 will receive a higher compensation in the case where exaggeration of M_1 for $s^r_{1.1}$ and $s^r_{1.2}$ is high enough to allocate all capital to M_1. Otherwise, he will receive the same compensation, but never less.

[19] See Baker (1999) for an overview.

[20] For example, the effect of changing participants can be observed far more easily in auctions (cf. Beckmann 2004) than in cartels. In general, the integration of different perspectives and research backgrounds also provides a means of coping with the problem of theory-ladeness: one must be aware that empirical studies do not state neutral "facts" but are already "theory laden", which can never be avoided. This point has already been made clearly by Popper:

"[O]bservations, and even more so observation statements and statements of experimental results, are always interpretations of the facts observed; they are *interpretations in the light of theories*." (Popper 1934/1959:107). "[W]e might say that these facts do not exist *as facts* before they are singled out from the continuum of events and pinned down by statements - the theories which describe them." (Popper 1946/1996:214). For a good discussion of the problem of "theory-ladeness", see also Hands (2001:91-93,102-109).

[21] Two objections can be made to such an inference. First, it is a specific type of inductive inference so that universality cannot be guaranteed. Secondly, this inference is based on the considerations and judgement of the authors of this paper. Ideally, the derivation of stylised facts would be accompanied by a discussion of a much larger group of experts. However, we tried to make the procedure as transparent as possible, which facilitates and hopefully encourages further discussion. Accordingly, this list should not be considered to be definitive, but as a first contribution to an ongoing discussion among experts in the field of collusion based on empirical research.

[22] Namely, Filson (2001), Doree (2004), Banerji and Meenakshi (2004), Kamerschen and Morgan (2004), Huck, Wieland and Normann (2001), Feinberg and Husted (1993), Fraas and Greer (1977), Slade (1987), Porter (1983), Beckmann (2004), Huck, Normann and Oechssler (2004) and Huck and Normann (1999). Our selection focussed mainly on recent refereed journal articles and was complemented by studies from journals with an accounted expertise for the area of industrial organization, where collusion is traditionally discussed. However, this is a very wide area of research and we do not claim to be exhaustive since this section has only an illustrative purpose. Moreover, the study of Waller and Bishop (1990) had to be excluded, because their experimental design does not allow for the investigation of the stability of collusion (cf. Waller and Bishop 1990:835).

[23] To give an overview of the origin of the results we name the methods used in the underlying analysis. Nevertheless, we face a given range of empirical studies and can therefore only state when a stylised fact is particularly well founded or particularly doubtful. The identified gaps can only be filled by further empirical studies from collusion researchers. As our analysis is based on 12 articles the chances for a complete triangulation of the stylised facts are limited.

[24] This implies that SF "Time Horizon" is not as well founded as SF "Group Size". Nevertheless, we state it, since it is a broad tendency we identified in the available empirical data. Future studies should strengthen the existing empirical evidence.

[25] F16 suggests that the stylised fact derived should be refined to account for exceptions like the reported one, e.g. by distinguishing different types of settings. Hence, a case study might be important in two ways, adding another empirical method and allowing a more detailed analysis of different factors that raise or reduce complications in the setting.

[26] We use "Benefits" as an overall perspective representing gains from collusion net of associated losses, risks etc. F10 reports the case where farmers that have no other markets to sell to, face a substantial risk, because any production exceeding the selling quota stipulated by the collusive agreement would turn into waste, potentially causing substantial losses (at least in terms of opportunity costs). In contrast, farmers that have at least one alternative market (beyond the reach of the collusive agreement) to sell their crops, do not face this risk, because they can offer any overproduction there (cf. Filson 2001:469-470).

[27] Because striving for benefits can be regarded as the driving force leading to collusion, SF "Benefits" is unique in the way that it is *assumed* to apply in most empirical studies: actors have an incentive to collude in order to gain the associated, substantial benefits. Therefore collusion is expected or observed. Statistical analyses and case studies cannot vary the potential benefits in a given setting and would therefore have to engage in a complicated comparison across settings. Furthermore, the benefits to be gained from collusion are difficult to determine, often

involving very noisy indicators. Experiments seem better suited to the purpose, but to our knowledge, there are no experiments addressing the influence of varying benefits on the stability of collusion. These aspects might explain the indirect empirical evidence we refer to when deriving SF "Benefits".

[28] More empirical evidence of enforcement issues would be desirable. In particular, case studies and experiments would complement the existing statistical analyses.

[29] The problem of theory-ladeness is of particular importance in this case (see note 20).

[30] From the perspective of collusion research, these stylised facts represent an attempt to outline the broad tendencies relating to the stability of collusion that emerge in the available empirical studies. Although our set of empirical studies is not exhaustive due to the illustrative purpose of this section, we are confident that the stylised facts identified represent a solid basis for a discussion of the characteristic features of the collusion phenomenon and its representation in theoretical models.

[31] Interestingly, they state this result without reference to any empirical evidence.

[32] Compared to a finite game with a fixed number of rounds, the uncertainty about the end does not allow for backward induction. The backward induction argument is as follows: If it is rational to defect in the last round, rational players anticipate this in the second last round. This will be anticipated in the round before and so one. Therefore, by using backward induction in a finite game, it can be shown that collusive agreements are not stable.

[33] Axelrod (1984:20) describes "tit for tat" as "the strategy which cooperates in the first move and then does whatever the other player did for the previous move."

[34] This concept has been developed to address credibility in games with the intention to overcome the restrictions posed by the subgame perfection solution concept. The original paper is Kreps and Wilson (1982). For a good introduction, see Gardner (1995:239-254).

[35] Furthermore, his analysis refers to the more general class of relative incentive schemes. In such a scheme, the compensation of a manager does not depend solely on his absolute achievements, but is also influenced by the achievements of others. Krapp shows that the collusion problem for relative incentive schemes has the same structure as for the Groves mechanism and therefore the results of his model can be applied to the Groves mechanism (Krapp 2000:262-266). In Krapp's model the collusion parameter is effort, but this does not change the structure of the arguments.

[36] For a complete list and discussion of the strategies used in the tournament, see Krapp (2000:268).

[37] In a round robin, each participant faces all the other participants individually playing a specified number of rounds, in Krapp's model 1000 (cf. Krapp 2000:269).

[38] This aspect conforms well to a more general discussion of the performance of the "tit for tat" strategy and particularly Axelrod's own work on this subject (cf. Wu and Axelrod 1995), although Krapp does not make this reference.

[39] The reasons for this can be found by considering in detail the process of the respective interactions: the strategy "reputation" suffers from large losses in the early interactions with an unfriendly player, which has a strong effect on performance subject to high discount rates (cf. Krapp 2000:263).

[40] For description of these characteristics, see Deliano (2000:49-51). Based on these variables,

Deliano (2000:54-61) develops a quite complex structural model describing how the respective behaviour is determined. For example, the managers' decision whether to report truthfully or not depends on his impatience (modelled as a probability). A higher impatience leads with a higher probability to a truthful report (cf. Deliano 2000:54-55).

[41] Cf. Deliano (2000:114,119,124), e.g. comparing the results for two and four involved managers the percentage of managers with exaggerated reports decreases from 57,7% to 34,1% (cf. Deliano 2000:131).

[42] Unfortunately, he does not even state which kind of observable pattern should be interpreted as collusive behaviour. He defines measures for "individual" and "collective" behaviour, but does not specify their meaning concerning collusion (cf. Deliano 2000:98-100).

[43] For a further elaboration on this conjecture, see e.g. Kreps (1990/1991:95-107).

[44] Although the SF "Benefits" remained unaddressed in the context of the problem under examination, several simulation models like Nachbar (1992), Fogel (1993), Kirchkamp (2000) and Mueller (1987) are analysing the impact of variations of the payoff tables for the PD on the effectiveness of specific strategies.

[45] Carley (1992) gives a good example on how the effect of exchanging agents can be analysed using a simulation model.

[46] Emphasis should be placed on "established". There is an extensive discussion on the empirical validity of game theory, but, for the purpose of this section, we focus on arbitrariness and a lack of robustness in models. As already noted, this cannot be a substitute for empirical testing.

[47] For details on the folk theorem, see Fudenberg and Maskin (1986).

[48] This is another indication that the result with respect to SF "Time Horizon" in the *threat* model can be considered a "by-product" of technical requirements.

[49] This information would be of particular importance because of possible path-dependency effects, e.g. for the "trigger" strategy, high variations in results can be expected, as the outcome of the strategy is very sensitive to the point in time when (permanent) punishment is triggered (cf. Krapp 2000:268).

[50] In addition, the chosen characteristics do not seem to be exhaustive as other characteristics like envy could also be supposed to influence the behaviour of managers.

[51] To give an example, with respect to the number of past and previous periods taken into account by an agent, he informs the reader that the value four is chosen. He mentions only a pretest, according to which an increase to eight does not produce "significant changes" (cf. Deliano 2000:85-86).

[52] This is not *necessarily* a problem inherent to the approach taken by Deliano. For better constructed models also addressing the issue of endogenous generation of strategies, see for example Macy (1996) or Axelrod (1997).

[53] In this respect, our derivation of stylised facts is only an initial step in an ongoing, critical discussion among the experts in a field.

References

ALBERT, H (1963/1967). Modell-Platonismus: der neo-klassische Stil des ökonomischen Denkens in kritischer Beleuchtung. In Topitsch, E (Ed.): *Logik der Sozialwissenschaften*. 4th ed. Köln: Kiepenheuer & Witsch, pp. 406-434.

AXELROD, R (1984). *The Evolution of Cooperation.* New York: Basic Books.

AXELROD, R (1997). Evolving new Strategies: the Evolution of Strategies in the Iterated Prisoner's Dilemma. In Axelrod, R (Ed.): *The Complexity of Cooperation : Agent-Based Models of Competition and Collaboration*. New Jersey: Princeton University Press, pp. 10-32.

BAKER, J (1999). Developments in Antitrust Economics*. The Journal of Economic Perspectives*, Vol. 13, No. 1, pp. 181-194.

BANERJI, A and MEENAKSHI, J V (2004). Buyer Collusion and Efficiency of Government Intervention in Wheat Markets in Northern India: an Asymmetric Structural Auctions Analysis*. American Journal of Agricultural Economics*, Vol. 86, No. 1, pp. 236-253.

BECKMANN, M (2004). Art Auctions and Bidding Rings: Empirical Evidence from German Auction Data*. Journal of Cultural Economics*, Vol. 28, No. 2, pp. 125-141.

BLAUG, M (1998). The Positive-Normative Distinction. In Davis, J B, Hands, D W and Mäki, U (Ed.): *The Handbook of Economic Methodology*. Cheltenham: Elgar, pp. 370-374.

BOLAND, L A (1987/1994). Stylized Facts. In Eatwell, J, Milgate, M and Newman, P (Ed.): *The New Palgrave Dictionary of Economics, Vol. 4*. repr. with corrections. London: Macmillan Press, pp. 535-536.

BUDDE, J, GÖX, R and LUHMER, A (1998). Absprachen beim Groves-Mechanismus: eine spieltheoretische Analyse*. Zeitschrift für Betriebswirtschaftliche Forschung*, Vol. 50, No. 1, pp. 3-20.

BURGESS, S and METCALFE, P (1999). *Incentives in Organisations: a Selective Overview of the Literature with Application to the Public Sector*. (CMPO) Working Paper: 99/016. University of Bristol, Bristol.

CARLEY, K (1992). Organizational Learning and Personnel Turnover*. Organizational Science*, Vol. 3, No. 1, pp. 20-46.

CHANG, M-H and HARRINGTON, J E (in press). Agent-Based Computational Economics. In Judd, K L and Tesfatsion, L (Ed.): *Agent-Based Models of Organisations: Handbook of Computational Economics II*. Amsterdam: North-Holland.

CHE, Y-K and KIM, J (2004). *Collusion-Proof Implementation of Optimal Mechanisms*. Working Paper. University of Wisconsin Madison.

COHEN, M (1999). Commentary on the Organization Science Special Issue on Complexity*. Organization Science*, Vol. 10, No. 3, pp. 373-376.

CRESWELL, J W (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches.* 2nd ed. Thousand Oaks: Sage Publications.

DELIANO, M (2000). *Computersimulation als Instrument zur Analyse von Anreizsystemen: eine Untersuchung am Beispiel des Groves-Mechanismus.* Munich: Privat Publishing House.

DOREE, A G (2004). Collusion in the Dutch Construction Industry: an Industrial Organization Perspective*. Building Research & Information*, Vol. 32, No. 2, pp. 146-156.

FEINBERG, R M and HUSTED, T A (1993). An Experimental Test of Discount-Rate Effects on Collusive Behavior in Duopoly Markets*. Journal of Industrial Economics*, Vol. 41, No. 2, pp. 153-

160.

FILSON, D (2001). Market Power and Cartel Formation: Theory and an Empirical Test. *Journal of Law and Economics*, Vol. 44, No. 2, pp. 465-480.

FOGEL (1993). Evolving Behaviors in the Iterated Prisoner's Dilemma. *Evolutionary Computation*, Vol. 1, No. 1, pp. 77-97.

FRAAS, A G and GREER, D F (1977). Market Structure and Price Collusion: an Empirical Analysis. *Journal of Industrial Economics*, Vol. 26, No. 1, pp. 21-44.

FRIEDMAN, M (1953/1989). The Methodology of Positive Economics. In Friedman, M (Ed.): *Essays in Positive Economics*. 15th ed. Chicago: Univ. of Chicago Press, pp. 3-43.

FUDENBERG, D and MASKIN, E (1986). The Folk Theorem in Repeated Games with Discounting or with Incomplete Information. *Econometrica*, Vol. 54, No. 3, pp. 533-554.

GARDNER, R (1995). *Games for Business and Economics.* New York: Wiley.

GROVES, T (1973). Incentives in Teams. *Econometrica*, Vol. 41, No. 4, pp. 617-631.

GROVES, T and LOEB, M (1979). Incentives in a Divisionalized Firm. *Management Science*, Vol. 25, No. 3, pp. 221-230.

HANDS, D W (2001). *Reflection Without Rules-Economic Methodology and Contemporary Science Theory.* Cambridge: Cambridge University Press.

HAUSMAN, D M (1994/1995). Why Look Under the Hood? In Hausman, D M (Ed.): *The Philosophy of Economics: an Anthology*. 2. ed., reprint. Cambridge: Cambridge Univ. Press, pp. 217-222.

HOLLAND, J H and MILLER, J H (1991). Artificial Adaptive Agents in Economic Theory. *American Economic Review*, Vol. 81, No. 2, pp. 365-370.

HUCK, S and NORMANN, H-T (1999). Learning in Cournot Oligopoly: an Experiment. *Economic Journal*, Vol. 109, No. 454, pp. 80-95.

HUCK, S, NORMANN, H-T and OECHSSLER, J R (2004). Two are Few and Four are Many: Number Effects in Experimental Oligopolies. *Journal of Economic Behavior & Organization*, Vol. 53, No. 4, pp. 435-446.

HUCK, S, WIELAND, M and NORMANN, H-T (2001). Stackelberg Beats Cournot: on Collusion and Efficiency in Experimental Markets. *The Economic Journal*, Vol. 111, No. 474, pp. 749-765.

KALDOR, N (1961/1968). Capital Accumulation and Economic Growth. In Lutz, F A and Hague, D C (Ed.): *The Theory of Capital*. Reprint. London: Macmillan, pp. 177-222.

KAMERSCHEN, D R and MORGAN, J E J (2004). Collusion Analysis of the Alabama Liquid Asphalt Markets. *Applied Economics*, Vol. 7, No. 36, pp. 673-693.

KIRCHKAMP, O (2000). Spatial Evolution of Automata in the Prisoner's Dilemma. *Journal of Economic Behavior & Organization*, Vol. 43, No. 2, pp. 239.

KRAPP, M (2000). Relative Leistungsbewertung im dynamischen Kontext: eine Analyse der Kollusionsproblematik bei wiederholter Delegation. *Zeitschrift für Betriebswirtschaftliche Forschung*, Vol. 52, No. 5, pp. 257-277.

KREPS, D M (1990/1991). *Game Theory and Economic Modelling.* Oxford: Clarendon Press.

KREPS, D M and WILSON, R (1982). Sequential Equilibria. *Econometrica*, Vol. 50, No. 4, pp. 863-894.

KUNZ, A H and PFEIFFER, T (1999). Investitionsbudgetierung und implizite Verträge: Wie resistent ist der Groves-Mechanismus bei dynamischer Interaktion? *Zeitschrift für Betriebswirtschaftliche Forschung*, Vol. 53, No. 8, pp. 203-222.

LAMBERT, R A (2001). Contracting Theory and Accounting. *Journal of Accounting & Economics*, Vol. 32, No. 1-3, pp. 3-87.

LAWSON, T (1989). Abstraction, Tendencies and Stylised Facts: a Realist Approach to Economic Analysis. *Cambridge Journal of Economics*, Vol. 13, No. 1, pp. 59-78.

LINDENBERG, S (1992). The Method of Decreasing Abstraction. In Coleman, J S and Fararo, T J (Ed.): *Rational Choice Theory: Advocacy and Critique*. Newsbury Park: Sage Publications, pp. 3-20.

MACY, M (1996). Natural Selection and Social Learning in Prisoner's Dilemma. In Liebrand, W B G and Messick, D M (Ed.): *Frontiers in Social Dilemmas Research*. Berlin, pp. 235-266.

MÄKI, U (1998). Realism. In Davis, J B, Hands, D W and Mäki, U (Ed.): *The Handbook of Economic Methodology*. Cheltenham: Elgar, pp. 404-409.

MARCET, A and NICOLINI, J P (2003). Recurrent Hyperinflations and Learning. *American Economic Review*, Vol. 93, No. 5, pp. 1476-1498.

MEYER, M (2004). *Prinzipale, Agenten und ökonomische Methode. Von einseitiger Steuerung zu wechselseitiger Abstimmung.* Tübingen: Mohr.

MILGROM, P R and ROBERTS, J (1992). *Economics, Organization and Management.* Englewood Cliffs, New Jersy.: Prentice-Hall.

MODELL, S (2005). Triangulation between Case Study and Survey Methods in Management Accounting Research: an Assessment of Validity Implications. *Management Accounting Research*, Vol. 16, No. 2, pp. 231-254.

MUELLER, U (1987). Optimal Retaliation for Optimal Cooperation. *Journal of Conflict Resolution*, Vol. 31, No. 4, pp. 692-724.

NACHBAR, J H (1992). Evolution in the Finitely Repeated Prisoner's Dilemma. *Journal of Economic Behavior and Organization*, Vol. 19, No. 3, pp. 307-326.

OKRUHLIK, K (1994/1998). Gender and the Biological Sciences. In Curd, M and Cover, J A (Ed.): *Philosophy of Science : the Central Issues*. New York: Norton, pp. 192-208.

OLSEN, W (2004). Triangulation in Social Research: Qualitative and Quantitative Methods Can Really Be Mixed. *Developments in Sociology*, Vol. 20, No. 1, pp. 103-121.

POPPER, K R (1934/1959). *The Logic of Scientific Discovery.* New York: Basic Books.

POPPER, K R (1946/1996). Why are the Calculi of Logic and Arithmetic Applicable to Reality? In Popper, K R (Ed.): *Conjectures and Refutations: the Growth of Scientific Knowledge*. 5th ed. (rev.). London: Routledge, pp. 201-214.

PORTER, R H (1983). A Study of Cartel Stability: the Joint Executive Committee, 1880-1886. *Bell Journal of Economics*, Vol. 14, No. 2, pp. 301-314.

ROBERTS, K (1987/1994). Collusion. In Eatwell, J, Milgate, M and Newman, P (Ed.): *The New Palgrave Dictionary of Economics, Vol. 4*. repr. with corrections. London: Macmillan Press, pp. 482-484.

SLADE, M E (1987). Interfirm Rivalry in Repeated Games: an Empirical Test of Tacit Collusion. *Journal of Industrial Economics*, Vol. 35, No. 4, pp. 499-516.

SOLOW, R M (1969/1988). *Growth Theory: an Exposition.* Paperback ed. New York, NY: Oxford Univ. Press.

WALLER, W S and BISHOP, R A (1990). An Experimental Study of Incentive Pay Schemes, Communication, and Intrafirm Resource Allocation. *Accounting Review*, Vol. 65, No. 4, pp. 812-836.

WU, J and AXELROD, R (1995). How to Cope with Noise in the Iterated Prisoner's Dilemma. *Journal of Conflict Resolution*, Vol. 39, No. 1, pp. 183-189.

Return to Contents of this issue